

Durham E-Theses

Spatiotemporal integration of spatial frequencies

Eugene McSorley

How to cite:

McSorley, Eugene (1999) Spatiotemporal integration of spatial frequencies. Doctoral thesis, Durham University.

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a <https://etheses.durham.ac.uk/id/eprint/4615/> is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full Durham E-Theses policy](#) for further details.

Author: Eugene McSorley

Abstract: Spatiotemporal integration of spatial frequencies

Ph. D. 1999

Spatial frequency processing delays suggest that their integration occurs from coarse to fine. However it may be the case that integration is flexible. The first aim of this thesis was to examine the integration of spatial frequencies and the second was to examine the delays involved in initial spatial frequency processing.

A coarse to fine and a fine to coarse presentation order of filtered natural images were shown. The former is judged to be of a higher quality (experiment 1) but neither show any increase in incorrect detection reports of a fullbandwidth image (experiment 2, 3 and 4). When the first three harmonics of a square-wave were shown progressively from coarse to fine and fine to coarse the former leads to more incorrect detection reports of a square-wave (experiment 5) but the latter is favoured in judgements of square-wave resemblance (experiment 6, 8, 9a and 9b), a masking (experiment 10) and a Vernier acuity task (experiment 12). These results suggest that the integration of spatial frequencies is flexible.

Despite this, the integration of spatial frequencies must occur from coarse to fine because of the delays involved in the processing of spatial frequencies. In Chapter 6 (experiments 13 – 18) it was found that two widely different spatial frequencies were judged to have simultaneous onset when they were in fact simultaneous.

The results of this thesis provided evidence that the integration of spatial frequencies is flexible and dependent upon the information to be derived by any particular processing. The lack of evidence for spatial frequency processing delays suggest that a flexible integration of spatial frequencies may be possible when spatial frequencies are presented simultaneously.

Contents

Abstract	i
Contents	iii
List of Figures	viii
List of Tables	xvi
Declaration	xxi
Statement of copyright	xxi
Acknowledgements	xxii
Dedication	xxiii
1. The spatio-temporal integration of spatial frequencies	1
1.1 Introduction	1
1.2 Integration of spatial frequencies	3
1.2.1 Spatial frequency channels	3
1.2.2 Spatial frequency integration	14
1.2.3 Summary	25
1.3 The temporal form of integration	26
1.3.1 Computational advantages for particular temporal forms of spatial frequency integration	27
1.3.2 Evidence for particular temporal forms of spatial frequency integration	30

1.3.3	The temporal form of the integration mechanism	56
1.3.4	Conclusions	66
1.4	Overall summary	67
1.5	The aims and scope of the thesis	69
2.	Spatial frequency integration and natural images	73
	Summary	73
2.1	Introduction	74
2.1.1	Interpretation of image sequences	75
2.1.2	Use of natural images in empirical research	75
2.1.3	Preview of the chapter	77
2.2	Experiment 1: natural image quality	78
2.3	Experiment 2: natural image detection	92
2.4	Experiment 3: natural image detection - less discriminable images	103
2.5	Experiment 4: natural image detection - images and power spectra	108
2.6	General Discussion	124
3.	Spatial frequency integration and square-waves	131
	Summary	131
3.1	Introduction	132
3.1.1	Advantages of Gabor patches	133

3.1.2	Square-waves as stimuli	134
3.1.3	Task considerations	135
3.1.4	Aims of experiments	138
3.1.5	Preview of chapter	140
3.2	Experiment 5: square-wave detection	141
3.3	Experiment 6: square-wave likeness	151
3.4	General discussion	157
4.	Spatial frequency interactions	160
	Summary	160
4.1	Introduction	161
4.1.1	Preview of Chapter	162
4.2	Experiment 7: natural image likeness	166
4.3	Experiment 8: square-wave likeness with two harmonics	168
4.4	Experiment 9: last frame	171
4.5	Experiment 10: masking	176
4.6	General discussion	184
5.	Spatial frequency interactions and Vernier acuity	189
	Summary	189
5.1	Introduction	190

5.1.2 Vernier acuity	192
5.1.3 Preview of chapter	194
5.2 Experiment 11a: Vernier acuity of low spatial frequencies	195
5.2.2 Experiment 11b: Vernier acuity of compounds	202
5.3 Experiment 12: Vernier acuity of High spatial frequencies	206
5.4 General discussion	212
6. Spatial frequency processing delays	216
Summary	216
6.1 Introduction	217
6.1.1 Preview of Chapter	220
6.2 Experiment 13: Parker & Dutch (1987) replication	220
6.3 Experiment 14: Temporal order judgements of LED and spatial frequencies at different durations	225
6.4 Experiment 15: Temporal order judgements of spatial frequencies equally above contrast threshold	228
6.5 Experiment 16: Temporal order judgements of spatial frequencies equally above contrast threshold – cycled	237
6.6 Experiment 17: Temporal order judgements of equally detectable spatial frequencies	242
6.7 Experiment 18: Temporal order judgements of equally detectable spatial frequencies – cycled	250
6.8 General discussion	256

7. General Discussion	259
7.1 Overview of thesis findings	259
7.1.1 The temporal nature of spatial frequency integration	260
7.1.2 Spatial frequency processing delays	269
7.2 Future avenues of research	271
7.3 Conclusion	275
Notes	276
References	280

List of Figures

- Figure 2.1: Images used in experiment 1. Top row = fullbandwidth image; 82
Middle = LP 0.5 and HP 0.5; Bottom = LP 1 and HP 1.
- Figure 2.2: Power spectra of the images used in experiment 1. 83
- Figure 2.3: showing the judged quality of the filtered and fullbandwidth images. 84
The vertical bars are standard error bars.
- Figure 2.4: Mean judged quality of the image triplets. The triplets are shown in 87
the legend from top to bottom and these correspond, from left to right, to the
bars shown on the graph. The triplets are grouped by spatial frequency content
regardless of order.
- Figure 2.5: Images used in experiment 2. Top row = LP 2 and HP 2; Bottom = 95
LP 4 and HP 4. For fullbandwidth image see figure 2.1.
- Figure 2.6: Power spectra of the images used in experiment 2. 96
- Figure 2.7: Mean number of fullbandwidth presented responses as a function of 97
image type. The vertical bars are the standard errors associated with each mean.

Figure 2.8: Mean number of fullbandwidth present responses out of 20 maximum as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that triplets containing the same images are together regardless of the order in which they are shown. The vertical bars are standard error bars associated with each mean.	100
Figure 2.9: Mean number of fullbandwidth presented responses as a function of image type. The vertical bars are the standard errors of the associated means.	105
Figure 2.10: Mean number of fullbandwidth present responses out of 20 maximum as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that triplets containing the same images are together regardless of the order in which they are shown. The vertical bars are standard error bars associated with each mean.	106
Figure 2.11a: Showing a picture of a cityscape (New York). The power spectra of this root image had an α value of 0.93. Top row = fullbandwidth; Middle = low pass cut off at 1.92 and high pass cut off at 2.15 cpd; Bottom = low pass cut off at 4.2 and high pass cut off at 4.62 cpd.	112
Figure 2.11b: Showing a picture of a building (University of Stirling). The power spectra of this root image had an α value of 2.13.	113

Figure 2.11c: Showing a picture of a face. The power spectra of this root image had an α value of 1.16.	114
Figure 2.11d: Showing a picture of a face. The power spectra of this root image had an α value of 1.5.	115
Figure 2.12: Power spectra of the images used in experiment 4. a) shallow sloped power spectra for cityscape scene; b) Steep sloped power spectra for building scene; c) Shallow sloped power spectra for face with shallow sloped power spectra; and d) Steep sloped power spectra for face with steep sloped power spectra. The legend shows the key to the power spectra for the fullbandwidth and filtered images for all images.	116
Figure 2.13: The four graphs show the results from the four images employed: the top row show the results for natural scenes with shallow and steep sloped power spectra respectively from left to right; the bottom row shows the results for facial stimuli, again with a shallow and steep sloped power spectra. The results for each image shows the mean number of fullbandwidth present responses out of 20 maximum as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that triplets containing the same images are together regardless of the order in which they are shown. The vertical bars are standard error bars associated with each mean.	118
Figure 3.1: shows the stimuli used in the experiments presented in this chapter.	143

The left hand column shows, running from top to bottom, the square-wave which has been convolved with a Gaussian, the fundamental frequency F , the $3F$ spatial frequency and the $5F$ spatial frequency. The right hand column shows the corresponding power spectra.

Figure 3.2: Shows the mean number of square-wave presented responses as a function of image type. Vertical bars are the standard errors. 145

Figure 3.3: shows the mean number of square-wave present responses as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that the two orderings of triplets containing the same images are paired together. Vertical bars are the standard errors. 148

Figure 3.4: showing the percentage of square-wave likeness responses as a function of spatial frequency triplet temporal order. Vertical bars are the standard error of the mean. 153

Figure 4.1: the percentage of natural image like judgements as a function of the temporal order of spatial frequencies. The Y vertical bars are the standard error bars. 167

Figure 4.2: shows the percentage of trials on which each temporal order of spatial frequencies was judged to more resemble a square-wave. The vertical bars are the standard error bars. 170

Figure 4.3: The average percentage square-wave likeness judgements as a function of presentation order. The vertical bars are the standard error bars. 174

Figure 4.4: The average percentage square-wave likeness judgements as a function of presentation order. The vertical bars are the standard error bars. 175

Figure 4.5: shows the contrast thresholds (in percent) for a 3F when presented alone, preceded by a F (F/ 3F) and succeeded by a F (3F / F). The vertical bars are the standard error bars. 182

Figure 5.1: Depicts the increasing sharpness of the boundary between progressively darker and progressively lighter portions of a sinusoid. It can be seen that the F + 3F compound has a sharper boundary than that found for the F or the 3F alone. This suggests that if the two individual spatial frequencies integrated then the Vernier acuity threshold for a horizontal displacement of two gratings would be lower than for either spatial frequency alone. 195

Figure 5.2: Vernier stimuli and their power spectra. The left hand column shows a detail of the power spectra of the Vernier stimuli shown in the right hand column. The images shown at the top are of a 2 cpd spatial frequency, while below this is shown the 6 cpd spatial frequency. Only a section of these power spectra is shown as the details were obscured in the full power spectra images due to their large scale. It can be seen that the predominant energy in the Vernier stimuli are at their respective spatial frequencies, however energy at all 197

spatial frequencies is present in the images because of the sharp edges introduced by the gap between the two grating halves.

Figure 5.3: The compound Vernier stimulus and its power spectra. The left hand column shows a detail of the power spectra of the compound Vernier stimulus shown in the right hand column. Only a section of these power spectra is shown as the details were obscured in the full power spectra images due to their large scale. It can be seen that the predominant energy in the Vernier stimuli are at their respective spatial frequencies, however energy at all spatial frequencies is present in the images because of the sharp edges introduced by the gap between the two grating halves. 203

Figure 5.4: Vernier stimuli and their power spectra. The left hand column shows a detail of the power spectra of the Vernier stimuli shown in the right hand column. The images shown at the top are of a 5 cpd spatial frequency, while below this is shown the 15 cpd spatial frequency. Only a section of these power spectra is shown as the details were obscured in the full power spectra images due to their large scale. It can be seen that the predominant energy in the Vernier stimuli are at their respective spatial frequencies, however energy at all spatial frequencies is present in the images because of the sharp edges introduced by the gap between the two grating halves. 208

Figure 6.1: shows the estimated onset of a grating of various spatial frequencies in msec, using a LED as a matching stimulus. The actual onset of the gratings is represented as occurring at zero. The data points denoted by black squares 223

represent those conditions where the LED and spatial frequency gratings had an equal duration of 40 msec. The black diamonds denote conditions where the LED was shown for 40 msec and the spatial frequency grating was shown for 400 msec.

Figure 6.2: shows the estimated onset of a grating of various spatial frequencies in msec, using a LED as a matching stimulus. The actual onset of the gratings is represented as occurring at zero. 226

Figure 6.3: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject. 234

Figure 6.4: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject. 240

Figure 6.5 : Showing detectability level on the ordinate and contrast level on the abscissa. If detectability level a (ordinate) is chosen as the level for stimulus equality then contrast for the stimulus in this case equals a (abscissa). However, a short shift down the ordinate produces a much larger shift on the abscissa thus suggesting that any small variability in the estimation of the detectability level produces a much larger error in the choice of contrast level. A point further down the psychometric function detectability level c produces contrast level c an equal amount of variability as in the case of a produces no change in the choice of contrast level. 247

Figure 6.6: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject. 248

Figure 6.7: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject. 255

List of Tables

Table 1.1: Summary of the response characteristics of the P and M pathways in the human and primate visual system (Derived from Kaplan et al ,1990; Lennie, 1993).	41
Table 2.1: Shows the ratings given for image triplets. The first column of data shows the mean rating given to the single images which make up the sequences as taken from experiment 1a. The second column of data shows the actual triplet image ratings found in this experiment.	89
Table 2.2: Shows the total percentage of trials on which the filtered images were not discriminated from their fullbandwidth versions in this experiment and Experiment 4 of Parker et al (1992). The percentages from Parker et al (1992) were not all explicitly reported and are thus calculated from the graph which they presented. The image names not in brackets are those used in the current experiment while those within brackets are those used by Parker et al (1992) for the corresponding image.	98
Table 3.1: Showing the predicted neural delays involved in the processing of spatial frequencies in msec following a conservative estimate of neural delay of	155

5 msec per octave (Parker & Dutch, 1987).

Table 3.2: Showing the predicted neural delays involved in the processing of spatial frequencies in msec following a liberal estimate of neural delay of 21 msec per octave (Breitmeyer, 1975). 156

Table 4.1: Showing the pooled mean contrast thresholds and their standard deviations by subject. Std dev = Standard deviation. 181

Table 5.1: shows the Vernier acuity thresholds for each of the 6 subjects. The values are in seconds. 199

Table 5.2: shows the Vernier acuity thresholds for each of the 6 subjects. The values are in seconds. 204

Table 5.3: shows the Vernier acuity thresholds for each of the 6 subjects. The values are in seconds. 210

Table 6.1: shows the contrast threshold (in percent) of six subjects for a 0.5 and a 16 cpd spatial frequency Gabor patch in percent. 233

Table 6.2: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = 235

negative confidence interval; Positive conf = positive confidence interval).

Table 6.3: Contrast thresholds (in percent) for the 0.5 and 16 cpd Gabor patches for six subjects. 238

Table 6.4: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval). 240

Table 6.5: the contrast thresholds (in percent) for the 0.5 and the 16 cpd Gabor patch 244

Table 6.6: Four contrast values (in percent) for the two spatial frequencies to be used in the temporal order judgements of experiment 19c. The contrast thresholds for both spatial frequencies were used and from this contrast value three other contrast levels were chosen. The contrast increased in log steps of 0.1 or 0.15 for the 0.5 cpd spatial frequency and the 16 cpd spatial frequency respectively. This enabled the detectability of a good spread of contrast values to be assessed. 245

Table 6.7: the detectability (in percent) of the 0.5 cpd and 16 cpd spatial frequencies as a function of their contrast in percent. 246

Table 6.8: shows the 90% detectability contrasts (in percent) for the 0.5 and 16 cpd spatial frequencies.	247
Table 6.9: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval).	248
Table 6.10: Contrast thresholds (in percent) of the 0.5 cpd and 16 cpd spatial frequencies for three subjects.	252
Table 6.11: Four contrast values (in percent) for the two spatial frequencies to be used in the temporal order judgements of experiment 20c. The contrast thresholds for both spatial frequencies were used and from this contrast value three other contrast levels were chosen. The contrast increased in log steps of 0.1 or 0.15 for the 0.5 cpd spatial frequency and the 16 cpd spatial frequency respectively. This enabled the detectability of a good spread of contrast values to be assessed.	253
Table 6.12: the detectability (in percent) of the 0.5 cpd and 16 cpd spatial frequencies as a function of their contrast.	253
Table 6.13: shows the 90% detectability contrasts (in percent) for the 0.5 and 16 cpd spatial frequencies.	254

Table 6.14: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval).

Declaration

The research contained in this thesis was carried out by the author between October 1996 and April 1999 whilst a postgraduate in the Department of Psychology at the University of Durham. None of the work contained in this thesis has been submitted in any form in candidature for any other degree.

Statement of Copyright

The copyright of this thesis rests with the author. No quotation from it should be published without their prior written consent and information derived from it should be acknowledged.

Acknowledgements

I would like to extend thanks to John Findlay for supervising this thesis. I also want to thank Iain Gilchrist for his patience and help. I would also like to thank the following people for their help throughout the undertaking of this thesis: Val Brown, Rachel Genn, Charlie Heywood, Suzanne Higgs, Ala Hola, Bob Kentridge, Simon Liversedge, Penny Machin, Bob Metcalf, Anne-Sophie Molkenboer, Fiona Newell, Matt O'Callaghan, Jamie Pearson, Paul Thompson, Roger Watt and all those who took part in my, sometimes gruelling, experiments. Thanks also go to my family for supporting me.

Finally I want to thank my wife, Jan, without the understanding of whom I doubt I would have kept my sanity.

This is dedicated to Jan

Chapter 1

The spatio-temporal integration of spatial frequencies

1.1 Introduction

This thesis concerns the initial temporal processing of spatial frequencies and its subsequent consequences for their integration, historically referred to as microgenesis. Microgenesis refers to the initial temporal period between the onset of a stimulus and a resultant integrated, stabilised cognitive form (e.g. thought or percept) (Werner, 1956). In the case of perception, Microgenesis is the process of the formation of an unitary integrated percept (Flavell & Draguns, 1957). It has been suggested that Microgenesis might progress from a blurred whole percept increasing in sharpness or fine detail as the duration of the stimulus increases (Sander & Jinuma, 1928; Undeutsch, 1942; Mantell, 1936; Morschner, 1940; all cited in Flavell & Draguns, 1957). In this thesis the notion of the temporal form of initial percept formation is examined in terms of spatial frequency content.

It has been well documented that visual information is initially processed via channels responsive to different spatial frequencies (Campbell & Robson, 1968; Blakemore & Campbell, 1969; Graham & Nachmias, 1971; DeValois & DeValois, 1988; Graham, 1989). Early experiments examining spatial frequency processing suggested the existence of band - limited independent spatial frequency channels

(Campbell & Robson, 1968; Blakemore & Campbell, 1969; Stromeyer & Julesz, 1972; Graham, 1980). Although the existence of independent channels has been widely accepted, it has not always been found. The results of a number of experiments suggest the existence of interactions across channels, where the activity in one channel affects the activity in another channel responsive to a different spatial frequency (Nachmias, Sansbury, Vassilev & Weber, 1973; Henning, Hertz & Broadbent, 1975; Nachmias & Weber, 1975; Tolhurst & Barfield, 1978), and others where the activity in one channel affected the activity in another channel responsive to the same spatial frequency but at a different spatial location (Polat & Sagi, 1993; Polat & Sagi, 1994; Field, Hayes & Hess, 1993; Hess & Dakin, 1997).

Evidence has also been found that spatial frequency channels interact through time such that spatial frequency information is integrated (Morgan & Watt, 1997). However, the temporal form of this integration is not clear: evidence has been found showing that spatial frequencies integrate more effectively when they are made available sequentially from low to high (coarse to fine)¹; and evidence has also been found showing that spatial frequencies integrate when they are available sequentially from high to low (fine to coarse) (Parker, Lishman & Hughes, 1992; Schyns & Oliva, 1997).

The aim of this thesis is to examine the temporal nature of spatial frequency integration, in particular whether the spatial frequency integration process improves its efficacy when spatial frequencies are presented sequentially in one temporal form (coarse to fine or fine to coarse) rather than another. Such work is essential for understanding further the nature of spatial frequency processing after their initial registration by spatial frequency sensitive channels.

As an introduction to the series of experiments presented in this thesis this chapter will address work relevant to the integration of spatial frequencies. Firstly, an overview of the nature of spatial frequency processing will be outlined (see section

1.2.1). Secondly, evidence which suggests that spatial frequencies integrate will be reviewed (see section 1.2.2). Finally, the research conducted which pertains to the temporal nature of spatial frequency integration will be presented (see section 1.3).

1.2 The integration of spatial frequencies

This section will examine the evidence which suggests that spatial frequencies integrate. Firstly an outline of some of the evidence concerned with early visual processing is given. In particular the existence of channels in the visual system responsive to various spatial frequencies is discussed followed by considerations of non-linearities and further processing stages. Secondly the evidence suggestive of interactions between the outputs of spatial frequencies channels is examined. This is followed by a review of the literature suggesting that there is non-independent access to individual filters by high level processing and more specifically that spatial frequency information is integrated.

1.2.1 Spatial Frequency Channels

First stage filtering

Visual images contain large amounts of information and can be represented in a number of ways. They may, for instance, be considered as an array of pixels and grey levels. A description in pixels is an array of discretely sampled intensity points approximating the continuous intensity changes present in the image. Each pixel therefore represents the average intensity of the areas of the image which are contained within it as in cathode ray tube displays. An alternative descriptive decomposition of an

image is to specify the spatial frequencies present within it. In any image there are a number of spatial frequencies present, that these can be individually identified was derived from Fourier's work on heat transference. It has been shown that any one dimensional waveform can be decomposed into component sine waves of specified phase, amplitude and frequency. This can be extended into two dimensions when the additional parameter of orientation is included. This is commonly referred to as a Fourier analysis or a Fourier transform (Bracewell, 1965). It is also the case that this operation can be reversed by the linear sum of sine waves of specified phase, amplitude, orientation and frequency to form any complex waveform.

The idea of pixelisation of images in terms of the intensity present fits neatly on the function performed by the retinal photoreceptors. They thus respond to the intensity changes present in the image and provide an approximation of the overall image luminance profile. It has also been suggested that the human visual system performs some kind of image analysis in the spatial frequency domain (Campbell & Robson, 1968; Sekuler & Levinson, 1974; DeValois & DeValois, 1980), however, it is now generally accepted that the visual system does not perform a pure Fourier analysis but rather a patch-wise one (DeValois & DeValois, 1988).

Evidence that the mammalian visual system carries information about images in the spatial frequency domain comes primarily from the identification of the differing sensitivities of human observers to a variety of spatial frequencies as shown by the contrast sensitivity function (CSF) (Schade, 1956). The CSF is a measure of sensitivity to spatial frequencies in terms of the contrast required to detect their presence. It was implicitly assumed for many years that the CSF was the result of the processing of a single channel in the visual system broadly tuned to the envelope of spatial frequencies represented in the CSF analysis (DeValois & DeValois, 1988). In 1968, however, Campbell & Robson found that a complex grating (one made up of more than one

frequency, for example those created from square-waves) will appear like a simple sinewave when only its most visible component alone reaches contrast threshold. It will appear complex when the contrast is raised to a level at which the other components would normally reach their individual thresholds when presented alone. This result suggests that the CSF, rather than being the product of a broadly tuned single channel, is representative of a number of narrowly tuned, independent, quasi-linear, band-pass channels which form this wider envelope (see also Graham & Nachmias, 1971).

Psychophysical methods, such as adaptation (Gilinsky, 1968; Pantle & Sekuler, 1968; Blakemore & Campbell, 1969; Blakemore & Sutton, 1969); subthreshold summation (Sachs, Nachmias & Robson, 1971); and masking (Carter & Henning, 1971; Stromeyer & Julesz, 1972; Harmon & Julesz, 1973) also suggest the existence of narrowly tuned independent channels, thus providing support for the result found by Campbell & Robson (1968). (Following Graham (1989) the term channel is adopted to represent a group of mechanisms, the psychophysical analogue of a set of single neurones, which are homogeneous along the dimensions of spatial frequency and orientation). These are briefly described in the following paragraphs.

Gilinsky (1968), Pantle & Sekuler (1968) and Blakemore & Campbell (1969) found that the sensitivity to specific spatial frequencies could be differentially affected by adaptation to a grating of that frequency. Viewing a grating of a specific frequency for a period of time results in the subject subsequently requiring more contrast to see a near threshold grating of the same spatial frequency. This does not affect sensitivity for spatial frequencies which are sufficiently removed from the adapted stimuli. The effect of adaptation was found not only to operate at the level of contrast thresholds, Blakemore & Sutton (1969) found that those gratings which are of lower spatial frequencies than of that adapted to appeared lower still, while those of higher spatial frequencies appeared even higher. This result suggests that the spatial frequency

information content of an image is encoded in terms of the distribution of activity amongst channels (Blakemore, Nachmias & Sutton, 1970).

Another line of evidence comes from subthreshold summation studies. It has been found that if two stimuli, which when presented alone would be undetectable, are simultaneously presented spatially they sometimes become detectable beyond that expected as a result of probability summation. This is interpreted as occurring when the two gratings are close in spatial frequency and stimulate the same channel. The findings from Sachs, Nachmias & Robson (1971) suggest that this does in fact occur.

Further evidence for the existence of multiple spatial frequency channels comes from masking studies. It has been found that when a grating of a particular spatial frequency is presented in conjunction with another of a similar spatial frequency it becomes more difficult to detect (Stromeyer & Julesz, 1972). This effect has been interpreted in much the same way as the adaptation and subthreshold results. Thus if the masking pattern has no effect on the detectability of the target pattern then this suggests that they are processed by independent channels. A number of studies have examined this phenomena and they all provide evidence that spatial frequency masks do disrupt the ability to perceive nearby spatial frequencies but not more distant ones (Carter & Henning, 1971; Legge & Foley, 1980).

An important aspect of the considerations of spatial frequency channels is their number and bandwidths. There are however difficulties when attempts are made to specifically map the sensitivities of these channels in terms of their number and their widths (Wilson, McFarlane & Phillips, 1983; DeValois, 1977; Blakemore & Campbell, 1969). For example Sachs et al (1971) found using a subthreshold summation paradigm that the bandwidth of spatial frequency channels were as small as 0.33 octaves, while those identified by Pantle (1974) were as large as 2.4 octaves. Thus depending upon the technique used different estimates are found. However, one of the most thorough studies

examining spatial frequency channel bandwidths was carried out by Wilson, McFarlane & Phillips (1983) using a masking paradigm. They showed that the bandwidths of spatial frequency channels vary inversely with spatial frequency sensitivity, such that at low spatial frequencies the bandwidth is approximately 2.5 octaves while at higher spatial frequencies this narrows to approximately 1.25 octaves. These bandwidth estimates agree well with those found by De Valois, Albrecht & Thorell (1982) who measured the spatial frequency sensitivity bandwidth of macaque striate cortical cells. These two studies, as well as showing approximately the same bandwidths of spatial frequency channels, also found that six separate spatial frequency sensitive channels were responsive over the entire spectrum.

It should also be emphasised that spatial frequency channels have also been found to be orientation specific such that the effect of one spatial frequency grating on another is found only within a limited range of orientations (Campbell & Kulikowski, 1966; Blakemore & Campbell, 1969; Gilinsky, 1968; Phillips & Wilson, 1984); and phase selective at supracontrast threshold levels (Burr, 1980; Lawden, 1983) but not at contrast threshold (Graham & Nachmias, 1971; Nachmias & Weber, 1975). Furthermore spatial frequency sensitivity is not homogenous across the retina. Swanson & Wilson (1985) showed that spatial frequency sensitivity decreases progressively in the peripheral retina where receptive fields become larger and thus more sensitive to lower spatial frequencies.

These studies all suggest that processing carried out in the visual system is performed, initially via multiple independent channels sensitive to different spatial frequencies, orientations and phase. It has been suggested that the visual system may be performing a pure Fourier analysis (Pollen, Lee & Taylor, 1971; Robson, 1975) one in which a visual image is transformed into the frequency domain via the visual system. This clearly does not happen, a pure Fourier analysis would require a set of bandpass

filters to be all infinitely narrow and would also require sinusoidal information which is infinite in extent. As can be seen from the evidence outlined previously this cannot occur. Thomas (1970) suggested that spatial frequency channels may reflect the operation of striate simple cell receptive fields which are found in the same area of space at a number of scales. This suggests that the visual system may be carrying out a localised spatial filtering operation at a number of scales, orientations and positions. This has been termed a patchwise Fourier analysis (De Valois & De Valois, 1988).

Non-linear visual processing

It has been pointed out that the bandwidths of channels responsive to spatial frequencies were found to be very narrow using a subthreshold summation technique (Sachs et al, 1971). This estimate has been suggested to be due to the introduction of a beat pattern into the subthreshold summation stimuli employed. The summation of the two spatial frequency gratings means that the stimuli were actually high spatial frequency patches whose amplitude was modulated at a low frequency (Stromeyer & Klein, 1975). Thus the bandwidth estimate of spatial frequency channels was narrowed by reducing spatial probability summation between channels. Once this was controlled for the bandwidth estimates increased. These sort of beat patterns found when two close spatial frequencies are added together have also been investigated by Burton (1973) and Henning, Hertz & Broadbent (1975). They found that the beat patterns introduced from adding two gratings of different spatial frequencies interacted with a luminance grating at the same periodicity as the beat periodicity, despite their being no power at this frequency. This represents one of many non-linearities found in the visual system.

There are a number of good reasons why non-linearities within the visual system should occur. Firstly, it is the case that a Fourier analysis is limited as the information that it conveys is simply a transform of the spatial information of an image into the

frequency domain. There is no analysis of the image as such only another representation. This shifts the problem of interpreting the content of a visual image into a problem of analysing the Fourier spectra of the image. Secondly, it has been pointed out that non-linear operations must occur in the visual system in order to perform anything useful. A Fourier analysis is a linear operation, if this were followed by a second linear system then it is the case this is equivalent to a single - stage linear system. Thus the operation of any further stages would gain nothing further in the analysis of visual information (Malik, 1994). Wilson & Wilkinson (1997) suggests therefore that the purpose of multi-scale representation within the visual system is so that non-linear interactions can occur. They argue that non-linear interactions between spatial frequency channels are necessary so that functionally important computations can be carried out enabling the visual system to meaningfully process visual information.

Non- linear interactions between spatial frequency channels have been found to occur in many different experimental paradigms. Interactions between spatial frequency channels have been found between spatial frequency channels located at the same region of space but were responsive to different spatial frequencies. Tolhurst (1972) found that adapting to a square-wave produced less of an adaptation effect to the third harmonic (3F) of the square-wave series than was found when the 3F was presented alone. If spatial frequency channels were independent as suggested by other experiments (Campbell & Robson, 1968; Graham & Nachmias, 1971) then this would not be predicted. This lessening of the adaptation effects of spatial frequencies when they were presented in compounds was also found by Nachmias, Sansbury, Vassilev & Weber (1973; see also De Valois, 1977). Further interactions were found by Nachmias & Weber (1975). They showed that a high contrast fundamental spatial frequency of a square-wave series facilitated the detection of a 3F. This effect was also found by Stromeyer & Klein (1975).

As well as spatial frequency interactions being found to occur within a region of space but across spatial frequency sensitivities, interactions have also been found at one spatial frequency sensitivity across space, that is non-linear interactions are found to operate within spatial frequency sensitive channels which are located at different points of space. Polat & Sagi (1993; Polat & Sagi, 1994) have found that the presence of two flanking Gabor patches (a Gabor patch is a spatial frequency grating multiplied with a Gaussian weighting function which Gabor (1946) showed maximises both spatial location and spatial frequency information (see section 3.1.1)) arranged to be collinear with a foveally presented Gabor patch facilitate its detection, that is the contrast threshold for the detection of the centrally presented Gabor patch is lower when it is flanked. This relationship is found to vary with distance between the target and the flankers, such that when the flanking targets are very close (overlapping) the threshold for detection is inhibited, this changes to a facilitatory interaction as the flankers are shifted to about 3 wavelengths away and extends to about a distance of 12 wavelengths. This effect required the local and global orientation as well as the spatial frequency of the Gabor patches to be the same. Thus the effects are strongly orientation and spatial frequency tuned. These interactions occur across a space too large to be mediated by the same spatial frequency channel thus they must involve non-linear interactions (although see Solomon, Watson & Morgan, 1999 for an alternative interpretation). Two models could account for the results: either a lateral interaction between spatial frequency channels; or a second stage pooling of spatial frequency channels across space such as collector or collator units (Moulden, 1994; Mussap & Levi, 1996). Mussap & Levi (1996) suggest that the output of spatial frequency channels are integrated into a second stage collator or collector unit (see also Moulden, 1994). They found, following Findlay (1973), that vernier acuity thresholds are maximally masked when shown on a spatial frequency grating background oriented 10 to 20° from vertical. The effect was found to

decrease in strength as a horizontal gap between the gratings increased. They found that large filters are contributing to the determination of the vernier threshold. In order to reconcile the high positional acuity for vernier offset with the contribution of large filters they suggest that smaller filters are initially coding positional information, these are then integrated along the axis parallel to their common orientation into a larger second stage filter.

Recently there has been a great deal of research examining what has been termed non-Fourier visual processing. This is best typified by the work investigating texture and motion processing. For example Chubb & Sperling (Chubb & Sperling, 1988; Chubb & Sperling, 1989; Chubb & Sperling, 1991; Sperling, 1989) have examined motion processing of a bar which moves across a grey background. They found that when this bar is contrast reversed on successive steps of movement that the Fourier sinewave components of this movement are to the left while the bar actually moves to the right. If this stimulus is full-wave rectified then both the stimulus and its Fourier components move to the right. When this bar stimulus is viewed from a short distance the bar appears to move rightward, however when viewed from a longer distance, peripherally or at very low contrast the bar appears to move leftward. This suggests that two different motion judgements are being made. One which involves the non-linear computation of full-wave rectification at short distances and one which involves a linear computation where the direction of motion is judged directly from the spatio-temporal Fourier motion components. In further experiments (Chubb & Sperling, 1991) involving “texture quilts” they showed that the non-Fourier stage of motion processing must be preceded by an initial linear stage of spatial filtering such as that described previously. These different types of motion processing have been labelled first order (linear or Fourier) and second order (non-Fourier) motion (Cavanagh & Mather, 1989).

A Second Wave of filtering

A further stage of visual processing is also suggested to occur involving linear filtering much as the first stage. For example in Malik & Perona's (1990) model of texture processing they showed that an initial spatial frequency processing stage followed by half-wave rectifying produced energy differences which were greatest at the texture boundary. The rectification stage of processing allows an edge detection process to be carried out which takes the form of a further linear filtering stage. These three stages (although technically a two stage model as rectification is not necessarily a stage as such) then are: filter - rectify - filter. These form the basis for many current models of early visual processing (Adelson & Bergen, 1985; Bergen & Landy, 1991; Wilson & Wilkinson, 1997; Graham & Sutter, 1998). For example Wilson, Wilkinson & Assad (1997) examined the detection of orientation information in Glass patterns (Glass patterns were introduced by Glass (1969) and are made up of random dots which together can form a large number of global patterns (Wilson & Wilkinson, 1997)) and found that a model of the form: filter - full-wave rectification - filter, could account for the results. The nature of these two filtering stages was such that the orientation of the filters were of an orthogonal orientation, displaced spatially in the orthogonal direction and their number doubled. Thus if the first stage filter was oriented horizontally then after rectification its output would be processed by two vertically oriented filters placed at equal distances above and below the first stage filter. This filter - rectify - filter process of a horizontally oriented first stage filter followed by two vertically oriented filters described by Wilson et al (1997) integrate to emulate the output of a pair of end stopped complex cells. In this particular model the output of this second filtering stage is then summed across orientation and thresholded.

The non-Fourier information processing carried out by the visual system then involves an initial filter stage, the output of which is passed through a non-linearity and

finally processed by a second filtering stage, the nature of which differs depending upon the task. It has been suggested in the previous example that this non-linearity is full-wave rectification, but it has also been suggested to take the form of a half-wave rectification (Watt & Morgan, 1985) and a squaring of the responses which is also a form of rectification (Morrone & Burr, 1988; Bergen & Landy, 1991). Graham & Sutter (1998) describe the operation of the different filter - rectify - filter models as being examples of “complex” or “higher order” channels. Complex channels are those where the second stage of processing takes responses from first stage filters of the same shape but at different positions, whereas higher order channels collect responses from first stage filters of different shapes at the same spatial position (Graham & Sutter, 1998).

Georgeson, Schofield, Guest & Anderson (1998) have also examined non-Fourier visual processing using contrast modulated gratings, akin to texture processing, comparing them with luminance modulated gratings. They found evidence suggesting that the non-Fourier information carried by the contrast modulated grating was processed by separate channels to that processing luminance modulated spatial frequency channels. Thus non-Fourier visual processing occurs in parallel with spatial frequency processing.

Summary

Early visual processing then seems to comprise of a bank of spatial frequency channels which differ in their spatial frequency, orientation and phase sensitivities. These have been suggested to perform a patchwise Fourier analysis. It is also apparent from the brief overview of non-linear interactions in visual processing that there is substantial evidence that such interactions do occur. These interactions are found both at different spatial frequency sensitivities within a single patch of space and spatial

frequency sensitivity across a limited patch of space. Further evidence of non-linear interactions suggests that the output of first stage filtering is passed through a non-linear operation such as half- or full-wave rectification and is followed by a second stage of filtering. Furthermore the nature of this non-linear processing does not appear to be homogeneous. Indeed the nature of the non-linear operation and the second stage filtering differ as the task changes. Graham & Sutter (1998) suggests a difference between the operation of complex channels and higher order channels.

Thus the introduction of at least two spatial frequency stimuli almost always involves the operation of some non-linearity (Wilson & Wilkinson, 1997). The evidence presented also suggests that non-Fourier visual processing occurs either subsequently to or in parallel with a spatial frequency filtering process. In the next section evidence is outlined which examines the suggestion that spatial frequencies are integrated.

1.2.2 Spatial frequency integration

It appears from the discussion of the nature of spatial frequency processing that not only does it involve an initial Fourier like processing but that it must also involve the operation of non-linearities following this initial filtering stage. In one of the examples outlined previously a model was described which was introduced by Wilson, Wilkinson & Assad (1997) to account for evidence arising from the use of Glass patterns. In the final stage of this model an integration stage was briefly alluded to. This stage appears after the filter - rectify - filter stages as a summation and threshold function. Thus the outputs from the second stage filtering process are integrated prior to any decision making process. It is not only in this model that such a stage of integration is found. In most models of early visual processing a stage of information integration is

found prior to decision processes (Watt & Morgan, 1985; Watt, 1988; Wilson, 1991; Georgeson, 1994; Georgeson & Meese, 1997; Wilson & Wilkinson, 1997; Wilson, et al., 1997; Morgan & Watt, 1997; Graham & Sutter, 1998). Indeed Wilson & Wilkinson (1997) state that "...current research suggests that early-visual-channel responses are almost always combined, but in a variety of task-dependent manners" (p.956).

The stage at which this integration takes place differs across models and as Wilson & Wilkinson (1997) point out may also change in a task-dependent manner. For example Graham & Sutter (1998) place the integration of visual information in a comparison and decision stage which occurs after the parallel operation of simple (linear) and complex (filter - rectify - filter) channels but prior to the observers rating, while the MIRAGE model of Watt & Morgan (1985) suggest that the integration should occur after a first stage simple filtering and a half-wave rectification stage. Thus, depending upon the model, the stage at which channel outputs are integrated changes. As these models were designed to account for results from experiments concerned with different aspects of visual (e.g. motion or texture processing) this suggests, echoing Wilson & Wilkinson (1997), that the pooling of channel outputs differ depending upon the task employed.

Although the exact locus of the integration of channel outputs is not clear, it has been pointed out by Watt & Morgan (1985), Watt (1988) and Morgan & Watt (1997) that the integration of spatial frequencies must involve a non-linearity. Thus the integrating (e.g. summing) of two different spatial frequencies reduces the resolution of the higher spatial frequency as the complex waveform will be dominated by the lower spatial frequency. In their MIRAGE model spatial frequencies are integrated only after they have been half-wave rectified. This acts to preserve the higher spatial frequency information. Thus Morgan & Watt (1997) suggest that spatial frequencies are integrated but only after the a non-linearity has been applied to the output of the initial spatial

frequency processing. However this does not go against the proposition that there are parallel simple, complex and higher order channels operating at an early stage of visual processing, as the problem of resolution would still be successfully addressed by the complex and higher order channels.

It should be made clear at this point that the integration of channel outputs will be referred to throughout this thesis as spatial frequency integration. Spatial frequency integration is used as the nomenclature because the filtering in both simple channels and the first stage filtering in complex / higher order channels is one where the spatial frequencies are initially separated. Although evidence for further stages in visual processing have been identified, the stage at which an integration of filter outputs takes place is not clear. Therefore as spatial frequency filtering underlies further visual processing the integration of channel outputs has been named spatial frequency integration. It should be noted then that no assumptions are being made about the locus of this integration. Furthermore, the word integration is used not in a mathematical sense but only to refer to the combination of filter outputs regardless of the mathematical form that combination may take.

In the following sections evidence is examined which suggests that spatial frequencies do indeed integrate. Firstly some computational reasons why a spatial frequency integration would be preferred are outlined. Secondly, some empirical evidence that spatial frequency integration does indeed take place in the human visual system is reviewed.

Computational advantages of spatial frequency integration

The decomposition of an image into its component spatial frequencies as a basis of initial image processing has been criticised as it provides only a re-description of the

initial image without providing any further information about the image itself. In order to make the properties of the image explicit further processing, such as that outlined in the previous section (1.2.1), must be performed. Another such process is that of spatial frequency integration. Computationally there are advantages to integrating information across spatial frequencies, for example there is an increase in reliability, a data reduction and a noise reduction.

The integration of spatial frequencies increases the reliability that the physical properties of the visual environment are represented. For example, it has been argued that edges are important in any representation of the visual image and thus their correct identification has been the subject of many computational models (Marr, 1982; Canny, 1986); one way to do this is to operate over a number of spatial scales and integrate the information found. Thus if something occurs in the image across a number of spatial scales then it is more likely to actually represents something in the physical world. This is best described by Marr's spatial coincidence assumption (1982: p. 70):

“If a zero-crossing segment is present in a set of independent $\nabla^2 G$ channels over a contiguous range of sizes, and the segment has the same position and orientation in each channel, then the set of such zero-crossing segments indicates the presence of an intensity change in the image that is due to a single physical phenomena (a change in reflectance, illumination, depth or surface orientation).”

Secondly, it has been suggested that the integration of spatial frequencies is a form of data reduction in that it involves distilling the visual information in order to make decisions. It has been found that the primate visual system is comprised of over 20 areas some of which project to further visuomotor systems (Felleman & Van Essen,

1991). Thus many areas show specialised visual and visuo-spatial functions and therefore must involve multiple representations of visual information. For example, information for the grasping of objects is derived from the visual information received in the parietal cortex (Sakata, Taira, Kusunoki, Murata & Tanaka, 1997). Visual information then is projected to many areas associated with action responses thus suggesting a distillation of the information from the initially large multi-dimensional spatial frequency representation to a much smaller representation which contains only the information necessary for certain actions. To continue the previous example then visual information should be distilled into the necessary information for hand movements, that is a 3 - dimensional space. This suggests an integration of visual information specifically for action control.

This idea of multiple representations involved in different visual processes was implicit in the previous discussion of early visual processing models. It was pointed out in filter - rectify - filter models that the non-linearity represented by the rectify stage differed depending upon the task. This is further compounded by Wilson & Wilkinson's (1997) review of non-linearities in early vision. The different models which account for evidence found in motion and texture analysis suggest that the processing of spatial frequency information changes as the task requirements change at the higher levels of non-linearities. This is also reflected in the differences found in integration rules which may take the form of a thresholded summation (Wilson et al, 1997) or a averaged summation (Morgan & Watt, 1997). Thus the evidence suggests that different tasks require different representations. This is not surprising as the different tasks may involve different areas of processing within the brain (e.g. motion processed by area MT). Thus it is not too fanciful to suggest that the integration of spatial frequency information provided by first stage filtering differs at higher levels of processing in different tasks. This suggests that spatial frequencies are integrated by multiple mechanisms which each

provide a representation optimal for the tasks carried out by different projection areas (cf. Allport, 1993 for a comparative discussion of attention).

The integration of smaller or narrowband channels or filters into larger, more broadband channels also reduces the noise involved with smaller filters. Mussap & Levi (1996) show that the pooling of narrowband channels into a larger collator mechanism acts to reduce the signal to noise ratio of the smaller filters. This is also suggested by Georgeson (1994).

Thus there are computational advantages to integrating spatial frequencies. In the next section some evidence which suggests that the human visual system actually integrates spatial frequency information is outlined.

Evidence of spatial frequency integration

It was suggested in early experimental work examining spatial frequency channels that they were independent in their operation (Campbell & Robson, 1968). This was forcibly supported by the adaptation studies of Blakemore & Campbell (1969) and was widely accepted as a principle of early vision (see also Graham & Nachmias, 1971). However, evidence has been outlined to show that spatial frequencies interact with each other both at a single location of space (Tolhurst, 1972; Nachmias, et al., 1973; Olzak & Thomas, 1981; Greenlee & Magnussen, 1988) and across a limited area of space, larger than the traditional receptive field, between channels sensitive to the same spatial frequencies (Polat & Sagi, 1993; Field, Hayes & Hess, 1993). This seems to be a function of contrast levels. The visual system approximates a linear system at contrast threshold where the outputs of independent channels are available to higher level processes, however at higher contrast levels the processing becomes highly non-linear and the outputs of independent channels are not accessible (Graham, 1989). This section outlines some evidence which suggests that the output of spatial frequency

channels are not independently accessible by higher visual processing, which is suggestive of spatial frequency integration, and then presents more direct evidence which indicates that the spatial frequency information integrated across channels at the same spatial location.

There is a great deal of evidence which indicates that spatial frequency channel outputs are not independently accessible. An example of non-independent access to filters is provided by the Abraham Lincoln stimuli created by Harmon & Julesz (1973). Here the image of a face is quantised, that is the image is segmented into a number of squares the luminance of which is computed by averaging the luminance from all the pixels within that square; once this average is found all the pixels within the square take on this average luminance value. Thus the low spatial frequency information of the image is still present but this is overlaid by spurious high spatial frequencies which have been introduced by the blocking process. Due to the introduction of this high spatial frequency information the image, despite the remaining presence of the low spatial frequencies, is made unrecognisable. It was originally suggested that this unrecognisability was due to the masking of the low spatial frequencies by the high spatial frequency noise introduced in the quantised process if this was within a critical bandwidth of the remaining low spatial frequencies of the image. However, Morrone, Burr & Ross (1983) added further high spatial frequency noise to a quantised picture of the Mona Lisa and found that recognisability level improved. Thus suggesting that the masking of the low spatial frequencies by a critical bandwidth of high spatial frequencies introduced in the quantised process is not the explanation. Morrone et al (1983) suggest that the added high spatial frequency noise reduces the organisation of the image into the blocks produced by quantisation, thus suggesting that the low spatial frequencies in the face is unable to be accessed independently of the high spatial frequencies.

Watt & Morgan (1983) found evidence for a performance level in the sensitivity of blur discrimination which could be accounted for by use of the largest filter available to the visual system despite a much improved performance which would have resulted if a much finer spatial scale was used. This suggests that the task could only be performed on the basis of the larger filter as the finer ones could not be independently accessed (Watt & Morgan, 1984; Morgan & Watt, 1984).

Jamar & Koenderink (1985) found evidence for non-independent filter access. They measured contrast thresholds for the detection of noise stimuli as a function of the bandwidth of the spatial frequencies within the noise. It would be expected that as the bandwidth increased more channels would become operative thereby reducing the contrast threshold of the noise target. This was not shown, they found that the visual system uses all the spatial luminance power in the target irrespective of its bandwidth (see also Quick, 1976).

As well as evidence showing that spatial frequencies are not independently accessible by higher processing levels, and thus suggestive of spatial frequency channel integration, there is also specific evidence which suggests that the outputs of spatial frequency channels are integrated. Early evidence suggesting that the output of spatial frequency channels were integrated was found by Kulikowski & King-Smith (1973), Nachmias & Weber (1975) and Klein & Stromeyer (1980) who all suggested the existence of edge or bar detectors. For example Klein & Stromeyer (1980) examined the suggestion of inhibition between spatial frequency channels shown by the lack of an equivalent adaptation effect when a spatial frequency was shown in a compound (Tolhurst, 1972). They showed that if the spatial frequencies are separately visible by separating the spatial frequencies, rapidly jittering the lower spatial frequency or drifting the gratings then adaptation to a third harmonic (3F) in the presence of a fundamental (F) is equivalent to the 3F presented alone. They suggested the existence of interactions

between edge detecting and line detecting mechanisms which respond to both the F and the 3F spatial frequencies. Suggesting that the spatial frequencies are combined across spatial frequency.

Bowne (1990) found that spatial frequency and temporal frequency discrimination judgements were independent of contrast levels (orientation judgements only showed an independence at higher spatial frequencies). In order to explain this Bowne (1990) suggested that the channel outputs must be combined within these dimensions prior to any decision level process and that the accuracy of these outputs are subject to a task-dependent noise process such as positional uncertainty.

Georgeson (1992) also suggested that spatial frequency information was combined from oriented spatial frequency channels. Evidence from the use of plaids suggested that their structure is not directly related to the output of two oriented filters but is that which would be predicted from the zero crossings of a circularly symmetric Laplacian of a Gaussian (LOG) suggested by Marr & Hildreth (1980). That the perceived structure of these plaids can be distorted by adaptation or masking suggests that the structure of plaids are derived from the linear summation of oriented first stage filters which emulate the operation of isotropic LOG filters (Georgeson, 1992). Further evidence for spatial frequency integration has been found by Georgeson & Meese (1997) using plaids. They found that when two identical spatial frequencies are presented at oblique angles of say $\pm 45^\circ$ they integrate to form a checkerboard pattern ("squares") rather than the expected diamonds which would be found if the two spatial frequencies were perceived as being separate. If these two obliquely oriented spatial frequencies are denoted as F, then if a 3F spatial frequency is presented at one of the oblique angles, that is there is one spatial frequency (F) at $+45^\circ$ while there are two (F + 3F) at -45° , then as soon as the 3F reaches its contrast threshold the pattern changes from one where a checkerboard is seen (as usual) to one where diamonds are seen. Thus it appears that the

introduction of the second spatial frequency interferes with the integration across orientation. They suggest that spatial frequencies can be integrated either across spatial frequency or across orientations but not both at the same time.

This suggestion is also found in the work of Olzak & Thomas (1992), Olzak & Wickens (1997) and Olzak & Thomas (1999). They have also shown that spatial frequency information is integrated across orientation at one spatial frequency and also across spatial frequency within an orientation band. They show that spatial frequency information is integrated prior to making orientation decisions and also that orientation information is integrated prior to making spatial frequency discriminations. Interestingly they found that both spatial frequency and contrast information was not integrated in other tasks. Their evidence suggests that an integration of spatial frequencies occurs only for the coding of orientation information, the results argue against explicit edge representations.

This seems to differ from the findings of Georgeson & Meese (1997) where it was suggested that spatial frequency information does integrate across scale. This may be due to two reasons. Firstly, it may be the case that this reflects task dependent integration. As mentioned previously, the nature of visual processing seems to differ with task requirements, this is compounded with the findings of Olzak & Wickens (1997) where spatial frequency information is integrated only for some decisions and not for others. This may be due to the possibility that filter information is both integrated and individually accessible depending upon the nature of the visual processing. Evidence has been outlined suggesting that early visual processing occurs in parallel via simple, complex and higher order channels. Therefore findings that the integration of spatial frequency information is task dependent is entirely possible given the models of early visual processing outlined previously, as they do not necessarily preclude each other, that is they are not mutually exclusive concepts. A second

possibility for the finding of a lack of spatial frequency integration may be a result of the large spatial frequency difference between the two components employed in Olzak & Wickens' tasks (e.g. 3 and 15 cpd) and thus they may be less integrated due to the lack of structure between the two components. For example, in Marr's (1982) conception of early visual processing, integration takes place on the basis of zero-crossings therefore it may be the case that the zero-crossings from the different spatial frequencies used here did not coincide and so are individually represented. This was also suggested by Georgeson & Meese (1997) as they state "There is clearly something special about the F, 3F relationship that leads immediately to the integration of frequency components at a common orientation, and to the breaking of links across other orientation" (p. 3264). It has also been argued that integration may rely on channels of progressively higher spatial frequencies, that is adjacent channels, being activated (Wilson et al, 1991) thus the successful integration of spatial frequencies may rely on the activation of contiguous channels. It is also noteworthy that the situation of widely separated spatial frequencies would not occur naturally. Thus it may be the case then that the results of Olzak & Thomas (1992) and Olzak & Wickens (1997; see Olzak & Thomas, 1999 for a review) are due to the wide differences between the spatial frequencies employed.

Thus the results of Olzak & Wickens (1997) and Olzak & Thomas (1996) may be a result of task dependent processing or the use of widely separated spatial frequencies. While this task dependency does seem to be a facet of visual processing, it does not seem likely that the results of Olzak & Wickens (1997) reflect this as the other evidence (e.g. Georgeson, 1992) outlined in this section strongly supports the integration of spatial frequencies. Thus it seems likely that the results are due to the use of widely different spatial frequencies which do not encourage their integration. It could be argued that if this were the case why the spatial frequencies are readily integrated to form an orientation code. However, the orientation information is contained within a narrow

bandwidth so that on the dimension of orientation, the one on which the information is integrated to code, the spatial frequency information is not widely separated. It remains to be seen whether this suggestion is the case.

Overall then, there is evidence then that spatial frequencies are not independently accessible and that they are integrated prior to processing by higher level processes. This integration seems to depend upon the activation of contiguous channels. Further, albeit indirect, evidence derives from single cell responses in the inferior temporal cortex. Cells here have been found to be responsive to complex stimuli (Perrett, Mistlin & Chitty, 1987). This responsiveness has been found to be both size and orientation insensitive (Miyashita & Chang, 1988; Lueschow, Miller & Desimone, 1994) suggesting a complex integration of spatial frequencies into a number of integrated representations much like the action representations discussed previously. It is interesting to note that the results of Olzak & Wickens (1997) suggests that evidence showing a lack of independent access to different spatial frequencies may not be due to the integration of spatial frequency information.

1.2.3 Summary

Evidence has been outlined which suggests that the visual system initially processes visual information via a bank of multi-resolution filters which are sensitive to different spatial frequencies, orientation and phase. The output of these channels are then passed through a non-linearity the form of which seems to differ across tasks before, in some cases, being passed through a second filtering stage the form of which again changes across task. Further evidence has been outlined which suggests that there are a number of non-linear interactions between the channels, one of which is spatial

frequency integration. The integration of spatial frequencies can be suggested to occur between spatial frequency channels which are very close on that dimension. The nature and the locus of spatial frequency integration has not been addressed. However it seems likely given the models of texture processing, motion processing and early feature identification that the integration of spatial frequencies occurs through many different processes. Therefore the nature and the locus may differ depending upon the visual processing involved. In the next section the temporal nature of this integration process is considered.

1.3 The temporal form of integration

This section outlines the evidence pertaining to the temporal form which spatial frequency integration takes. It has been suggested that spatial frequencies are processed from coarse to fine through time (Watt, 1987; Wilson, Blake & Halpern, 1991; Parker, et al., 1992), however other evidence has suggested that spatial frequencies are processed temporally from fine to coarse (Smallman, 1995; Parker, Lishman & Hughes, 1996). In this section consideration of the computational advantages for particular temporal forms of spatial frequency integration will be outlined first. Second evidence which suggests particular temporal forms will be outlined. This is followed finally by a discussion of the literature specifically examining temporal forms of spatial frequency integration.

1.3.1 Computational advantages for particular temporal forms of spatial frequency integration

Computationally it has often been suggested that it would be advantageous for a system capable of multi-resolution representation, to process spatial frequencies from coarse to fine through time, where the coarse representation either integrates with or constrains the processing carried out with the finer ones (Marr & Poggio, 1979; Burt & Adelson, 1983; Asada & Brady, 1986; Lindeberg, 1993; Palagi & Guerin-Dugue, 1995). Equally, it has been suggested that the processing carried out in the fine spatial frequency channels should precede that carried out in the coarser spatial frequency channels (Marr, 1982: p.70-71; Canny, 1986).

The computational advantage to processing spatial frequency information from coarse to fine is essentially the reduction of time involved in forming an integrated representation. The visual system must process information at a speed sufficient to allow it to be available for an organism to function adequately in its environment. Thus any visual system cannot involve a large amount of pre-processing before allowing visual information to be available to decision making processes. Watt (1987; 1988) showed that in any global image operation which requires iteration, the more iterations that are required to form an adequate representation the longer the time lag between the initial presentation of a stimulus and its subsequent representation. Watt (*ibid.*) suggests that in order to make a representation available to an organism as quickly as possible an intermediate representation could be formed which is always available for higher level decision processes and at each iteration this is updated so the time lag between stimulus and representation is minimal (Ferrero & Foster, 1986).

Evidence has been presented suggesting that the visual system processes stimuli across a number of scales. One of the ways in which to further reduce the time taken to

produce an incremental representation is to integrate these scales from coarse to fine. Watt (1987; 1988) argues that any representation of an input stimulus must involve the derivation of spatial position and that as the spatial scale of analysis becomes finer this becomes more difficult due to the increase in the number of elements (It has been shown that the number zero-crossings in an image decreases as the spatial scale becomes coarser (Koenderink, 1984)). It follows that as the number of elements increase so does the time necessary to compute their spatial position. This is due to the nature of grouping, at the coarser scales the number of groups whose spatial positions need to be computed are much lower than at finer scales thereby reducing the time related to the computation. Thus a visual system where matching at progressively finer scales was carried out would allow an incremental representation to be available at a shorter time. Thus a coarse to fine incremental representation would reduce the time necessary to produce a representation upon which an organism can act (see Watt, 1987 for a mathematical derivation of this). This is echoed by Lindeberg (1993) who outlines a coarse to fine strategy for the identification salient image structures by firstly detecting edges in the coarse scales and using this information to track edges found at higher scales. As the number of zero-crossings are smaller at the coarser scales (Koenderink, 1984) this reduces the number of computations and therefore the time involved in their processing.

However there have been suggestions that a fine to coarse integration of spatial frequencies would be more computational advantageous (Marr, 1982; Canny, 1986). Marr (1982) suggests that a fine to coarse integration is necessary as the zero-crossings in the smaller channels need to be encoded explicitly in order to correctly denote edges which would be blurred together by the larger channels. That is, there is less chance that neighbouring image structures would interfere with each other. The larger channels may also incorrectly denote the location of edges because of their blurring function. In order

to combat these problems Canny (1986) suggested that the marking of edges in the smaller channels could then be used to predict where edges should be found in the coarser channels. The prediction and the actual output of the coarser channel are then compared and are marked with edges only if the actual output was greater than the predicted output. Thus edges found in the smaller channels which are also found in the larger channels are marked only at the smaller channel output. Those not accounted for in the coarser channel by smaller channel activity are compared with an even coarser channel output. If a match is found it is marked in the smaller of the two channel outputs and those not matched are carried on to the next larger channel. This is then an iterative process progressing from fine to coarse where the overall output is an edge map².

There are problems with both a coarse to fine and a fine to coarse temporal form for integration. For the coarse to fine integration process there is the problem of initially imprecise localisation. For the fine to coarse integration process there is the problem of the large amount of smaller channels signalling edges which do not correspond to edges found in the physical stimuli. Thus the two processing strategies reflect two different constraints which must occur when spatial frequencies are integrated, one of time of processing and one of improved localisation.

In the following section empirical evidence is outlined which suggests the existence of a particular temporal form of spatial frequency integration: either coarse to fine or fine to coarse. These are then discussed in terms of a flexible task-dependent process.

1.3.2 Evidence for particular temporal forms of spatial frequency integration

In this section evidence will be outlined which is either suggestive that the integration of spatial frequencies takes a coarse to fine or a fine to coarse temporal form. First, psychophysical and physiological evidence from spatial frequency processing delays is reviewed. Second evidence from empirical studies which suggest a coarse to fine integration process are outlined, these include early experiments examining Microgenesis, motion and stereopsis. It will be seen that the evidence from this section predominantly favours a coarse to fine temporal form of spatial frequency integration.

Differential processing speeds of spatial frequencies

There is a substantial amount of evidence which shows a temporal difference in the processing of spatial frequencies (Breitmeyer, 1975). This suggests that in the initial period immediately after the onset of a visual image there is a temporal latency difference between the processing of the lower and higher spatial frequencies present within that image. This differential perceptual latency in the processing of spatial frequencies has been found to increase with the spatial frequency.

The processing time of spatial frequencies has been inferred using a variety of measurement methods, both psychophysical and electrophysiological. The findings from these studies will now be outlined under these broad headings of psychophysical: reaction times and temporal order judgements; and electrophysiological: visual evoked potentials and neuromagnetic imaging. Followed by a very brief consideration of the underlying mechanisms which may be responsible for spatial frequency processing delays.

*Psychophysical:**Reaction Times*

Reaction time as a measure has often been suggested to comprise of two or more stages: a perceptual stage which involves processing of the stimulus by the visual system; a motor stage which is the time involved in eliciting the motor response. Reaction times have been used to infer the processing times of spatial frequency channels by assuming that any motor time involved does not vary as a function of spatial frequency (Gish, Shulman, Sheehy & Leibowitz, 1986; Ejima & Ohtani, 1987). A number of experiments find that reaction times do vary with spatial frequency, it is therefore suggested that the differential processing times involved result from delay in the initial perceptual stage.

It has been shown that reaction times increase as a function of spatial frequency. For example, Breitmeyer (1975) presented vertical sinusoidal gratings ranging from 0.5 cpd to 11 cpd for exposure duration of 50 msec. Subjects were asked to respond as soon as they were detectable. Contrast was set at 50 %. He found that reaction time increased with increasing spatial frequencies at a rate of approximately 29.7 msec per octave. Parker (1980) also measured reaction times to spatial frequencies, but in this case at their onset, offset and contrast reversal. It was also found that reaction times increased as spatial frequencies did in all conditions.

When reaction times to spatial frequencies have been measured experimenters have either equalised physical contrast or used some method to attempt to equalise the detectability of spatial frequencies. These will be discussed in turn.

In experiments where the physical contrast of the spatial frequencies was the same (Breitmeyer, 1975; Breitmeyer, Levi & Harwerth, 1981; Tartaglione, Goff & Benton, 1975; Vassilev & Mitov, 1976) reaction times were found to increase with higher spatial frequencies. As outlined previously Breitmeyer (1975) found reaction

times increasing as spatial frequency increase when the physical contrast of the stimuli were set at 50%. Vassilev & Mitov (1976) examining the reaction times to square-waves found an increase of 21.7 msec per octave when the stimuli was shown at 60% contrast and 21.2 msec per octave when shown at 20% contrast. Breitmeyer et al (1981) found reaction times to onset, offset and contrast reversed spatial frequencies to increase at a rate of 45.9, 48 and 45.9 msec per octave respectively.

In order to ensure that it is perceptual processing delays which are the underlying cause for the differences in reaction times, various attempts have been made to assume that the detectability of the spatial frequencies that are shown have been equalised. Methods which have been used for attempting to equalise the detectability of stimuli at suprathreshold levels are: the contrast matching of the stimuli; the setting of the spatial frequency contrast as a multiple of threshold; and the equalising of detectability rates regardless of contrast.

First, equalising the apparent contrast of the suprathreshold spatial frequencies allows the assessment of the level to which reaction times increases may be due to the increase in contrast threshold found as spatial frequencies become higher (Breitmeyer, 1975). For example, Breitmeyer (1975) in a second experiment had subjects equate the contrasts of all test gratings to the contrast of the 11 cpd grating which was fixed at 66%. The reaction times to these gratings were than measured and although the reaction time advantage shown for low spatial frequencies did reduce, it was not completely abolished. However a problem with this method is that while the apparent contrast of the spatial frequencies are equalised it does not necessarily follow that the spatial frequencies are equally detectable.

The second method, finding the contrast threshold values for the spatial frequency gratings and increase these threshold levels by some constant multiplier (e.g. 1 log unit), assumes that the detectability and apparent contrast increases equally for all

spatial frequency as physical contrast increases (Gish et al, 1986). Lupp, Hauske & Wolf (1976) measured reaction times to stimuli whose contrast threshold was multiplicatively increased by a factor of 5. They found that reaction times increased as spatial frequency increased at a rate of 19.2 msec per octave. Given the results of Georgeson & Sullivan (1975) and Kulikowski (1976; Georgeson, 1991), however, it seems unlikely that the detectability of the spatial frequencies are equal in these cases. Georgeson & Sullivan (1975) presented subjects with two gratings of which the contrast of one was set while the other was manipulated. Subjects were asked to match the contrasts of the gratings. In this way subjective contrast equalisation could be measured. They found that the subjective contrast of low spatial frequencies of low contrast increased faster than high spatial frequencies of low contrast. This suggests that multiplicatively increasing contrast thresholds would not equalise subjective suprathreshold levels.

A further method for equalising stimuli has been presented by Gish et al (1986). Rather than manipulate contrast levels to provide equality they compared gratings which were equally detectable regardless of their actual contrast. Thus reaction times to sinusoidal gratings of 1, 4, and 10 cpd were measured at contrasts at individual thresholds and at three levels above spaced at 0.1 log steps. Detectability was also measured during the experiment. This was determined as the amount of stimuli reacted to in comparison to the number of stimuli presented and was z score transformed. Thus reaction times to gratings were compared not as a function of contrast level but rather as a function of detectability. Therefore assuming that gratings which are equivalently detectable produce the same level of evidence in the visual system and the motor response is static then the reaction time differences observed would be a result of perceptual latencies involved in the processing of spatial frequencies. On the basis of this system of comparative spatial frequencies they found similar results to those obtained by the previously discussed experiments, in that reaction times increased as

spatial frequency increased despite gratings being equally detectable from chance levels to near perfect detection levels. The reaction time delays were on the order of approximately 21 msec per octave.

Despite the varying methods of contrast equalising employed, a reaction time advantage for lower spatial frequencies has still been found (Breitmeyer, 1975; Lupp, Hauske & Wolf, 1976; Gish et al, 1986). Furthermore when the detectability of different spatial frequency gratings is equalised (Gish et al, 1986) the spatial frequency latency differences are not abolished.

This reaction time advantage has also been found under a variety of contrast levels from threshold to very high suprathreshold conditions where detection is near perfect (Harwerth & Levi, 1978; Gish, et al., 1986; Felipe, Buades & Artigas, 1993). Felipe et al (1993) presented vertical sinusoidal gratings between 1 and 40 cpd for a large number of contrasts from threshold to 95%. They found that as contrast increased reaction times decreased for both low and high spatial frequencies but the difference is not abolished in line with Breitmeyer's (1975) results. This method of assessing reaction times at a large number of spatial frequencies and contrast allows the plotting of the entire reaction time surface. Harwerth & Levi (1978) measured reaction times to stimuli of spatial frequency from 0.5 to 16 cpd and a wide range of contrasts. They found, as would be expected, that reaction time increased as spatial frequency increased and that reaction time decreased as contrast increased. They also found that as contrast increased the resultant functions, plotted with spatial frequency as a parameter, showed an inflexion as contrast increased. This was not found for either low or high spatial frequencies only for intermediate ones. From these results they suggest that the responses are derived from a single channel with two response characteristics dependent upon contrast. However, this result is easily interpretable as showing evidence for two parallel channels overlapping their operation over a significant range of spatial

frequencies. More recently Felipe et al (1993) show that as contrast increases the reaction time plot changes from one mirroring the contrast threshold function to one where the function is flat except for a shift around spatial frequencies of 4 - 8 cpd. This suggests that the reaction times are a function of two underlying mechanisms (See *Transient and sustained mechanisms* and table 1.1).

Ejima & Ohtani (1987) suggest that there may also be an increase in motor processing times associated with spatial frequency processing. When partialled out, with the assumption that the exposure duration necessary for the detection of different spatial frequencies equals the perceptual integration time which reflects spatial frequency processing delays, their results still show an increase in processing for different spatial frequencies of approximately 14.6 msec per octave.

In all the experiments discussed in this section different viewing distances and therefore sizes of stimuli are used, exposure durations, viewing conditions: monocular and binocular, and DC levels also differ. It is remarkable that reaction time increases with increasing spatial frequency regardless of the large vagaries in the methods employed .

Recently Thomas, Fagerholm & Bonnet (1999) have shown that the increased reaction time to low and intermediate spatial frequencies can be accounted for by a single mechanism with a low pass filter. However these data are derived from eccentrically presented stimuli and it is not clear at this time if these results reflect the same processes as those underlying reaction times to centrally presented spatial frequencies.

Temporal Order Judgements

A further dependent variable used in the measurement of latency differences in the processing of spatial frequencies is that of temporal order judgements. The logic in

these experiments has been to present two stimuli at varying stimulus onset asynchronies (SOA) and ask subjects to judge which stimulus was perceived as being presented first. An alternative is to use a method of adjustment and allow subjects to adjust the onset of the test stimuli so that it appears to co-occur with the onset of the standard. Thus if there is a latency difference in the processing lag with increasing spatial frequencies then the matching SOA should provide a measure of this.

Parker & Dutch (1987) present subjects with test gratings of spatial frequencies from 0.5 to 9.0 cpd and, as a standard, a flashing LED. Subjects were asked to match the onset of the peripherally presented LED to the onset of the centrally presented gratings. It was found that the temporal order judgements differed as a function of spatial frequency. As spatial frequency increased so did the deviation of the onset judgement from zero. They found that the LED latency increased as the spatial frequencies of the gratings increased, from a latency of 8.35 msec at 0.5 cpd to 29.3 msec at 9 cpd showing a visual latency increase of 21 msec over the 4 octaves, when the duration of the LED and grating were identical. This condition is not wholly satisfactory as it does not control for the judgement of order based on the other temporal aspects of the stimuli especially the offset. A second condition ran by Parker & Dutch (1987) addresses this problem by increasing the duration of the grating stimuli to 400 msec while the LED duration remained at 40 msec. They found, as in the previous condition, that as spatial frequency increased so the visual probe latency did likewise. In this case the increase ranged from 6 msec at 0.5 cpd to 32 msec at 9 cpd, an increase of 26 msec over the 4 octave period. This latter condition was not carried out at all spatial frequencies but only at the lowest (0.5 cpd) and the highest (9 cpd) spatial frequencies.

In a second experiment carried out by Parker & Dutch (1987) they examined the perceptual latency differences between the two principal components of a square-wave, presented separately. Thus LED latencies were measured to a grating of spatial

frequency F and contrast m and to a grating of spatial frequency $3F$ and contrast $m/3$.

They find that as in experiment 1 the $3F$ pattern has an increased LED latency and interestingly when F is increased the LED matching latency increases but the slope remains the same, the average increases were 21 msec at $F = 3$ cpd and 25 msec at $F = 4.5$ cpd. This suggests that the perceptual latency of a square-wave's fundamental and third harmonic differs by 21 - 25 msec. Further it suggests that the appearance of square-wave changes overtime due to dynamic nature of the visual system, that is the coarse spatial frequencies are registered prior to the fine.

Barr (1986) carried out a number of temporal order judgement experiments with spatial frequencies. Using spatial frequency gratings as test and as standard stimuli it was found that the SOA at which their onset was perceived as simultaneous was zero. In this experiment he segregated the screen horizontally presenting the spatial frequencies at the top or bottom of the screen separated by a neutral strip with a fixation spot located centrally. He repeated this experiment using a staircase methodology and found most subjects onset judgement was very close to zero, but two did place the high spatial frequency 20 msec prior to the presentation of the low spatial frequencies, a variation in the expected direction. As a comparison in both of these experiments the reaction times to the stimuli were also measured. The usual increase of reaction time with spatial frequency was found.

There is no clear explanation why the studies presented here differ so widely. Parker (1996: personal communication) has suggested that Barr may have expected the temporal order judgements differences to be large in line with the reaction time data and set sample SOA points correspondingly large. As a result of this the fine variations of SOA with spatial frequency found by Parker & Dutch (1987) may have been overlooked.

*Electrophysiological**Visual Evoked Potentials*

A third dependent variable used in the measurement of conduction velocity and spatial frequency is that of the visually evoked potential (VEP). VEP's are recorded via the scalp and are suggested to be the direct neural manifestation of the visual input. Thus a VEP elicited in response to a spatial frequency of 3 cpd is said to represent that 3 cpd spatial frequency. It is in this manner that latencies can be recorded. It has been shown that VEP's do vary as a function of spatial frequency (Kulikowski, 1977; Parker & Salzen, 1982; Russell, Kulikowski & Murray, 1987).

The overall finding with regard to latency has been much the same as that outlined in the reaction time literature, it is found that the latency of VEP's increase as the spatial frequencies from which they are elicited also increase in frequency content (Parker & Salzen, 1977; Plant, Zimmern & Durden, 1983; Musselwhite & Jeffreys, 1985; Russell, et al., 1987; Kulikowski, 1991; Vassilev, Stomonyakov & Manahilov, 1994). This has been found when the contrasts are equated (Vassilev & Strashmirov, 1979; Kulikowski, 1977), to pattern onsets, offsets and reversals (Parker & Salzen, 1982) and when VEP's with equal amplitudes are elicited (Jones & Keck, 1978).

Musselwhite & Jeffreys(1985) suggested that the increase in the latencies of VEPs may be due to the reduced contrast sensitivity to high spatial frequencies. They measured VEP components C1 and reaction times to different spatial frequencies and found that although reaction times did increase with spatial frequency the latency of the VEP component did not. However Vassilev et al (1994) measured VEP's in the presence of flicker masking which is known to increase reaction times to low spatial frequencies. They found a analogous increase in VEP latency which cannot be explained by reduced contrast sensitivity to high spatial frequencies.

It has also been pointed out by Musselwhite & Jeffreys (1985) that in some cases no latency differences are evident (Kulikowski, 1977) and suggest the evidence for latency differences may be due to the variable composition of peaks across the VEP studies. Parker & Dutch (1987) and Regan (1989) have also identified the difficulty of peak identification in the measurement of VEP's.

Overall then, there is great difficulty in interpreting the VEP findings with regard to spatial frequency latencies, although there is a great deal more evidence which does suggest that VEP latency varies with spatial frequency such that the lower spatial frequencies are processed faster than higher ones.

Neuromagnetic Imaging

Williamson, Kaufman & Brenner (1978) measured the visually evoked magnetic field which is associated with the flow of information in the visual system. They find that as spatial frequency increases so does the magnetic field responses. In line with the reaction time data they show a remarkable fit between their own data and that observed by Breitmeyer (1975) when a constant is added. This fit suggests that their measurements reflect the latency of spatial frequencies in the visual system while the necessary constant added to reflect Breitmeyer's (1975) data reflects a constant motor component. However a major problem with the use of this study for comparative purposes is the temporal variation of the stimuli which either flickered or were contrast reversed at a rate of 8 - 20 Hz.

Transient and Sustained mechanisms

Evidence has been presented suggesting that the low spatial frequencies are perceptually available prior to the high spatial frequencies. This is suggested by the results of the increased reaction times, temporal order judgement differences, and VEP

latency increases to increasing spatial frequencies. This increase in the latency involved in these measures when spatial frequencies increase suggest that when a visual stimulus is presented the lower spatial frequencies are processed faster than the higher ones, which in turn suggests that the lower spatial frequencies are available for higher level processing prior to the higher spatial frequencies.

It has been suggested that these spatial frequency latency effects reflect the operation of two systems whose response sensitivity overlaps. These two systems have been termed the transient and sustained systems. It has been proposed that the initial processing of the visual image by the visual system is carried out by these two underlying transient and sustained mechanisms. The distinction between the two was first suggested on the basis of the examination of the temporal characteristics of the X and Y cells in the cat (Enroth-Cugell & Robson, 1966; Cleland, Dubin & Levick, 1971) where X cells were found to respond throughout the duration of a stimulus (sustained) while the Y cells showed a response at the onset and the offset of a stimulus (transient). An analogous distinction between a transient and sustained system has been identified with the M and P pathways of the monkey visual system (Lennie, 1993). There is an anatomical distinction of cell types in the primate which is analogous to the X and Y cell distinction outlined previously in the cat, from the distinct populations of bipolar cells which drive the M and P pathways, the distinct ganglion cell populations, different layers in the LGN, up to the different layers of the striate cortex V1 (Kaplan, Lee & Shapley, 1990). A summary of the response characteristics of the two pathways can be seen in table 1.1.

	P	M
Response to stimulus onset	Tonic (sustained)	Phasic (transient)
Motion sensitivity	Lower	Higher
Conduction Velocity	~ 15 msec	~ 6 msec
Receptive Field size	Smaller	Larger
Spatial Frequency sensitivity	Higher	Lower
Linear Summation	Linear (X)	75% X & 25% Y
Contrast Gain	Lower	Higher (8-10 times as high)
Percentage Ganglion cells	80	10
Number of cells (in millions)	1.2	0.15
Distribution on Retina	Densest in fovea	Densest in fovea?
Contrast Sensitivity	Lower	Higher

Table 1.1: Summary of the response characteristics of the P and M pathways in the human and primate visual system (Derived from Kaplan et al ,1990; Lennie, 1993).

It can be seen from table 1.1 that the conduction velocity of M and P pathways do not seem to differ by a great enough degree to account for the findings of speed differences of different spatial frequencies. However, differences of 20 msec have been found by Nowak, Munk, Girard & Bullier (1995; although see Maunsell, Ghose, Assad, McAdams, Boudreau & Noerager, 1999). This coupled with the different contrast and temporal sensitivity and contrast gain of the cells probably lead to a much larger differential speed advantage. Thus the threshold for eliciting a reliable response from an M cell will be much lower and thus much faster (Lennie, 1993).

An illustration of this is given by lesion studies. Ibotenic acid or neurotoxins can be used to selectively lesion the P and M cells which comprise the P and M pathways. Merigan, Katz & Maunsell (1991) using ibotenic acid and Merigan & Eskin (1986) using neurotoxins damaged the P pathway and found that macaques were maximally sensitive to stimuli of low spatial frequency and high temporal frequency. Conversely lesions of the M pathway isolating the P pathway resulted in a loss of sensitivity to those stimuli (low spatial frequency and high temporal frequency) and a maximal sensitivity to higher spatial frequencies and lower temporal frequencies. Thus it may be the case that the M and P pathways provide an underlying explanation of the temporal processing advantage enjoyed by the low spatial frequencies (Merigan, Byrne & Maunsell, 1991).

Although this consideration of anatomically and physiologically defined pathways suggests the existence of two separable systems it need not necessarily be the case that this distinction is identifiable at a perceptual level. For example Maunsell et al (1999) have recently found that the latency differences between the M and P pathways are as little as 10 msec. Furthermore they suggest that this latency difference may not last into the cortex primarily because of the larger number of parvocellular neurons in the visual system where their convergence could reduce the latency differential such that under some circumstances the P pathway would be faster than the M pathway. There are however a number of strands of psychophysical evidence which can be interpreted as reflecting a transient / sustained distinction at a perceptual level. Evidence from spatio-temporal contrast (Robson, 1966; Kulikowski & Tolhurst, 1973; Tolhurst, 1973), subthreshold summation (King-Smith & Kulikowski, 1975; Tolhurst, 1975b), temporal waveform (Breitmeyer & Julesz, 1975; Tolhurst, 1975a), temporal impulse response (Watson & Nachmias, 1977; Watson, 1986), exposure duration (Schober & Hilz, 1965; Nachmias, 1967; Breitmeyer & Ganz, 1977b; Legge, 1978; Georgeson, 1987), masking (Legge, 1978; Breitmeyer & Ganz, 1977a; Breitmeyer, et al., 1981; Mitov, Vassilev &

Manahilov, 1981; Green, 1981a; Breitmeyer, 1984), and finally adaptation (Green, 1981b) all suggest the existence of two separable systems which seem to reflect the response characteristics of the M and P pathways as outlined in table 1.1. The results from these studies suggest the following characteristics for transient and sustained channels. The transient channel shows a relatively better sensitivity to stimuli of low spatial frequencies and high temporal frequencies, this coupled with a biphasic temporal impulse response function may also reflect a preference for changes in the image (temporal waveform). The transient channel also has a shorter temporal integration period and a faster conduction velocity which may underlie the faster latency responses, such as reaction times, to low spatial frequencies. The sustained channel is more sensitive to high spatial frequencies and low temporal frequencies. Also its monophasic temporal impulse response function may be reflected in its preference for steady state stimuli. The sustained channel also has a longer temporal integration period showing steady improvement over long periods with a longer conduction velocity.

It must be pointed out however, that, following Breitmeyer (1984) and others, a clear psychological dichotomy is not the general picture, but rather an overlapping of response properties of channels is envisaged with relative differences between them such as contrast sensitivity, which leads to one channel being more responsive than the other under some circumstances. Thus although initially heralded as a form and motion distinction, where the sustained carried form information of the visual image and transient transmitted motion information (Tolhurst, 1973), this has largely been replaced by a conception of their being different relative sensitivities such that the transient channel is relatively more sensitive to motion and the sustained channel more sensitive to high spatial frequency which carry more detailed aspects of the image. This then allows a role for the sustained channel in motion perception and the transient channel in form perception. Thus the segregation between the transient and sustained systems at a

psychophysical level can be largely ignored as a continuation between the two is often found. This means essentially that the transient and sustained systems mimic the operation of a single system (Watson, 1986). Watson (1986) points out that the hard-line theory suggesting two mechanisms would require evidence of two mechanisms responding to the same spatial frequency, such as that provided by Harwerth & Levi (1978) and Felipe et al (1993); and evidence showing them to be functionally distinct, such as most of the evidence outlined in the preceding paragraph.

It can be seen, following the discussion of the properties of the transient and sustained mechanisms, that the latency experiments are fully interpretable on the basis of the existence of these two channels. Thus the reason for low spatial frequencies producing faster reaction times (Breitmeyer, 1975), VEP's (Vassilev et al, 1994), Neuromagnetic imaging potentials (Williamson et al, 1978) and different temporal order judgements (Parker & Dutch, 1987) is because they are transmitted via the transient channel which shows a faster conduction velocity and shorter temporal integration period.

Conclusions

In this section the latencies involved in the processing of spatial frequencies have been considered. Evidence has been presented from reaction time, temporal order judgement, VEP's and neuromagnetic imaging studies which show that spatial frequencies are processed at different speeds, such that as spatial frequencies increase so does the time taken to process them. It has been suggested that the operation of transient and sustained channels mediate these differential spatial frequency processing delays. It has further been suggested that the sustained and transient channels reflect the operation of the M and P pathways. While it is difficult to ensure that the M and P pathways are the cause of the psychophysical operation of the sustained and transient channels they do

share a number of common characteristics. However, for the present argument it is not necessary to ensure the relationship between the two. It is sufficient that there is great deal of evidence for differential delays in the processing of spatial frequencies such that lower spatial frequencies are processed faster than higher ones suggesting that coarse precedes fine in spatial frequency integration. Thus it can be suggested from these spatial frequency processing delays that the integration of spatial frequencies must proceed from coarse to fine.

In the next section other empirical evidence will be outlined which suggests the existence of a coarse to fine integration. These are suggested to be a consequence of the differential delays involved in the processing of spatial frequencies outlined in this section.

Empirical evidence suggestive of a coarse to fine integration

Microgenesis

Microgenesis refers to the initial formation of an unitary integrated percept (Flavell & Draguns, 1957). Early experiments examining the Microgenesis of perception concerned themselves with temporally degrading the stimuli by varying the exposure duration of tachistoscope presentations upwards until the stimuli was completely perceptible or downwards until it could no longer be perceived. (Kruger, 1928; Sander, 1930; Undeutsch, 1942: all cited in Flavell & Draguns, 1957). The stimuli employed in these studies have varied from famous paintings to geometrical figures, while the experimental designs have varied from increasing or decreasing exposure durations, size or eccentricity and measures used have varied from introspection to the reproduction of the stimuli presented.

The results from these wide ranging studies led Sander and his group to suggest Microgenesis as progressing from a diffuse whole percept which becomes sharpened through the increase of fine detail with increasing duration (Sander & Jinuma, 1928; Undeutsch, 1942; Mantell, 1936; Morschner, 1940; all cited in Flavell & Draguns, 1957). For example Stein (1949; cited in Flavell & Draguns, 1957) and Phillips & Framo (1954) found that there is a tendency to form percepts which are initially integrated over the whole stimulus and as time progresses more detailed aspects of the stimulus are added.

There are, however, a number of methodological problems with these early studies such as few subjects, inadequate statistics (if any), and the informal nature of data collection, but this body of work does show the first indication of a coarse to fine integration in perception.

The notion of Microgenesis then can be seen to be an extension of the Gestalt claim that the global aspect of a figure has primacy. This is captured in the notion of holistic processing where it is suggested that the processing of a visual image proceeds in a hierarchical manner from global aspects to more fine local details of the figure with increasing time.

A further example of the notion of microgenesis can be found in the work of Navon (1977) and others which examined the global and local properties of image information. Navon (1977) presented stimuli made up of small letters which together formed a larger letter which may or may not differ from small letters (following Kinchla, 1974), these he termed the local and global aspects of the figure. In order to control for the differential properties inherent in the global and local aspects of a stimulus such as complexity, familiarity, recognisability, salience and relative diagnosticity he constructed stimuli which were controllable at both a global and local level and which ensured the independence of these two levels, that is neither is

predictable from the other. The logic of these stimuli, then, is that the two constituent parts are equal except for the level at which they stand.

Given the background to this section it follows that the basic hypothesis is that the global letter formed out of the local letter units will be processed first followed by an analysis of the local units themselves. In a number of experiments Navon (1977) found two basic conclusions, these are generally grouped under the term Global precedence effect:

a) The global aspects of the stimuli were perceived before the local level. Navon presented subjects with the hierarchical stimuli and asked them to report either the global or the local level letter depending upon the condition. All stimuli were presented at brief exposure durations. He found that reaction times to the global letter were significantly faster than those to the local letter (See Navon, 1977: Experiment 3). This is termed the global temporal advantage effect

b) The perception of the local level stimuli is adversely affected by the global level but not vice versa. In the same experiment as described above, Navon found that when the local was inconsistent with the global level the latencies for the identification of the global letter were not significantly different, while the latencies for the identification of the local letters increased significantly. This has been termed the Asymmetrical interference effect (See also Boer, 1982; Grice, Canham & Burroughs, 1983 for replications of these findings).

This then provides evidence for the Microgenetic approach to early visual perception. These effects form the basis of what Navon termed the Global Precedence Hypothesis. This states that the global information in an image is processed prior to the

local information, thus Navon claimed that there is no parallel processing of this information, one is completed before the other starts (For discussion of this issue see Miller, 1981 and Navon's reply (Navon, 1981)).

Since the Navon study was carried out many more experiments have examined various aspects of this affect and most have found confirmation of either the temporal advantage or the asymmetric interference effect. Variables which have been manipulated are: retinal location of the stimuli (Pomerantz, 1983; Grice, Canham & Burroughs, 1983); varying visual angle of the stimulus (Kinchla & Wolfe, 1979; Lamb & Robertson, 1989); attentional allocation (Ward, 1982); spatial uncertainty (Lamb & Robertson, 1988); the number of local elements (sparsity) (Martin, 1979); exposure duration (Paquet & Merikle, 1984); the Gestalt notion of a goodness of form (Hoffman, 1980; LaGasse, 1993). See Kimchi (1992) for a review.

It can be seen that a possible explanation for the two basic findings from Navon's study could be in terms of spatial frequencies involved in the two levels of the stimuli thus suggesting that for some reason the low spatial frequencies are dominant in early image derivation. Whether spatial frequency processing underlie this effects remains a moot point, although there is general agreement that spatial frequency processing underlies the global precedence effect (Hughes, 1986; Shulman, Sullivan, Gish & Sakoda, 1986; Shulman & Wilson, 1987; Badcock, Whitworth, Badcock & Lovegrove, 1990; Hughes, Fendrich & Reuter-Lorenz, 1990; Lovegrove, Lehmkuhle, Baro & Garzia, 1991; Lamb & Yund, 1993; Lamb & Yund, 1996a; Lamb & Yund, 1996b; Hughes, Nozawa & Kitterle, 1996), while the asymmetric interference effect has been found not to be spatial frequency dependent and may be the result of differential hemispheric processing (Lamb & Yund, 1993; Lamb & Yund, 1996a; Lamb & Yund, 1996b; Badcock, et al., 1990; LaGasse, 1993; Lovegrove, et al., 1991; Sergent, 1982; Lamb & Robertson, 1989; Lamb & Robertson, 1990; Robertson, Lamb & Knight, 1988;

Kitterle, Christman & Conesa, 1993; Christman, 1993; Fink, Halligan, Marshall, Frith, Frackowiak & Dolan, 1996; Grabowski & Nowicka, 1996) or may reflect visual field processing (Previc, 1993; Christman, 1993; Lamb & Yund, 1996a).

The evidence outlined in this section lends some support to the notion that the underlying basis of the Microgenetic formulation of early perception is the differential primacy given to spatial frequencies. Despite the uncertainties about the mechanisms underlying the asymmetric interference effect, which is not of primary importance in a Microgenetic formulation, it is clear that the overall global temporal advantage seems to suggest a form of processing in line with Microgenesis, where the global aspect of the stimuli (the low spatial frequencies) is processed immediately after stimulus onset, with increasing fine scale aspects of the stimuli (the higher spatial frequencies) available successively.

Motion Processing

The differential delays in the processing of spatial frequencies would be an important facet in any model of motion processing. It can be envisaged that a delay in the processing of high spatial frequencies relative to low would actively distort a moving image as the phase relationships between the constituent waves would be affected and the salient aspects of the image, such as edges, would no longer be correlated. For example, it has been found in motion perception that an illusory square-wave may be seen when a drifting triangle wave grating is presented. Anderson (1993) suggested that this illusion was due to the processing delays in the initial processing of spatial frequencies and the actively changing phase relationship. Thus a drifting complex waveform could be presented which, because of the differential delays involved in the processing of spatial frequencies, would integrate in a phase relationship dependent upon the spatial frequencies and the velocity at which the waveform was

drifting. Anderson varied these parameters of a drifting square-wave and triangle-wave and found that the phase of the harmonics must be advanced relative to the fundamental in order to compensate for the phase lags. This was also found to depend upon the waveform used. The square-wave stimulus needed little phase advancement suggesting less temporal delay than the triangle-phase stimuli. It was also found that the phase distortion in the square-wave stimuli was more tolerable than a disruption in the phase relationship of the triangle-wave stimulus.

Stereopsis

Stereopsis is concerned with the problem of fusing each eye's monocular view into a single image. From the work by Julesz (Julesz, 1960; Julesz, 1971) it has been suggested that this is carried out not globally but locally point by point. That this is so is well illustrated by the work with Random Dot Stereograms. However this conception of a point by point matching process leads to the correspondence problem, that each point in the visual image of one eye can be potentially matched with any point in the visual image of the other eye. This is well illustrated by Marr & Poggio (Marr & Poggio, 1976; Marr & Poggio, 1979; Marr, 1982) where they showed that even with a situation of 4 dots presented to each eye there are 16 possible (4x4) matches, thus with two 10x10 dot arrays there 10^{10} possible solutions.

Marr & Poggio (1979) developed a model of stereopsis which depends on a coarse to fine integration of spatial frequencies. (That spatial frequencies are involved in the process of stereopsis has been shown by Julesz & Miller (1975) and Mayhew & Frisby (1976)). In the model, each eye processed their respective visual images into a number of spatial frequency channels, thus there are representations at each scale. Binocular matching then proceeds by matching the zero-crossings in the low spatial frequency representation first, these then provide a constraint on the higher spatial

frequency representations which have far more possible zero-crossing matches, and this algorithm is applied progressively to higher spatial frequencies.

It has been found that the low spatial frequencies constrain the range of binocular fusion possible with the high spatial frequencies. Wilson, Blake & Halpern (1991) measured diplopia thresholds for high spatial frequencies in the presence of low spatial frequency gratings. They found that when the frequencies were within 2 octaves of each other the fusion range was severely reduced, however as this difference increased to a 4 octave difference the two frequencies were always seen transparently at different depths. From these results it seems a coarse to fine integration of stereo matching does occur but only over a small range (~2 octaves) beyond this (4 octaves) the two frequencies are perceived as transparent. In natural scenes of course spatial frequencies across the spectrum are represented in a continuous manner thus this transparency would not result. A hierarchical model can be conceived where each individual scale constrains the processing in the next adjacent scale. This has implications for the use of widely separated stimuli such as that employed by Olzak & Thomas (1999; see section 1.2.2)

Recently however, Smallman (1995) has presented evidence that the disambiguation of stereo-matching can occur from fine to coarse, that is the fine scale information can be used to select the depth at which the coarse scale information is shown when the depth of that coarse scale information is ambiguous. For example in one experiment a low spatial frequency was presented in anti-phase to the two eyes and its end points did not betray which depth the spatial frequency should be matched to. When a higher spatial frequency random dot stimulus was presented with the phase separated low spatial frequencies the low spatial frequency sine-wave was judged to be at the same depth as the unambiguous random dot stimulus. This fine to coarse scale disambiguation was confirmed in a second depth matching experiment. Therefore

Smallman showed that the information held in the higher spatial frequencies can influence the perceived depth of the lower spatial frequencies contrary to the predictions from the Marr & Poggio (1979) model. It should be noted that the results from Smallman (1995) also show that the low spatial frequencies can disambiguate the depth of the high spatial frequencies. The results show that both a coarse to fine and a fine to coarse matching strategy can take place³.

Hyperacutities

Further evidence suggesting a coarse to fine processing of spatial frequencies has been shown by Watt (1987). Watt (1987) presented a number of hyperacuity judgements where the exposure duration and size of the stimuli varied. It was found that thresholds decreased with increasing exposure duration for orientation, linewidth, curvature and stereoacuity tasks, while a dot resolution task showed no variation with increasing exposure duration. The results were interpreted as showing an increase in the spatial scale of analysis with increasing time of exposure. In order to reconcile the results of the experiments Watt (1987) suggested that all spatial scale information is available at very early exposure durations but for the processing of geometric tasks the information held in the lowest available filter is used first. The improvement in these geometric tasks improves through time due to a switching off of the outputs of the lower spatial channels thus allowing the next lowest spatial scale to be used to make the judgement. In this way it is possible to improve thresholds through time for geometric tasks and account for the unaffected thresholds of the dot resolution task, which is a non-geometric task, and therefore not subject to this coarse to fine integration.

However there are problems with this account. Burbeck & Yap (1990b), using the geometrical task of line separation, found that even at short exposure durations high spatial frequency information is utilised. Furthermore Waugh & Levi (1993) found that

when, in a vernier acuity task, the targets are made equally visible at all exposure durations the effect of improving thresholds with increasing exposure duration disappears, suggesting that the results of Watt (1987) are due to the unequal visibility of the target used. Burbeck & Yap (1990) suggest that the use of high spatial frequencies in these geometrical tasks and the improvement shown over time reflects the longer temporal integration times of high spatial frequency filters (Waugh, 1998). These results suggest that the results from Watt (1987) do not provide clear evidence of a coarse to fine process occurring in initial visual processing.

Furthermore it can be seen that the conception of this process is not of a coarse to fine integration of spatial frequency information but of a coarse to fine turning off of channel outputs. In the model outlined by Watt (1987) the integration of spatial frequencies has already occurred by the time this coarse to fine processing takes place.

Overall then, it is not clear that the results of Watt (1987) provide evidence for a coarse to fine processing of spatial frequencies information and furthermore according to his own MIRAGE model it is not clear that these results have a bearing on whether a coarse to fine integration of spatial frequencies takes place.

Conclusions

Much of the preceding discussion suggests that the integration of spatial frequencies proceeds in a coarse to fine manner. That is the outputs from those channels responsive to the low spatial frequencies are available prior to the outputs from the spatial frequency channels responsive to the higher spatial frequencies, therefore the integration of spatial frequencies must proceed from coarse to fine. This temporal form of spatial frequency integration seems to be a direct result of the differential delays in the processing of spatial frequencies which are suggested by reaction time, temporal order judgement, and VEP studies examining the latencies associated with different

spatial frequencies. Thus the evidence from spatial frequency processing delays suggests that the integration of spatial frequencies is constrained to occur from the low spatial frequencies to the high spatial frequencies through time from the very first filtering stage. The evidence for microgenesis or the global precedence effect, effects of spatial frequency matching in motion processing, stereopsis and hyperacutities can also be suggested to be a result of these spatial frequency processing delays. Thus the evidence suggests that the processing of visual information involves a default coarse to fine integration of spatial frequencies.

Some evidence was presented suggesting that a fine to coarse integration of spatial frequencies could take place. Barr (1986), examining temporal order judgements of different spatial frequencies, showed that there was no evidence for spatial frequency processing delays, which suggests that any temporal form of integration could take place. Furthermore Smallman (1995) showed that information held in the high spatial frequencies could be used to guide the processing carried out in the low spatial frequencies. While Smallman's (1995) evidence does not address the temporal form of spatial frequency integration as such, the results suggest that the high spatial frequencies can impact on the processing carried out in the low spatial frequencies despite the evidence suggesting that the high spatial frequencies are processed at a slower rate than the lower spatial frequencies.

From the evidence outlined in this section and in Section 1.2.1 a model of the temporal form of integration can be suggested. Section 1.2.1 outlined evidence for the existence of spatial frequency processing and the form that processing can take. It is apparent that a basic model of early visual processing carried out by the visual system is one of an initial filtering followed by a non-linearity applied to the output of this stage. It can be assumed that prior to the integration of spatial frequencies it is necessary to include at least these two stages, although a further second stage of filtering can be

added if required (e.g. as suggested by some texture models (Graham & Sutter, 1998)). Thus in this very basic generic model there are two stages with an intermediate non-linearity: filter - non-linearity - integrate. This is directly analogous to the first stages of the MIRAGE model of early visual processing suggested by Watt & Morgan (1985), where they filter - half-wave rectify - integrate. Following much of the evidence outlined in this section, such as spatial frequency processing delays, the initial filtering stage will involve constraining the spatial frequency availability to the rest of the model so that a coarse to fine pattern of activation will occur. Thus the integration of spatial frequencies is by default a coarse to fine process.

In section 1.3.1, in which considerations of the computational advantages of different temporal forms of spatial frequency integration were outlined, it was suggested that a spatial frequency integration process should proceed from coarse to fine (Watt, 1987). A great deal of the evidence presented supports this suggestion. However it can be seen from a consideration of the integration model (filter – rectify – integrate) that the empirical evidence and the computational advantages address two different levels in that model. Thus the evidence, such as spatial frequency processing delays, suggests that the integration of spatial frequencies proceeds from coarse to fine because of the spatial frequency processing delays suggested to occur in the first stage filtering. The computational considerations suggest that the integration stage itself should be constrained to process spatial frequencies from coarse to fine. It could be suggested, because of the computational advantages for a coarse to fine integration of spatial frequencies, that the visual system has evolved in such a way so that the delays involved in the processing of spatial frequencies occur so that a coarse to fine integration is possible. Thus the argument is that it is advantageous for the integration of spatial frequencies to proceed in a coarse to fine manner and because of this the delays in the processing of spatial frequencies occur.

However, it is clear from consideration of the two stage model of spatial frequency integration that all of the evidence outlined above does not address whether the integration of spatial frequencies proceeds from coarse to fine because of an obligatory coarse to fine integration strategy or is simply a function of the processing delays. It has been pointed out that it is computationally advantageous for the integration to occur from coarse to fine. However the existence of the first stage processing delays does not address whether this is in fact the case. The evidence only gives indirect support for an integration of spatial frequencies which occurs more efficiently or effectively when the spatial frequencies are made available from coarse to fine. Thus it may be the case that the integration of spatial frequencies occurs from coarse to fine simply as a function of the spatial frequency processing delays introduced at the initial filtering stage and if these delays were removed the integration of spatial frequencies would occur just as well.

It is this difference between an integration stage whose efficacy depends upon the coarse to fine availability of spatial frequencies or one which integrates spatial frequencies as and when they become available which will be addressed in this thesis. In the next section evidence which pertains to this issue will be outlined.

1.3.3 The temporal form of the integration mechanism

Despite the evidence presented in the preceding section which is suggestive of a coarse to fine integration mechanism, none of it is directed at examining such a mechanism. The evidence cited thus can be said to be purely circumstantial and supportive of an integration mechanism which operates with greater efficacy when spatial frequencies are made available from coarse to fine, and one which proceeds

irrespective of the temporal order of spatial frequencies. Thus there are two concepts of how the integration of spatial frequencies operates which emerge from this line of thinking: First, one where the integration of spatial frequencies is improved if the spatial frequencies are constrained to be made available from coarse to fine, any disruption of this availability would lead to a corruption of the integration process; or second, a model where the integration of spatial frequencies proceeds just as effectively on the basis of the spatial frequencies which are available at any particular time. It just so happens to proceed in a coarse to fine manner because spatial frequencies are initially made available due to delays which occur in initial spatial frequency processing. Thus in this view there is no inherent bias in the integration mechanism itself for a coarse to fine integration of stimulus information, but rather spatial frequencies are accumulated as and when they become available. These two models have been termed the temporally anisotropic mechanism and the temporally isotropic mechanism respectively (Parker, et al., 1992).

Parker et al (1992) carried out experiments employing images of natural scenes where they examined subjective ratings of stimuli and the discrimination of stimuli as a function of the presentation order of their spatial frequency components. They presented filtered versions of the images sequentially either from coarse to fine or fine to coarse with each version being shown for 40 msec making up a sequence of 120 msec, thus seeking to stretch the delivery of spatial frequency information over a longer than normal interval. In the first instance they asked subjects to rate the filtered image combinations in terms of their perceived quality. They found that those presented coarse to fine were rated as being of higher quality than those present fine to coarse whether or not a fullbandwidth image was presented within the image triplets. In a second experiment (their experiment 5) subjects were required to discriminate the presence of the fullbandwidth version of an image when it was presented in a coarse to fine or fine

to coarse image sequence. They found that subjects were more likely to indicate the presence of a fullbandwidth image target when it was presented in a coarse to fine sequence of the filtered images than if it was embedded in the same sequence of filtered images presented from fine to coarse. It was also found that a coarse to fine sequence of filtered images which did not contain a fullbandwidth image target was also more likely to elicit more reports of a fullbandwidth image target being present than a fine to coarse sequence of filtered images, despite the target not actually having been presented at all. This was shown not to be a function of the subjects response being determined by the last frame shown by comparing conditions in which the fullbandwidth image was shown as the first or final image.

Parker et al (1992) interpret these results as being supportive of a temporally anisotropic integration mechanism. In the first experiment the results can be interpreted as reflecting coarse to fine presentation of information being more readily integrated thus leading to higher quality ratings. In the second experiment the results suggest that the increased likelihood of reporting a fullbandwidth image when images are presented coarse to fine also reflects the fact that the filtered images are more readily integrated. Conversely Parker et al (1992) suggest that the fine to coarse sequence of images is less integrated thus allowing a sequence of sub-optimal images to be discriminable. However it must be pointed out that this is not a valid conclusion without appealing to some additional process which reduces the visibility of the fullbandwidth image in the fine to coarse sequence. If it was the case that a fine to coarse image sequence was not readily integrated then the fine to coarse sequence should elicit as many "fullbandwidth image present" responses as the coarse to fine when a fullbandwidth image is indeed present, as it follows from the logic presented previously that the fullbandwidth image should be easily discriminable from the filtered image versions therefore leading to an increase in correct discrimination responses. This is not what was found. However it

remains the case that the sequence of filtered images which did not contain a fullbandwidth image target elicited more errors when it was presented from coarse to fine rather than the fine to coarse suggesting a temporally anisotropic integration of spatial frequencies.

A possible problem with the natural image discrimination task was reported by Parker, Lishman & Hughes (1997). They point out that the image presentation sequences were prone to flicker. They suggest that the subjective impression of the filtered image sequences was not one of three distinct images but rather of only one image coupled with flicker. In order to perform the task they propose that subjects firstly used the flicker as a cue to discriminating a degraded sequence from a fullbandwidth sequence, where all three images are the fullbandwidth image. Once this discrimination has been made subjects then make image quality judgements to assess the likelihood that a fullbandwidth image was present in the image sequence rather than discriminating the fullbandwidth image target within the filtered image sequence. Parker et al (1997) replicated the discrimination experiment of Parker et al (1992) adding a further condition of flicker to the fullbandwidth image sequence by varying the contrast of all three fullbandwidth images, so that one was normal contrast, the second was darker than normal and the third brighter. Their results are much the same as that of their earlier experiments. The coarse to fine presentation of spatial frequency information resulted in more fullbandwidth images sequences being reported whether the fullbandwidth image was present or not. While it was found in Parker et al (1992) that the last frame shown did not determine the response, this effect was more equivocal in this experiment with the difference between the two conditions specifically testing this not reaching significance.

In a second experiment Parker et al (1997) attempted to generalise the findings in experiment 1 to other images, as the same image had been used in all experiments of

Parker et al (1992) and experiment 1 of Parker et al (1997). They filtered and presented a facial image as in the previous experiments and replicated the fullbandwidth image discrimination experiment. They found that the presence of a fullbandwidth image is more often reported in a coarse to fine sequence of filtered image only on two occasions, one when a fullbandwidth image target was present and one when it was not present in the filtered image sequence. Otherwise there was no significant difference between the two coarse to fine and fine to coarse image sequences. Thus a coarse to fine integration was supported but not as dramatically as in their experiment 1. Indeed the results were far more equivocal.

Further evidence for a coarse to fine integration was found by Schyns & Oliva (1994). They also utilised natural images which were presented as fullbandwidth, low pass, highpass or hybrid stimuli. Hybrid stimuli are made up of the low spatial frequencies of one natural scene and the high spatial frequency of another which are combined into single images. They briefly present one of these stimuli in one interval and a fullbandwidth target in another. Subjects were required to make a yes / no matching response, i.e. was the same image presented in both intervals. They find that subjects show a better matching performance on the basis of the low spatial frequency information in the hybrid stimuli at brief exposure durations. As exposure duration increased matching performance was found to better to the fine spatial frequency information of the stimuli. This was shown not to be due to the poor detection of the high spatial frequency information at short exposure durations as d' in the low pass and high pass comparison cases showed that both stimuli were detected, suggesting that the high spatial frequency information was available as a basis for matching. In a second experiment an animated sequence of the hybrid stimuli were presented. Thus a coarse to fine sequence of one scene was shown concurrently with the fine to coarse sequence of another. This consisted of two hybrids being presented consecutively for a brief period

of time. For example if the fine information in the first hybrid represented a highway scene and the coarse information was that of a city scene then in the second hybrid these would switch such that the highway would be represented in the coarse spatial frequencies and the city in the fine. They found that subjects categorised the stimuli on the basis of the coarse to fine information more frequently than on the basis of the fine to coarse information.

The results of Parker et al (1992; 1997) and those from Schyns & Oliva (1994) suggests that the integration of spatial frequencies proceeds from coarse to fine. However Parker, Lishman & Hughes (1996) have provided evidence which suggests that the integration of spatial frequencies may also occur for a fine to coarse image sequence. In these experiments they used images of faces and objects. In the first experiment filtered images of faces were presented in two intervals. In one interval a fullbandwidth facial image was presented for 500 msec while in the other either a composite face or another fullbandwidth image was shown. The composite image comprised of a 100 msec presentation of a filtered face and a 400 msec presentation of a fullbandwidth face. The first facial image was either a filtered version of the fullbandwidth image succeeding it or was a filtered version of different face. Thus these can be considered as low and high spatial frequencies relevant and irrelevant cues. Further the fullbandwidth image presented in the latter part of the composite was either the same or different from that presented in the other interval. Thus the task was a same / different judgement. It can be hypothesised on a coarse to fine integration basis that the initial filtered facial image presented would guide the processing which succeeds it, thus the condition where this stimulus was the same as the fullbandwidth image presented after it would prove the least disruptive and elicit faster reaction times. They found that when fullbandwidth comparison stimuli are used reaction times are faster than when composite images are used suggesting that not having fullbandwidth information

impairs processing efficacy. Furthermore, they found no difference between the reaction times elicited when a low spatial frequency relevant cue precedes the fullbandwidth image and when it is preceded by a high spatial frequency relevant cue. They did find a difference in the reaction times when a high spatial frequency irrelevant cue preceded the fullbandwidth image. This condition elicited longer reaction times indicating an increased disruption in the processing efficiency of the succeeding fullbandwidth image in comparison to a low spatial frequency irrelevant cue. This result then provides evidence against the coarse to fine temporally anisotropic integration mechanism and seems to suggest a fine to coarse one in its stead. This finding was confirmed in a second experiment employing more varied object images which were specifically chosen as they contained more differences on a coarse scale, that is they specifically biased the sample to allow a coarse scale cueing to have a maximal effect. They found high pass irrelevant information to be more disruptive to the reaction times to fullbandwidth same / different judgements than the low pass irrelevant information. This was significantly different in the same judgements and a trend was found in the different judgements. In a further experiment these findings were shown not to be a result of the characteristics of the precues alone. They found that reaction times and errors to the cues presented alone and conditions when they were coupled with a fullbandwidth image differed significantly in a same categorisation task. The reaction times and errors were found to increase when the cues were presented alone. No similar enhancement of performance was found when a different categorisation task was employed. It is worth noting that reaction times to high spatial frequency cues alone were lower than low spatial frequency cues alone when in this same / different matching task, the exact opposite of that which would be expected given the spatial frequency processing delays (See Section 1.3.2). Overall then this set of experiments provides evidence for an advantage for the high spatial frequencies in the guiding of further processing in the

cases of face and object identification.

Parker et al (1996) suggest that these findings reflect Parish & Sperling's (1991) proposal that some spatial scales contain more useful information than others for performing different tasks. This suggestion can be seen to be a reflection of the different models of early visual processing which are required to process different types of information (See section 1.2.1). It is also reminiscent of the model of spatial scale selection put forward by Burbeck & Yap (1990) in which the scale ideal for performing the task is selected prior to visual processing, suggesting a match between spatial frequency processing and the task to be performed. The implication being that the integration of spatial frequencies could differ depending upon the task demands.

This idea of task constraints has also been suggested by Schyns & Oliva (Schyns & Oliva, 1994; Schyns & Oliva, 1997; Oliva & Schyns, 1995; Oliva & Schyns, 1997). An example of this is provided by Oliva & Schyns (1995; 1997). They presented hybrids of images in which two images are presented together, one as a low pass filtered image and the other as a high pass filtered image. In this experiment one group of subjects was presented with a low pass image and high pass structured noise, the other was presented with a high pass image and low pass structured noise in a 3 frame animation. The first frame of this animation contained spatial frequencies below 2 cpd and above 6 cpd. In the second frame the spatial frequencies were brought closer together with the low pass cut off point being shifted to 3 cpd and the high pass cut off point shifted to 5 cpd. Finally the third frame brought the cut off points of both the low pass and high pass filtered images in correspondence at 4 cpd. Subjects were asked to categorise these images into one of a number of predetermined categories such as highway or cityscape. This is carried out for six trials. Immediately after these six trials and without any discontinuity in the experiment further animated hybrids of two images was presented, e.g. a highway and a cityscape, one as the low pass filtered images and the other as the

high pass filtered images. Thus the structured noise presented on the preceding phase was replaced by a natural image. Subjects were again asked to categorise them. It was found that depending upon which temporal form of spatial frequency presentation subjects categorised images on in the first phase of the experiment dictated which temporal form they categorised on in the second phase. Thus if the subject categorised on the basis of the coarse to fine low pass filtered images in the first phase then they categorised on the basis of the coarse to fine low pass filtered images in the second phase; and if the subject categorised on the basis of the fine to coarse high pass filtered images in the first phase they categorised on the basis of the fine to coarse high pass filtered images in the second phase. This finding suggests that it possible to bias the temporal nature of the spatial frequency integration so that it takes place on spatial frequency information available from coarse to fine or fine to coarse depending upon which contained the information utilised for categorisation in the previous instances.

On the basis of these results Oliva & Schyns (1995) suggest that recognition is carried out by a flexible perceptual system which is constrained by the information demands of the task or the goals of the organism and the perceptually available information. This conception of matching the spatial scale of processing with the tasks demands has also been suggested from the results of other experiments by Oliva & Schyns (1995; 1997) where it has been found that uncategorised high pass images e.g. a highway, presented in hybrids with categorised low pass images e.g. a cityscape, can influence the speed at which a low pass version of the same image as shown in the high pass image of the preceding trial e.g. the highway, speeds the reaction time for that “highway” categorisation. Thus they found that the uncategorised filtered images of one trial can facilitate the categorisation of the next trial despite their being a change of scale. In a further experiment this was found to be perceptual (e.g. edge similarities) not conceptual (e.g. bedroom to different bedroom) level priming. Thus these results

suggest that the filtered images forming the coarse to fine and fine to coarse temporal presentation orders are integrated and that these integrated representations can influence processing across scale.

However, there are problems with both the flexible integration and task dependent conclusions. First, the results do not necessarily require the integration of the three hybrid filtered images presented to occur for the completion of this task. The results could be a function of the final frame only. Thus in the final frame of any of the ambiguous hybrids a low pass image is presented with high pass image both at a cut off point of 4 cpd. Either image provides enough information which can be utilised in order to successfully categorise it (Ginsburg, Cannon & Nelson, 1980; Fiorentini, Maffei & Sandini, 1983). Therefore the performance on this task does not necessarily reflect the temporal nature of spatial frequency integration across images presented from coarse to fine or fine to coarse as the final frame alone could be employed. Furthermore, it is the case that the final images presented have the full spectra of spatial frequencies and thus will involve spatial frequency processing delays. The final frame, containing two images, is processed from coarse to fine. In terms of the model presented in the preceding section both the low pass and high pass images are selectively filtered and processing delays are imposed such that the spatial frequencies become available to the second stage from coarse to fine. Further if it is assumed that the two images are separately represented, as would be suggested from a consideration of the edge detecting algorithms of Marr (1982) and Canny (1986), then both the spatial frequency content of the low pass and high pass images are available in this coarse to fine manner. Therefore the results of Oliva & Schyns (1995; 1997) may not reflect a flexible spatial frequency integration process but the operation of a coarse to fine one. Furthermore the biasing of spatial frequency use produced in the first phase of presentation also may not change the nature of spatial frequency integration but rather selectively biases the decision on to the

low pass or the high pass image depending upon which has been ideal for task completion on previous trials. This may reflect a shift in the decision process, or strategy employed by the subject, from one image to the other⁴, but does not reflect a flexible integration process. This concept of a final frame based explanation with no fine to coarse integration can account for both the biasing and priming results of Oliva & Schyns (1995; 1997).

Second the argument for the use of scales being matched to the task has not been shown as the use of coarse to fine and fine to coarse presented spatial frequencies had to be artificially biased which may have purely produced a strategy change for the completion of the experiment but which may not reflect general visual processing. Indeed, in a control experiment carried out by Oliva & Schyns (1997: see their footnote number 2) they found that when subjects were presented with hybrid animations which contained two scenes with no preceding sensitisation phase to the low pass or high pass filtered image sequences, that the categorisation of the images was approximately equal. Thus subjects categorised images on the basis of the low pass and the high pass image sequences equally often (averages were 53% categorised on the low pass images and 45% categorised on the high pass images). This suggests that categorisation of images can take place on the basis of the high or the low pass image which are both integrated from coarse to fine. Thus it is not clear that these tasks show any evidence for a flexible visual processing which matches the information presented in an image to that which would be optimal for the completion of the task.

1.3.4 Conclusions

In this section two concepts of spatial frequency integration have been introduced: temporal anisotropy and temporal isotropy. Where temporal anisotropic spatial frequency integration is one where spatial frequencies are integrated more effectively from coarse to fine and temporal isotropic spatial frequency integration is one where spatial frequencies are integrated as and when they become available. The evidence outlined suggests that the integration of spatial frequencies is anisotropic rather isotropic. In terms of the model presented in the preceding section, the spatial frequency processing delays introduced in the first stage filtering could ensure that this temporal form of spatial frequency availability occurs in the integration stage. The results of Oliva & Schyns (1995; 1997) suggest that the integration of spatial frequencies may be more flexible than this although there are problems with the interpretation of these experiments. However, these results coupled with the results of Parker et al (1996) are suggestive that spatial frequencies may be integrated more flexibly so that the integration of spatial frequencies occurs in which ever form is best suited for the completion of the task. That this may be so is also suggested from the early spatial frequency processing models outlined in section 1.2.1 where the nature of processing differs depending upon the tasks involved e.g. motion detection (Chubb & Sperling, 1989) and texture processing (Graham & Sutter, 1998). This reflects a more refined version of the temporally anisotropic mechanism in that the temporal form of the integration of spatial frequencies would be chosen to be most efficient depending upon the information required. For example, the integration of spatial frequencies could take a different form for the extraction of motion and the extraction of texture.

1.4 Overall Summary

The main points raised in this literature review have been concerned with the integration of spatial frequencies and its temporal form. Evidence has been presented which suggests that the initial stages of visual processing involves a spatial frequency filtering of the visual input on a patchwise basis (De Valois & De Valois, 1988; Graham, 1989). This filtering stage may then be followed by non - linearities and a second stage filtering prior to any decision processes (Wilson & Wilkinson, 1997). Evidence was then outlined suggesting that after the breakdown of the visual input into these component channels their information is then integrated (Georgeson & Meese, 1997). The temporal form which this integration takes was then examined. Evidence was presented showing that there are delays involved in the processing of spatial frequencies which may constrain the temporal form of spatial frequency integration so that the integration of spatial frequencies proceeds sequentially from the low spatial frequencies to the higher ones. However it was pointed out that the evidence for an initial coarse to fine availability of spatial frequencies does not mean that the optimal integration of these spatial frequencies only occurs when spatial frequencies are available from coarse to fine. The integration process may take place on any spatial frequencies which become available, but integration is constrained to occur in a coarse to fine manner because of the processing delays involved in the initial processing stage. Evidence was presented which examined the integration of spatial frequencies by temporally “stretching” their availability so that on successive frames spatial frequencies were presented from coarse to fine or fine to coarse. Initial experiments showed that spatial frequencies are more effectively integrated when they are made available from coarse to fine rather than fine to coarse suggesting a temporally anisotropic integration process (Parker et al, 1992; 1997). However, evidence was also outlined suggesting that integration may take place with spatial frequencies which were available sequentially from fine to coarse (Parker et al, 1996; Oliva & Schyns, 1997). Overall the results

suggest that the integration of spatial frequencies may be flexible and be task dependent but it was pointed out that the interpretation of some of the evidence supporting this flexibility was debatable.

The main aim of this thesis is to further examine the nature of the temporal form which the integration of spatial frequencies takes. In particular the existence of a temporally anisotropic integration process which integrates spatial frequencies more effectively when they are available from coarse to fine will be examined.

1.5 The aims and scope of the thesis

The main aims of the thesis are twofold. The first aim is to examine the form which the integration of spatial frequencies takes and the second is to examine the delays involved in initial spatial frequency processing. This work extends that carried out by Parker et al (1992; 1997) and Schyns & Oliva (1994; 1997; Oliva & Schyns, 1995; 1997) examining spatial frequency integration with natural images and band - limited Gabor patches. This work has implications for early visual processing models and assumptions made about the temporal form of spatial frequency integration.

The evidence reviewed examining the integration of spatial frequencies suggests two different forms which spatial frequency integration may take. One is where spatial frequencies are integrated more effectively when they are available from coarse to fine, a temporally anisotropic integration, and the other is where spatial frequencies are integrated in the temporal form which is optimal for the recovery of specific information from the input, a flexible integration. Spatial frequency integration as a temporally anisotropic process is suggested from the delays involved in the processing of spatial frequencies and the more specific evidence provided by Parker et al (1992; 1997) and

Schyns & Oliva (1994). The flexible integration of spatial frequencies is suggested from a consideration of early visual processing models which involve different processes for the extraction of different information from the visual input, suggesting that the integration processes involved may also differ. However, the experiments which have specifically suggested this form of integration are few and their interpretations are not clear. The next four chapters will further examine the temporal nature of spatial frequency integration. In particular the concepts of spatial frequency integration as a temporally anisotropic process or a flexible process will be examined.

The spatial frequency stimuli used in the next four chapters (chapters 2 – 5) are presented sequentially from coarse to fine or from fine to coarse as Parker et al (1992; 1997). In the next chapter the image quality and image detection experiments of Parker et al (1992) are carried out using natural images. The first experiment assesses the image quality judgements of sequentially presented triplets of filtered images whose spectral content progresses from coarse to fine or fine to coarse. The second experiment examines the detectability of a fullbandwidth image target which may or may not appear within a temporal sequence of filtered images. The third experiment examines the role of the discriminability of the filtered images from the fullbandwidth target for target detectability. The fourth experiment is a preliminary investigation into the impact that different images and power spectra slopes have on spatial frequency integration.

The third chapter extends the work utilising natural images with the same paradigms of image quality and image detection but with band limited Gabor patches as stimuli. The first experiment (Experiment 5) examines the detectability of a square-wave when presented in a sequence of the first three spatial frequencies of a square-wave series (F, 3F and 5F: the fundamental, third harmonic, and fifth harmonic respectively) shown sequentially from coarse to fine or fine to coarse. The second experiment (Experiment 6) examines the resemblance of the F, 3F and 5F to a square-

wave in a two alternative forced-choice experiment where subjects are presented with a coarse to fine then a fine to coarse sequence (or vice-versa) and indicate which appeared to most resemble a square-wave.

The fourth chapter examines the results of the square-wave resemblance experiment (Experiment 6) of Chapter 3 in more detail: examining the nature of the task (Experiment 7), the role of the final frame (Experiment 9 and 9a) and the role of masking (Experiment 10). Experiment 8 showed that the number of spatial frequencies presented was reducible to two with no significant difference from the results obtained from the use of three spatial frequencies, this allowed the examination of the interactions between the spatial frequencies carried out in experiment 9 and experiment 10.

The fifth chapter examines the integration of spatial frequencies with a vernier acuity task in three experiments (Experiment 11a, Experiment 11b and Experiment 12).

The conclusion from these four chapters (chapters 2 – 5) is that the integration is flexible and dependent upon the nature of the information which is to be derived from the visual input in order to optimally complete the task at hand.

The secondary aim of the thesis is to examine the delays involved in spatial frequency processing. This is examined in chapter 6. This extends the work examining the temporal form of spatial frequency integration assessing whether any other form of spatial frequency integration is possible in visual processing when spatial frequencies are not artificially made available in different temporal orders. This suggestion follows as the evidence for delays involved in spatial frequency processing would seem to constrain further processing to occur from coarse to fine. Therefore the findings for a spatial frequency integration flexibility will not have any general applicability. As this is an important constraint on the applicability of the research into spatial frequency integration further experiments examining spatial frequency processing delays were

carried out using a temporal order judgement technique introduced by Barr (1986) and Parker & Dutch (1987). Experiment 13 and 14 used the paradigm introduced by Parker & Dutch (1987) where temporal order judgements are made using Gabor patches as a match and an LED as a standard. Experiment 15 examined the temporal order judgements of two widely different spatial frequencies at multiples of contrast threshold. In experiment 16 the paradigm of experiment 15 is changed to a cycled presentation. Experiment 17 and 18 repeated the methods of experiments 15 and 16 with equally detectable stimuli.

Chapter 7 presents a review of the findings of the thesis and a discussion of their implications for the temporal nature of spatial frequency integration. Problems with the experiments presented and future experimental work is also discussed.

Spatial Frequency Integration and Natural Images

Summary

In this chapter the existence of a temporal anisotropy in the integration of spatial frequencies was examined using natural images. Two paradigms were employed to examine spatial frequency integration: Natural image quality; and Natural image detection. In both of these paradigms the natural images were low and high pass filtered and presented sequentially so that the spatial frequency information temporally progresses from coarse to fine or from fine to coarse. In the first paradigm the pictorial quality of the image sequence was rated. It was found that subjects rated the quality of image sequences whose spatial frequency content progressed from coarse to fine as being better than the same spatial frequency content sequences progressing from fine to coarse. This supports previous findings that the efficacy of the integration of spatial frequencies increases when presented from coarse to fine. However the data were well accounted for by a temporally isotropic weighted average model of spatial frequency integration. In the second paradigm the presence of the fullbandwidth image target appearing within an image sequence was assessed. On trials in which the target was not actually presented, it was found that target present responses were more likely when the spatial frequency content of the image sequence progressed from coarse to fine rather than fine to coarse, again suggesting a temporal anisotropy in the integration of spatial

frequencies. However, this finding was not evident in all conditions as has been previously reported (Parker, Lishman & Hughes, 1992). Reasons for this are explored. It is concluded that whilst the Natural image quality paradigm promotes the integration of spatial frequencies as the judgement of quality is made upon the whole sequence of filtered images the Natural image detection paradigm does not.

2.1 Introduction

As outlined in section 1.3.1 there are computational and empirical reasons for delaying high spatial frequencies and building an incremental representation of the image from the low spatial frequencies to the higher ones. It was pointed out that a great deal of evidence was supportive of a coarse to fine integration process while not actually examining the nature of spatial frequency integration itself (see section 1.3.2). The question to be addressed in this chapter is whether the integration of spatial frequency information depends upon the coarse to fine presentation of spatial frequencies for improved efficacy. In this chapter the two paradigms of natural image quality and natural image detection introduced by Parker et al (1992) will be employed and examined.

The evidence presented in section 1.3.3 from experimental work carried out by Parker et al (1992; 1996; 1997), Schyns & Oliva (1994; 1997) and Oliva & Schyns (1995; 1997) suggests either that the integration of spatial frequencies depends upon their coarse to fine availability for improved efficacy, or that the spatial frequency integration process is flexible depending upon task demands. These are the temporally anisotropic spatial frequency integration process and the flexible integration process introduced in section 1.3.3. The experiments in this chapter specifically examine the temporally anisotropic integration process, that is whether the integration of spatial

frequencies occurs most effectively when they are available from coarse to fine. In order to do this the paradigms of natural image quality and natural image detection, introduced by Parker et al (1992), are employed.

2.1.1 Interpretation of Image sequences

Both the natural image quality paradigm and the natural image detection paradigm involve conditions where the fullbandwidth image is presented within the image sequences. It has been found by Parker et al (1992) that subjects are more likely to indicate the presence of a fullbandwidth image target when these sequences are presented from coarse to fine thus supporting a temporally anisotropic integration process, but they suffer from interpretative problems. This is due to three reasons. First there is a large amount of information held in the triplets over and above that necessary for reconstruction. Second in all of these conditions a fullbandwidth image is included which contains all the spatial frequency information required in precisely the temporal sequence which is normally encountered. Third, it is not clear why a fine to coarse sequence which included a fullbandwidth image target in a natural image detection experiment would lead to a decrease in the targets detected. Given a temporally anisotropic system biased toward a coarse to fine integration a fine to coarse sequence of spatial frequency information should not be integrated at all well and because of this a fullbandwidth image target would be easily detected as there would be little interference from the filtered images presented sequentially from fine to coarse. Thus the only conditions which can be interpreted as representing a clear test of integration are those when no target was present.

2.1.2 Use of natural images in empirical research

There have been many arguments for the use of ecological stimuli in vision research (Gibson, 1950; 1979) and that to best understand the functioning of the human visual system would involve the examination of its performance in the natural environment in which evolved and developed (Barlow, 1961; Srinivasan, Laughlin & Dubs, 1982; Marr, 1982; Atick, 1992). The assumption is that the human visual system has evolved and developed to efficiently code the natural environment in order to optimise interactions with that environment given the demands placed upon the human species in the ecological niche which it occupies. Thus the nature of the sensory coding would reflect the structures inherent in natural images. The presumed unstructured nature of natural images has acted as a deterrent with most research concentrating on contrived tightly controlled unnatural stimuli. However recent close examination of the structure of natural images has revealed highly consistent properties in the underlying statistics of natural images thus supporting their use as stimuli in empirical research (Burton & Moorhead, 1987; Field, 1987; 1993; 1994; Tolhurst, Tadmor & Chao, 1992; Ruderman, 1994; 1997).

Areas within natural images which are in close proximity have been shown to be more likely to have properties which are related than those areas which are far apart (Field, 1987; Hancock, Baddeley & Smith, 1992). This necessarily involves the examination of the second-order statistics of natural images usually by analysing the power spectra of natural images which is related to the pairwise correlation (autocorrelation function) between pixels (Field, 1987; van der Schaaf & van Hateren, 1996; Ruderman, 1997). It has been found that the amplitude spectra of natural images have been found to approximate a $1/f$ slope or a $1/f^\alpha$ power spectrum slope (Field, 1987), where f is the spatial frequency and α is an exponent whose value has been found to vary between 1.88 ± 0.43 (Schaaf & van Hateren, 1996) and 2.4 ± 0.26 (Tolhurst et al, 1992) (Power spectra are discussed here, but these are related to amplitude spectra

by taking the square root of the power spectra). This similar exponent value shows a similar degree of spatial correlation within natural images which suggests that there is a similar amount of redundancy within all natural images. This redundancy of information in natural images matches theories of visual coding in which it is suggested that a primary goal of early visual processing is to reduce redundant information from sensory information (Barlow, 1961; 1989).

The commonalties between the statistical properties of natural images have also been found to match the coding carried out by the visual system thus suggesting that the link between the human visual system and the natural environment in which it evolved and developed is a causal one where the functioning of the visual system is a product of the environment in which it operates (Field, 1987; Olhausen & Field, 1996; Bell & Sejnowski, 1997; van Hateren & van der Schaaf, 1998). Thus the use of natural images in an empirical setting has recently been justified in terms of its underlying statistical regularity and match to the coding carried out by the visual system.

2.1.3 Preview of the chapter

The aim of this chapter was to examine whether the efficacy of spatial frequency integration improved when spatial frequencies were available from coarse to fine. The existence of a temporal anisotropy in the integration of spatial frequencies was examined utilising the paradigms of natural image quality and natural image detection, introduced by Parker et al (1992). In both of these paradigms a fullbandwidth natural image was employed along with four filtered versions of it (two low pass and two high pass). These images were initially presented singly in order to assess the discriminability of the filtered images from the fullbandwidth image and to allow subjects to familiarise themselves with the stimuli. These images were then grouped into triplets and presented

sequentially with their spectral content progressing from coarse to fine or from fine to coarse.

The first experiment reported involved an assessment of the apparent quality of these image sequences. Thus a triplet of images, which may or may not include the fullbandwidth image, was presented and subjects rated the quality of that image from 1 to 4. The spatial frequency content of those sequences and the order in which that content was presented was varied.

The second experiment assessed the detectability of a fullbandwidth image when presented in a sequence of filtered versions of itself. As with the first experiment, the spatial frequency content of those sequences and the order in which that content was presented was varied. The third experiment examined the effect that decreasing the discriminability of the filtered images from the fullbandwidth had on fullbandwidth image detection. The fourth experiment examined the effect that the use of a small number of other natural image exemplars with different power spectra slopes had upon detection performance.

2.2 Experiment 1: Natural image quality

The first experiment reported below was designed to investigate whether the integration of spatial frequencies is temporally anisotropic utilising the image quality paradigm employed by Parker et al (1992). In the first part (Experiment 1a) the fullbandwidth image and the filtered versions of it were individually rated for quality. This ensured that the filtered images were discriminable from the fullbandwidth image and also allowed the subjects to become familiar with the stimuli employed. In the second part (Experiment 1b) the images were grouped into triplets such that the spectral content of the triplets was presented sequentially from coarse to fine or from fine to

coarse. Given the results of Parker et al (1992) it was expected that a coarse to fine sequence of filtered images would lead to an increase in the quality ratings associated with those images when compared with the same images shown from fine to coarse.

2.2.1 *Experiment 1a: Quality of single images*

2.2.1.1 *Method*

Subjects. Ten observers, six male and four female with an age range of 23 to 34, took part in this experiment. All subjects had normal or corrected to normal vision. One subject was the author who was fully aware of the purpose of the experiment, the remaining nine were naive.

Apparatus and Materials. The main stimulus used in this experiment was a natural image, the remaining stimuli were filtered versions of this image. The root image was a picture of the foot of a hill¹ (See figure 2.1). The amplitude spectra of natural images have been found to approximate a $1/f$ slope or a $1/f^\alpha$ power spectrum slope (Field, 1987), where f is the spatial frequency and α is an exponent whose value has been found to vary between 1.88 ± 0.43 (van der Schaaf & van Hateren, 1996) and 2.4 ± 0.26 (Tolhurst, Tadmor & Chao, 1992). The exact value of α has been shown to be relatively unimportant unless it deviates too far from 2 (van der Schaaf & van Hateren, 1996). The scene image used in this experiment had an α value of 1.5, which although low is well within the expected variations found over a large number of natural images. This value was computed using Matlab by evaluating a least squares curve fitting procedure using the function:

$$A/f^\alpha$$

where f is the spatial frequency value, α is the exponent and A is a scaling constant which determines the overall image contrast (Ruderman, 1997).

The four filtered versions (two low pass and two high pass) were created using Matlab (See figure 2.1 for all images and figure 2.2 for image spectra). The root image (denoted as F) was transformed into the frequency domain, convolved with low pass or high pass Gaussian filters with a half-width half-height of 0.5 cpd (cycles per degree) (7.5 cycles per picturewidth)² (low pass denoted as $L0.5$ and high pass denoted as $H0.5$) and 1 cpd (13 cycles per picturewidth) (low pass denoted as $L1$ and high pass denoted as $H1$), then inverse Fourier transformed back into the spatial domain. The Gaussian filters are described as follows: The low pass filter:

$$G(r,\theta,f) = \exp(-r^2/2f^2),$$

and the high pass filter:

$$G(r,\theta,f) = 1 - \exp(-r^2/2f^2),$$

where r is the distance from the centre of the Gaussian, θ is the direction and f is the space constant or standard deviation of the normal distribution. The Gaussian filter is isotropic and for this reason the polar co-ordinates, θ , specifying direction from the centre are irrelevant and thus not included on the right hand side of the equation (Watt, 1988).

This procedure differs from that of Parker et al (1992). Therefore the actual spatial frequency content of the filtered images vary from those utilised by Parker et al

(1992). However, given the fact that these images are being employed to examine the integration of spatial frequencies the actual spatial frequency content should not be important.

The stimuli in this experiment were shown at a resolution of 256 x 256 pixels on a black background. They were displayed using a VSG display system (Cambridge Research Systems Ltd) and were presented on a Gamma corrected 21 inch (1024 x 768 pixels) monochrome Eizo flexi-scan 6600-M monitor with a refresh rate of 69 Hz. Stimuli were displayed for 42 msec (millisecond) each always with a step temporal onset and offset. They were displayed centrally and viewed binocularly at a distance of 0.4m. The subject's head was placed in a chin rest throughout the duration of the experiments.

There were two main differences between this study and that of Parker, Lishman & Hughes (1992). Firstly the viewing distance used here differed from that used by Parker, Lishman & Hughes (1992). This was due to the differences in the resolution of the monitors used. In order to maintain the 13 degrees of visual angle subtended by the image³ used by Parker et al (1992) the viewing distance had to be decreased. Secondly the number of filtered versions tested was decreased from the seven used by Parker et al (1992) to five. This is due the redundancy of the two images, one low pass and one high pass, which, although used by Parker, Lishman & Hughes (1992) in their discrimination experiment, played no further role in the detection experiment. As this experiment is designed to ensure that the filtered images are discriminable from the fullbandwidth target and to familiarise subjects with the test materials there was no purpose served by including these stimuli.

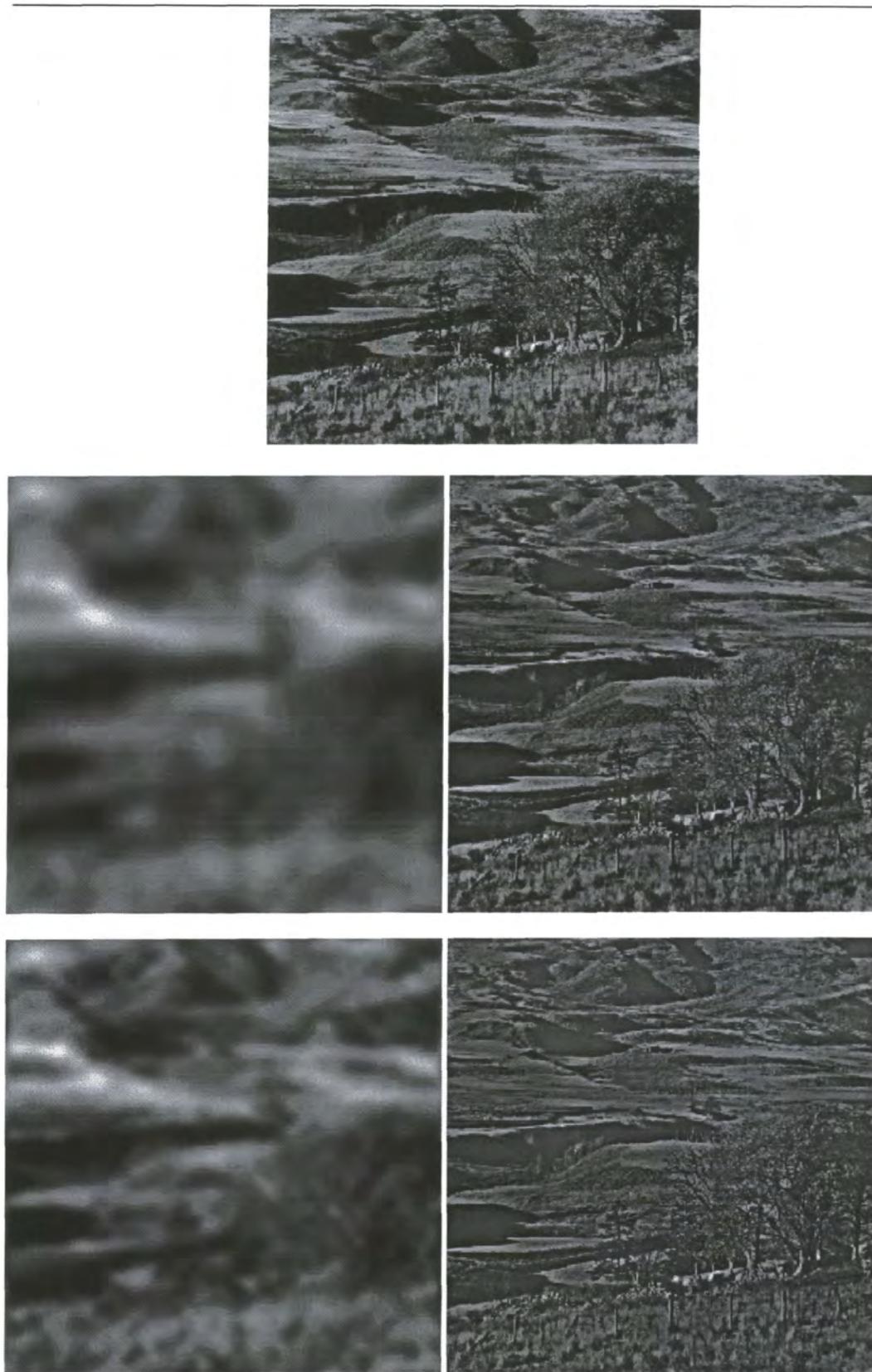


Figure 2.1: Images used in experiment 1. Top row = fullbandwidth image; Middle = LP 0.5 and HP 0.5; Bottom = LP 1 and HP 1. See text for details

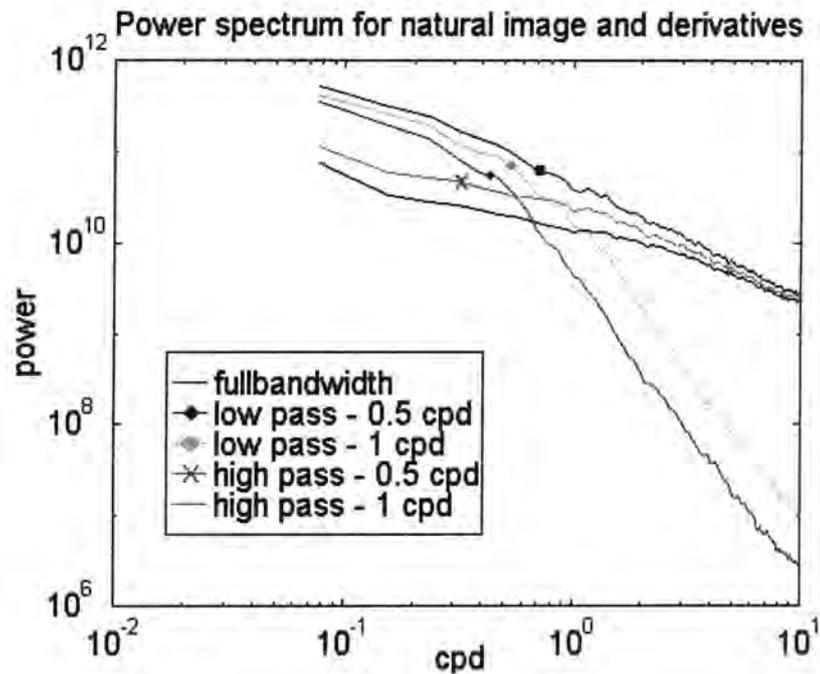


Figure 2.2: Power spectra of the images used in experiment 1. See text for details

Design and Procedure. The design of the experiment was a repeated measures one. The following images were presented singly in a random order: Fullbandwidth target, L0.5, L1, H0.5 and H1. Subjects were fully informed about the nature of the task and the stimuli involved. Example images and trials were shown. The subject initialised the experiment with a key press and also indicated the quality of the image presented on a scale from 1 to 4 where 1 is the lowest or worst quality and 4 is the highest or best quality. Each image was presented twenty times, equalling 100 trials overall. The experiment took place in a blacked out room.

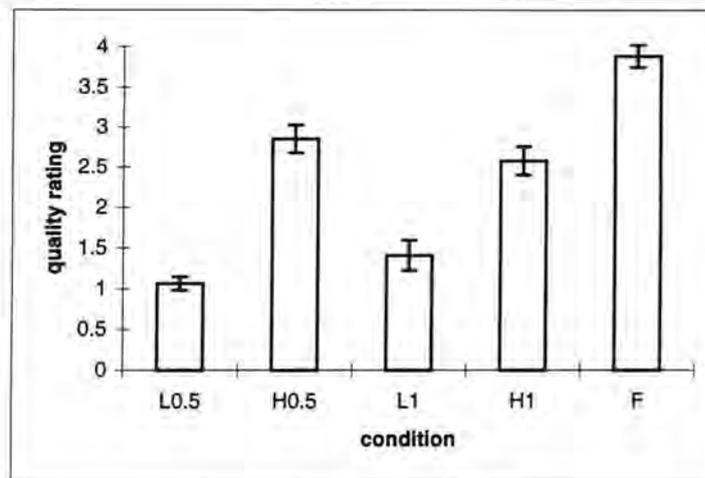


Figure 2.3: showing the judged quality of the filtered and fullbandwidth images. The vertical bars are standard error bars.

2.2.1.2 Results

The results can be seen in figure 2.3. This shows the mean quality ratings as a function of spatial frequency content of the image. An analysis of variance shows an overall effect of spatial frequency content on quality ratings ($F(4, 36) = 508.9; p < 0.01$). A series of planned comparisons was conducted. Overall the fullbandwidth image was rated as being of the highest quality (3.89). This was significantly different from the highest rated filtered version, H0.5 which was rated as 2.86 ($F_1 = 207.7; p < 0.01$). The low pass images, L0.5 and L1, were rated as being of significantly less quality than the high pass images, H0.5 and H1 ($F_1 = 857.8; p < 0.01$). Further the filtered images which contained the most spatial frequency information, L1 and H0.5, were rated as being of better quality than those having the least spatial frequency information, L0.5 and H1 (L0.5 vs. L1: $F_1 = 23.3; p < 0.01$ and H0.5 vs. H1: $F_1 = 14.8; p < 0.01$). It was also found that the highest rated low pass image was rated as being of significantly poorer quality than the highest rated high pass image ($F_1 = 408.8; p < 0.01$).

2.2.1.3 *Discussion*

This experiment shows that the filtered images were discriminable from the fullbandwidth image in terms of their quality. Subjects also seemed to make judgements of quality which were based upon the spatial frequency content of the image. The fullbandwidth image was rated as being of the highest quality with the high pass images being rated as higher quality than the low pass images. Within these groups of filtered images the highest content of spatial frequency information also dictates the quality rating given. This result replicates the findings of Parker et al (1992) using a different sample image. It provides further support for Parker et al's (1992) suggestion that this gives a quantitative basis on which high pass images can legitimately be said to be of better quality than the low pass images. Thus the edge information best typified by high pass images was a more important factor in image quality judgements than broader contrast changes provided by the low spatial frequencies. These quality ratings form the base from which the results of the integration of spatial frequencies found in the following experiment can be interpreted.

2.2.2 *Experiment 1b: Quality of triplet images*

In this experiment the stimuli used in experiment 1a were grouped into triplets. The subjects task was to rate the image quality of the entire sequence. Experiment 1a then provided the opportunity to rate the images which form the triplet sequences and allowed a comparison to be made between those ratings and those obtained for all three.

2.2.2.1 *Method*

Subjects, apparatus and materials are as experiment 1a. Differences are noted below.

Design. The design used in this experiment was a repeated measures one, using the same quality rating scale as experiment 1a. The stimuli from experiment 1a were grouped into triplets (126 msec per triplet). The five image sequences were: L0.5, L1, H1; L0.5, H0.5, H1; L0.5, L1, F; F, H0.5, H1; (and their converses) and finally F, F, F. The particular groupings used showed a progression of spatial frequency information sweeping from the coarse to fine or fine to coarse. For example, the grouping L0.5, L1, H1 if shown from coarse to fine showed low spatial frequencies initially followed by those low spatial frequencies plus medium spatial frequencies followed by the high spatial frequencies. This sequence of spatial frequency information presentation mirrors the natural progression of spatial frequency processing carried out by the human visual system. Overall then there were nine conditions. Subjects made twenty decisions on each.

Procedure. Subjects were fully informed about the nature of the task and were familiar with the stimuli involved. Example images and trials were shown. The subject initialised the experiment with a key press. Each trial consisted of a single presentation interval containing three of five possible images. Once the triplet had been shown subjects indicated their quality judgement of the image from 1, the worst quality to 4, the best quality. The term “quality” was not defined in experimental instructions. The experiment took place in a blacked out room.

2.2.2.2 Results

Figure 2.4 shows the mean quality ratings of the triplets as a function of temporal sequence. A two way analysis of variance on the factors image content and the presentation order of that content showed a significant effect of image content regardless of the order ($F(3, 27) = 118.2$, M.S.e.(mean square error) = 10.1; $p < 0.001$) and an

effect of temporal order regardless of content ($F(1, 9) = 21.1$, $M.S.e. = 4.9$; $p < 0.01$). A significant interaction was also found ($F(3, 27) = 18.4$, $M.S.e. = 9.6$; $p < 0.001$). This indicates that subject performance depended upon the spatial frequency content of the image triplet shown and the order in which the sets were shown, while the interaction shows that the effect of the temporal sequence varies depending upon the triplet shown. Planned comparisons were also carried out for the four critical temporal sequences (e.g. L0.5, L1, H1 vs. H1,

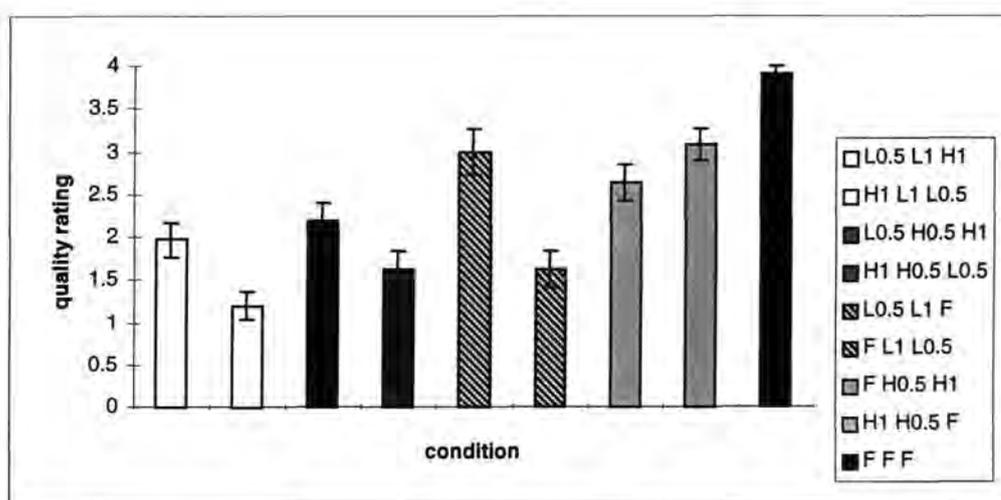


Figure 2.4: Mean judged quality of the image triplets. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The triplets are grouped by spatial frequency content regardless of order.

L1, L0.5). It was found that L0.5, L1, H1 ($F_1 = 38.9$; $p < 0.01$) L0.5, H0.5, H1 ($F_1 = 29.0$; $p < 0.01$) and L0.5, L1, F ($F_1 = 126.7$; $p < 0.01$) sets showed significant order effects with the coarse to fine progression of spatial frequencies being rated higher while F, H0.5, H1 ($F_1 = 13.5$; $p < 0.01$) showed significant order effects for a fine to coarse progression.

The ratings for the single images conducted in experiment 1a provide a comparative basis on which to judge these results. This can be understood by

considering that each triplet in this experiment was formed from the images presented in the first experiment each of which has an independent rating. The independent ratings can be averaged to allow a comparison to the rating for the triplet which they formed. Thus L0.5 L1 H1 were given ratings of 1.07, 1.41 and 2.58 respectively which average to 1.69 and L0.5 H0.5 H1 were given ratings of 1.07, 2.86, 2.58 which average to 2.17. The remaining conditions, L0.5 L1 F and F H0.5 H1 had average quality ratings of 2.12 and 3.11 respectively. These mean ratings can be correlated with the mean ratings found when they are grouped as triplets as a function of their temporal order. The actual ratings given were: L0.5 L1 H1 = 1.98; H1 L1 L0.5 = 1.21; L0.5 H0.5 H1 = 2.19; H1 H0.5 L0.5 = 1.63; L0.5 L1 F = 3; F L1 L0.5 = 1.62; F H0.5 H1 = 2.64; H1 H0.5 F = 3.08 (These values are shown alongside the mean ratings derived from their single image quality ratings in table 2.1). If significant relationships are found then it suggests that the ratings of the images when presented singly are crucial for the quality judgements when they are grouped. The mean quality ratings for the single images were correlated with the coarse to fine and fine to coarse progressions separately to give respective values of $r = 0.4619$ ($N = 4$; $p > 0.05$) and $r = 0.9914$ ($N = 4$; $p < 0.01$).

2.2.2.3 Discussion

The judged quality of image triplets was found to significantly improve when they were presented from coarse to fine rather than fine to coarse. This suggests that the coarse to fine presentation of spatial frequencies was more effectively integrated than the fine to coarse presentation order. This finding supports the model of early human visual processing of spatial frequencies involving a temporal anisotropy in the integration of spatial frequencies.

Sequence	Mean of single image ratings	Triplet ratings
L0.5 L1 H1	1.69	1.98
H1 L1 L0.5	1.69	1.21
L0.5 H0.5 H1	2.17	2.19
H1 H0.5 L0.5	2.17	1.63
L0.5 L1 F	2.12	3
F L1 L0.5	2.12	1.62
F H0.5 H1	3.11	2.64
H1 H0.5 F	3.11	3.08

Table 2.1: Shows the ratings given for image triplets. The first column of data shows the mean rating given to the single images which make up the sequences as taken from experiment 1a. The second column of data shows the actual triplet image ratings found in this experiment.

Further support for this was also found when comparing the mean quality ratings for the constituent images with their overall quality when presented as triplets. Parker et al (1992) found that the quality ratings of the image triplets presented from coarse to fine correlated significantly with the mean of their constituent image quality ratings. They suggested that this significant correlation showed that the quality of the individual images was being integrated to form the overall rated quality. However this may not be the case. If the image triplet quality is dependent upon the quality of the individual images then this suggests that it is the quality of those single images which determine

the image triplet quality rating. If the quality of the image triplet is higher than the mean of the individual images this suggests that the image triplet formed by the single images is of a quality over and above that determined by the single images. Thus it is when the image triplet quality is not correlated with the mean of constituent images and in fact is higher than their individual quality ratings that the single images can be proposed to be integrated. In this experiment the mean ratings of the single images were found to correlate strongly with the ratings for the image triplets if they progressed from fine to coarse but not if they progressed from coarse to fine.

Thus the relationship between the rated quality of the image triplets and their constituent images was found to be significant only for the fine to coarse presentation order. This significant correlation suggests that the quality rating given to the fine to coarse sequences depended upon the quality ratings given to the images which made up the triplet when presented individually. The lack of a significant correlation found when the spatial frequency content of the image progressed from coarse to fine showed that the quality ratings did not vary as a function of the content of the triplet. If the actual ratings are examined, in three of the four cases the quality rating given to the triplet was greater than that produced by taking the average of the single image ratings. Thus the ratings given for those coarse to fine conditions are higher than would be suggested from the individual ratings. It can be suggested then that the triplets are rated as being of better quality due to the integration of the spatial frequency content. The images presented in the coarse to fine sequences are integrated to form a better quality image than when those same images are shown from fine to coarse.

However such a conclusion should be taken as an indicator only as this analysis included conditions in which the fullbandwidth image was present. As pointed out in section 2.1.1 the conditions in which the fullbandwidth image was present are difficult to interpret. Therefore this lack of a correlation between the averaged quality ratings of

the coarse to fine sequence and the actual quality rating that that sequence elicited is suggestive of a coarse to fine anisotropy in the integration of spatial frequencies but not conclusive.

Despite this questioning of the correlation the results suggest that overall the successful integration of spatial frequencies seems to depend upon a temporal anisotropy where spatial frequencies are available from coarse to fine. The presentation of spatial frequencies from fine to coarse showed no evidence of integration taking place.

An alternative model, however, is one where the filtered images of the triplet sequences are integrated regardless of their temporal order, that is an isotropic model where spatial frequencies are integrated as and when they become available. If it is assumed that the signal derived from each filtered image presented within the triplet sequences decays over time then it will be the decaying signals which will be integrated. A simple way to depict this decay is in terms of weighting each of the filtered images relative to its position in the triplet sequences such that the first image shown has the largest decay and therefore contributes the smallest signal to the overall sum, that is it has the smallest weight. The second filtered image is given a correspondingly higher weight as it has a short time to decay and finally the last filtered image shown has the largest weight. In terms of the quality rating experiment this weighting model transfers easily to the quality ratings given to each filtered image where those ratings are weighted depending upon when they presented in the sequence. These weighted quality ratings are then summed and averaged. If this simple operation is carried out, with weights of 0.3, 0.6 and 1 for the first to final image shown respectively, the resulting quality ratings are: L0.5 L1 H1 = 2.23; H1 L1 L0.5 = 1.27; L0.5 H0.5 H1 = 2.71; H1 H0.5 L0.5 = 1.75; L0.5 L1 F = 3.09; F L1 L0.5 = 1.31; F H0.5 H1 = 2.80; H1 H0.5 F = 3.63. This can be compared with the actual quality ratings given to each of these triplets

in table 2.1. The weighted average and actual quality ratings were found to correlate highly ($r = 0.94$; $n=8$). This weighted sum model accounts remarkably well for the data without assuming any temporal anisotropy.

This concept of the weighted average model differs from the simple averaging model, where the quality ratings are integrated regardless of when they were presented, in that it suggests that when a significant correlation is found then the weighted average model is a good predictor of performance. However when a significant correlation is found between the triplet ratings and the simple averaging model then this is not a good predictor of performance. It was argued that the integration of the three filtered images presented as triplet sequences could be expected to give a better quality rating than is predicted from the simple averaged sum of the triplet ratings given to the filtered images when they were presented individually, that is their integration produces something which is not inherent in the individual images. The weighted average model does not suggest this it simply predicts that the quality rating given to a triplet sequence will be a weighted average of the quality ratings given to the component filtered images. It allows an easily quantifiable predictor of quality ratings. For this reason and the better account for the overall data it is this model which is to be preferred. Thus the results of experiment 1b can be accounted for much better by a temporally isotropic model where the quality ratings given to the triplet sequences are simple weighted averages of the quality ratings given to the component filtered images.

2.3 Experiment 2: Natural image detection

This experiment examined the detectability of the presence of a fullbandwidth image within a triplet of spatial frequencies. This is the second paradigm introduced by Parker et al (1992) where it was found that a coarse to fine sequence of spatial frequencies increased the likelihood of a target being detected whether it was in fact

present within the sequence or not. The experiment follows the pattern set out by the quality judgement experiment. First, the discrimination of the individual filtered images from the target fullbandwidth image was assessed. These single images were then grouped into triplets which may or may not have included the target. The task of the subject was to detect the presence of the target image at any point in the sequence. Given the results of Parker et al (1992; 1997) it would be expected that the coarse to fine sequence of filtered images should lead to an increase in the error rates associated with those images when compared with the same images shown from fine to coarse. The results from the quality judgement experiment also suggest that a coarse to fine presentation of spatial frequencies will increase target present responses. This experiment also allows a chance to replicate the finding of Parker et al (1992; 1997) for order preference in the availability of spatial frequencies with a different image exemplar.

2.3.1 Experiment 2a: Discrimination

2.3.1.1 Method

Subjects. Ten observers, six male and four female with an age range of 23 to 30, took part in this experiment. One subject was the author who was fully aware of the purpose of the experiment, the remaining nine were naive.

Apparatus and Materials. All apparatus and materials are equivalent to those used in experiment 1. The four filtered versions (two low pass and two high pass) differed in their spatial frequency cut off point. The root image (denoted as F) was transformed into the frequency domain, convolved with low pass or high pass Gaussian filters (as described previously) with a half-width half-height of 2 cpd (26 cycles per picturewidth)

(low pass denoted as L2 and high pass denoted as H2) and 4 cpd (52 cycles per picturewidth) (low pass denoted as L4 and high pass denoted as H4), then inverse Fourier transformed back into the spatial domain. (See figure 2.5 for images and figure 2.6 for spectra).

Design and Procedure. The purpose of this experiment was to measure the discriminability of the filtered images from the fullbandwidth image when they were individually presented. This ensures that when the images were presented sequentially in triplet sequences subjects were not basing their detection responses on individual filtered images which were easily confused with the fullbandwidth image target. A repeated measures design was adopted. The following images were presented randomly and sequentially: Fullbandwidth target, L2, L4, H2 and H4. Subjects were fully informed about the nature of the task and the stimuli involved. Example images and

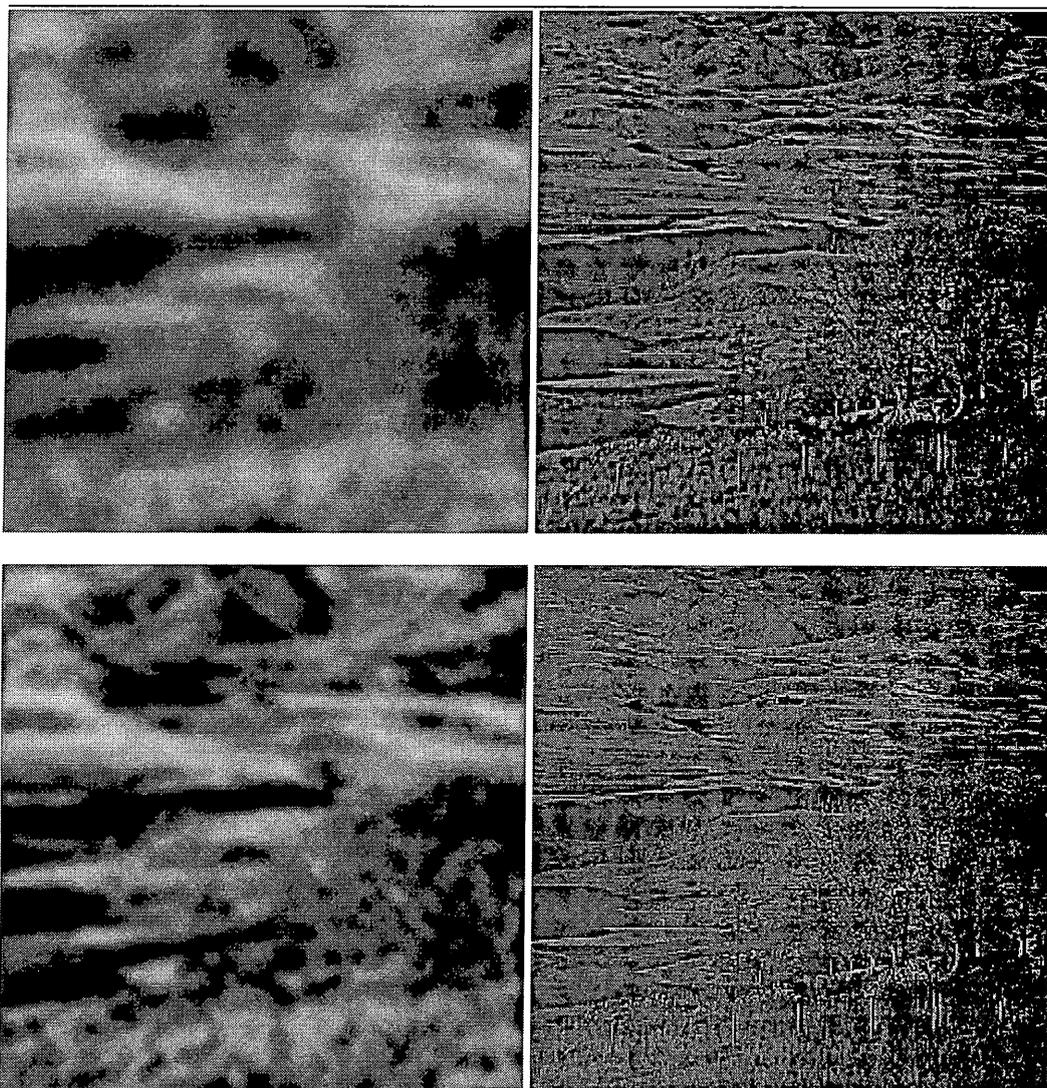


Figure 2.5: Images used in experiment 2. Top row = LP 2 and HP 2; Bottom = LP 4 and HP 4. For fullbandwidth image see figure 2.1. See text for details

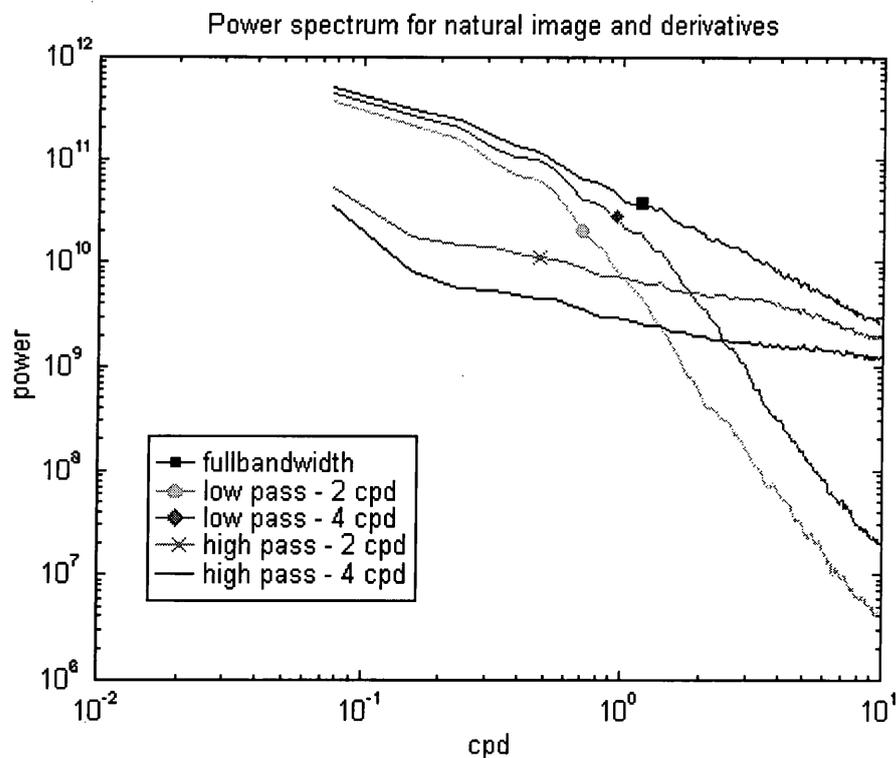


Figure 2.6: Power spectra of the images used in experiment 2. See text for details

trials were shown. The subject started the trials with a key press and also indicated which image type had been presented (either the target or a filtered distracter) by a key press. Each image was presented twenty times, equalling 100 trials overall. The experiment took place in a blacked out room.

2.3.1.2 Results

The results can be seen in Figure 2.7. This shows the mean number of trials in which subjects reported a fullbandwidth image as a function of the image type.

Overall the filtered images are rarely mistaken for the fullbandwidth image: on average the filtered versions were mistaken for the fullbandwidth image on 1.125% of trials. The filtered version H2 which contains the largest bandwidth of high spatial frequency information was found to be the most likely to be reported as a

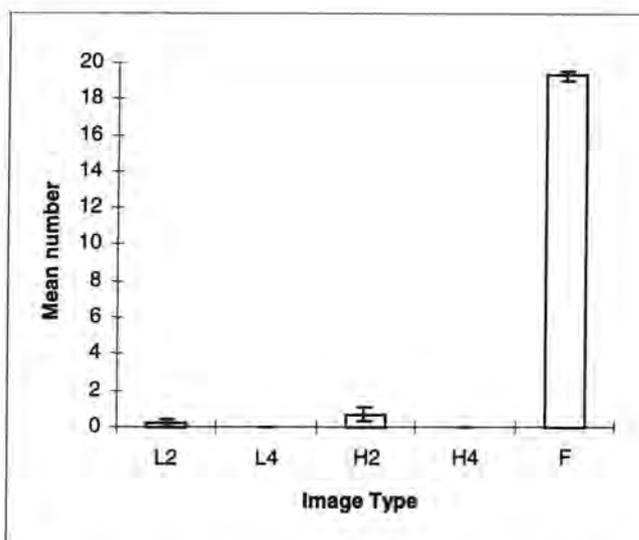


Figure 2.7: Mean number of fullbandwidth presented responses as a function of image type. The vertical bars are the standard errors associated with each mean.

fullbandwidth image, although this was still very infrequent at 3.5%. The fullbandwidth image itself was identified correctly on 96% of the trials.

2.3.1.3 Discussion

This experiment shows that the filtered versions of the fullbandwidth image target are very rarely mistaken for the target itself. Thus subjects are able to discriminate the fullbandwidth image and this provides a baseline on which to judge the results of experiment 2b. The image and filtered versions used in this experiment show a different performance to that found in Parker, Lishman & Hughes (1992). A comparison with Table 2.2 showing the discriminability of the stimuli used in Parker et al (1992) shows that the stimuli employed here were very easily discriminable from the fullbandwidth

Image	Current Experiment	Parker et al (1992)
L2 (L1)	1	3.7
L4 (L3)	0	7.4
H2 (H3)	3.5	38
H4 (H1)	0	18.5
F	96.5	92

Table 2.2: Shows the total percentage of trials on which the filtered images were not discriminated from their fullbandwidth versions in this experiment and Experiment 4 of Parker et al (1992). The percentages from Parker et al (1992) were not all explicitly reported and are thus calculated from the graph which they presented. The image names not in brackets are those used in the current experiment while those within brackets are those used by Parker et al (1992) for the corresponding image.

image. In their experiment the high pass filtered versions were much more confusable. The filtered image which contained the most high spatial frequency information was again one of the images most mistaken for the fullbandwidth image but on a much higher number of occasions (38%). In fact all the images used by Parker et al (1992) were more often confused with the target than the stimuli used here. Indeed the most discriminable image used by Parker et al (1992), their low pass filtered image with the lowest spatial frequency content, showed a similar number of errors evoked (3.7%) to the least discriminable image employed here, the high pass filtered image with the highest spatial frequency content, which had an error rate of 3.5%. Thus the filtered images were found to be very easily discriminable from the fullbandwidth image from which they were derived.

In the following experiment the detection of the fullbandwidth image is assessed when presented within the image triplets which vary in their spatial frequency content and the order in which that content is presented. The conditions of interest are those which contain no target. Experiment 2a, then, served both as a measure of the discriminability of the filtered images and as an opportunity for subjects to familiarise themselves with the target and variations of the distracters.

2.3.2 *Experiment 2b: Detection*

In this experiment the stimuli used in Experiment 2a were grouped into triplets. The subjects task was to judge the presence of the fullbandwidth image target in the triplet sequence. The groupings of the filtered images allowed the comparison of the detection performance as a function temporal order. It was predicted on the basis of previous results that the coarse to fine temporal sequence would elicit higher error rates in the detection of the fullbandwidth target.

2.3.2.1 *Method*

Subjects, apparatus and materials were as experiment 1a. Differences are noted below.

Design. A repeated measures design was used in this experiment using a simple yes / no decision. The stimuli from Experiment 2a were grouped into triplets (126 msec per triplet). The five image sequences were: L2, L4, H4; L2, H2, H4; L2, L4, F; F, H2, H4; (and their converses) and finally F, F, F. Each of the five image sequences show spatial frequencies progressively from coarse to fine or fine to coarse. Overall there were 9 conditions including the condition where all three images shown were of the fullbandwidth image F, F, F. Each stimulus was presented twenty times.

Procedure. Subjects were given all the information they needed about the task until they were completely satisfied that understood the form the experiment would take. This included showing examples of the images to be used, with which they were already familiar, and the trials with which they would be presented. It was pointed out that the images which they encountered in the discrimination experiment would again be shown but this time as triplets. The target fullbandwidth image was presented and subjects were told that this might or might not appear within the triplet of images which would be presented on each trial. The subject started the trials with a key press. Each trial consisted of a single presentation interval containing three of five possible images. Once the triplet had been shown subjects indicated whether the target image had been presented at any time in the sequence by further key presses. The experiment took place in a blacked out room and lasted approximately 15 minutes.

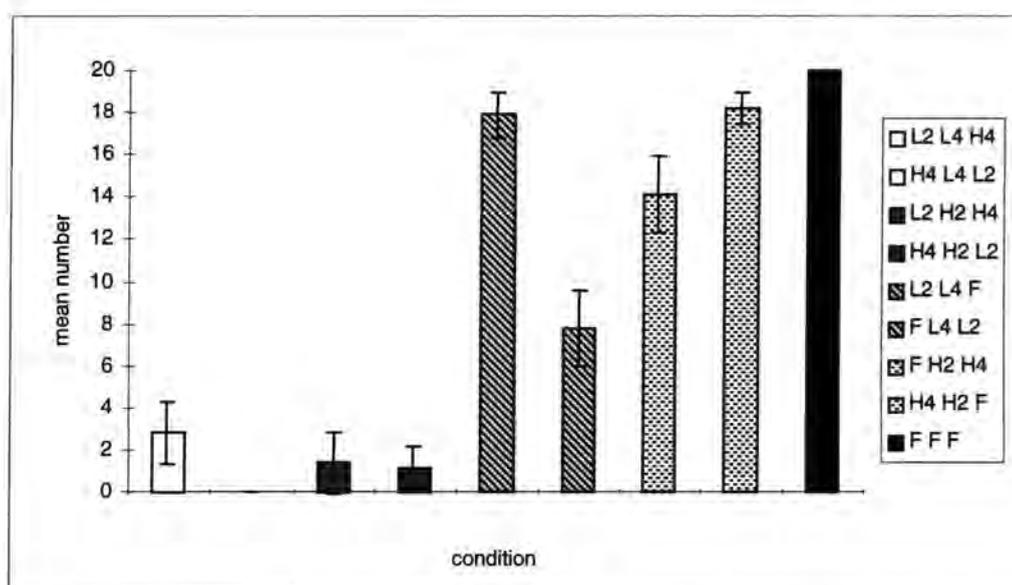


Figure 2.8: Mean number of fullbandwidth present responses out of 20 maximum as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that triplets containing the same images are together regardless of the order in which they are shown. The vertical bars are standard error bars associated with each mean.

2.3.2.2 Results

Figure 2.8 shows the mean number of reports of the presence of a fullbandwidth image as a function of temporal sequence. Overall there was a larger number of correct target present reports when the fullbandwidth image was presented at a point during the sequence especially when it was presented as all three components (F, F, F). As with Parker et al (1992) a two way analysis of variance on the factors image content and the order of that content was carried out. This showed a significant effect of image content regardless of the order ($F(3, 27) = 118.2$, M.S.e. = 10.1; $p < 0.001$) and an effect of temporal order regardless of content ($F(1, 9) = 21.1$, M.S.e. = 4.9; $p < 0.01$). A significant interaction was also found ($F(3, 27) = 18.4$, M.S.e. = 9.6; $p < 0.001$). This indicates that subject performance depended upon the spatial frequency content of the image triplet shown and the order in which the sets were shown, while the interaction shows that the effect of the temporal sequence varies depending upon the triplet shown. Planned comparisons were carried out for the four temporal sequences (e.g. L2, L4, H4 vs. H4, L4, L2). It was found that L2, L4, H4 ($F_1 = 4.1$; $p = 0.05$), L2, L4, F ($F_1 = 53.0$; $p < 0.01$) and F, H2, H4 ($F_1 = 8.7$; $p < 0.01$) sets showed significant order effects while L2, H2, H4 ($F_1 = 0.05$; $p > 0.05$) did not.

2.3.2.3 Discussion

Unlike Parker et al (1992) the results, across all conditions, showed no evidence that a coarse to fine temporal order of spatial frequency increased the detection of a fullbandwidth natural image. Of primary importance here was the comparison between the coarse to fine and fine to coarse image orders when the target was not present. The coarse to fine sequence L2, L4, H4 clearly showed an increase in the number of errors evoked when compared to the fine to coarse temporal order H4, L4, L2. However, the



coarse to fine sequence L2, H2, H4 showed no increase in the number of errors (i.e. false fullbandwidth present reports) evoked when compared to its fine to coarse temporal order. It should be noted that these results mirror those found by Parker et al (1997) in experiment 2 using a face stimulus where it was found that the two conditions on which no target was presented showed a coarse to fine bias for the L2, L4 H4 condition (their L1, L2, H1 condition) but no bias at all for the L2, H2, H4 condition (their L1, H2, H1 condition).

Thus the evidence from conditions in which no target appears could be interpreted as supporting either a temporally anisotropic or a temporally isotropic integration mechanism. In condition L2, L4, H4 the temporally anisotropic process of coarse to fine integration was clearly found while in the L2, H2, H4 condition both the coarse to fine and fine to coarse temporal order evoked similar error rates suggesting an isotropic integration process. However given the support for a temporally anisotropic integration mechanism found by Parker et al (1992; 1997) and experiment 1 presented in section 2.2 of this chapter, it seems likely that although evidence for an isotropic integration mechanism was found in this experiment this may be due to the very low error rates found in this experiment as compared to Parker et al (1992; 1997). Indeed this was the explanation hinted at by Parker et al (1997) for precisely the same pattern of results found in their experiment 2 where the overall lower pattern of errors was suggested as an explanation for the lack of consistent support found for a coarse to fine integration process. Overall then the results show support for a temporally anisotropic integration process dependent upon a coarse to fine presentation of spatial frequencies. This was found not to be as strong as has been previously found which may be a result of the low error rates over all conditions.

The results from the first part of this experiment show a clear difference between the ratings of images used here and those used by Parker et al (1992) which may account

for the differences in the results found in experiment 2b. It is possible that the larger error rates found by Parker et al (1992) were due to the reduced discriminability of the filtered images which formed the image triplets. The reduced error rates found in this experiment would then be a function of the increased discriminability of the individual filtered images found in experiment 2a and not just spatial frequency integration. Thus if this discriminability was reduced, then it would be expected that the error rates evoked by the filtered images would increase. Interestingly if this was the case then the error rates found in this experiment reflect the operation of the spatial frequency content integration process alone.

In order to examine the possibility that the discrimination error rates had an effect on the detection errors made, the next experiment replicates the design of the first experiment utilising less discriminable filtered images. It would be expected that the error rates evoked by the image triplets would rise in accordance with the error rates evoked by the filtered images. If the discrimination error rates of the individual filtered images do not have an influence on the number of detection errors made then no difference in the levels of integration found in this experiment should be found in the next.

2.4 Experiment 3: Natural image detection - decreased discriminability

The replication in the previous experiment of the discrimination and detection experiments of Parker, Lishman & Hughes (1992) resulted in no clear support for an integration mechanism with a coarse to fine temporally anisotropic preference for spatial frequency information. This may have been due to the ease by which subjects could discriminate between the target fullbandwidth natural image and the distracting filtered

versions. In order to test this suggestion the natural image discrimination and detection experiments were repeated with refiltered versions of the image used in experiment 1.

2.4.1 *Experiment 3a: Discrimination*

2.4.1.1 *Method*

All experimental details are identical to experiment 2a except where noted.

Subjects. Ten observers, four male and six female with an age range of 23 to 34, took part in this experiment.

Apparatus and Materials. The natural image target was the same as in experiment 2. The filtered versions differed. The same technique for filtering was employed as were the same filters however the space constants of the Gaussians differed. The half height of the filters was shifted from 2 and 4 cpd back to 0.5 and 1 cpd used in experiment 1 to assess judged image quality. Thus the four filtered images were low pass 0.5 cpd and low pass 1 cpd (6.5 and 13 cycles per picturewidth respectively) (denoted as L0.5 and L1) and high pass 0.5 cpd and high pass 1 cpd (denoted as H0.5 and H1)(See figure 2.1 and figure 2.2).

2.4.1.2 *Results*

The results are shown in figure 2.9. It can be seen that the confusion between the filtered versions and the fullbandwidth image has increased to a level more comparable to that of Parker et al (1992). The high pass images both show false positives with H1 resulting in 11% errors and the higher spatial frequency content image, H0.5, resulting in 26.5% errors. The fullbandwidth image is recognised as being so in 96.5% of trials.

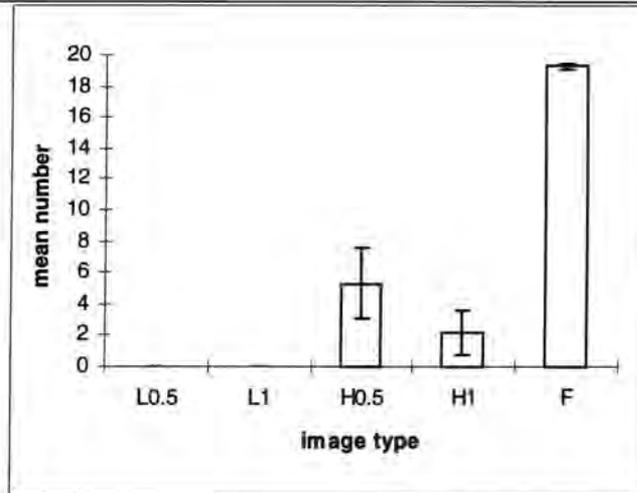


Figure 2.9: Mean number of fullbandwidth presented responses as a function of image type. The vertical bars are the standard errors of the associated means.

2.4.1.3 Discussion

The results show that the refiltering of the images has increased the confusion rates between the distracters and the targets without detracting appreciably from the detection rate elicited by the target itself. The goal of increasing the confusion rates has been achieved, although this is still not to the level shown by the stimuli employed by Parker et al (1992) where the high pass filtered version with the largest amount of spatial frequency information resulted in being mistaken for the target 38% of the trials (see table 2.2).

2.4.2 Experiment 3b: Detection

2.4.2.1 Method

The methods were as in experiment 2b. The stimuli used were those employed in experiment 3a. The image sequences were: L0.5, L1, H1; L0.5, H0.5, H1; L0.5, L1, F; F, H0.5, H1 (and their converses) and F, F, F.

2.4.2.2 Results

The mean number of reports of the presence of a fullbandwidth image are shown in figure 2.10 as a function of temporal sequence. As with experiment 2b there was a larger number of correct target present reports when the fullbandwidth image was presented at a point in the image sequence especially when it was presented as all three components (F, F, F). As with Parker et al (1992) a two way analysis of variance carried out on the factors image content and the order of that content shows a significant effect of image content regardless of the order ($F(3, 27) = 106.3$, M.S.e. = 6.6; $p < 0.05$) and an effect of temporal order regardless of content ($F(1, 9) = 7.7$, M.S.e. = 11.8; $p < 0.05$). A significant interaction was also found ($F(3,27) = 120.9$, M.S.e. = 2.9; $p < 0.05$). This indicates that subject performance depended upon the spatial frequency content of the image triplet shown and the order in which the sets were shown, while the interaction shows that the effect of the temporal sequence varies depending upon the triplet shown.

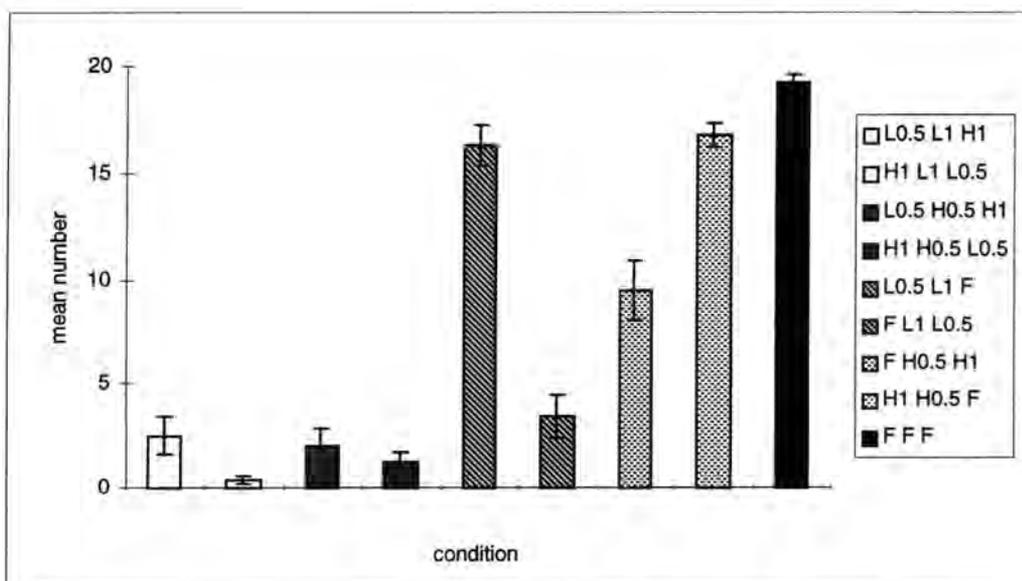


Figure 2.10: Mean number of fullbandwidth present responses out of 20 maximum as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that triplets containing the same images are together regardless of the order in which they are shown. The vertical bars are standard error bars associated with each mean.

The same planned comparisons as those performed in experiment 2b were carried out for the four critical temporal sequences (e.g. L0.5, L1, H1 vs. H1, L1, L0.5). It was found that L0.5, L1, H1 ($F_1 = 7.7$; $p < 0.05$) and L0.5, L1, F ($F_1 = 292.0$; $p < 0.05$) sets showed significant order effects for a coarse to fine order. While F, H0.5, H1 ($F_1 = 93.5$; $p > 0.05$) showed a significant order effect for a fine to coarse order. L0.5, H0.5, H1 ($F_1 = 1.1$; $p > 0.05$) did not show any significant order effect.

2.4.2.3 Discussion

The results show no evidence for an overall increase in the error rates elicited by the image triplets suggesting that the difference between the results found here and those found by Parker et al (1992; 1997) are not due to the decreased discriminability of the images used. Furthermore, evidence was found for both a coarse to fine and a fine to coarse effect of order upon the detection of a natural image target. When the target was present the same pattern of results as experiment 2b was found: the order of the spatial frequency information shown does not seem to have a consistent effect upon target detection. Both coarse to fine and fine to coarse order effects were found. However, as suggested before these conditions suffer from interpretative problems. When the target is not present no order effect is shown in one condition (L0.5, H0.5, H1) but in the other condition (L0.5, L1, H1) a coarse to fine order increases the number of mistaken target detections. Thus the conditions where the target is not present show support for an integration mechanism with a temporally anisotropic coarse to fine preference for spatial frequency presentation. However, the results do not show the overall pattern shown by Parker et al (1992) nor do the effects show the same strength. Therefore it is difficult to conclude from these results that the integration of spatial frequencies depends upon the coarse to fine availability of spatial frequency information.

The difference between the results found in experiment 2 and 3 presented in this chapter and those found by Parker et al (1992;1997) may have been a result of the stimuli employed in the study. Parker et al (1997) have suggested that the differences they found in their experiments using an image of a natural scene and face may be due to the differences inherent in the processing of images and faces or be a function of spectral composition of the images themselves. In the following experiment both of these factors were examined using a limited set of natural image exemplars with different power spectra slopes. The natural image detection methodology was again employed using both faces and scenes. Two examples of each image type was used one with a shallow power spectra slope and one with a steep power spectra slope. This allows both the effects of variations in image type and the underlying spectral composition of the image on the order preferences for spatial frequency integration to be assessed. It was predicted that the semantic content of the image (that is what the image is actually of) would have no effect as in order for image type to have an effect the image must be categorised, but in order to categorise it some spatial frequency integration must have occurred. It was predicted that power spectra would have an effect as the slope of the power spectra deviates from that found normally in natural images.

2.5 Experiment 4: Natural image detection - images and spectra

The weak support shown for a coarse to fine temporally anisotropic integration of spatial frequencies shown in experiment 2 and 3 and the weak support shown in Parker, Lishman & Hughes (1997) when a facial stimulus is employed, suggest that this coarse to fine preference may not generalise to all images. Parker, Lishman & Hughes (1997) suggested that the differences found in the results when employing different

natural images may be due to differences in their processing, however it seems unlikely that the semantic content of the natural image could affect how spatial frequencies are integrated as there may be a causal relationship where some integration must precede image categorisation. It is more likely that the spectral content of the image is playing a part. Thus the following experiment will explore a limited set of natural image exemplars and power spectrum slopes as a pilot experiment to assess whether it is the case that the type of image and power spectra of the images modulate the temporal form of spatial frequency integration.

It has been found that the power spectra of natural images are approximately equal in their slope (Field, 1987; Ruderman, 1997). It may be the case then that the experiments where strong support was found employed images, the slope of whose power spectra approached this normal value. Conversely then the experiments where weak support for a coarse to fine integration of spatial frequencies involved images whose power spectra deviated strongly from this slope.

In order to assess this, images with different power spectra were used in separate natural image detection experiments. The images used were of both scenes and faces both with widely different power spectra slopes. It would be expected that the type of image should have no effect on order preferences for spatial frequency integration, however as the slope of the power spectrum approached that normally found evidence for a coarse to fine integration of spatial frequencies would be found.

2.5.1 *Method*

All experimental details are identical to experiment 2b except where noted.

Subjects. Ten observers, four male and six female with an age range of 23 to 39, took part in this experiment.

Apparatus and Materials. As outlined in section 2.1.2, Field (1987) amongst others (Burton & Moorhead, 1987; van der Schaaf & van Hateren, 1996; Ruderman, 1997) has shown that the power spectra of natural images follows the function, a / f^α , where a is a scaling factor, f is the spatial frequency and α is the exponent which determines the slope of the function. The value of α when determined from a large number of natural images has been found to approximate 2 (for example α values of: 1.88 ± 0.43 (van der Schaaf & van Hateren, 1996) and 2.4 ± 0.26 (Tolhurst, Tadmor & Chao, 1992) have been found).

In order to test the effect of varying the slope of natural image power spectra functions extreme α values were chosen. With the set of face images available the range of α values was found to be small with the extremes being 1.16 and 1.5, while the value of α for scenes were found to vary from 0.93 to 2.13 (these were the most extreme found from a sample of 20 faces and scenes; It should be noted that the α value for the image used in the previous two experiments was 1.5). These four images constituted the stimuli of this experiment. The images are shown in figure 2.11 and their spectra are shown in figure 2.12⁴. The natural images employed in this detection experiment were of two different faces with α values equalling 1.16 and 1.5 and two different scenes where α equalled 0.93 and 2.13. The same technique for filtering was employed as in experiment 1, however the space constants of the Gaussians differed. The half height of the filters was shifted from 2 and 4 cpd or 0.5 and 1 cpd to match those used in Parker et al (1997). These cut off points were specified in cycles per picture width as low pass cut off points of 25 and 54 cycles per picture and high pass cut off points of 28 and 60 cycles per picture. The match was made on the distal scale of cycles per picture width because there is a large amount of evidence showing scale invariance in the processing

of faces (Riley & Costall, 1980; Ginsburg, et al., 1980; Fiorentini, et al., 1983)². These translate to 1.92, 4.2, 2.15 and 4.62 cycles per degree respectively. In order to reduce confusion between these images and those used in the previous experiments the same shorthand is used, thus the filtered images are denoted L2, L4, H2 and H4 respectively

Design. The design of the experiment was the same as experiment 2b and 3b. The trials were blocked so that each image was utilised as the target in separate blocks performed on different occasions.

2.5.2 Results

The mean number of incorrect target present responses are shown in figure 2.13. The trials on which a target was actually presented within the image triplets were excluded from analysis as it is not clear theoretically how these conditions could reflect any order processing preference shown by the visual system. A 2 x 2 x 2 x 2 way ANOVA was performed with the order of the image triplets shown (2 levels: coarse to fine and fine to coarse), image type (2 levels: scene and face), power spectra (2 levels: shallow and steep); and the spatial frequency content of the image triplet (2 levels: Set 1 = L1, L2, H2 and Set 2 = L1 H1 H2) as factors.

The results show that the order in which spatial frequencies are presented does affect the number of errors made by the subject, with more being made when spatial frequencies are presented progressively from coarse to fine rather than fine to coarse ($F(1, 9) = 5.8$, M.S.e. = 9.05, $p < 0.05$). However, the number of errors evoked was also found to depend upon image type and the spatial frequency content of the image triplets suggesting that the order effect found is not so simply interpreted.

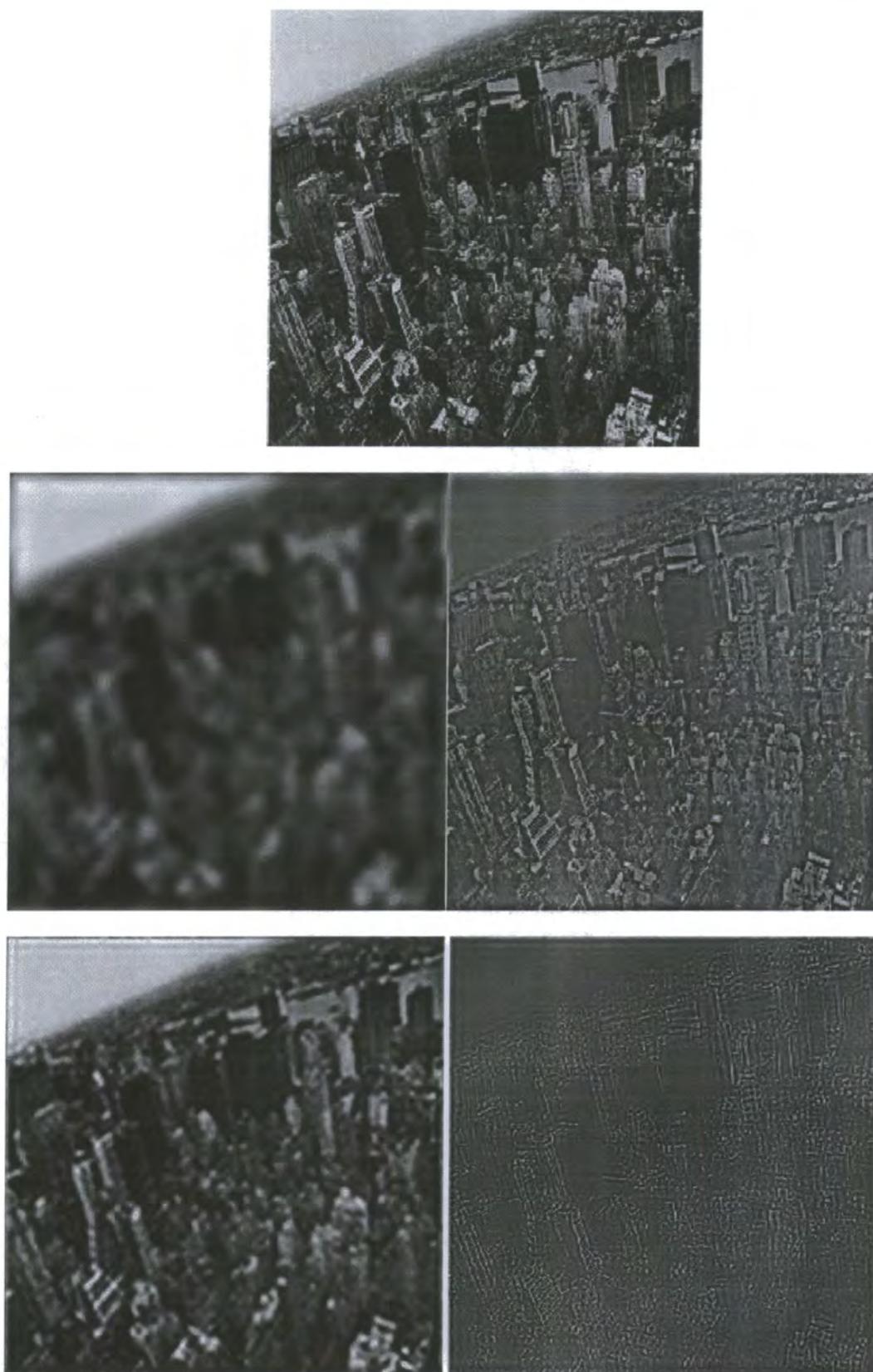


Figure 2.11a: Showing a picture of a cityscape (New York). The power spectra of this root image had an α value of 0.93. Top row = fullbandwidth; Middle = low pass cut off at 1.92 and high pass cut off at 2.15 cpd; Bottom = low pass cut off at 4.2 and high pass cut off at 4.62 cpd. See text for details.

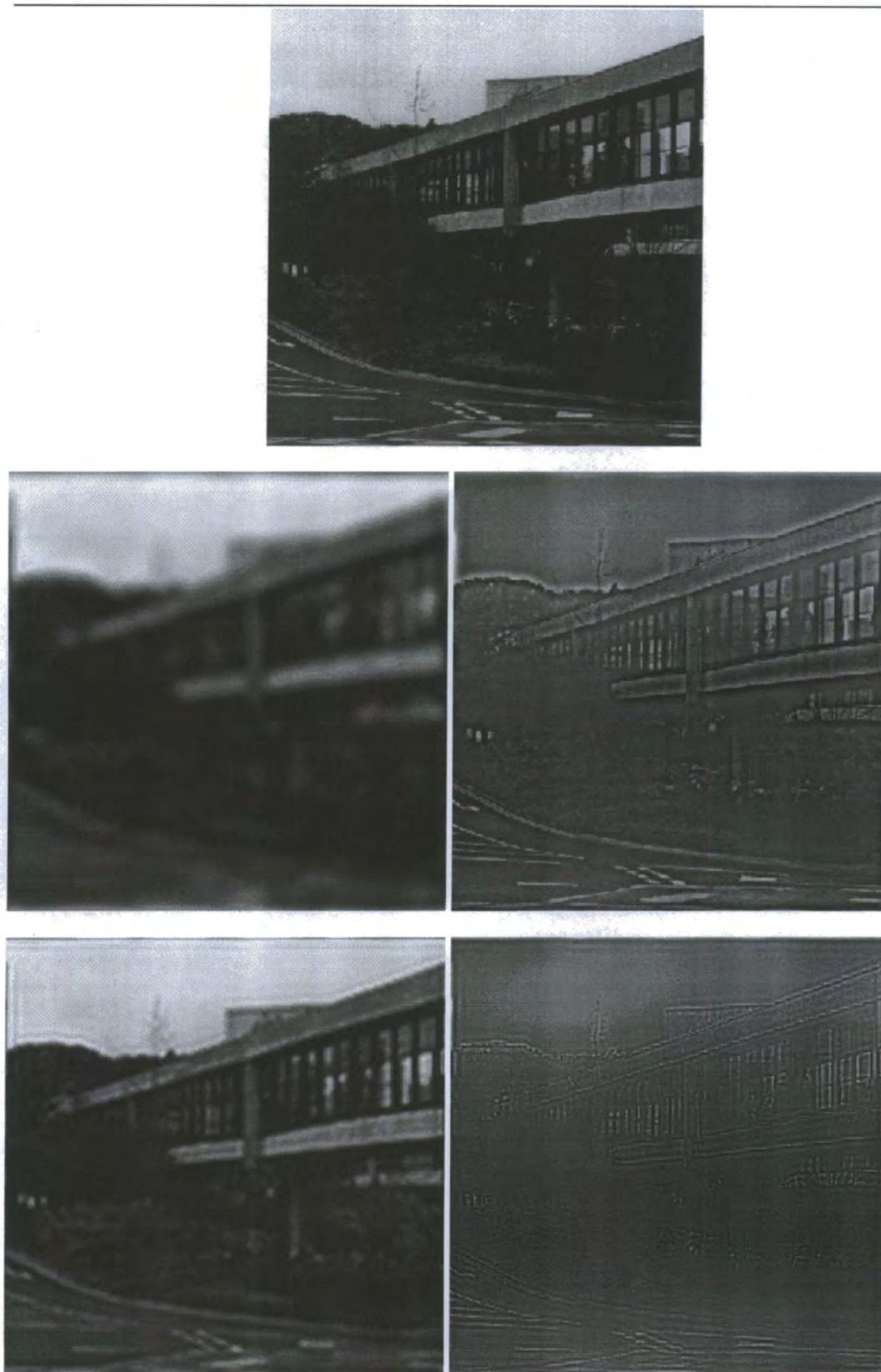


Figure 2.11b: Showing a picture of a building (University of Stirling). The power spectra of this root image had an α value of 2.13. See text for details.



Figure 2.11c: Showing a picture of a face. The power spectra of this root image had an α value of 1.16.
See text for details.

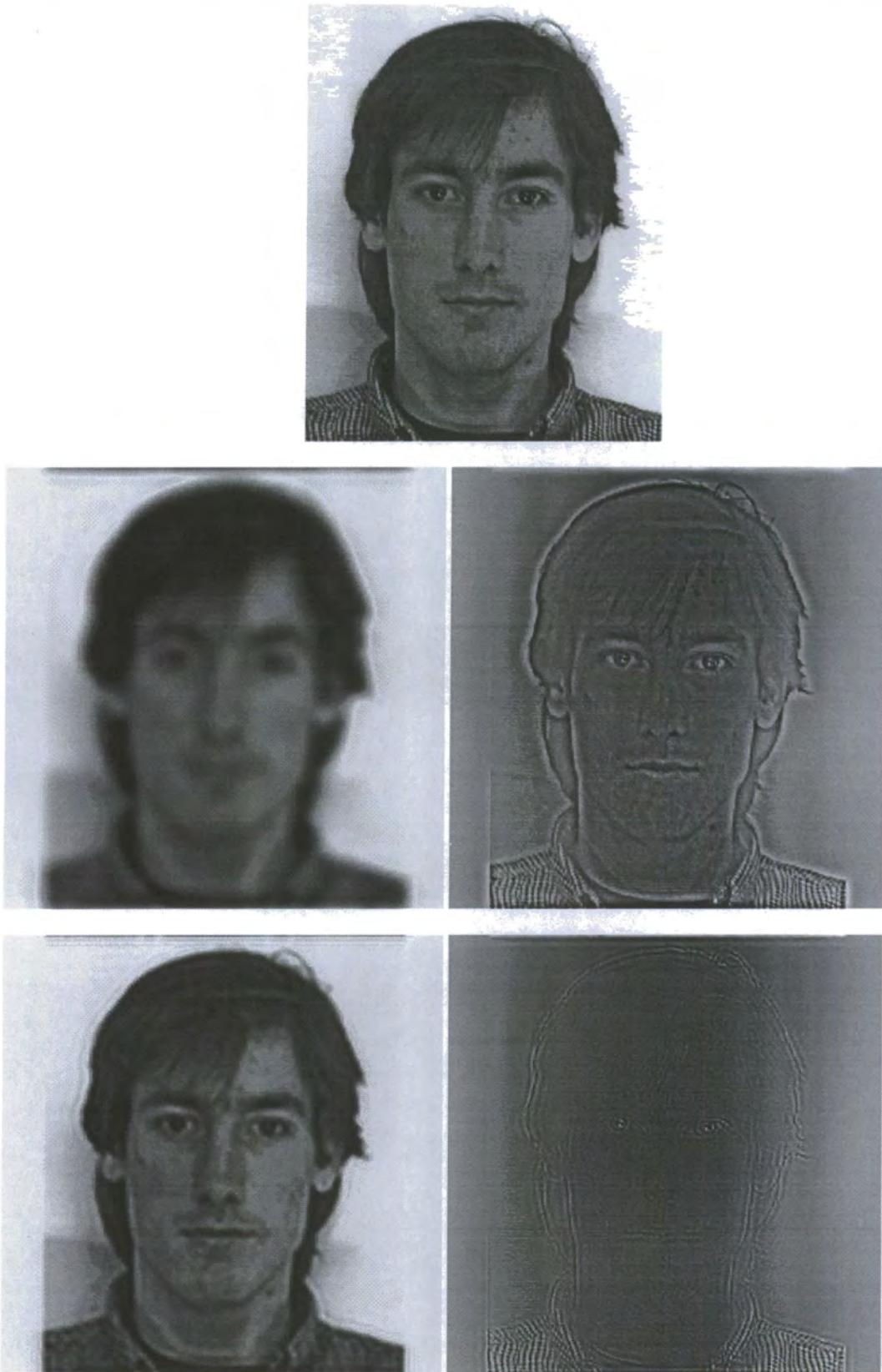
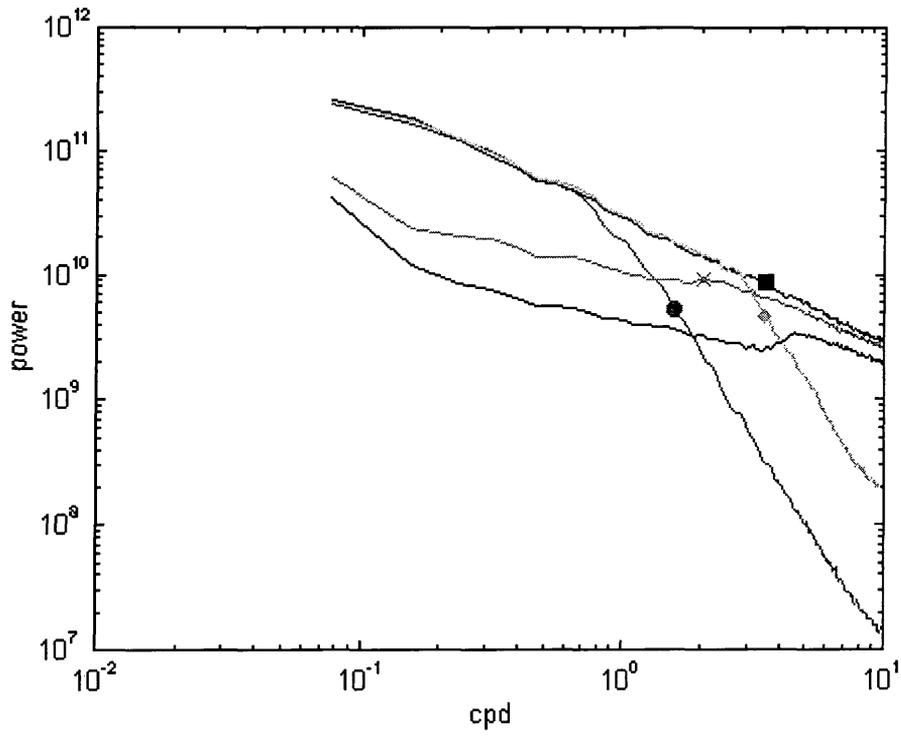


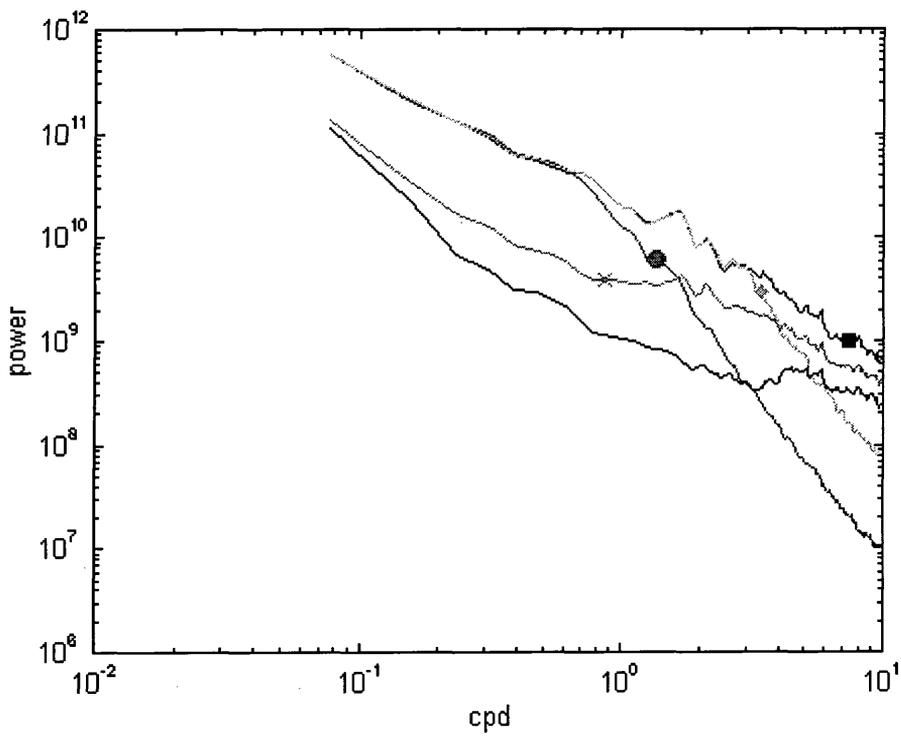
Figure 2.11d: Showing a picture of a face. The power spectra of this root image had an α value of 1.5.

See text for details.

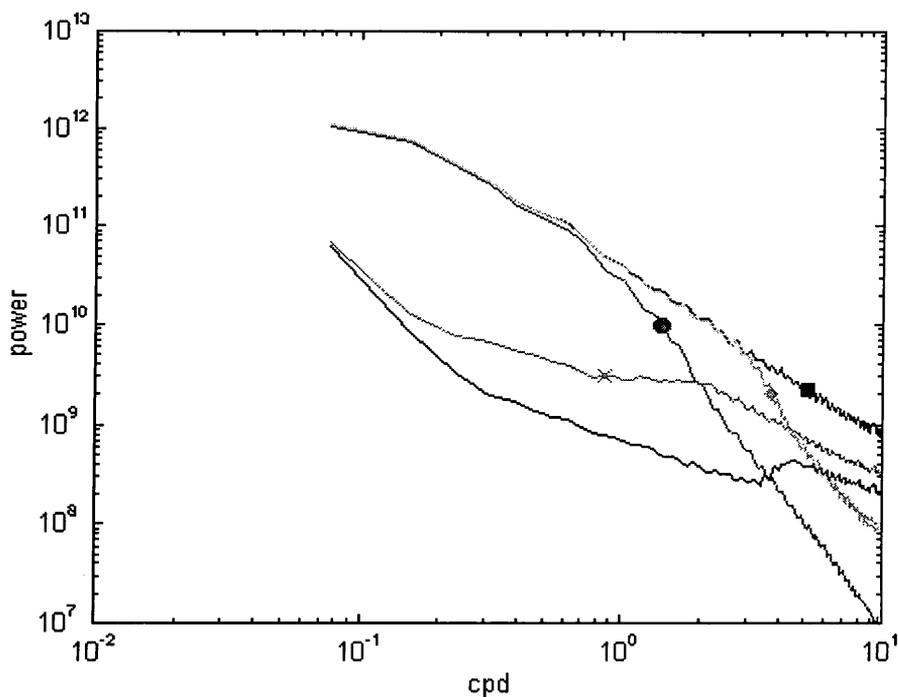
a)



b)



c)



d)

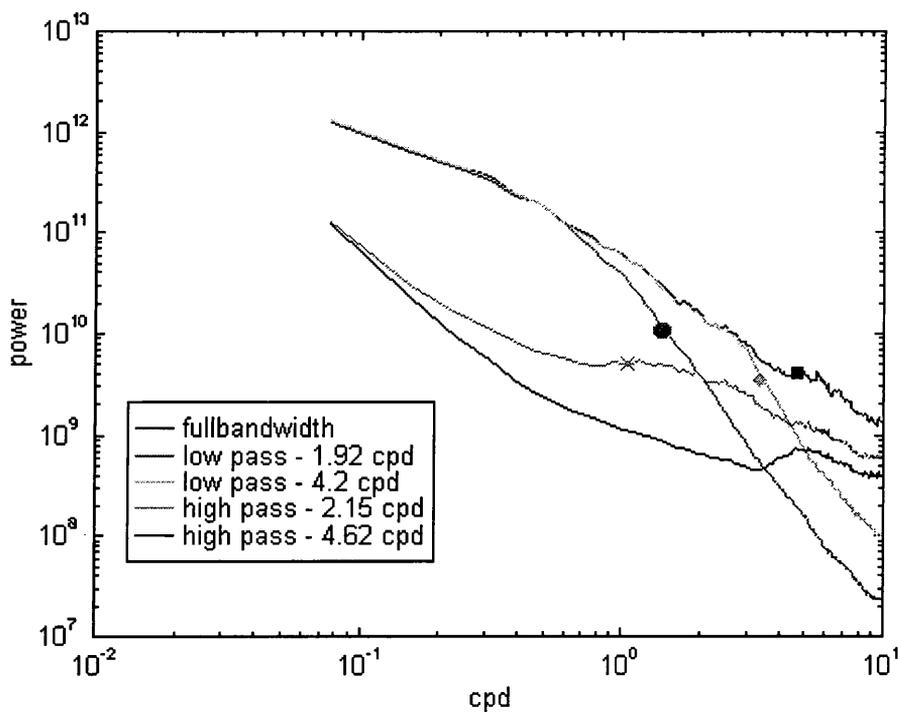


Figure 2.12: Power spectra of the images used in experiment 4. a) shallow sloped power spectra for cityscape scene; b) Steep sloped power spectra for building scene; c) Shallow sloped power spectra for face with shallow sloped power spectra; and d) Steep sloped power spectra for face with steep sloped power spectra. The legend shows the key to the power spectra for the fullbandwidth and filtered images for all images. See text for details.

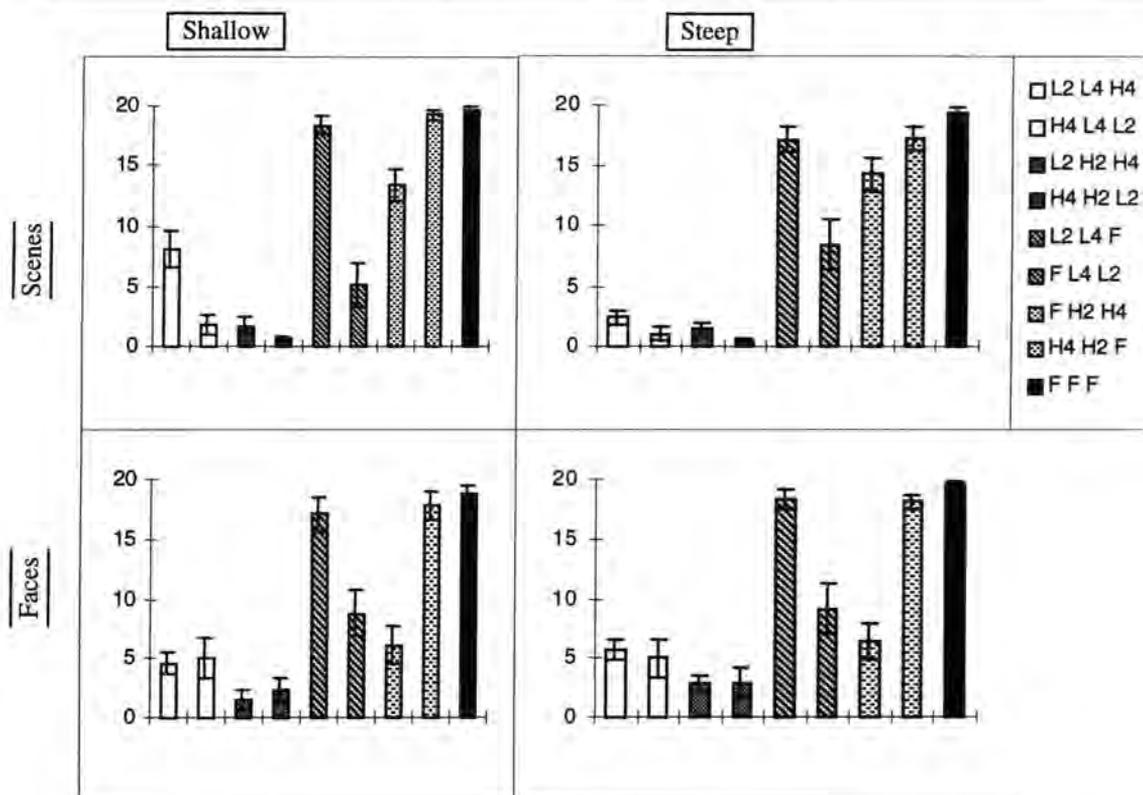


Figure 2.13: The four graphs show the results from the four images employed: the top row shows the results for natural scenes with shallow and steep sloped power spectra respectively from left to right; the bottom row shows the results for facial stimuli, again with a shallow and steep sloped power spectra. The results for each image shows the mean number of fullbandwidth present responses out of 20 maximum as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that triplets containing the same images are together regardless of the order in which they are shown. The vertical bars are standard error bars associated with each mean.

It was found that the number of target present errors varied as function of the image used ($F(1, 9) = 11.4$; $M.S.e. = 8.45$, $p < 0.01$), with face stimuli leading to more false positives than scenes. The type of image used was also found to interact with the order in which the spatial frequency content of the image triplets was presented ($F(1, 9) = 7.5$, $M.S.e. = 8.71$, $p < 0.05$).

The power spectra of the image was found not to have an effect ($F < 1$), but there was found to be an interaction between the power spectrum of the images utilised and

the content of the image triplets ($F(1, 9) = 6.16$, $M.S.e. = 4.97$, $p < 0.05$). A three way interaction was also found between the power spectra of the images, the spatial frequency content of the image triplets and the order in which the spatial frequencies were presented ($F(1, 9) = 6.8$, $M.S.e. = 1.78$, $p < 0.05$).

The spatial frequency content of the image triplets was also found to have an effect on the number of incorrect target present responses, with Set 1: L2, L4, H4, producing more errors than Set 2: L2, H2, H4 ($F(1, 9) = 40.2$, $M.S.e. = 6.09$, $p < 0.01$). An interaction between the spatial frequency content of the image triplets and the order in which they are presented was also found ($F(1, 9) = 9.06$, $M.S.e. = 3.19$, $p < 0.02$).

A set of planned comparisons were also carried out which mirrored those carried out in the previous natural image detection experiments examining the order effects within each set of spatial frequencies employed for each image and power spectra. Only with a shallow sloped power spectra scene type image and set 1 spatial frequencies (L2 L4 H4) was an order effect found where a coarse to fine presentation of spatial frequencies lead to more mistaken square-wave present responses being made (scene with shallow sloped power spectra, set 1 (L2 L4 H4), coarse to fine vs. scene with shallow sloped power spectra, set 1 (H4 L4 L2), fine to coarse: $F = 41.54$; $p < 0.01$).

2.5.3 Discussion

Support has been found for a temporally anisotropic integration of spatial frequencies as the number of incorrect target present responses has been shown to increase when spatial frequencies are presented from coarse to fine. However, the finding that the type of image used, the power spectrum of the images and the spatial frequency content of the image triplets also have an effect on target present responses shows that the interpretation of the results demonstrating an order preference for spatial frequency integration is not so simple.

The results show that an order effect is only found under very specific circumstances: with a shallow power spectrum slope with one particular set of spatial frequencies. Notably this result is equivalent to those found with experiment 2 and 3. Those experiments found an order effect with the equivalent set of spatial frequencies as used here (L2, L4, H4) comprising the image triplet.

Contrary to expectation, the type of image employed did have an effect on the number of incorrect target present responses made, suggesting that the content of the image, that is what the image is actually of, affects how it is integrated. It could be suggested from the effect of image type upon the order of spatial frequency integration that this reflects differences in the processing of scenes and faces. This could be due to the fact that a face is a single object while a scene is a larger number of objects against a background. Thus the results found here may reflect these differences in processing. This interpretation of the results suggests that scenes are integrated from coarse to fine in spatial frequency content while single objects are not.

However, the findings of an effect of the set of filtered images used upon the integration of spatial frequencies suggests that the results do not seem to reflect spatial frequency integration alone. If the methodology of natural image detection reflects the integration of spatial frequencies only, then it is clear that the spatial frequency content of image triplets should not vary the manner in which spatial frequencies are integrated as long as the sweep of spatial frequencies includes the entire spatial frequency content of the image. Thus it appears that the number of incorrect target present responses made by the subject are not only a result of a spatial frequency integration stage in visual processing but seem to be due to the introduction of other factors.

This suggestion that the results found in this experiment are not solely due to the integration of spatial frequencies is also indicated from a consideration of the power spectra effects. It was predicted that the power spectra of the images would affect the

integration of spatial frequencies, such that as the power spectra approached that normally found (that is $\alpha = 2$; Field, 1987; van der Schaaf & van Hateren, 1996; Ruderman, 1997) the integration of spatial frequencies would take place more effectively when shown from coarse to fine. However the results show that the only coarse to fine effect was found when the α value of the image was the furthest away from 2. This can be compared with the results found in experiment 2 and 3. The images used in these experiments had a power spectrum slope with an α value of 1.5 which suggests that as the power spectra steepens the order effect disappears. Thus the integration of spatial frequencies was found to be more effective when they are available from coarse to fine when images with shallow sloped power spectra of α value 0.93 and 1.5 are employed but this disappears at an α value 2.13. However due to the small number of image exemplars used the results may be due particular images employed in this experiment.

If the results found here are taken as reflecting the operation of spatial frequency integration then they show that spatial frequency integration does not occur when the power spectra slopes of images approach that normally found, showing that the integration of spatial frequencies occurs only when the spatial frequency content of images deviate from that normally found. However, if it is considered that the visual system evolved in tandem with the environment then it follows that advantage would be taken of the statistical properties underlying natural scenes (section 2.1.2). Thus the finding that the number of errors increase when spatial frequency information is presented from coarse to fine and images with shallow sloped power spectra are used again suggests that the order dependence, although consistent in these cases, is not due solely to a spatial frequency integration process.

Overall then the results, if taken as reflecting the operation of a spatial frequency integration process, suggest that the temporal form of spatial frequency integration depends upon the image used in terms of its power spectra and its semantic content i.e. scene or face, and the particular filtered images employed in the triplet sequences. However, it seems unlikely that the operation of spatial frequency integration would depend upon any of these, especially the set of filtered image used, as it is difficult to understand why one set of spatial frequencies presented sequentially would be integrated while another set would not when they both contain the same overall spatial frequency content.

It can be suggested that the results found here may be due to differences in the integration not of spatial frequencies themselves but of spatial frequencies being presented in triplets. Coupled with the fact that the spatial frequency content of the filtered images which constitute the triplets greatly overlap, the nature of the method may introduce an interaction at a second level. It follows that when the image triplets are presented they are in fact three large bandwidths of spatial frequency information presented one after the other. Thus within any one image a large bandwidth of spatial frequency information is presented simultaneously. It can be suggested that the integration of spatial frequencies takes place within the filtered images and due to the delays introduced by the spatial frequency processing this integration takes place from coarse to fine (see section 1.3.2). Thus when a triplet of filtered images is presented would be the case that the spatial frequency content of each filtered image is processed from coarse to fine. Therefore even in a fine to coarse sequence the integration of spatial frequencies within each filtered image progresses from coarse to fine. Given this conception of the presentation method it follows that the spatial frequency content of the initial image presented is integrated from coarse to fine. When the second image is presented its spatial frequency content is also presented from coarse to fine. The

question is how does this new spatial frequency information interact with the already integrating spatial frequencies of the initial image. The same line of thought can be extended to the third image. However, given this coarse to fine processing which occurs within each of the filtered images presented it is still the case that the overall presentation of spatial frequency content is one dependent upon the pattern of image presentation. Thus in the case of the fine to coarse presentation order of filtered images, the spatial frequency content of each filtered image is processed from coarse to fine but the overall spatial frequency content is shifting from fine to coarse.

Thus there are two problems with the methodology as conceived here: first, the initial filtered image of each triplet sequence is integrated from coarse to fine in spatial frequency content; and second the additional presentation of other filtered images introduces a secondary spatial frequency interaction where the spatial frequency content of the second filtered image interacts with the first images spatial frequency content which has already been integrated, this follows also for the third filtered image presented.

Thus it may be the case then that the pattern of results found here are due to a secondary interaction of spatial frequencies across the large bandwidths of the filtered images. Thus the differences in results found here between images, sets and power spectra might be due to the abnormal presentation method of one large bandwidth of spatial frequencies after the other. Thus suggesting that the natural image detection paradigm does not directly measure the low level integration of spatial frequencies. This interpretation points to an initial integration of spatial frequencies followed by a second order interaction of spatial frequencies which depends upon image type.

Overall then the results show that while an order effect is found it seems to depend upon a highly specific stimulus being used. This specificity suggests that the results cannot be due to a generalisable mechanism which integrates spatial frequencies

from coarse to fine across time. It is proposed that the results are due to the methodology employed which may lead to second order integration strategies being employed which vary from image to image depending upon type and / or power spectra slope. However, as the stimulus set was limited no firm conclusion can be made. The results may be due to the particular images used. Therefore the interpretation of the results, while suggesting a number of possibilities are constrained by this limited set of images. Further work employing a larger set of images and a less complicated design would be of great value in addressing the issues raised by this experiment.

2.6 General Discussion

The results from the two sets of experiments (image quality and image detection) presented here show that evidence that the visual system integrates spatial frequencies more effectively when they are presented from coarse to fine is mutable. In the quality judgement experiments there was a clear difference in the rated quality of spatial frequencies presented from coarse to fine which was higher than those presented from fine to coarse. However a temporally isotropic weighted average model was found to account for the data very well. The detection experiments also show some evidence for a coarse to fine integration of spatial frequencies however the results of experiment 4 show that this effect depends upon the image used, the power spectra of that image and the set of spatial frequencies presented as the triplet. This effect dependence was suggested to be a result of a spatial frequency interaction introduced by the method employed.

2.6.1 *Task considerations*

The results of experiment 4 were suggested to be due in part to secondary interactions between the filtered images presented as triplets. This suggests that the quality judgement experiments are also subject to the same interactions as the method of presentation is identical. However, there are fundamental differences involved in the tasks of detection and quality which along with the secondary interactions introduced by the sequential presentation of filtered images, supports the quality judgement task as reflecting the operation of spatial frequency integration.

There are differences between the tasks of detection and quality which mean that they can be respectively regarded as being indirect and direct measures of spatial frequency integration. In a natural image detection experiment a fullbandwidth image is to be discerned from a sequence of filtered images. (This can be equivalently regarded in the spatial frequency domain as follows: a set of spatial frequencies with a particular delay structure imposed by spatial frequency processing delays (the fullbandwidth image) is to be discerned from other, smaller, sets of the same spatial frequencies with partially the same delay structures but presented a different times (the filtered images)). Thus the optimal strategy for performing the detection experiment would involve processing the spatial frequency bandwidths provided by the individual images separately. The results, under this theoretical viewpoint, would represent the success by which this strategy has been employed, thus if a coarse to fine presentation leads to more incorrect fullbandwidth image target present responses being made, when the target was not in fact present, this temporal order can be said to have disrupted this separatist strategy more so than the fine to coarse temporal order. It is the nature of this disruption which is questionable. It could be argued that the disruption of this separatist strategy is due to the integration of spatial frequencies and thus the detection task can be seen to be an indirect measure of spatial frequency integration. However since image

type and the particular set of spatial frequencies used modulate this disruption (experiment 4), it is likely that factors other than spatial frequency integration are playing a part.

In the quality task a judgement on the basis of the overall image sequence is required, that is, the degree to which a sequence of spatial frequencies appears to be like or resemble the natural image target. It can be suggested that the quality judgement task actively encourages the integration of spatial frequencies by demanding that the judgement be made on the overall quality of the triplet. The quality task then is a direct measure of spatial frequency integration.

This identification of the fundamental differences between the tasks suggests not only why the results from the two tasks might differ but also supports the quality judgement experiment as directly measuring an early spatial frequency integration preference. It was pointed out in section 2.5.3, that the spatial frequency content of the initial image of every sequence of filtered images can be considered to be integrated from coarse to fine and that the sequential presentation of filtered image introduces a second order interaction between the spatial frequency content of this image and those that follow. From the consideration of the tasks it can be suggested that in the quality judgement task the second order interaction is encouraged to take the form of an integration of spatial frequency content; while in the detection task this second order interaction may take the form of masking of the spatial frequency content of the filtered images.

Thus if it is assumed that the spatial frequency content of the filtered images is being integrated when the quality of the overall triplets is judged then the result showing the increase in rated quality for a coarse to fine presentation sequence suggests that the integration of spatial frequencies takes place more effectively from coarse to fine, that is the triplets are more like or resemble the target when presented from coarse to fine. That

integration is taking place is suggested by the finding that when spatial frequency information was presented from coarse to fine the resulting representation was more target-like than would be predicted from the quality ratings of the individual images, when presented singly, which formed the triplets. However the results are well accounted for by a temporally isotropic weighted average model. Thus while the quality task does encourage integration of the image triplets to take place, integration occurs equally well regardless of order.

2.6.2 *Last Frame*

A very clear concern is the results from the conditions when the fullbandwidth image is present. These conditions have been discounted from all discussion of the results as it is unclear how they can be interpreted. However as the results from conditions where the fullbandwidth image is present do show a consistent effect across all conditions they clearly merit discussion.

The results from all experiments presented in this chapter show the same pattern of results in conditions when the fullbandwidth image is present: the low pass, low pass, fullbandwidth conditions (e.g. L0.5 L1 F) all show a coarse to fine advantage for the triplet quality or the detection of a fullbandwidth image; while the fullbandwidth, high pass, high pass conditions (e.g. F H0.5 H1) show a fine to coarse advantage. This is different to the results found in Parker et al (1992) where a coarse to fine advantage is found in both conditions, however Parker et al (1997) found no significant difference between the latter conditions. A possible explanation for the results found here but discounted by Parker et al (1992) is one where the subject's performance depends on the last frame shown. Thus the reason for a coarse to fine and fine to coarse order effect when the target is present is that in the L0.5, L1, F condition the coarse to fine sequence

has as its last frame the fullbandwidth target and when the F, H0.5, H1 is shown from fine to coarse it also has as its last frame the fullbandwidth image.

It follows then that this last frame effect has influenced the results found in those conditions where the target is not present. The natural image quality experiments provide evidence showing that the results may be due to quality judgements being based on the last frame alone. The conditions where a fullbandwidth image was not present and a coarse to fine presentation order lead to a improved quality rating could be due to the last frame of the image triplet, as it was more confusable with the fullbandwidth image. Thus the two explanations of an order preference for spatial frequency integration and the quality judgement being based upon the last frame shown lead to the same predictions. The results from the conditions where no fullbandwidth image was presented in the natural image detection experiments show either evidence for a last frame explanation or no clear evidence either way. However as these results seem to be subject to many influences, as discussed previously, it would be difficult to glean clear support either way for this last frame explanation.

In Parker et al (1992) the finding that the F, H0.5, H1 (F, H3, H1 in their terminology) condition resulted in an improved quality rating when compared to the converse suggests that a last frame hypothesis cannot explain quality ratings, however this finding was not shown in Parker et al (1997). Furthermore, clear support for a last frame explanation was found here and this explanation has the advantage that it could tie together the results from the conditions where the fullbandwidth image is present and those when it was not present, as simply basing decisions upon the last frame shown does not involve the theoretical complications which are introduced when considering the integration of spatial frequencies and the presence of the fullbandwidth image itself (Section 2.1.1).

Another possibility, and one discussed in section 2.2.2.3, is that of a weighted average model where the quality ratings of filtered images presented as triplets were found to correlate very well with the weighted average of the quality ratings elicited by the filtered images when presented individually. Thus either a last frame explanation or a weighted average explanation could account for the results found here. As the last frame account makes no assumption about spatial frequency integration taking place within the triplets then this is to be preferred as it is a simple explanation which accounts for the data found in experiment 1 for conditions on which a fullbandwidth image was or was not present within the image triplets.

2.6.3 *Conclusions*

In this chapter the integration of spatial frequencies was assessed using natural image quality and natural image detection paradigms introduced by Parker et al (1992). It was found that the support for a temporally anisotropic model of spatial frequency integration was limited. The image quality experiment shows support for a temporally isotropic weighted average model of spatial frequency integration while very weak support for a temporally anisotropic coarse to fine integration of spatial frequencies was found in the natural image detection experiments.

It was pointed out that the paradigm of stretching the presentation of the spatial frequency content of an image by sequentially presenting filtered images was problematic mainly because it introduced a secondary interaction across the large spatial frequency bandwidths of the filtered images. It was suggested that the form of this secondary interaction may vary depending upon the task employed. It was shown that the tasks of detection and quality differ fundamentally and as a consequence of this it was suggested that the lack of spatial frequency integration found in the detection task may be a result of this secondary interaction across the filtered images. The results of

the quality judgement task were suggested to be a reflection of spatial frequency integration in that the secondary interaction between filtered images is encouraged to take the form of integration.

Finally, a second possible interpretation of the results was outlined in which the subject's response is based on the last image shown alone. Under this interpretation no interactions occur between the spatial frequencies presented. If the results of the natural image detection experiments are discounted, then the results of the natural image quality experiment could be explained on the basis of this.

In the next chapter both the problems of employing large bandwidths of spatial frequencies so that the initial image of every sequence is integrated and the use of the final image shown as a basis for performance in the natural image tasks are addressed. An alternative means by which to assess the integration of spatial frequencies is to use simple spatial frequency stimuli. As the integration mechanism examined here is suggested to perform on the basis of spatial frequency information it seems sensible to extend the experiments carried out here and those carried out by Parker, Lishman & Hughes (1992; 1997) to stimuli which can be controlled in spatial frequency content.

Spatial Frequency Integration and Square-waves

Summary

The existence of a temporal anisotropy in the integration of spatial frequencies is examined using Gabor patches which peak at specific spatial frequencies. In this way it is possible to isolate specific spatial frequency channels and examine their interactions across time. A square-wave grating and the first three harmonics of that square-wave were employed as stimuli. The first experiment (experiment 5) is based on the natural image detection paradigm used in experiments 2, 3 and 4. The harmonics of the square-wave were presented sequentially from coarse to fine or fine to coarse. Within this sequence a square-wave might be present. Subjects indicated the presence of the square-wave within the sequence. It was found that the square-wave was more likely to be falsely reported in a coarse to fine sequence than in a fine to coarse sequence. This supports the suggestion that spatial frequency integration takes place more effectively when they are shown from coarse to fine thus approximating that normally found in the human visual system. In experiment 6 the suggestion that a coarse to fine sequence of spatial frequency information integrates to approximate the representation formed when a stimulus is presented normally was further examined. The first three harmonics of a square-wave were presented from coarse to fine or fine to coarse in a temporal two alternative forced choice experiment. Subjects were asked to indicate which interval

appeared to resemble the fullbandwidth image more. It was found that a fine to coarse sequence of spatial frequencies was judged to more resemble the target than the coarse to fine sequence. This suggests that the results of the experiment 5, the square-wave detection experiment, was not due to one temporal order of spatial frequencies appearing to resemble a square-wave more than the other. Possible reasons for this discrepancy were examined.

3.1 Introduction

In the previous chapter (see section 2.6) it was suggested that the use of specific spatial frequencies would be better suited to examining the existence of a temporal anisotropy in the integration of spatial frequencies. This suggestion was based upon the finding in Chapter 2 that the use of natural images as stimuli resulted in the introduction of other factors on the integration of spatial frequencies, such as image type, masking between blocks of spatial frequencies and task employed, which may obfuscate any order preference for the integration of spatial frequencies.

In this chapter the existence of a temporal anisotropy in the integration of spatial frequencies will be examined employing Gabor patches, which approximate the spatiotemporal weighting function of receptive fields found in the primary visual cortex (V1) (Graham, 1989). First the advantages of using spatial frequency Gabor patches will be outlined followed second by a discussion of suitable complex stimuli which would take the place of the fullbandwidth natural image used in Chapter 2. Third, the nature of the detection and quality judgement tasks introduced in Chapter 2 and their relation to the use of Gabor patches will be outlined. Fourth, the aims of the experiments will be discussed. Finally, a outline of the experiments to be performed in this Chapter will be presented followed the experiments themselves.

3.1.1 Advantages of Gabor patches

A problem with a purely Fourier account of visual processing is that a spatial frequency grating or sinusoidal waveform is only localised in the frequency domain if it is infinite in extent this leads to the problem of an infinitely extended grating not being localised in space. Gabor (1946) showed that it is possible to offset the two localisation problems to maximise their localisation in both the spatial and frequency domains simultaneously. Thus the Gabor function has the advantage of minimising location uncertainty in both the spatial and frequency domain.

A Gabor patch is a carrier wave of a certain spatial frequency convolved with a envelope or window Gaussian function. This is defined as:

$$G(x, y) = e^{-(x^2+y^2)/2\sigma^2} * (\cos(2\pi f(x \cos \theta + y \sin \theta))) * C$$

where C is the contrast, f is the spatial frequency, θ is the orientation of the Gabor patch, and σ is the space constant or standard deviation of the Gaussian. An example is shown in figure 3.1.

As with infinitely extended sinusoids, a Gabor function has the property of being able to fully describe any complex waveform. Marcelja (1980) and Daugman (1980) have shown that Gabor functions can be utilised in the spatial domain in both one and two dimensions. When extended two dimensions, Gabor functions can be used to specify any visual stimulus maximising both frequency and location optimally.

Gabor functions have been suggested to mirror the spatiotemporal weighting function of the receptive fields found in the primary visual cortex (Graham, 1989) however, they are not the only models of the filters involved in human visual

processing. Others have suggested somewhat different models of filters a discussion of which is beyond the scope of this thesis (Young, 1987; Stork & Wilson, 1990; Yang, 1992; Klein & Buetter, 1992; Daugman, 1993). It is sufficient to note that the Gabor function is a good approximation of the filters involved in early human visual processing and that they also optimally maximise the localisation of information in both the spatial and frequency dimension. Indeed Daugman (1993) states that “early vision can do no better than to encode image structure by using 2D Gabor primitives” (p. 376).

In terms of the experiments to be presented here a Gabor function provides a stimulus which minimises the role of other factors found when examining the integration of spatial frequencies using natural image stimuli. They have the advantage of containing a small band of specific spatial frequencies which, given the methodology of stretching the presentation of spatial frequencies used in Chapter 2, would reduce the spatial frequency interactions which occur within each temporal period or frame of spatial frequency information presented and isolate the interactions which occur across the frames. The reduction of spatial frequency interactions within each spatial frequency presentation period would result as a direct consequence of the selective activation of spatial frequency channels which the use of Gabor patches affords.

3.1.2 *Square-waves as stimuli*

In order to examine the order effects in the interactions between spatial frequencies they must integrate to form a stimulus not inherent in their individual structure. Square-waves are ideal for this. A square-wave is a pattern of sharply delineated black and white bars, the Fourier analysis of which reveals that it is decomposable into individual sinusoids which are harmonically related. Thus a square-wave can be formed by integrating a set of harmonically related sinusoid according to the following function:

$$4 / \pi(\sin(2\pi fx) + \frac{1}{3} \sin(3(2\pi fx)) + \frac{1}{5} \sin(5(2\pi fx)) + \dots + \frac{1}{n} \sin(n(2\pi fx)))$$

where f is the fundamental frequency of the square-wave and n indicates that the series is infinite. It can be seen that as the series progresses the spatial frequencies added to the fundamental are higher odd harmonics and their amplitude is divisively reducing by the order of the harmonic. This is illustrated in figure 3.1. Thus the spatial frequencies which integrate to form a square-wave can be employed instead of the natural image stimuli of Chapter 2. For example the first three harmonics of a square-wave series, F , $3F$ and $5F$ can be presented sequentially from coarse to fine or fine to coarse. Detection and quality judgements can then be made.

3.1.3 Task Considerations

In the previous chapter two methodologies were employed to examine the order effect of spatial frequency integration: natural image quality and natural image detection. It was argued that the task of detecting a fullbandwidth image within a sequence of filtered images was not ideal for examining integration effects as it did not demand the decision be based on the overall sequence of filtered images. Rather the ideal strategy would be to segregate the blocks of spatial frequencies presented (i.e. the filtered and fullbandwidth images). The quality task however was suggested to be better suited for examining the existence of a temporal anisotropy in the integration of spatial frequencies as the decision needed to be based on all the images presented within the sequence (see section 2.6).

Detection

The strategy of segregation said to underpin the detection task can still be interpreted in terms of spatial frequency integration. The segregation strategy can be seen to be an attempt not to integrate across spatial frequency blocks. Thus if the subject is attempting to dissociate the incoming blocks of spatial frequencies then the conditions for which no error was found (i.e. a mistaken target present response was not made) did not disrupt this strategy. The conditions in which an error was found can be said to have disrupted the effectiveness of this strategy, that is, these conditions can be taken as evidence for the integration of spatial frequencies as despite the attempt to hold the blocks of spatial frequency separate they interact. Therefore in spite of the detection task not encouraging the integration of spatial frequencies, it occurs regardless. The results of experiment 2, 3 and 4 of Chapter 2 show that, under this interpretation, there is some evidence that this disruption does take place. However it is clear from experiment 4 that this disruption of the segregation strategy depends upon other influences such as image type and possibly masking between spatial frequencies.

It is exactly these influences which the use of specific spatial frequencies is designed to circumvent. Therefore this suggests that the detection task can be utilised as an indirect measure of spatial frequency integration if it is assumed that the increase in errors or target present responses indicates a disruption of a strategy due to the process of spatial frequency integration.

Quality

The image quality paradigm, whilst encouraging decisions to be based upon the entire sequence of spatial frequencies presented, is not suited for use with Gabor patches. This is because of the differences in the structural relationship between filtered natural images and the fullbandwidth image from which they are derived, and the

harmonics of a square-wave and the square-wave itself. Filtered natural images depict the same broad structures as a fullbandwidth image only certain details will be lessened such as fine details or broad contrast information. An individual spatial frequency Gabor patch does not resemble the square-wave target which it integrates to form, unless that spatial frequency is the fundamental as this has the same periodicity and phase of the square-wave. As a result of this it is difficult to judge the quality of the spatial frequency harmonics as it is not clear upon which dimension it is possible to judge that quality. Further, it has also been suggested that the quality judgements are nebulous in that it is difficult to be clear upon which aspects quality is being judged for natural images and whether those aspects are the same for all subjects (Parker et al ,1992; 1997).

An alternative is to use an equivalent holistic task which is based upon the quality of the overall triplet sequence without involving the quality ratings of the spatial frequency Gabor patches. As the quality rating of image sequences is suggested to be a direct measure of how much the sequence of filtered images appears to resemble the fullbandwidth image it was decided to simply ask for a judgement of the resemblance of the overall image triplet to the fullbandwidth image. This involved showing two sequences of images in a two temporal interval forced choice and judging which of the two most resemble the target.

Conclusions

Thus the considerations of the tasks used in Chapter 2 show that they can be employed to examine the existence of a temporal anisotropy in the integration of spatial frequencies using Gabor patches. The image detection task was shown to be well suited for use with Gabor patches and it was suggested to measure the integration of spatial frequencies in an indirect manner. The quality judgement, a direct measure of spatial frequency integration, was suggested to be less well suited in the form used with natural

images, however an equivalent task was suggested involving a direct judgement of resemblance to the fullbandwidth image.

3.1.4 *Aims of experiments*

Both of the experiments presented in this chapter are concerned with examining the effect of a particular temporal availability of spatial frequencies for the increase efficacy of their subsequent integration. The experiments presented will examine this by both the indirect and direct methods of detection and quality as outlined in the preceding section. Experiment 5 will also examine the suggestion in the last chapter that the results could be accountable by assuming that the subjects are simply basing their decisions upon the last image or frame shown.

Last Frame explanation

It was suggested in section 2.1.1 that the method of stretching spatial frequency presentation by showing images sequentially with a fullbandwidth image within the sequence was difficult to interpret for a number of reasons. However it was found that over all the experiments carried out there was a similar pattern of results in these conditions. It was suggested that this may be a result of subjects' decisions being based purely on the last frame presented. Thus in the image quality and detection task when a fullbandwidth image was presented on the last frame that sequence was rated as being of higher quality or the target was more detectable than when the fullbandwidth image was presented as the first image. This pattern could also account for the results when a fullbandwidth image was not present in the sequence. Thus the sequence in which the highest rated quality single image (a high pass filtered image with the greatest bandwidth) was presented last of all was rated as being of higher quality than if it was

presented as the first frame. Similarly the target was mistakenly reported more often if the most confusable image was presented last.

This suggests that the results of Chapter 2 may not be a result of a weighted average temporal isotropy in the integration of spatial frequencies but rather the result of a much simpler mechanism where the subject bases their judgement upon the most recent image presented. In order to test whether the last frame is the most important it is necessary to identify a stimulus which appears to more similar to its low spatial frequencies than it high spatial frequencies. For example, if the low spatial frequencies of a stimulus are rated as being of higher quality than the high spatial frequencies and the subject did base their decision on the last image shown then it would be expected that a fine to coarse sequence would be rated as being of higher quality than a coarse to fine image sequence. If however a coarse to fine sequence is rated as being of better quality then this shows that only a temporal anisotropy in the integration of spatial frequencies could account for the results. A similar argument also follows for the image detection paradigm.

As has been previously mentioned a square-wave is a complex waveform comprising of a fundamental frequency which has the same periodicity and phase as the square-wave and odd harmonics of this fundamental. It follows then that the fundamental component of a square-wave is far more similar to it than the higher harmonics as these have greater periodicity and lower contrast. Thus in this stimulus the low spatial frequencies are more similar to the square-wave target, precisely the stimulus which would form a good test of the last frame explanation.

The first experiment to be presented investigates the existence of an order preference and also examines whether this preference is due to the integration of spatial frequencies or simply a function of the last frame shown to the subject.

3.1.5 *Preview of Chapter*

This chapter will follow directly on from the previous one. The existence of a temporal anisotropy in the integration of spatial frequencies will be examined using both the image detection and image quality paradigms. However specific spatial frequencies in the form of Gabor patches will be presented instead of fullbandwidth and filtered natural images, in order to reduce the other influences on the integration of spatial frequencies found to operate with natural images.

In the first experiment to be presented the image detection methodology is employed. The first three harmonics of a square-wave are presented from coarse to fine or fine to coarse and within this sequence a square-wave stimulus may appear. The subject's task is to detect the presence of the square-wave target. This is directly analogous to the image detection paradigm utilised in the previous chapter. This experiment provides a direct test of a last frame explanation for the results found in Chapter 2. If a temporal anisotropy for the integration of spatial frequencies does exist then it would be expected, given the results of the natural image quality task of experiment 1 and Parker et al (1992), that the coarse to fine sequence of the harmonics of a square-wave would lead to an increase in the number of incorrect target present responses made. However if the subjects base their decision upon the last image shown then the number of mistaken target present responses will increase when the harmonics are shown from fine to coarse.

The second experiment to be presented is a variation of the image quality paradigm of experiment 1. In this experiment the resemblance of the first three harmonics to a square-wave is judged. These harmonics are again presented in different temporal orders: from coarse to fine or fine to coarse. It would be expected that the

coarse to fine order would be judged as resembling a square-wave more than a fine to coarse temporal sequence.

3.2 Experiment 5: Square-wave detection

This experiment was designed as a test of whether the findings for an order effect using natural images are due to a temporal anisotropy in the integration of spatial frequencies or subjects basing their decision on the last image shown. A square-wave stimulus is ideal for this as the harmonic which appears to resemble a square-wave most is its fundamental. If a last frame explanation could account for the results of the natural image experiments then the fine to coarse sequence of the first three spatial frequencies of a square-wave series would result in the greater number of target present mistakes as the fundamental frequency is presented last. However if the same coarse to fine order effect is found then it can be concluded that there is a temporal anisotropy in the integration of spatial frequencies.

Two experiments are presented in this section. First a discrimination task in which a square-wave target and the harmonics of that square-wave are presented singly. This primarily allows the subject to familiarise themselves with the target and the harmonics but also measures the discriminability of a square-wave when presented for brief periods of time. This experiment is a baseline on which to judge the second experiment. In the second experiment the individual spatial frequencies are grouped into triplets and presented from coarse to fine or fine to coarse. Subjects detect the presence of a square-wave target which may be present within the sequence.

3.2.1 *Experiment 5a: Discrimination*

In order to familiarise subjects with the stimuli and examine the detectability of the targets a discrimination task was performed based upon the methodology of the natural image discrimination experiments outlined in Chapter 2 experiment 2a and 3a. The first three spatial frequencies of a square-wave series and the square-wave target themselves are presented singly and subjects indicate whether or not a square-wave had been shown.

3.2.1.1 *Method*

Subjects. Ten observers, five male and five female with an age range of 25 to 34, took part in this experiment. One subject, the author, was fully aware of the purpose of the experiment the remaining nine were naive.

Apparatus and Materials. The stimuli used in this experiment took the form of spatial frequency Gabor patches of the first three harmonics of a vertically oriented square-wave with a fundamental of 1 cpd and a circular spatial envelope standard deviation of 1.74 degrees (See formula in section 3.1.1). The stimuli also took the contrast relationships of that dictated by their harmonic relationship, specifically the fundamental (F) was shown at 93.75 % (Michelson contrast = $(I_{\max} - I_{\min}) / (I_{\max} + I_{\min}) * 100$; where I_{\max} is the highest luminance value; and I_{\min} is the lowest luminance value) while the third harmonic (3F) was shown at a third of this at 31.25% and the fifth harmonic at 18.75%.

Also used were vertically oriented square-waves convolved with a Gaussian of the same size and fundamental frequency as the Gabor patches:

$$S(x, y) = e^{-(x^2+y^2)/2\sigma^2} * [4 / \pi(\sin(2\pi fx) + \frac{1}{3} \sin(3(2\pi fx)) + \frac{1}{5} \sin(5(2\pi fx)) + \dots + \frac{1}{n} \sin(n(2\pi fx))]$$

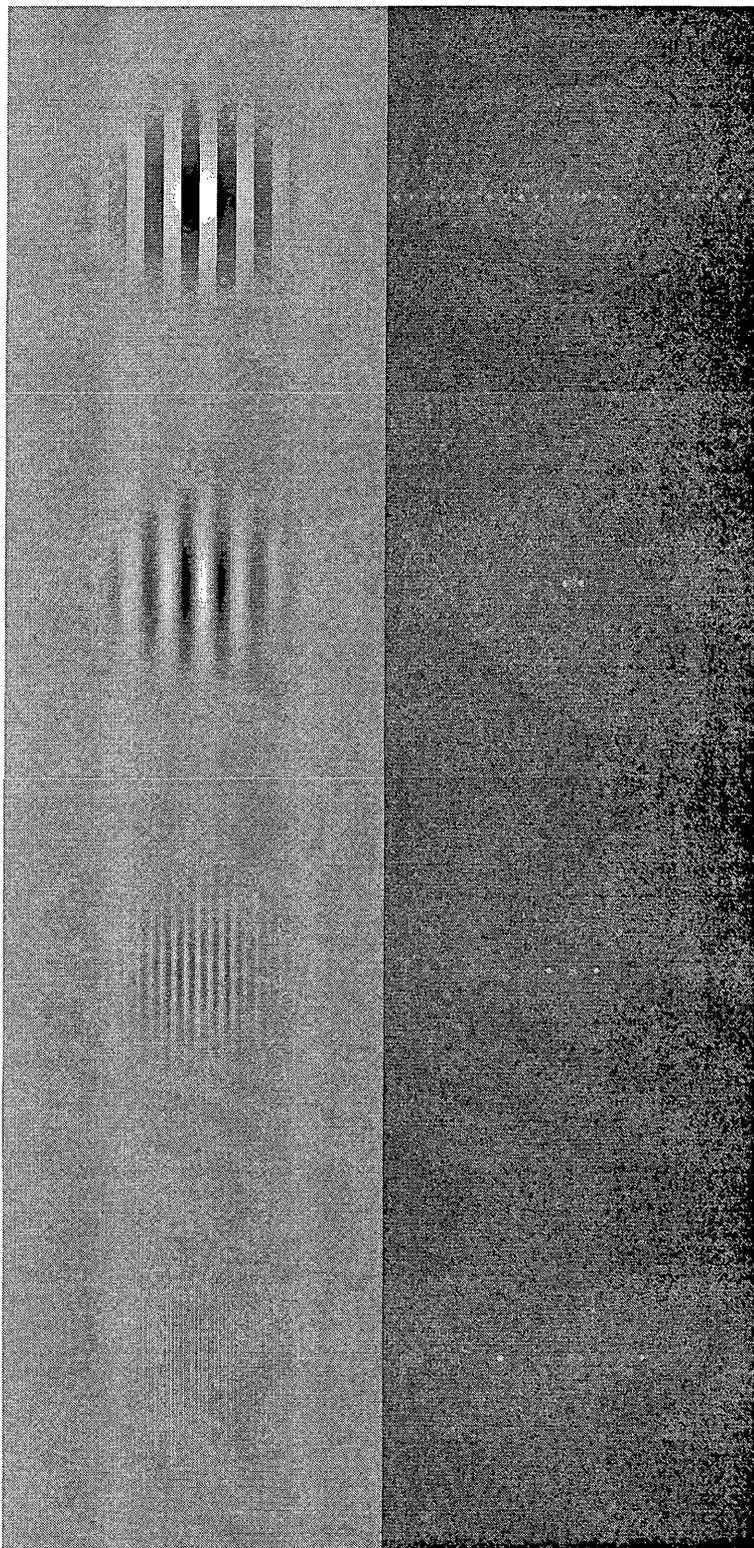


Figure 3.1: shows the stimuli used in the experiments presented in this chapter. The left hand column shows, running from top to bottom, the square-wave which has been convolved with a Gaussian, the fundamental frequency F , the $3F$ spatial frequency and the $5F$ spatial frequency. The right hand column shows the corresponding power spectra.

This operation was carried out so that the square-wave appeared in the same format as the Gabor patches. The contrast of the square-wave target was varied depending upon its position within the image sequence. This deviates from the Parker et al (1992) experiment as they used full contrast stimuli as replacements at all times. However a pilot study indicated that the higher contrast square-wave patch ($\pi/4$ (93.75) = 73.6%) was quite easily detectable when presented within image triplets thus subjects were less inclined to indicate the presence of a square-wave unless it was clearly seen. As one of the aims of the experiment is to examine the error rates made by subjects when the target is not present it was decided that the square-wave contrasts match that of the missing components. This increased the overall error rates found and therefore allowed the examination of an order effect. Thus square-waves were shown at 93.75, 31.25 and 18.75%.

All stimuli were created through software which utilised a VSG display system (Cambridge Research Systems Ltd) and were presented on a Gamma corrected 21 inch (1024 x 768 pixels) monochrome Eizo flexi-scan 6600-M monitor with a refresh rate of 69 Hz and a mean luminance of 12.9 cd/m². It was ensured that the spatial integration of the stimuli employed in these experiments did not cause a deviation away from this mean luminance. Stimuli were displayed for 42 msec each all with a step temporal onset and offset. These were displayed centrally and viewed binocularly at a distance of 1m. The subject's head was placed in a chin and forehead rest throughout the duration of the experiments.

Design. The experiment was a repeated measured design with each of the six components (F, 3F, 5F, SQ (F contrast), SQ (3F contrast), SQ (5F contrast)) being shown singly in a random order. There were 20 trials per condition.

Procedure. The subjects were fully informed about the nature of the task, including being shown examples of the images in the sequences, the form the trials took and the identity of the target at all contrasts. Once the subjects were satisfied they initialised the experiment with a key press, after which a single image was presented. Subjects indicated “yes” for the presentation of a square-wave or “no” for the presentation of a harmonic by different response keys. This took place in a blacked out room

3.2.1.2 Results

The results in figure 3.2 show the mean number of square-wave discriminations as a function of temporal order. All the square-wave targets are discriminated with over 93% accuracy, with both the square-wave shown at full and one-third contrast showing 99.5% detectability. The 3F and 5F harmonics are never detected as a square-wave and only 3% of trials show F being mistaken for a square-wave.

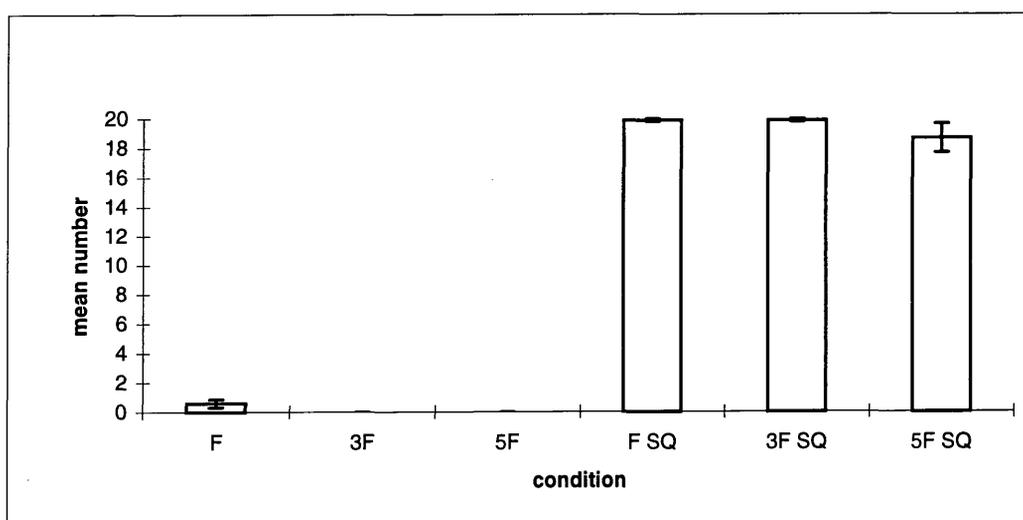


Figure 3.2: Shows the mean number of square-wave presented responses (out of 20 maximum) as a function of image type. Vertical bars are the standard errors.

3.2.1.3 *Discussion*

This experiment shows that the targets are easily detectable as being square-waves while the harmonics almost never are. This experiment shows that a square-wave of any contrast is easily discriminable from the first three harmonics of its square-wave series. Furthermore, the experiment provided an opportunity for the subjects to become familiarised with the targets and distracters to be employed in the next experiment.

3.2.2 *Experiment 5b: Detection*

In this experiment the images were again grouped into triplets as in Chapter 2 experiment 1b, 2b, 3b and 4. The first three harmonic components of a square-wave were presented in either coarse to fine or fine to coarse temporal order. A square-wave target replaced one of the spatial frequency components. It was expected, if a temporal anisotropy in the integration of spatial frequencies existed, that the coarse to fine presentation of the three harmonics would evoke more square-wave present responses than the fine to coarse presentations. This finding would discount the last frame explanation which predicts that a fine to coarse temporal order would evoke more errors. If, however, the subjects' decisions were based upon the last frame shown then the results should show a varying pattern: in the conditions where no target was presented the fine to coarse sequence should be reported as having contained a square-wave more than the coarse to fine sequence; while if the square-wave target was present the results should change with the position of the target within the spatial frequency sequence, such that in both the SQ/ 3F / 5F and F / SQ / 5F sequence a target should be reported more often in the fine to coarse temporal order, while in the F / 3F / SQ sequence a target should be reported more often in the coarse to fine temporal order. The predictions from a last frame explanation suggest that the results should show both a coarse to fine preference and a fine to coarse preference.

3.2.2.1 *Method*

Methods are as experiment 5a with the following exception.

Design. The experiment was a repeated measures design employing a simple yes / no decision task. The stimuli were grouped into triplets (126 msec overall). There were six possible sequences built from an initial two containing no square-wave, F / 3F / 5F or 5F / 3F / F; where one component was replaced by a square-wave with a contrast equivalent to it, for example F / Square-wave / 5F or 5F / 3F / Square-wave. Subjects made 50 decisions per sequence condition, thus 400 overall.

Procedure. The subjects were fully informed about the nature of the task, including being shown examples of the stimuli which they were familiar with from experiment 5a, the temporal order which they would take, the form the trials took and the identity of the target at all contrasts. Once the subjects were satisfied they proceeded with the experiment. Once the subjects initialised the experiment with a key press a triplet of images was shown for 126 msec after which it was replaced by a screen of mean luminance of the Gabor patch. Subjects responded with a key press either yes or no to indicate the presence of the square-wave target. This took place in a blacked out room.

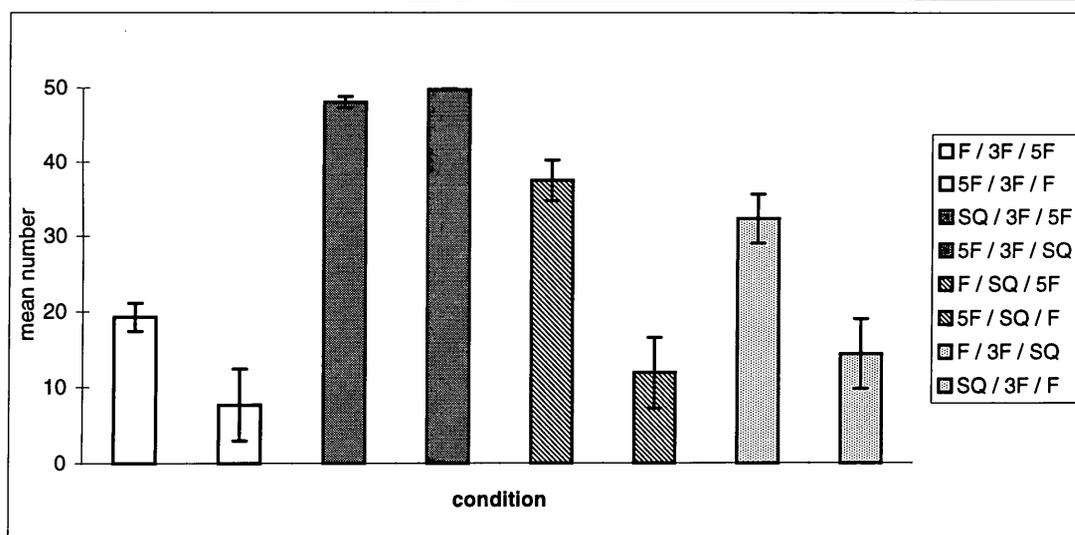


Figure 3.3: shows the mean number of square-wave present responses as a function of image triplet. The triplets are shown in the legend from top to bottom and these correspond, from left to right, to the bars shown on the graph. The columns are paired so that the two orderings of triplets containing the same images are paired together. Vertical bars are the standard errors.

3.2.2.2 Results

The results are illustrated in figure 3.3, showing the detectability of square-waves as a function of presentation order. It can be seen that performance is dependent upon the sequence of spatial frequencies presented and the order in which they appear. If the square-wave target appeared at any point in the degraded image sequences it was significantly more likely to be reported than when it was not presented ($F_1 = 115.8$; $p < 0.01$). A two-way ANOVA examining image set (spatial frequencies present) and image order shows two main effects (image set: $F_{1,9} = 72.86$; $p < 0.01$; image order: $F_{1,9} = 16.34$; $p < 0.01$) and a significant interaction between them ($F_{3,27} = 14.53$; $p < 0.01$). There were four planned critical comparisons in this analysis, these reflect performance differences resulting from the same stimuli presented in the different temporal orders of coarse to fine and fine to coarse. F, 3F, 5F ($F_1 = 14.6$; $p < 0.01$) F, Square-wave, 5F ($F_1 = 71.3$; $p < 0.01$) and F, 3F, Square-wave ($F_1 = 34.8$; $p < 0.01$) all show significant order effects while Square-wave, 3F, 5F ($F_1 = 0.3$; $p > 0.05$) does not.

The analysis carried out shows that coarse to fine presentation of spatial frequencies improves the detection of a square-wave target in two of the three conditions when a square-wave target was actually present. This trend carries over to the two conditions when no square-wave is present where a square-wave is more likely to be mistakenly detected if that sequence of spatial frequencies was shown progressively from coarse to fine.

3.2.2.3 *Discussion*

The results show that a coarse to fine presentation of spatial frequencies leads to an increase in square-wave present responses even if the target was not in fact present. This provides support for the temporally anisotropic hypothesis put forward by Parker et al (1992) and examined in Chapter 2 using simple spatial frequency stimuli rather than natural images. The pattern of the results is different from that expected if subjects' responses were based solely on the last image shown or any concept of a temporally isotropic weighted average model of spatial frequencies integration. First the conditions on which the target was not present shows a significant increase in the number of target present responses made when spatial frequencies are presented from coarse to fine which is not explicable in terms of a last frame explanation. Second the trials on which a target was present also do not show the variation in the pattern of results predicted if subjects were basing their decisions purely on the last frame shown.

This suggests that neither the last frame explanation nor a temporal anisotropy in the integration of spatial frequencies can explain the results found when a fullbandwidth target was presented in the image sequences whether this sequence comprises of natural images or Gabor patches. First, while it was found in Chapter 2 that the pattern of results found when a fullbandwidth image was presented in the temporal sequences could be explained if the subject was basing their decision on the last frame alone, the

predicted pattern from this last frame explanation was not found here. Nor was this pattern was found in Parker et al (1992; 1997). Second, the pattern of results found when a fullbandwidth natural image is present within a natural image sequence cannot be explained on the basis of a temporal anisotropy in the integration of spatial frequencies. Therefore the conditions in which a target is present within the image triplet cannot be safely interpreted as reflecting the integration of spatial frequencies.

In conclusion, the conditions of primary importance in a detection task are those in which the target was not actually present. The results from these conditions in the square-wave detection experiment show that the coarse to fine presentation of spatial frequencies was more likely to be mistakenly identified as having had a square-wave present in its sequence than the fine to coarse temporal presentation of the same harmonics. This supports the temporally anisotropic model of spatial frequency integration suggested by Parker et al (1992; 1997) and Schyns & Oliva (1994).

It must be pointed out however, that there is some difficulty to be had in interpreting these results safely. While the results are well accounted for by a temporal anisotropy in the integration of spatial frequencies there are other possible interpretations. For example, subjects may be basing their square-wave detection judgements on the fundamental component F alone. In the sequence F $3F$ $5F$, the fact that the F is succeeded by the $3F$ and $5F$ will serve to obscure its form and may increase the likelihood that some false square-wave present reports are made. This is unlikely to be due to masking as this has been found to behave such that lower spatial frequencies backward mask higher spatial frequencies more than vice versa, therefore this masking of the F by the $3F$ and $5F$ would not be expected. However, the point is that there are other explanations for the pattern of results found which do not involve the integration of spatial frequencies.

There is also a difficulty of interpretation when the task is considered. It is not clear exactly what the subjects are doing in this square-wave detection task. It was suggested that because subjects know the identity of the target they are actively seeking its form in the triplet sequence. However it may be the case that the subjects are actually looking for specific patterns of a square-wave target plus other harmonics. Thus when the target is not present the three harmonics integrate to form a square-wave like representation they do not integrate to form the pattern of a square-wave plus other harmonics which the subject is actively seeking. It may be the case then that the conditions where the target is not present do not reflect the integration of spatial frequencies at all but rather the bias of the subjects to judge that a sequence of harmonics appears to be a sequence of square-wave plus harmonics.

3.3 Experiment 6: Square-wave likeness

It has been suggested that the results of experiment 5 are due to the operation of a temporally anisotropic process in the integration of spatial frequencies: That is spatial frequencies are integrated more effectively if they are presented from coarse to fine thus mirroring the natural availability of spatial frequencies due to the delays involved in their processing. Therefore the reason more square-wave present responses are made when a target is not present and spatial frequency information is shown progressing from coarse to fine is that the information is more effectively integrated to form some sort of square-wave representation. To examine this further an experiment was carried out testing a simple prediction of this result namely that a coarse to fine sequence of spatial frequencies should appear to resemble a square-wave more so than a fine to coarse sequence. This prediction formed the basis of Experiment 6.

The experimental method of quality judgements (experiment 1) was not employed in this experiment as it is difficult to see on what dimension spatial

frequencies are qualitatively related, for example in what way is a 3F qualitatively related to a F spatial frequency. Thus when natural images are employed the filtered images and fullbandwidth image are clearly related as they depict the same structures. However it is clear that individual spatial frequencies do not (see section 3.1.1). Therefore it was decided a judgement of resemblance or likeness would be employed as this is still a qualitative judgement and it explicitly examines the same aspect of the overall image triplet which quality judgements were designed to reflect.

This experiment took the form of a temporal 2AFC. In one interval either a F / 3F / 5F or 5F / 3F / F pattern appeared; in the second interval the opposite appeared, i.e. if F / 3F / 5F in the first interval then 5F / 3F / F in the second or if 5F / 3F / F in the first interval then F / 3F / 5F in the second. The subject had to indicate which appeared to be more like a square-wave. This is a strange task for the subject as there are in fact no square waves at all in the experiment just the first three harmonics presented one after the other, therefore there is no right answer. However subjects did not find this task uncomfortable to complete.

3.3.1 Method

Subjects. There were ten subjects, three female and seven male with an age range of 25 to 34. One of the subjects was the author and so was fully aware of the purpose of the experiment, the remaining nine were naive.

Apparatus and Materials. Methods are the same as those outlined in Experiment 5. The stimuli used in this experiment took the form of vertically oriented spatial frequency Gabor patches of the first three harmonics of a square-wave exactly the same as those used in Experiment 5.

Design. The task is a judgement of likeness task employing a temporal two - alternative - forced - choice paradigm. Subjects were shown either a sequence of F / 3F / 5F and 5F / 3F / F, or one in the reverse order, separated temporally by a blank screen of mean luminance. Subjects were asked to indicate which interval they perceived to be more like a square-wave Gabor patch. This was a within-subjects design as each subject performed all combinations of the variables in a single test period. Performance was measured as the number of times a particular presentation resulted in a square-wave likeness judgement as a function of the order of their presentation. Subjects were required to make 200 decisions, 100 per condition, with both conditions randomly presented.

Procedure. The experiment was carried out in a blacked out room. Subjects were informed what the task required. Subjects were shown examples of the stimuli set, including a square-wave of the same size and fundamental frequency as the experimental stimulus F, and examples of the temporal form that the trials were to take. The subject proceeded with the experiment when fully satisfied. All subjects' data were collected in a single block session.

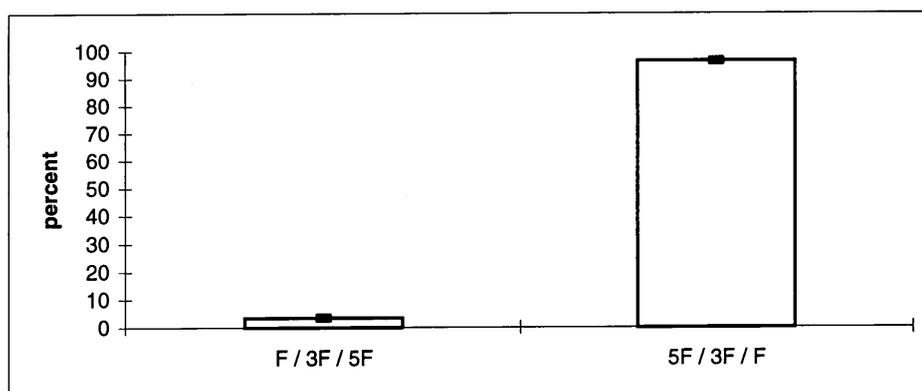


Figure 3.4: showing the percentage of square-wave likeness responses as a function of spatial frequency triplet temporal order. Vertical bars are the standard error of the mean.

3.3.2 Results

The results can be seen in figure 3.4. It was found that the temporal order of spatial frequencies did have an effect upon subjects' square-wave like responses. Subjects predominately chose (96.45%) the presentation order of the triplet of Gabor patches which showed the 5F initially followed by the lower spatial frequencies of 3F and F. On only 3.55 percent of the trials was the coarse to fine presentation indicated as being more square-wave like.

3.3.3 Discussion

The results clearly show that there is a definite preference for the 5F / 3F / F triplet of spatial frequencies as being more square - wave like than their converse. This is a surprising result as it would be expected, from the results of the square - wave detection experiment (see also experiment 1; Parker et al, 1992; 1997), that the F / 3F / 5F triplet would be indicated as being more square-wave like as it represents a coarse to fine progression of spatial frequency information and has been suggested to mirror the natural progression of spatial frequency integration shown by the visual system (Parker et al, 1992).

Integration and spatial frequency overlaps

It could be argued that the results of experiment 6, supporting a fine to coarse integration of spatial frequencies, may be due to the overlap between the constituent spatial frequencies when they are presented from fine to coarse due to the temporal delays involved in the processing of spatial frequencies.

It has been found that there are differences in the processing times of different spatial frequencies, such that processing times increase with spatial frequency

(Breitmeyer, 1975; Legge, 1978; Vassilev et al ,1979; Gish et al, 1986; Parker & Dutch, 1987; see section 1.3.2). When this evidence is applied to the stimuli presented in the square-wave likeness experiment it follows that there is a qualitative difference between the presentation orders which is not evident physically. Thus, although the three spatial frequency components are presented physically at time $t_1 = 0$, time $t_2 = 42$ msec and time $t_3 = 84$ msec, due to the neurally imposed delays this physical time outline may not stand. It is possible, with reference to the outline of spatial frequency processing delays presented previously (See section 1.3.2), to estimate the neural onset and offset times of the stimuli (assuming onset and offset response decay to be the similar. See Parker, 1982; Parker & Dutch, 1987; Baro, Brzezicki, Lehmkuhle & Hughes, 1992). If the initial processing delay of low spatial frequencies is assumed to be 4 msec estimated physiologically (Lennie et al ,1990) and smallest delay per octave for spatial frequencies beyond this is taken as being 5 msec (Parker & Dutch, 1987). The differences between the physical and neural times can be estimated and are shown in table 3.1.

Presentation Order	Physical		Neural	
	Onset	Offset	Onset	Offset
5F	0	42	19.25	61.25
3F	42	84	53	95
F	84	126	88	130
F	0	42	4	46
3F	42	84	53	95
5F	84	126	103.25	145.25

Table 3.1: Showing the predicted neural delays involved in the processing of spatial frequencies in msec following a conservative estimate of neural delay of 5 msec per octave (Parker & Dutch, 1987).

Thus it can be seen that the stimuli presented in the first set of experiments may not have the sharply defined physical onset and offset times but may reflect the neural presentation times. If a liberal estimate of the delay per octave, 21 msec (Gish et al , 1986), is taken then the processing times would be those shown in table 3.2.

It is apparent that in both estimates, when the neural processing times are examined, that there occurs an overlap or a gap between the components which is not physically present. Thus when the stimuli were presented 5F / 3F / F there would occur an overlap between the components; whilst in the F / 3F / 5F presentation order case there would be a gap between the two components. The length of this overlap and gap can be seen to vary depending upon the estimate of the processing delay per octave which is taken.

Presentation Order	Physical		Neural	
	Onset	Offset	Onset	Offset
5F	0	42	50.2	92.2
3F	42	84	77.5	119.5
F	84	126	88	130
F	0	42	4	46
3F	42	84	77.5	119.5
5F	84	126	134.2	176.2

Table 3.2: Showing the predicted neural delays involved in the processing of spatial frequencies in msec following a liberal estimate of neural delay of 21 msec per octave (Breitmeyer, 1975).

Thus it may be the case then that the 5F / 3F / F presentation sequence is chosen to be more square-wave like because of this temporal overlap in their neural presentation. Thus a fine to coarse sequence may be chosen as more like a square-wave because the three components are neurally available simultaneously and integrate more readily to form a square-wave. However, it is unclear, if this is the case, why the square-wave detection results (experiment 5) and the results from the image quality experiment (experiment 1) show evidence for a coarse to fine integration of spatial frequencies. Furthermore, visual processing of spatial frequencies in normal circumstances would lead to a coarse to fine presentation of spatial frequency information because of the delays involved in their processing. Therefore it must be this temporal nature of spatial frequency processing which the visual system has evolved to deal with (See Chapter 6 for further discussion of spatial frequency processing delays). Because of these reasons a neural delay explanation for the results of experiment 6 is discounted.

3.4 General Discussion

The results of experiment 5 show that a square-wave target was reported as having been present within a sequence of spatial frequencies in which it was not actually shown. This was found to depend critically upon the order in which the spatial frequencies were presented, with more square-wave present reports being made when the spatial frequencies were shown from coarse to fine than when the same spatial frequencies were shown from fine to coarse. This was suggested to reflect the operation of a temporal anisotropy in the integration of spatial frequency such that integration of spatial frequencies improves if they are presented to the visual system from coarse to fine over a period of time. It was proposed that this temporal anisotropy is a result of the natural progression of spatial frequencies resulting from the inherent delays involved in

their processing. Thus a coarse to fine sequence of spatial frequencies is more effectively integrated because that is the temporal order in which they are naturally available to the integration process. For example, the reason a square - wave was mistakenly detected more often in the $F / 3F / 5F$ triplet was because this triplet was integrated more effectively than its converse order and appeared to be more square - wave like.

A coarse to fine preference was also expected in the square - wave likeness experiment. Experiment 6 shows that, contrary to expectation, the fine to coarse sequence of spatial frequency information is judged to resemble the target or square-wave than the coarse to fine sequence. Thus the results of experiment 5 suggest that the reason a square-wave is more often mistakenly reported as being present in a coarse to fine sequence of spatial frequencies was because that sequence integrates more effectively to form a more target like representation than the fine to coarse sequence. However experiment 6 shows that it is the fine to coarse triplet that appears to be more square - wave like, suggesting the exact opposite temporal anisotropy in the integration of spatial frequencies. This result suggests that the reason that the $F / 3F / 5F$ presentation order is more often mistaken as containing a square - wave is not because it forms a more target-like representation.

Overall then, it does not seem that the results of the square-wave detection experiment are a function of a temporal anisotropic preference of coarse to fine signal integration. However, it could be argued that the integration model is not fixed to one temporal anisotropy, coarse to fine, but rather is a flexible mechanism which processes spatial frequency information as a function of task demands (see section 1.3.3). This is a form of spatial frequency integration whose temporal form varies as a function of the information to be derived from the image. This interpretation gains support from the work by Schyns & Oliva (1997; Oliva & Schyns, 1997) and Parker, Lishman & Hughes

(1996)(see section 1.3.3) and a consideration of the various models for posited for different aspects of visual processing (see section 1.2.1).

Another simpler explanation for the results showing both a coarse to fine and a fine to coarse temporal anisotropy in the integration of spatial frequencies is that the results of the square-wave likeness experiment do not reflect the integration of spatial frequencies at all. Given the substantial empirical evidence showing differential delays involved in the processing of spatial frequencies (Breitmeyer, 1975; Lupp, Hauske & Wolf, 1976; Parker & Salzen, 1977; Gish et al, 1986; Parker & Dutch, 1987; see section 1.3.2) and evidence for a coarse to fine preference in the processing of spatial frequencies (Parker et al, 1992; 1997; Schyns & Oliva, 1994), it seems likely that the square-wave likeness result of experiment 6 does not involve the integration of spatial frequencies at all. The following chapter will examine possible explanations for the square-wave likeness result of experiment 6 which do not involve the integration of spatial frequencies.

Spatial frequency interactions

Summary

In this chapter the anomalous finding from the square-wave likeness task of experiment 6 is investigated. First the underlying nature of the task and its relationship with the natural image quality task of experiment 1 is examined. Filtered versions of a natural image are presented in a likeness two alternative forced choice paradigm and subjects judge which appears to be more fullbandwidth natural image like (Experiment 7). It was found that the coarse to fine sequence of images is judged to more resemble the fullbandwidth image than the fine to coarse image triplet. This pattern of coarse to fine order preference was also found in the image quality experiment presented in chapter 2 (experiment 1). This suggests that the two tasks of image quality and likeness reflect the same process. Second, it was shown that the fine to coarse sequence of two, rather than three, spatial frequencies was also judged to be more square-wave like than a coarse to fine sequence of two spatial frequencies (Experiment 8). This shows that it is possible to examine the square-wave likeness result by assessing the interactions between two spatial frequencies. The first interaction examined was the possibility that the judgement of square-wave likeness was based on the last frame shown (Experiment 9a and 9b). It was found that the spatial frequency shown in the penultimate frame did have an effect of the judgement of square-wave likeness, with the fine to coarse

sequence of the first two spatial frequencies of a square-wave series being chosen more often to resemble a square-wave. The second interaction examined was that the preference for a fine to coarse spatial frequency sequence in judgements of square-wave likeness may be due to masking such that the fundamental frequency masks the high harmonics more when it succeeds them than when it precedes them (Experiment 10). This would lead to more fine to coarse preferences as the high harmonics would be more obvious in the coarse to fine sequences. This was indeed found to be the case. Thus it may be case that the square-wave likeness preference for a fine to coarse sequence of spatial frequencies may be the result of masking interactions.

4.1 Introduction

The results from experiment 5 and experiment 6 examining square-wave detection and square-wave likeness showed completely different outcomes. The square-wave detection experiment found that a square-wave was more likely to be detected when the first three spatial frequencies were presented sequentially from coarse to fine and the square-wave target was not present. This suggested that the spatial frequency triplet was more effectively integrated when it was presented from coarse to fine rather than fine to coarse, that is the coarse to fine triplet resulted in more mistaken square-wave present responses because the triplet integrated to form a degraded square-wave like representation. The square-wave likeness experiment, in which subjects judged which spatial frequency triplet more resembled a square-wave, either the coarse to fine or fine to coarse sequence, found that, contrary to expectation, the fine to coarse triplet was judged to be more square-wave like. It was suggested that this may be due to the operation of a flexible integration mechanism where the temporal preference for spatial

frequency integration varies with the task as a function of the information to be derived from the image (see section 1.3.4 & 1.4). However, a simpler explanation is that the square-wave likeness result did not reflect the operation of a spatial frequency integration process, but rather the subject was basing their decision upon some other aspects of the stimuli. The possible explanations for this fine to coarse preference in the integration of spatial frequencies will be individually examined in this chapter.

4.1.1 Preview of Chapter

The purpose of this chapter is to investigate the result of the square-wave likeness task of experiment 6 which suggests that spatial frequencies are preferentially integrated when they are available from fine to coarse. This goes counter to the results of experiments 1 and 5 and those carried out by Parker et al (1992; 1997) which both suggest that the temporal order preference of spatial frequencies is from coarse to fine. It is also anomalous to that expected given the delays involved in the processing of spatial frequencies. It is for these reasons that the task of square-wave likeness is to be examined more closely in this chapter.

Experiment 7: Natural image quality and Square-wave likeness

Experiment 1 found that when subjects judge the quality of a triplet of filtered images, those shown sequentially progressing from coarse to fine in spatial frequency content are rated as being of better quality than those shown progressing from fine to coarse. If these stimuli are changed from filtered natural images to spatial frequency Gabor patches which when summed form a square-wave then the image sequence judged to most resemble a square-wave is the fine to coarse sequence. This result is the exact opposite to that found when filtered natural images are employed. It was suggested (section 3.1.3) that these two tasks, quality ratings and likeness, reflect the same

underlying integration process: how much a sequence of spatial frequencies are judged to be like or similar to the fullbandwidth image, whether it be a natural image or a square-wave. However, it may be the case that the two paradigms do not reflect same process.

In order to examine this it is necessary to run the same paradigm with the two stimuli. It has been pointed out (section 3.1.3) that it is not sensible to rate the quality of the sequences of spatial frequencies. Therefore it was decided to present sequences of filtered natural images using the two alternative forced paradigm of the square-wave likeness experiment. Thus filtered images were presented from coarse to fine and fine to coarse in two temporal intervals, subjects indicated which appeared to more resemble the fullbandwidth natural image standard.

Experiment 8: Interactions between spatial frequencies

In order to further investigate the preference for a fine to coarse presentation of spatial frequencies found in the square-wave likeness experiment, it is necessary to examine whether the interactions between the spatial frequencies is of an integrative nature. To make this a more viable venture the number of spatial frequencies must be reduced from three to two. The experiment presented in this section examined the square-wave likeness judgements with two spatial frequencies. If the pattern of results are the same under these conditions as the results found when three spatial frequencies are employed then it can be suggested that the same interaction is responsible.

Experiment 9: Last Frame

In section 3.1.4 it was suggested that the results from experiment 1 to 4 employing natural images were a consequence of the subject basing their decision upon the last image shown. It was shown in experiment 5, using a square-wave detection

paradigm, that this last frame explanation could not explain the results found. It remains a possibility, however, that the results of the square-wave likeness experiment may be due to the last frame dictating the subjects response. Thus, the reason the $5F / 3F / F$ sequence was chosen as resembling the square-wave more than the $F / 3F / 5F$ sequence is because the $5F$ does not appear to resemble a square-wave at all, while the F has the same periodicity and phase as the square-wave standard.

If this is the case then it follows that the spatial frequency presented in the penultimate frame should have no effect on the subject's decision. Therefore a simple test of this last frame explanation is to again assess the square-wave likeness of a sequence of spatial frequencies. However in these two experiments two fine to coarse sequences were presented. In the first experiment, experiment 9a, one of these the square-wave harmonics of $3F / F$ were presented, while in the other a $4F / F$ sequence was shown. If the two spatial frequencies are interacting to form some sort of square-wave like representation then the $3F / F$ should be chosen to resemble the square-wave more often than the $4F / F$. If on the other hand the last frame explanation can account for the square-wave likeness result then both fine to coarse sequences should have equal probability of being chosen. In the second experiment, experiment 9b, the phase relationship between the F and the $3F$ was varied rather than the frequency relationship as in the first experiment. Two fine to coarse sequences of $3F / F$ were presented, one with the $3F$ in peaks subtract phase and the other with the $3F$ phase shifted by 180° in peaks add phase. As with experiment 9a, if the square-wave likeness judgement is based purely on the last frame shown then both fine to coarse sequences should have equal probability of being chosen.

Experiment 10: Masking

Finally, it may be the case that the reason that a fine to coarse sequence results in more square-wave like responses is because of interactions between the spatial frequencies other than spatial frequency integration. It has been shown that interactions between different spatial frequencies do occur (Nachmias, et al, 1973; Nachmias & Weber, 1975; Olzak & Thomas, 1981; Georgeson & Georgeson, 1987; Georgeson & Meese, 1997).

Thus it may be the case that the fine to coarse sequence is judged as being more square-wave like not because it is integrated to form a square-wave like representation but rather the fine spatial frequencies become less visible because they are backward masked by the coarse spatial frequency. Thus the $5F / 3F / F$ sequence is chosen to be more square-wave like than the $F / 3F / 5F$ sequence because the $5F$ and $3F$ are less visible in the fine to coarse sequence because the F has more of a backward than forward masking effect on the contrast of the $3F$ and $5F$. So the high spatial frequency components are more visible in the coarse to fine sequence thus it appears to be less like a square-wave than the fine to coarse sequence where the high spatial frequencies are less visible. Thus the judgement of resemblance is a relative one and it depends not upon the integration of the spatial frequencies but upon the masking interactions that occur between them.

In order to examine this the contrast thresholds of a $3F$ Gabor patch were measured when it was preceded or succeeded by an F under the same contrast and temporal conditions as employed in all the square-wave likeness experiments.

4.2 Experiment 7: Natural image likeness

It may be the case that the quality judgements made for the natural image stimuli are not directly related to the resemblance measure employed in experiment 6. In order

to investigate this link, the resemblance of natural image triplets to a fullbandwidth image target was examined. The triplets were grouped into opposite temporal orders triplets which contained exactly the same three filtered images. Subjects decided which appeared to more resemble the fullbandwidth image standard.

4.2.1 *Method*

Subjects. There were ten subjects, three female and seven male with an age range of 25 to 34. One of the subjects was the author and so was fully aware of the purpose of the experiment, the remaining nine were naive.

Apparatus & Materials. The stimuli and apparatus used in this experiment are identical to those used in experiment 1 and 3 of Chapter 2. The low and high pass half height cut off points are at 0.5 and 1 cpd. No fullbandwidth image was used.

Design. The stimuli were grouped into triplets as with experiment 3: L0.5 L1 H1; H1 L1 L0.5; L0.5 H0.5 H1; H1 H0.5 L0.5. The four image triplets were grouped according to which set of spatial frequencies they contain. These two sets were presented randomly in a two temporal interval forced choice paradigm as used in experiment 6. For example if L0.5 L1 H1 was presented in the first temporal interval, H1 L1 L0.5 was presented in the second. Subjects made 50 decisions per set and temporal condition, totalling 200 decision overall.

Procedure. The experiment was carried in a blacked out room. Subjects were informed that the task required them to indicate which of two temporal sequences of filtered natural images, one coarse to fine and the other fine to coarse containing exactly the same spatial frequency content, most resembled a fullbandwidth version of that image.

Subjects were shown examples of the stimuli set, including the fullbandwidth image from which the filtered images were derived and the filtered images themselves.

Examples of the temporal form that the trials were to take were also given. The subject proceeded with the experiment when fully satisfied. All subject's data were collected in a single block session.

4.2.2 Results

Figure 4.1 shows the percentage of trial on which an image triplet was judged to resemble the fullbandwidth image from which it was derived. It was found that the temporal order of spatial frequencies did have an affect upon subjects fullbandwidth natural image-like responses. Subjects predominately chose the presentation order of the triplet of filtered images which presented spatial frequencies progressively from coarse to fine (L0.5 L1 H1 = 96.6%; L0.5 H0.5 H1 = 86%). On only a small percent of the trials was the fine to coarse presentation indicated as being more fullbandwidth image-like (H1 L1 L0.5 = 3.4; H1 H0.5 L0.5 = 14). These differences were found to be significantly different (L0.5 L1 H1 vs H1 L1 L0.5: $t = 66.35$; d.f. = 9; $p < 0.01$; L0.5 H0.5 H1 vs H1 H0.5 L0.5: $t = 9.85$; d.f. = 9; $p < 0.01$).

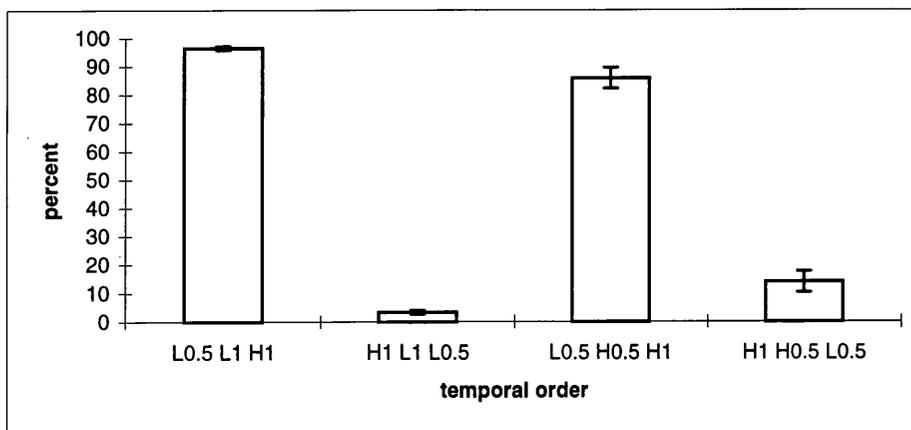


Figure 4.1: the percentage of natural image like judgements as a function of the temporal order of spatial frequencies. The Y vertical bars are the standard error bars.

4.2.3 Discussion

The results clearly show that a coarse to fine sequence of filtered images are judged to resemble the unfiltered natural image from which they were derived, more so than the same image sequence shown from fine to coarse. This result shows the same pattern of results as that found with the natural image quality paradigm (Experiment 1, Chapter 2). This suggests that the task of resemblance judgements reflects the same processes as the task of rating image quality. Thus the differences in task cannot account for the fine to coarse temporal order preference. Consequently it appears that different temporal anisotropies exist in the integration of spatial frequencies depending upon the stimulus employed. This may reflect the flexible integration process which has been suggested in section 1.4.

4.3 Experiment 8: Square-wave likeness with two harmonics

The preference for a coarse to fine temporal order in the integration of spatial frequencies found using a natural image has been shown to reflect the same processes as those underlying the fine to coarse preference found when subjects carried out the same task using Gabor patches in a square-wave likeness experiment (experiment 6: chapter 3). This suggests that the likeness task and the image quality judgements reflect the operation of the same underlying processes. These results indicate that spatial frequencies are integrated from either coarse to fine or fine to coarse. However it remains the case that the fine to coarse integration of spatial frequencies found in the square-wave likeness experiment (experiment 6) may be due to some other interaction between the spatial frequency components rather than integration. In order to uncover what form this interaction takes it is necessary to investigate the relationship between

the spatial frequencies shown. A problem with this is that there are three spatial frequencies which go to make up the image triplets, therefore the interactions between them are difficult to assess. In order to assess these spatial frequency interactions then it is necessary to reduce the number of spatial frequencies shown in the image sequences to two.

The following experiment repeated the square-wave likeness paradigm of experiment 6 with two spatial frequencies. The image sequences now consisted of the fundamental (F) and the third harmonic (3F) of a square-wave series being shown from coarse to fine (F / 3F) or from fine to coarse (3F / F) in a temporal two alternative forced choice design. Subjects indicated which appeared to resemble a square-wave more. Following from experiment 6, it was expected that the fine to coarse temporal order of spatial frequencies would be judged as resembling a square-wave more than a coarse to fine temporal order.

4.3.1 *Method*

Subjects. There were six subjects, four female and two male with an age range of 25 to 34. One of the subjects was the author and so was fully aware of the purpose of the experiment, the remaining five were naive

The remaining methods are the same as Experiment 6. The only difference was the reduction in the number of spatial frequencies being presented. Rather than the first three spatial frequencies of a square-wave series being presented only the fundamental frequency (F) and the third harmonic (3F) were shown. These were shown in the same form of temporal order as in experiment 6: coarse to fine (F / 3F); or fine to coarse (3F / F). Again all subject data was collected in a single block session.

4.3.2 Results

Figure 4.2 shows the percentage of trials on which each of the spatial frequency temporal orders were chosen to resemble a square-wave. As can be seen the choice of square-wave resemblance is dependent upon the presentation order of the harmonic pairs. The 3F / F presentation order was chosen 97.4% of the trials, while on only 2.6% of the trials were the coarse to fine pair chosen. This was found to be a significant difference ($t = 52.27$; $d.f. = 5$; $p < 0.01$).

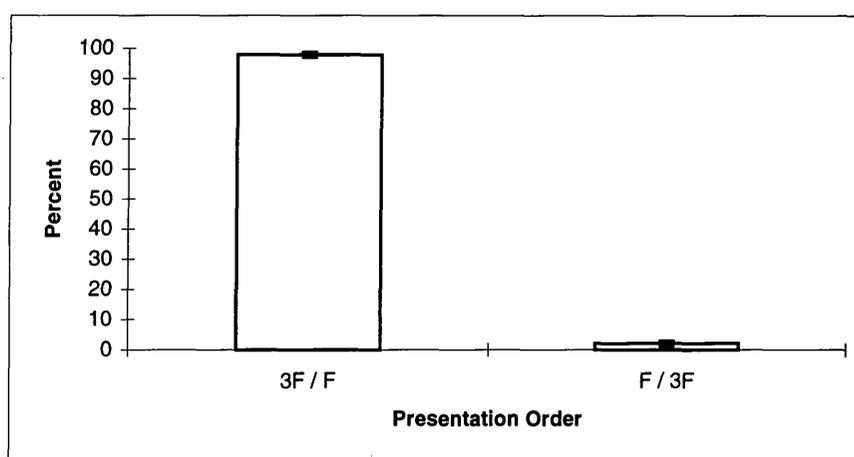


Figure 4.2: shows the percentage of trials on which each temporal order of spatial frequencies was judged to more resemble a square-wave. The vertical bars are the standard error bars.

4.3.3 Discussion

The results show clearly that the fine to coarse temporal sequence of spatial frequencies is judged to more resemble a square-wave than a coarse to fine sequence, replicating the pattern of results found in experiment 6. This experiment was carried out in order to examine the viability of reducing the number of spatial frequencies shown in experiments 6 from three to two. This would allow the interactions which occur between the three spatial frequencies in experiments 6 to be reduced to two and thus enable the nature of these interactions to be investigated. Performance in this experiment shows

exactly the same pattern as that found in experiment 6 when three spatial frequencies are employed. This suggests that the underlying processing which is involved in the interactions between three spatial frequencies is also occurring when two spatial frequencies are shown. Thus this allows a practicable means by which the interactions that occur between spatial frequencies shown temporally from coarse to fine or fine to coarse can be investigated.

4.4 Experiment 9: Last Frame

The results supporting a fine to coarse integration of spatial frequencies could be explained by suggesting that the spatial frequencies are not interacting at all but rather the decision is being carried out on the basis of the last spatial frequency shown. It has been shown that subjects basing their decisions on the last frame shown does not hold as an explanation of the detection experiments (both natural images and square-wave), but this explanation may still account for the square-wave likeness methodology. In the square-wave likeness experiments a fine to coarse image sequence is chosen to resemble the square-wave more than a coarse to fine sequence because the former temporal order shows the fundamental as its last frame while the latter sequence shows the $5F$ or $3F$ as the last frame. Indeed subjective reports from subjects completing experiment 8 involved such comments as “one of the stimuli had more lines than the reference square-wave”. Thus the coarse to fine sequence does not appear to resemble the square-wave standard and it may be the case that it is for this reason that the fine to coarse temporal sequence is chosen to more resemble a square-wave. Thus it may be the case that the fine to coarse sequence does not appear to resemble a square-wave but because the fundamental has the same periodicity and phase as the square-wave standard this image

sequence is chosen to be relatively more square-wave like than the coarse to fine sequence.

While the results of experiment 5 (square-wave detection: see section 3.2) show that a last frame explanation cannot account for the results found there, the last frame explanation may account for the results of experiment 1, 7 and Parker et al (1992) where quality judgements of the filtered images were ascertained. It was found that the high spatial frequency images were judged as being of higher quality than the low spatial frequency images. Thus the coarse to fine bias they found in these experiments could be as a result of the high spatial frequency image being shown last and that being of higher quality or more similar to the full bandwidth image.

In order to empirically address this issue it is possible to compare performance when the last frame remains the same across conditions but the preceding frame differs. If the preceding frame information is not interacting with the last frame then its content should make no difference to performance.

In the following experiments two fine to coarse sequences are shown as in a square-wave likeness experiment only the penultimate frame differs. Thus if the penultimate frame is not interacting with the final frame then each sequence has an equal probability of being chosen as most resembling a square-wave. In the first experiment, experiment 9a, a 3F / F pair is compared to a 4F / F in a square-wave likeness experiment. It would be expected that if subjects are basing their decision upon the spatial frequency shown in the last frame then both of the image sequences should be chosen equally often as most resembling a square-wave. If, however, the spatial frequencies are integrating then the 3F / F pair should be chosen to resemble a square-wave more often than the 4F / F pair as the F and 3F spatial frequencies are the first two of a square-wave series. In a second experiment, experiment 9b, a 3F / F pair is shown

in both intervals only the phase of the 3F has been shifted to form either a peaks subtract or a peaks add phase relationship with the fundamental.

4.4.1 Experiment 9a

4.4.1.1 *Methods*

The apparatus and materials, design and procedure are as Experiment 8.

Variations are noted in the following sections.

Subjects. There were five subjects, three female and two male, with an age range of 25 to 34. One of the subjects was fully aware of the purpose of the experiment the remaining four were naive.

Stimuli. The stimuli employed in this experiment were much the same as those used in experiment 8, with the addition of a 4 cpd Gabor patch (4F), of the same size and contrast as the 3 cpd, which took the place of the 3 cpd in one of the stimulus pairings. In one interval a 3F / F pair were shown and the other a 4F / F pair were shown. Thus the square-wave judgements were compared as a function of the preceding spatial frequency information presented.

4.4.1.2 *Results*

Figure 4.3 shows the percentage of square-wave like responses as a function of penultimate frame shown. There was found to be a large and significant difference in the results dependent upon the preceding frame with 81.2 % of the trials eliciting a square-wave like judgement to the 3F / F pair ($t=-4.46$; d.f. = 4; $p<0.05$).

4.4.1.3 Discussion

The results clearly show that performance differed as a function of the penultimate frame, thus the judgement of square-wave likeness does not depend upon the spatial frequency shown in the last frame alone. This is further supported when it is considered that it is the harmonically related pair which elicits the most square-wave like judgements. Thus this suggests that there are interactions between the spatial frequencies shown fine to coarse and that this interaction is the integration of the spatial frequencies presented.

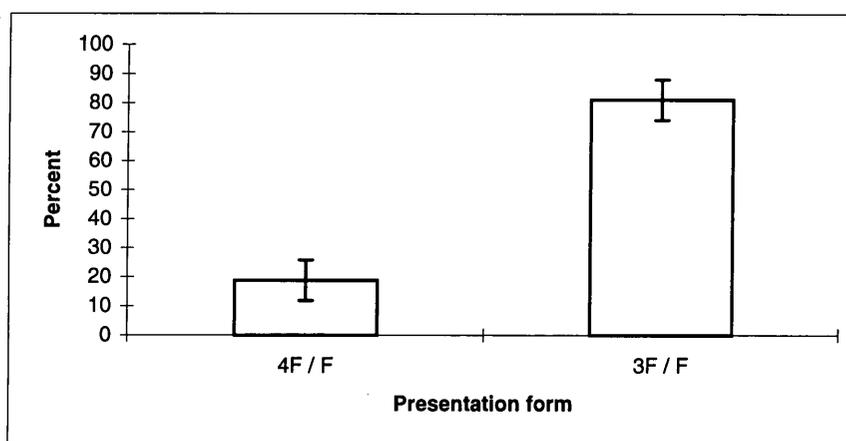


Figure 4.3: The average percentage square-wave likeness judgements as a function of presentation order.

The vertical bars are the standard error bars.

4.4.2 Experiment 9b: Phase

4.4.2.1 Methods

The apparatus and materials, design and procedure are as Experiment 8.

Variations are noted in the following sections.

Subjects. There were six subjects, two female and four male, with an age range of 21 to 40. One of the subjects was fully aware of the purpose of the experiment the remaining five were naive.

Stimuli. The stimuli employed in this experiment were much the same as those used in experiment 8, with the addition of a phase shifted 3F so that its phase relationship with the F was one of peaks add. In one interval a 3F / F pair in peaks subtract phase was shown and in the other a 3F / F pair in peaks add phase was shown. Thus the square-wave judgements were compared as a function of preceding spatial frequency information presented.

4.4.2.2 Results

Figure 4.4 shows the percentage of square-wave like responses as a function of penultimate frame shown. There was found to be a significantly large difference in the results dependent upon the preceding frame with 69.4 % of the trials eliciting a square-wave like judgement to the 3F / F pair in peaks subtract phase ($t=3.07$; d.f. = 5; $p<0.05$).

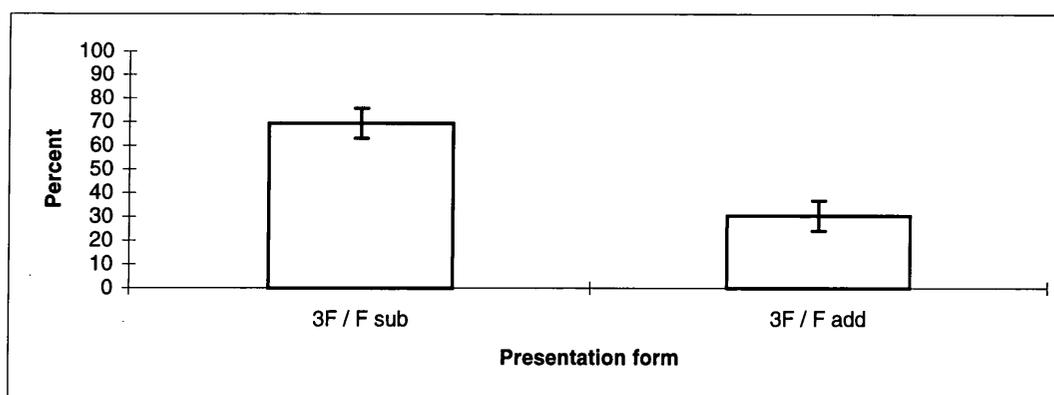


Figure 4.4: The average percentage square-wave likeness judgements as a function of presentation condition. The vertical bars are the standard error bars.

4.4.2.3 Discussion

As with experiment 9a the results clearly show that performance differed as a function of the penultimate frame, thus the judgement of square-wave likeness does not depend upon the spatial frequency shown in the last frame alone. This is further supported when it is considered that it is the square-wave phase related pair which elicits the most square-wave like judgements. Thus this suggests that there is an interaction between the 3F and the F shown fine to coarse. Furthermore the results show that this interaction is one of spatial frequency integration.

4.5 Experiment 10: Masking

Experiments 9a and 9b showed that the finding of a fine to coarse preference in the judgement of square-wave resemblance of a sequence of spatial frequencies cannot be explained on the basis of the last spatial frequency shown. The results must therefore reflect an interaction between the spatial frequencies presented. While it has been suggested that the interaction is one of integration, i.e. one where the spatial frequency information is integrated through time, it is possible that the interaction between the spatial frequencies is a masking one. Thus it may be the case that there is a differential masking of one of the spatial frequencies shown in the image sequence.

It has been found that the detectability of a spatial frequency is differentially affected by the presence of another spatial frequency (Carter & Henning, 1971; Stromeyer & Julesz, 1972; Nachmias & Weber, 1975; Legge & Foley, 1980; Ross & Speed, 1991; Foley & Yang, 1991; Watson & Solomon, 1997; see Breitmeyer, 1984 for an extensive review), such that the detectability is much more difficult or much easier (Interactions between different spatial frequencies have been discussed in section 1.2.1).

The masking effects of specific interest here are those found when a masking spatial frequency precedes or succeeds the presentation of the test spatial frequency. This has been termed forward and backward masking respectively (Rogowitz, 1983; Breitmeyer, 1984; Georgeson & Georgeson, 1987; Foley & Yang, 1991). For example, Georgeson & Georgeson (1987) examined the interactions between various spatial frequencies when separated at a number of stimulus onset asynchronies. They found evidence of facilitatory and masking effects between spatial frequencies. They found facilitation between spatial frequencies when they were separated only very briefly in time. This facilitatory effect disappeared at a stimulus onset asynchrony between the spatial frequencies of 50 msec, at this point masking effects were found. Of interest to the experiments here are the conditions involving two spatial frequencies whose frequency differs by a factor of 3. When such F and 3F spatial frequencies are employed they found facilitation for the detection of the 3F at simultaneous presentation times when the F masker contrast is high. When the F mask has a high contrast and is presented 50 msec prior to and 50 msec after the 3F, the contrast thresholds were raised. Interestingly the contrast thresholds were raised more when the mask succeeded the 3F than if it preceded it. Thus showing evidence for a greater backward than forward masking of the 3F by the F.

In terms of the results reported in this chapter the spatial frequency interactions of importance are those found from experiments examining the forward and backward masking effects of spatial frequencies. Following from these results it can be suggested that the preference shown for the fine to coarse sequence in judgements of square-wave resemblance could be a result of the F spatial frequency in the image sequence having a greater backward masking effect than forward masking effect on the 3F component. This would result in the 3F being more visible or rather having a higher perceived contrast in the coarse to fine sequence than in the fine to coarse sequence and therefore

the coarse to fine sequence would appear to resemble a square-wave less than the fine to coarse sequence. That this may be the case was suggested from the results of Georgeson & Georgeson (1987; see also Rogowitz, 1983) previously outlined, and by the spontaneous observer reports that the reason they chose one pair over the other in experiment 1 was due to the visibility of the many black and white bars of the 3F in one pairing then in the other. Thus the 3F component of the F / 3F pair seemed to be more visible than that in the 3F / F pair. This suggests that the F may be differentially masking the 3F. This masking explanation for the results of the square-wave likeness results of experiment 6 and 8 could also explain the results found in experiment 9a and 9b. Thus the 3F / F pair may be chosen to resemble a square-wave more than a 4F / F pair because the 3F is masked more than the 4F. It has been shown that as spatial frequencies become more widely separated the masking which occur between them decreases (Georgeson & Georgeson, 1987). Furthermore the 3F / F in peaks subtract phase may be chosen to resemble a square-wave more than a 3F / F in peaks add phase because of a phase sensitivity in masking interactions. Thus the 3F / F may be chosen as more resembling a square-wave simply because the higher spatial frequency is less visible.

In this experiment then the forward and backward masking effect of the F on the 3F was measured. If it was the case that the 3F is less visible when succeeded by the F and when preceded by it then the contrast threshold should be higher.

It could argued that the results from a contrast threshold experiment have no bearing on the experimental conditions where both spatial frequencies are shown at suprathreshold contrasts. However, it has been found that the apparent contrast of a suprathreshold spatial frequency is a function of its threshold (Kulikowski, 1976; Georgeson, 1991). Kulikowski (1976), for example, finds that the apparent contrast of a suprathreshold spatial frequency can be predicted by subtracting its threshold contrast. If it is assumed that this underlying contrast threshold mechanism determines the

suprathreshold appearance of the spatial frequencies employed in the square-wave likeness experiment then it follows that the 3F which is presented in the pair F / 3F has a higher apparent contrast than when it is shown in the 3F / F pairing. This is a simplification of the models suggested for the contrast matching data of Georgeson & Sullivan (1975) as most recent models include not just a threshold corrected element but also an attenuation, compression and normalisation factor to account for contrast constancy (Georgeson, 1991; Georgeson & Shackleton, 1994). However the outcome would be similar, as the thresholding ($C - C_0$) of the spatial frequency contrast included in these models would result in a differential response which would remain throughout the remaining aspects of the contrast models. Thus it may be on the basis of this that a fine to coarse temporal order of spatial frequencies results in more judgements of square-wave resemblance. Given the well documented relationship between the contrast thresholds and those at suprathreshold (Kulikowski, 1976; Georgeson, 1991). It follows that the perceived suprathreshold contrast of the 3F will be reduced when the F follows the 3F as opposed to those occasions on which the F precedes the 3F spatial frequency.

Thus it was expected, following the findings of masking interactions between widely different spatial frequencies, that the fine to coarse preference found for judgements of square-wave resemblance are not due to the integration of spatial frequencies but rather they reflect the existence of masking interactions between the spatial frequencies, such that the masking effect increases when the low spatial frequencies succeed the higher spatial frequencies rather than precede them.

4.5.1 *Method*

Subjects. There were six subjects, five male and one female, with an age range of 25 to 31. One of the subjects was the author while the other subjects were naive as to the purpose of the experiment.

Apparatus and Materials. The same stimuli and display apparatus were used as in experiment 9a. The contrast of the low spatial frequency, F , was held constant at 93.75% while that of the $3F$ was allowed to vary depending on the subjects previous responses.

Design. The experiment was a temporal 2 - alternative - forced - choice experiment using a 2 up, 1 down double staircase as described by Cornsweet (1962). Initially the contrast threshold for the $3F$ when presented alone was determined. In the second part of the experiment the contrast threshold for the $3F$ was determined when preceded and succeeded by an F in a peaks subtract phase relationship. In this second part both intervals contained a Gabor patch of spatial frequency F and the subjects task was to identify in which interval they detected the presence of a high spatial frequency, $3F$. As the contrast threshold for both the $F / 3F$ pair and the $3F / F$ pair were measured simultaneously there were two double staircases one starting at zero and the other at 29.3 % contrast. Initially the step changes in contrast were large at 1.83 % to allow a rapid convergence to a threshold value. After ten decisions this contrast change step was changed to 0.37%, this allowed for a fine tuning of the resulting threshold value. The experiment was a within subjects design. The independent variables in this experiment were the presentation order of the component pairs, while the dependent variable was the contrast of the $3F$ component.

Procedure. Subjects were fully informed as to the nature of the task, including examples of the stimuli set and the temporal form the experiments were to take. Once subjects were satisfied with the experimental procedure they then adapted to the luminance of the screen for two minutes before initiating the first experimental block. In

the first part of the experiment the contrast threshold for the 3F alone was determined.

In the second part the contrast threshold of the 3F was determined when it was preceded and succeeded by an F. All experiments were carried out in a blacked out room and data was collected in two sessions. The contrast threshold for the 3F alone was determined in one session, while the contrast threshold for the 3F when presented with the F was determined in the other.

Data. The contrast thresholds were identified by taking the last five reversal pairs of eight from each staircase and finding their mean and standard deviation. The mean and standard deviations of the two staircases were averaged.

4.5.3 Results

The results of subjects mean contrast thresholds (in percent) for a 3F when presented alone and when preceding and succeeding a F Gabor patch are shown in table 4.1, along with their standard deviations.

Subject	3F		F / 3F		3F / F	
	mean	std dev	mean	std dev	mean	std dev
1	3.59	0.35	9.77	1.95	11.88	0.80
2	3.88	0.37	9.22	2.40	10.41	2.86
3	3.78	0.43	7.66	0.68	8.77	0.61
4	4.26	0.46	6.66	0.66	13.73	0.51
5	3.32	0.88	10.68	2.77	14.06	2.21
6	3.14	0.47	6.74	0.55	7.23	0.73
mean	3.14	0.42	7.25	1.29	9.44	1.1

Table 4.1: Showing the mean contrast thresholds and their standard deviations by subject. Std dev = Standard deviation.

The results clearly show an increase in the contrast necessary in order to detect the presence of the 3F when either preceded or succeeded by the F. Furthermore, the contrast threshold of the 3F when succeeded by the F was found to be higher than when the 3F was preceded by the F, indicating a greater interaction between the two spatial frequencies when the lower spatial frequency was shown after the higher spatial frequency. It can be seen that there is a consistent difference between the contrast thresholds derived across subjects between the presentation orders and that found when the 3F was shown alone. This is shown graphically in figure 4.5. This difference between the contrast thresholds for the 3F detection in the two presentation orders of F / 3F and 3F / F was found to be significantly different ($t = -2.58$; d.f. = 5; $p < 0.05$).

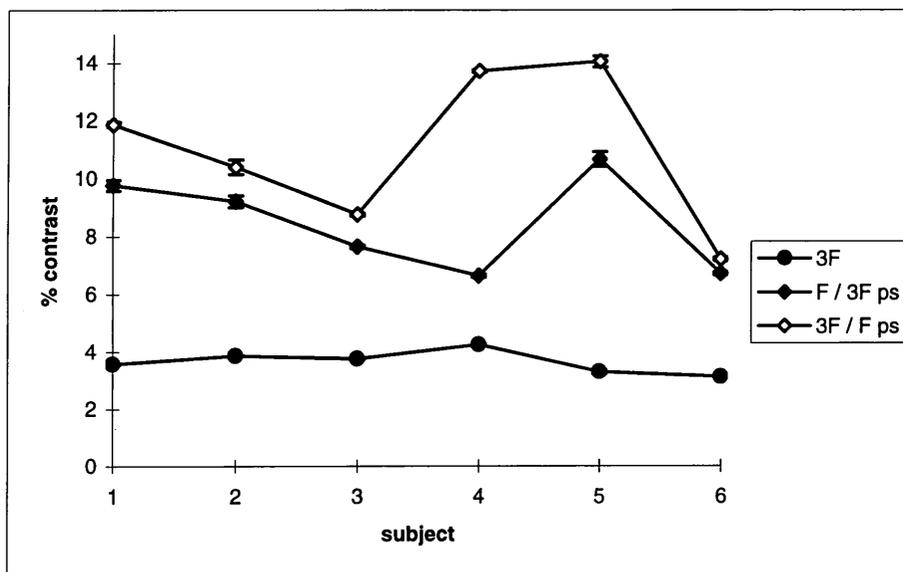


Figure 4.5: shows the contrast thresholds (in percent) for a 3F when presented alone, preceded by a F (F/3F) and succeeded by a F (3F/F). The vertical bars are the standard error bars.

4.5.4 Discussion

The results show that the contrast threshold for detecting the presence of a 3F is significantly greater when the F succeeds the 3F temporally than when it precedes its presentation. Thus the backward masking effect of an F on the contrast threshold of the 3F is greater than its forward masking effect.

This differential masking effect can be suggested to be a function of single channel activation or the inhibitory interactions found operating between channels responsive to different spatial frequencies (see section 1.2.1) and the neural delays involved in the processing of spatial frequencies (see section 1.3.2). It may be the case then that the masking effects reflect the operation of inhibition of the 3F by the F. The strength of this inhibition is varied by the degree to which the spatial frequencies overlap in time (see section 3.3.3). It was pointed out in section 3.3.3 that a 3F / F presentation order although physically presented as separate components they overlap neurally due to their processing delays, while a F / 3F presentation of spatial frequencies introduced a gap between the neural representations of the two spatial frequencies. Thus the increase in the contrast threshold of the 3F when succeeded by a F can be suggested to be due to the temporal overlap of the F and the 3F components, thus increasing the inhibition of the 3F channel by the F channel.

This finding that there is significant differential masking effect could account for the predominance of choices of the 3F / F pair in the first experiment. Thus when presented in peaks subtract phase (as in experiment 6, 8, 9a and 9b) the 3F is masked more by a succeeding F than a preceding one. If it assumed that the suprathreshold contrast of spatial frequencies are intimately related to their thresholds the decrease found in the contrast threshold of the 3F when succeeded by the F could explain the square-wave likeness result.

4.6 General Discussion

The experiments presented in this chapter were all directed to interpreting the finding of the square-wave likeness experiment 6 presented in chapter 3. The results from experiment 6 show that a fine to coarse presentation of spatial frequencies was judged to be more square-wave like than a coarse to fine sequence. This was an unexpected result given the findings of Parker et al (1992; 1997) and the supporting evidence outlined in sections 1.3.2 and 1.3.3 suggesting a coarse to fine integration process. Indeed the results of experiment 5 show that a square-wave is more likely to be mistakenly detected in a coarse to fine sequence of spatial frequencies suggesting that this sequence of spatial frequencies is more effectively integrated to form a square-wave like representation. It was suggested that the results may reflect the operation of a flexible integration process where factors such as stimuli, task and attention may all play a role (Schyns & Oliva, 1997; Oliva & Schyns, 1997), however a simpler explanation is that the square-wave likeness result is due to aspects of the stimuli presented other than integration.

Experiment 7 examined the underlying nature of the likeness task and its relationship with the natural image quality experiment (experiment 1). Both were designed to examine the quality of the phenomenal appearance of the image triplets whether they were filtered versions of natural images or spatial frequency Gabor patches. It was argued that, unlike the natural image quality paradigm (experiment 1; Parker et al, 1992), the square-wave likeness experiment (experiment 6) may not be assessing the integration of spatial frequencies, but rather some other qualitative aspect of the stimulus. It was found that when employing the same task with natural images rather than Gabor patches, that is a natural image likeness experiment, the coarse to fine

sequence of filtered images was judged to have more resemblance to the fullbandwidth natural image. This matched the result found in the natural image quality experiment suggesting that the two quality judgement tasks are assessing the same aspects of quality. Thus it cannot be on the basis of task differences that the results between experiment 1 and experiment 6 differ.

Experiment 8 showed that it was possible to assess the interactions between the spatial frequencies presented. This experiment showed that the same fine to coarse preference in square-wave likeness was found with two spatial frequencies (F and $3F$), making it practicable to examine the interactions which occurred between them.

Experiment 9a and 9b examined the role of the last spatial frequency shown in dictating the square-wave likeness judgement. Experiment 9a and 9b showed that the spatial frequency information presented in the penultimate frame did play a role in the determination of square-wave likeness suggesting that the judgement was not based on the final frame alone and must reflect some form of interaction between the spatial frequencies presented.

Experiment 10 examined the nature of these interactions. It was argued that the results of the square-wave likeness experiment (experiment 6) may be due to the masking interactions between the spatial frequencies presented and not the integrative interactions between them. It was found that the increase in the contrast threshold of a $3F$ when succeeded by a F could account for the fine to coarse sequence of spatial frequencies being judged more square-wave like as, following suggested models of perceived contrast (Georgeson, 1991), the $3F$ would appear to be of lower contrast than the same spatial frequency presented in a coarse to fine sequence. Thus the coarse to fine sequence would appear to be less square-wave like simply because the $3F$ appeared to be of higher contrast.

Thus the results from Experiment 10 suggest that the interaction which occurs between the spatial frequencies presented is not one of integration but one of masking. Thus it appears that the square-wave likeness experiment does not reflect the operation of a flexible integration process. However it difficult to see how this masking explanation connects with the square-wave detection experiment (experiment 5), but if the nature of the square-wave likeness task is considered then this difficulty disappears. The square-wave experiment is a two alternative forced choice one and the judgement, as suggested by the masking experiment, is a relative one, that is one sequence is less like a square-wave than the other, even if both resemble a square-wave very little the most square-wave like will be chosen. The results from experiment 10 suggest that the stimuli sequence in which the 3F had a higher perceived contrast, or more a clear definition, was chosen as the stimulus which looked least like a square-wave. It can be envisaged that the fine to coarse sequence of spatial frequencies was judged to be more square-wave like, not because it actually did appear square-wave like, but rather the fine to coarse sequence appear to be more square-wave like than the coarse to fine sequence in which the 3F appeared with a higher perceived contrast. This is markedly different from the square-wave detection experiment where this forced relative choice is not encountered.

It is illuminating to examine the predictions of integration and masking in the square-wave likeness experiment. Rather than reflecting the forward and backward masking of spatial frequencies it could be argued that the finding of increased contrast threshold found when a low spatial frequency follows a high spatial frequency sequentially may be a result of the improvement in integration which occurs when spatial frequencies are presented in this fashion. Thus it may be the case that rather than reflecting the operation of inhibitory interactions between spatial frequency channels it may be the case that there are excitatory interaction between the channels such that their

outputs are combined in some fashion. The increased contrast threshold for the 3F could be a result of the detection being based on the output of this combined signal rather than on the output of a 3F channel. There is a great deal of evidence that decision making processes do not have access to the outputs of individual spatial frequency channels (see section 2.2.2). For example, following a model of early spatial vision suggested by Georgeson & Meese (1997), it may be the case that the detectability of the 3F is limited by the spatial structure of the stimuli rather than the individual spatial frequencies, such that the assignment of zero crossings to the combined signal (e.g., 3F / F) do not significantly differ from those assigned when the F alone is presented until a higher contrast of the 3F is reached.

Thus the reason a 3F has an increased contrast threshold when succeeded by a F is because they are more effectively integrated and thus the 3F needs more contrast in order to be visible over the F. It may be the case then that the masking result of experiment 10 reflects the integration of spatial frequencies. Both a masking explanation or a integration explanation of the results would predict the same pattern of performance as found, that is both would predict that a 3F would have a higher contrast threshold when succeeded by a F than when succeeded by it. Therefore the square-wave likeness finding showing an increase in square-wave like judgements for spatial frequencies when presented from fine to coarse may be due to integration not masking. (This distinction between masking and integration is similar to that between masking by integration and masking by interference found by Turvey (1973).)

In summary, the square-wave likeness result, showing that a fine to coarse sequence of spatial frequencies is found to be more square-wave like than the coarse to fine sequence, seems to be due to masking interactions which occur between the spatial frequencies presented in the sequence. However, it has been pointed out that the results of the masking experiment (experiment 10) could also be accounted for if integration

was taking place between the spatial frequencies. In the next chapter these two possibilities are examined further.

Spatial frequency interactions and Vernier acuity

Summary

In this chapter two experiments were carried out which were designed to address whether the evidence indicative of a fine to coarse integration of spatial frequencies (see chapter 4) was due to the integration or the masking of spatial frequencies. In both of these experiments the Vernier acuity threshold for two grating halves was determined for a variety of different grating types notably a fundamental spatial frequency (F), its third harmonic ($3F$), a F followed by $3F$ ($F / 3F$) and finally $3F$ followed by F ($3F / F$). It was predicted that if the F and $3F$ integrate only when presented from fine to coarse then the reference edge from which the Vernier judgement was made would be sharper and thus would lower the Vernier threshold beyond that determined for either spatial frequency when presented alone and when they were presented in a coarse to fine temporal order. If, however, the F masks the $3F$ more when presented from fine to coarse then the Vernier acuity should never be lower than when the F is shown alone. In the first experiment (experiment 11a) the Vernier acuity thresholds for all four conditions did not differ and therefore the masking hypothesis was supported. However, it was pointed out that because the Vernier acuity thresholds for the F and the $3F$ did not differ when each was presented alone then the integration of their positional information would not lower the Vernier acuity threshold for the $3F / F$ pair as each component

provided the same level of information (assuming that the noise associated with both signals is uncorrelated). To examine this the Vernier acuity threshold for the compound of the F and 3F was determined in experiment 11b. It was found that the Vernier acuity threshold did not differ from that determined from the F, 3F, F / 3F or the 3F / F. This suggested that Vernier sensitivity was limited by the positional information provided by each spatial frequency and that it was this which was integrated.

In order to assess whether it was the positional information provided by the spatial frequencies which was integrated the Vernier acuity thresholds of higher spatial frequencies were determined in the same four conditions as experiment 11: F, 3F, F / 3F and 3F / F (Experiment 12). The Vernier acuity threshold determined when the spatial frequencies were presented from fine to coarse were found to be significantly lower than that determined from either spatial frequency when presented alone or when presented from coarse to fine. It was proposed that the improvement in the Vernier acuity threshold found in the fine to coarse condition was due to the fine to coarse integration of the different positional information provided by the two spatial frequencies.

5.1 Introduction

In the last chapter a number of possible explanations for the anomalous results of experiment 5 and experiment 6 were examined. In particular the finding from experiment 6 showing an improved integration when spatial frequencies were presented from fine to coarse was investigated. The experiments presented in chapter 4 (experiments 7, 8, 9a and 9b) showed that the evidence for a fine to coarse integration of spatial frequencies was not attributable to the nature of the task or the result of the decision being based on the last frame alone. The masking interactions between the

component spatial frequencies which comprise the temporal sequences were also examined (Experiment 10). It was found that the 3F in the sequence 3F / F (fine to coarse) had a lower apparent contrast than the 3F shown in the F / 3F (coarse to fine) sequence. Thus it may be the case that the 3F / F sequence was chosen to resemble a square-wave more often because the 3F had a lower apparent contrast when compared with the 3F shown in the F / 3F pair.

However, while this apparent contrast argument is valid it is not clear what caused this reduction in apparent contrast. It seems likely that the interaction between the spatial frequencies is a masking one where the fundamental frequency backward masks the third harmonic more than forward masks it. This has been found in similar experiments carried out by Georgeson & Georgeson (1987). On the other hand it may be a result of the integration of the spatial frequencies. If it is assumed that a fine to coarse sequence of spatial frequencies is judged to resemble a square-wave because the spatial frequencies integrate to form a square-wave like representation, then it follows that the contrast threshold of a 3F would be higher in the fine to coarse sequence. This is because the 3F and the F would integrate more effectively when presented 3F / F than F / 3F and the 3F would need more contrast for it to be visible above the 3F + F compound it is more likely to form. Thus both of these explanations for the increase in contrast thresholds found for the 3F when succeeded by an F, masking and integration, would predict the same pattern of results.

In order to examine this further it is necessary to design an experiment where the two types of spatial frequency interaction would predict different results. Vernier acuity of spatial frequencies was chosen as it provides just such a task.

5.1.2 Vernier acuity

Vernier acuity is one of many sensitivities to where human observers perform better than seems possible given the spatial resolution limit of human vision. It has been found that grating resolution is limited to about 30'' under optimal conditions, this corresponds to a spatial frequency of 60 cpd (Campbell & Gubisch, 1966; Campbell & Robson, 1968; Parker & Hawken, 1985). However, much finer resolution limits than this have been found with a number of tasks. These tasks have been called hyperacuities by Westheimer (1975) and include curvature discrimination (Watt & Andrews, 1982; Watt, 1984a), orientation acuity (Westheimer, 1981) and Vernier acuity amongst others (Watt, 1984b; Morgan, 1991). A Vernier acuity task generally refers to the ability to discriminate the offset of two abutting lines. For example, two lines are presented one above the other and one of these is systematically horizontally displaced from perfect alignment until the displacement can no longer be discriminated. Resolution in this task has been found to be as little as 2'' - 5'' (Westheimer, 1979). There are a number of theories which have been suggested to explain these hyperacuities, these include channel models such as the spatial frequency model of Wilson (1986; Burbeck, 1987) and Local sign models (Hering, 1920; Westheimer & McKee, 1977). The channel model involves the implicit coding of hyperacuities via narrowband channels responsive to different spatial frequencies. The acuity is limited to the largest channel which responds to both objects shown. For example, in a temporal two-alternative forced choice experiment a Vernier offset could be signalled by the difference in the activation of a single channel across presentation interval. The local sign model on the other hand explicitly signals positional information of the stimuli by coding their position with local signs or place tags (Westheimer, 1981), subsequent comparisons are then able to make judgements about their relative positions. There are also hybrid models of the two, such as an orientation detector model (Findlay, 1973) or more recently the Collator model (Morgan

& Hotopf, 1989; Mussap & Levi, 1996). These suggest that the differential responses from small filters are integrated across space to provide positional information. This combines the filter responses from the channel model with local averaging from local sign models.

In a variation on the standard Vernier acuity task spatial frequency gratings have been employed as stimuli (Bradley & Skottun, 1987). For example a grating is presented in which the bottom half is shifted to the right or to the left of the top half. When this is carried out subjects can resolve a horizontal displacement of approximately 5'' - 10'' of arc at spatial frequencies of between 6 and 16 cpd (Whitaker & MacVeigh, 1991; Bradley & Skottun, 1987; Morgan, 1984 reported in Wilson, 1986). The Vernier acuity threshold has been found to vary as a function of spatial frequency (Bradley & Freeman, 1985), contrast (Bradley & Skottun, 1987) and separation between the two halves of the spatial frequency grating (Whitaker & MacVeigh, 1991; Whitaker, 1993). Vernier acuity thresholds are found to improve with increasing spatial frequency to an optimum between 4 and 16 cpd and then decrease at spatial frequencies higher than this (Bradley & Skottun, 1987). However, when contrast thresholds for the spatial frequencies are taken into account the Vernier acuity thresholds vary with spatial frequency such that there appears to be roughly three regions: at low spatial frequencies the Vernier acuity thresholds improve constantly when considered in terms of the phase shift of the gratings; a second region at intermediate frequencies where decreasing Vernier thresholds are offset by increasing contrast thresholds thus serving to plateau the acuity threshold; and finally, at higher spatial frequencies, an increase in acuity thresholds as contrast threshold is approached. Vernier acuity thresholds are also found to vary with the separation or gap between the grating halves such that as the separation between the two grating halves increase the Vernier acuity thresholds also increase, that is they become worse. Whitaker & MacVeigh (1991) found that this relationship varied in

strength as a function of spatial frequency: high spatial frequencies show a sharper increase in acuity threshold as separation increased than was found in the low spatial frequencies at the same separations.

Such studies employing spatial frequency gratings have been interpreted in terms of the spatial frequency channel model of hyperacuties and unless additional assumptions were made the local sign model cannot account for them. However there are several findings which cannot be explained by the spatial frequency channel model (Morgan & Ward, 1985; Morgan, Ward & Hole, 1990; Burbeck & Yap, 1990a). Thus it may be the case that a hybrid model, such as the Collator model, could best account for all of the hyperacuity results. However a discussion of the merits of each theory is outside the scope of this thesis as the models underlying Vernier acuity with spatial frequency gratings are not investigated. Rather the methodologies which employ specific spatial frequency gratings and the results which have been found are of primary interest as it is the integration of these spatial frequencies which is to be examined.

5.1.3 Preview of the chapter

In the experiments to be reported the assessment of a Vernier acuity threshold was carried out with specific spatial frequency gratings of F and $3F$. Using these stimuli it is possible to examine the nature of the interaction which occurs between the two spatial frequencies. It was suggested that the interaction may be a masking one where the F spatial frequency interferes with or inhibits the $3F$ channel, or an integration interaction where the two spatial frequencies sum to form a compound. The predictions from these two hypotheses are very different when applied to a Vernier acuity task employing spatial frequency gratings. If the interaction between the spatial frequencies is a masking one then the performance when the pair are presented should never exceed

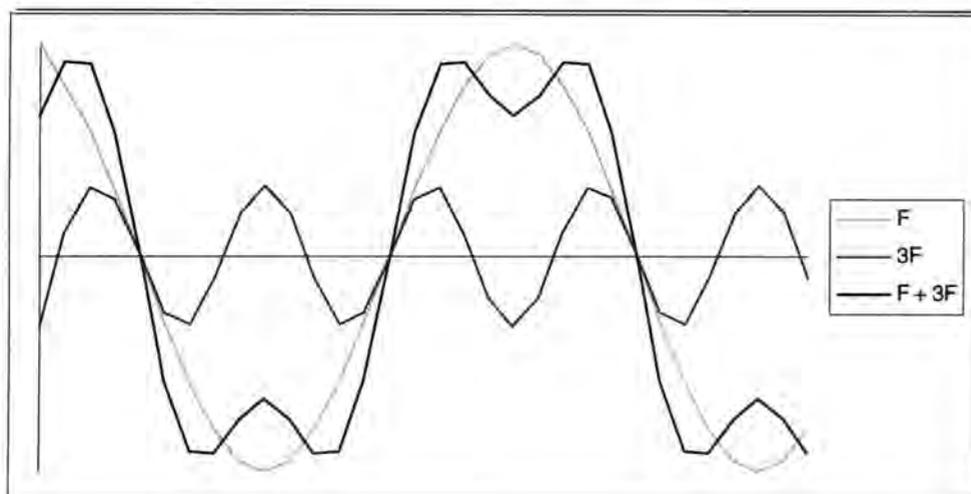


Figure 5.1: Depicts the increasing sharpness of the boundary between progressively darker and progressively lighter portions of a sinusoid. It can be seen that the $F + 3F$ compound has a sharper boundary than that found for the F or the $3F$ alone. This suggests that if the two individual spatial frequencies integrated then the Vernier acuity threshold for a horizontal displacement of two gratings would be lower than for either spatial frequency alone.

that determined from the F when presented alone. Thus if the F spatial frequency gives the lowest Vernier acuity threshold of x then the threshold found when the F and the $3F$ when presented sequentially should never exceed this. If the spatial frequencies of $3F$ and F are integrated to form a square-wave like representation then the edges between the bars of the sinusoids would be sharper in the complex $3F + F$ wave (See figure 5.1). Thus the Vernier thresholds for the offset of a grating would be lower than that found when the F or the $3F$ are presented alone.

5.2 Experiment 11

5.2.1 Experiment 11a: Vernier acuity of Low spatial frequencies

Vernier acuity thresholds were measured for Gabor patches consisting of a single spatial frequency and when each was presented as a $F + 3F$ sequence running either

from coarse to fine or fine to coarse. It would be expected if the spatial frequencies were integrated that the fine to coarse sequence would improve thresholds compared with those determined when either the F or the 3F were presented alone. If however the interaction between spatial frequencies was a masking one then it would be expected that the threshold found for the paired spatial frequencies, shown progressively from coarse to fine or fine to coarse, would not improve beyond that determined from the F when presented alone.

5.2.1.1 *Method*

Subjects. Six subjects, three female and three male, with an age range of 21 to 40 took part in this experiment.

Apparatus & Materials. The stimuli used in these experiments were vertically oriented spatial frequency Gabor patches of 2 and 6 cpd and a spatial envelope of standard deviation 0.47 degrees (see figure 5.2). It should be noted that whilst the overall spatial frequency patch was a Gabor the top and bottom half were treated as separate entities because of the horizontal displacements introduced between them. In effect the stimuli were two vertically oriented gratings. The contrast relationship between the spatial frequencies was held in a square-wave relationship of 3:1; the lower spatial frequency was shown at a contrast of 93.75% and the higher spatial frequency was shown at a contrast of 31.25%. Purpose written software was used to create the stimuli employing a VSG display system (Cambridge Research Systems Ltd). In a pilot study with abutting gratings subjects performance at judging the direction of the bottom half of the Gabor patch was at ceiling. In an attempt to increase subjects thresholds to a measurable level a central gap of 3.29' was introduced into the stimulus thus creating a clear top and the bottom half. This has been shown to increase the thresholds for spatial frequency

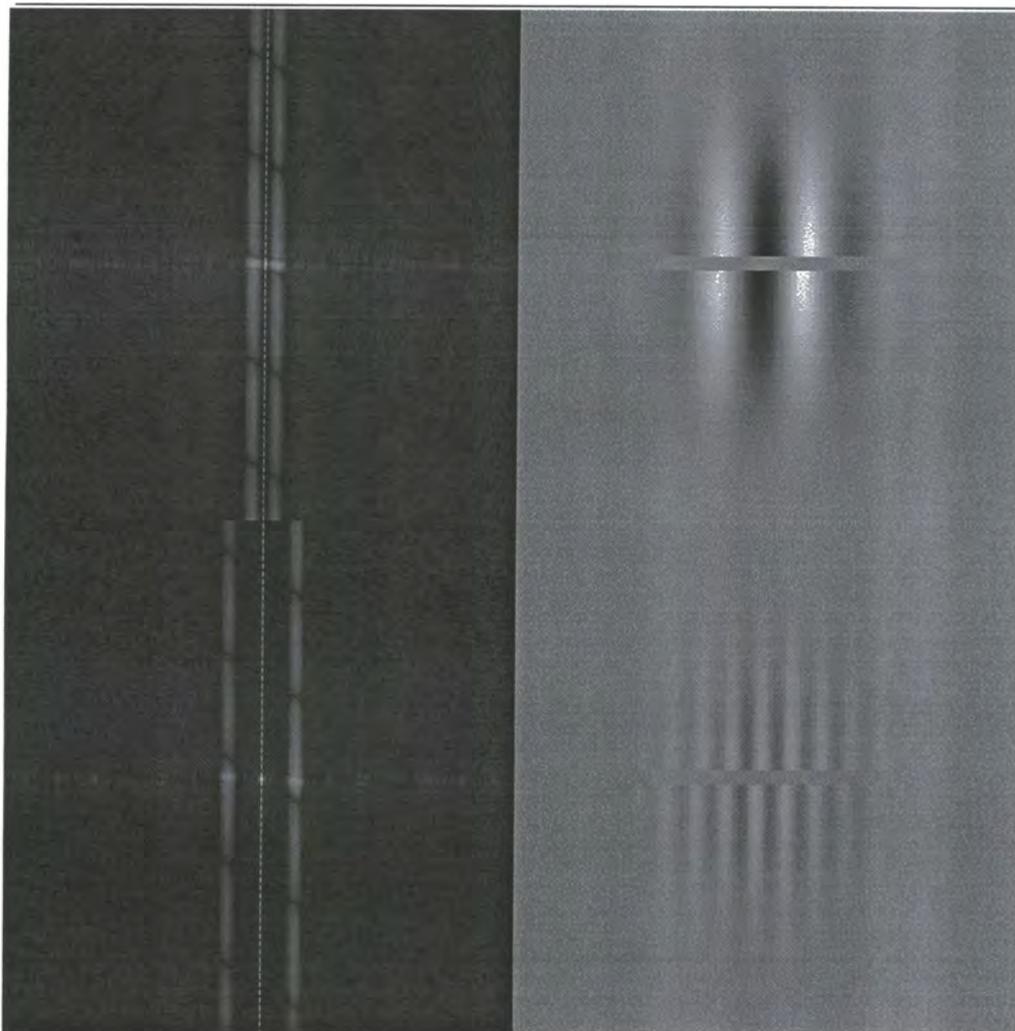


Figure 5.2: Vernier stimuli and their power spectra. The left hand column shows a detail of the power spectra of the Vernier stimuli shown in the right hand column. The images shown at the top are of a 2 cpd spatial frequency, while below this is shown the 6 cpd spatial frequency. Only a section of these power spectra is shown as the details were obscured in the full power spectra images due to their large scale. It can be seen that the predominant energy in the Vernier stimuli are at their respective spatial frequencies, however energy at all spatial frequencies is present in the images because of the sharp edges introduced by the gap between the two grating halves.

gratings in Vernier acuity experiments (Whitaker & MacVeigh, 1991). Thus this degree of separation was introduced to increase the Vernier acuity thresholds to a medium range so that improvements could be measured. In this experiment the top half of the Gabor patch was stationary while the bottom half varied to the left or to the right by 119,

59, 30, and 10 seconds. Prior to the onset of the grating stimuli a small white fixation box was presented centrally for 350msec. This was included to keep the subjects eyes steady during the trial and thus enable more reliable thresholds to be determined. The Vernier stimulus followed this fixation box by a further 350 msec. All stimuli were presented on a 21 inch(1024 x 768 pixels) monochrome Eizo flexi-scan 6600-M monitor with a refresh rate of 69 Hz and a mean luminance of 12.9 cd/m². Stimuli were viewed binocularly at 7430mm for 42 msec each and were presented with a square-wave temporal waveform. The subjects head was placed in a chin rest throughout the duration of the experiments.

Design. The design of the experiments was a within subjects one using the method of constant stimuli. There were three conditions: Vernier acuity thresholds for the low spatial frequency alone (condition 1); Vernier acuity thresholds for the high spatial frequency alone (condition 2); and the Vernier acuity thresholds for the both patches presented from low to high or high to low (condition 3). Subjects were assigned a course of conditions by a Latin square so as to combat possible order effects. These were: a, b, c; b, c, a; and c, a, b (where a = condition 1; b = condition 2 and c = condition 3). Because of the arduous nature of this experiment the conditions were carried out separately and each condition was split into two blocks, with 400 decisions per block. There were equal number of decisions for each of the 8 horizontal displacements conditions, 50 per block, thus 100 overall. Condition 3 determined the Vernier acuity of both the low to high and high to low presentation orders of spatial frequencies thus there were 16 conditions, because of this there were 25 decisions per block and 50 decisions overall.

Procedure. Subjects were informed what the task required including being shown examples of the stimuli set both in terms of their spatial frequency and horizontal displacements. The subject proceeded with the experiment when fully satisfied. All subject data was collected in three sessions, each of which was separated into two 15 minute blocks. The experiment was carried in a blacked out room.

5.2.1.2 Results

Table 5.1 shows the Vernier acuity thresholds found as a function of the spatial frequencies shown. The values are in seconds. It can be seen that overall the 3F spatial frequency when presented alone produces higher thresholds than any other condition for five out of six subjects.

Subject	F	3F	F / 3F	3F / F
1	38.3	43.0	26.0	25.9
2	71.1	232.9	65.5	66.8
3	60.4	51.2	64.2	54.6
4	26.1	49.4	23.4	29.7
5	41.9	55.1	36.2	39.8
6	33.6	48.1	47.3	46.9
mean	38.77	68.96	37.51	38.1

Table 5.1: shows the Vernier acuity thresholds for each of the 6 subjects. The values are in seconds.

However, a one way ANOVA examining image condition: F alone; 3F alone; F / 3F; and 3F / F failed to reach significance ($F(3, 15) = 1.81$; $M.S.e = 1055.24$; $p > 0.05$).

5.2.1.3 Discussion

It was suggested that the results of experiment 10 (differential masking) reflected an integration of spatial frequencies which operates more effectively if spatial frequencies are presented from fine to coarse. Thus it was predicted that the Vernier acuity threshold for the alignment of two gratings would improve if the spatial frequencies were presented from fine to coarse relative to a coarse to fine sequence. It was also predicted that the Vernier acuity threshold for the 3F / F pair would also improve relative to the lowest Vernier acuity threshold found for a single spatial frequency of the sequence. The results clearly show that the Vernier acuity threshold elicited by a fine to coarse pair of spatial frequencies is not significantly lower than that found when spatial frequencies are presented from coarse to fine and these thresholds are also not lower than that found when either of the spatial frequency components are shown alone. This does not support a integration hypothesis but it does support a masking hypothesis as the Vernier acuity thresholds for the fine to coarse and coarse to fine temporal orders of the F and the 3F do not show any deviation from the Vernier acuity threshold found for the F alone.

However, it may be the case that the lack of evidence for the integration of spatial frequencies reflects a limit on the sensitivity to the horizontal displacement of the two halves of the gratings given the positional information provided by the two spatial frequencies. Thus although the compound of the F and the 3F does provide a sharper reference edge it is not possible to use this positional information as this is limited by the positional information derived from the individual spatial frequencies.

In order to make this clearer it is useful to consider the positional information which can be derived from the Vernier stimulus in order to complete the task. It can be suggested that the Vernier task as employed in this experiment involves two forms of positional uncertainty: the first is locational uncertainty which concerns the absolute localisation of either grating in the space of the display; the second is relational uncertainty which concerns the localisation of the two gratings with respect to each other. It is obvious that the higher the spatial frequency of a grating the lower its locational uncertainty as the spatial frequency channels responsive to such a grating have a small receptive field and so are more localised in space than a low spatial frequency channel. As pointed out by Waugh & Levi (1995) this also follows from a consideration of statistical theory where the accuracy for localising the mean of a Gaussian distribution is limited by its standard deviation. Thus the larger the receptive field the greater the standard deviation of the distribution of activation and therefore the more difficult it is to localise the target. However, because of the gap introduced between the two halves of the gratings, channels responsive to higher spatial frequencies will give poorer relational information as their receptive fields cease to span the gap between the two halves of the gratings. This will result in a matching ambiguity between the two grating halves. Thus channels responsive to low spatial frequencies will give excellent relational information but poor localisation of the absolute position of the grating bars themselves, while the higher spatial frequencies will give better localisation of the grating bars but poor relational information.

These two positional uncertainties may account for the results found here. Both spatial frequencies utilised in this experiment (2 and 6 cpd) are relatively low and therefore provide poor locational information but good relational information. It may be the case then that when the spatial frequencies are presented as a compound the integration of the two spatial frequencies involves the maximisation of positional

information from both spatial frequencies and as both spatial frequencies provide good relational information and poor locational information no improvement in Vernier sensitivity is found assuming the noise associated with both signals is uncorrelated. This pattern of maximising the positional information from both would also extend to the conditions where spatial frequencies are presented from coarse to fine and fine to coarse and predicts that the Vernier acuity threshold elicited from a compound of the F and 3F spatial frequencies would not differ from that elicited from either spatial frequency alone. This forms the basis of the next experiment.

5.2.2 Experiment 11b: Vernier acuity of Compounds

The results of experiment 11a show that the Vernier acuity thresholds for the F, 3F, F / 3F and 3F / F did not significantly differ. This suggests that the F and 3F spatial frequencies are not integrating to form a sharper reference edge from which the Vernier judgements are made. It was pointed out that the Vernier acuity thresholds of the F and 3F when presented alone do not differ and therefore if the positional information which they provide was being integrated a lower Vernier threshold would not be expected as it would be limited to the positional information carried by the component spatial frequencies. In order to test whether this is the case the Vernier acuity thresholds for the compound of the F and 3F was determined. If the F and 3F did integrate to form a sharper reference edge then it would be expected that the Vernier acuity threshold would be lower than that for the F or the 3F when they are presented alone. However, if the positional information provided by the F and 3F was integrated then it would be expected that the Vernier acuity threshold for the compound would not differ from that determined for either component as their Vernier acuity thresholds were not significantly different from each other.

5.2.2.1 Methods

All methods are as experiment 11a. The stimuli was a compound of the F and 3F utilised in experiment 11a and are shown in figure 5.3. Only the Vernier acuity thresholds for the compound were determined.

Subjects. Six subjects, three female and three male, with an age range of 21 to 40 took part in this experiment.

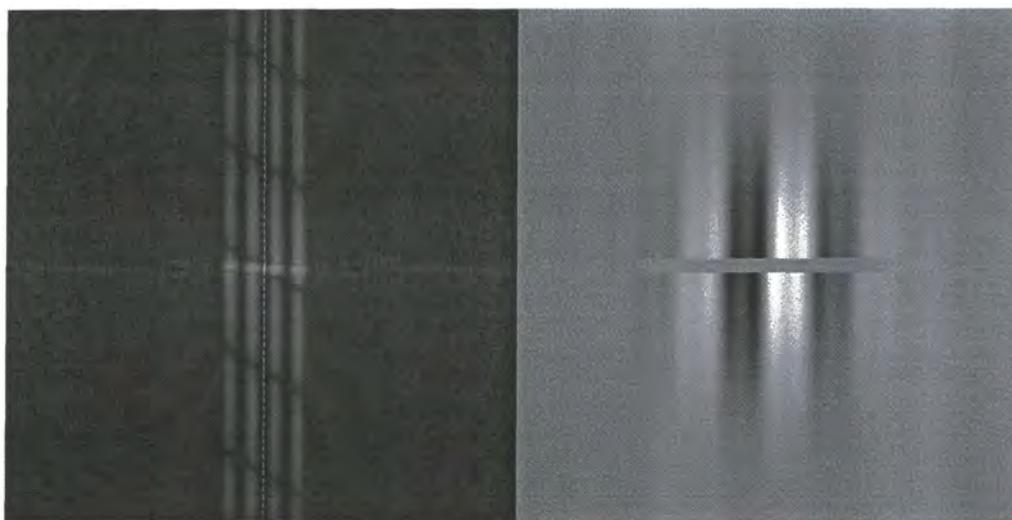


Figure 5.3: The compound Vernier stimulus and its power spectra. The left hand column shows a detail of the power spectra of the compound Vernier stimulus shown in the right hand column. Only a section of these power spectra is shown as the details were obscured in the full power spectra images due to their large scale. It can be seen that the predominant energy in the Vernier stimuli are at their respective spatial frequencies, however energy at all spatial frequencies is present in the images because of the sharp edges introduced by the gap between the two grating halves.

5.2.2.2 Results

Table 5.2 shows the Vernier acuity thresholds found as a function of the spatial frequencies shown. The values are in seconds. As different subjects took part in this experiment as participated in experiment 11a an independent subjects t -test was carried out with the averaged results from the four conditions (F, 3F, F / 3F, 3F / F) of experiment 11a. There was found to be no significant difference ($t = 0.30$; d.f. = 10; $p > 0.05$).

Subject	F
1	27.4
2	101.0
3	58.9
4	31.2
5	34.3
6	36.9
mean	41.39

Table 5.2: shows the Vernier acuity thresholds for each of the 6 subjects. The values are in seconds.

5.2.2.3 Discussion

The results clearly show that the Vernier acuity thresholds for the compound of F and 3F did not differ from the Vernier acuity thresholds determined in experiment

11a. This matches the predicted results if the Vernier acuity threshold is a function of the integration of the positional information provided by the spatial frequencies. Thus it suggests that the F and 3F are integrated but because the positional information they provide individually is equivalent no improvement is possible.

However, it may be the case that the Vernier acuity thresholds are determined on the basis of only one of the spatial frequencies presented. If, for example, the Vernier acuity threshold is based upon the F only then the results found in experiment 11a and those presented here would also be predicted. Evidence that Vernier acuity thresholds to compounds are based on only one of the spatial frequencies presented has been found by Barrett & Whitaker (1998). This suggests that even if the F and 3F spatial frequencies did integrate to form a sharper reference edge from which positional judgements could be made the judgement is limited by the positional information provided by the individual spatial frequencies. The results then do not provide evidence against the notion that spatial frequencies are integrated but rather the evidence suggests that the improved positional information given by this edge is not being used in making the Vernier judgement.

Overall then, the results from experiment 11a and 11b show that neither the Vernier acuity threshold for the coarse to fine and fine to coarse presentation of the F and 3F spatial frequencies significantly differ from that threshold determined when either spatial frequency was presented alone. This would be consistent with the proposal that evidence for a fine to coarse integration of spatial frequencies is actually a result of the masking interactions which occur between the F and the 3F. However, it was also proposed that the results could be understood if the positional sensitivity to the two spatial frequencies was limited by the positional information provided by each alone. The lack of a significant difference between the Vernier acuity thresholds elicited by the individual spatial frequencies and the compound of them indicates that this is indeed the

case. However, it was pointed out that the results from both this experiment and those of experiment 11a could be explained if the Vernier judgement was based only on one of the component spatial frequencies.

The concept of two different forms of positional information, locational and relational, being involved in the Vernier task, as employed here, suggests that it is possible to maximise both types of information and therefore lower the Vernier acuity thresholds of either spatial frequency component when presented alone. In order to maximise both types of positional information it is necessary to use a relatively low spatial frequency to maximise relational information and a high spatial frequency to maximise locational information. If the positional information from the two spatial frequencies is integrated then the Vernier acuity threshold determined from this integrated signal would be lower than that determined from either spatial frequency when presented alone. This would not be predicted if the Vernier judgement was based purely on one of the component spatial frequencies. This forms the basis for the next experiment to be presented.

5.3 Experiment 12: Vernier acuity with higher spatial frequencies

In the previous experiment it was found that both a coarse to fine and a fine to coarse sequence of spatial frequencies resulted in a Vernier acuity thresholds which were not significantly different from each other and both could be accounted for if the offset of the gratings was judged on the basis of either spatial frequency component alone. There was no significant worsening or improvement of the Vernier thresholds regardless of the order in which the spatial frequencies were shown. It was also found that when the Vernier threshold for the compound of the F and 3F shown

simultaneously there was no improvement beyond that shown by the Vernier acuity thresholds of the F alone. It was proposed that the lack of improvement in the Vernier acuity thresholds elicited by the compound and the two temporal orders of coarse to fine and fine to coarse was a result of the similar positional information provided by both spatial frequencies. The relatively low spatial frequencies provided excellent relational information but relatively poor locational information about the absolute spatial positions of the gratings. Therefore, assuming the noise associated with both signals was uncorrelated, the integration of the two spatial frequencies did not provide any further positional information which could be employed to improve Vernier acuity thresholds. However, it was pointed out that the results could also be a function of the Vernier judgement being based on only one of the component spatial frequencies.

This concept of two types of positional information limiting Vernier sensitivity predicts that acuity thresholds would improve if the integration of two spatial frequencies did provide more positional information. One way to maximise positional information is use relatively low and high spatial frequencies, where the low spatial frequency would provide good relational information and the high spatial frequency would provide good localisation information. As the results of experiment 11a show that the Vernier acuity thresholds found for the 2 and 6 cpd spatial frequencies were not significantly different from each other, the low spatial frequency utilised in this experiment was from this same area and set at 5 cpd (F). The relationship between this spatial frequency and the higher spatial frequency was the same as in the previous experiment, that is the high spatial frequency is set at 15 cpd (3F) with one third the contrast of the F. If the positional information from the F and 3F are integrated then the Vernier acuity thresholds for the F and 3F when presented from fine to coarse should be lower than if either spatial frequency was presented alone or if the F and 3F were presented from coarse to fine. If, however, the Vernier acuity thresholds are subject to

masking interactions between the two spatial frequencies or were based upon one spatial frequency only then no improvement would be possible. The experiment to be presented here assessed the Vernier acuity thresholds for these stimuli using exactly the same methodology as experiment 11a. It was predicted that the Vernier acuity thresholds for

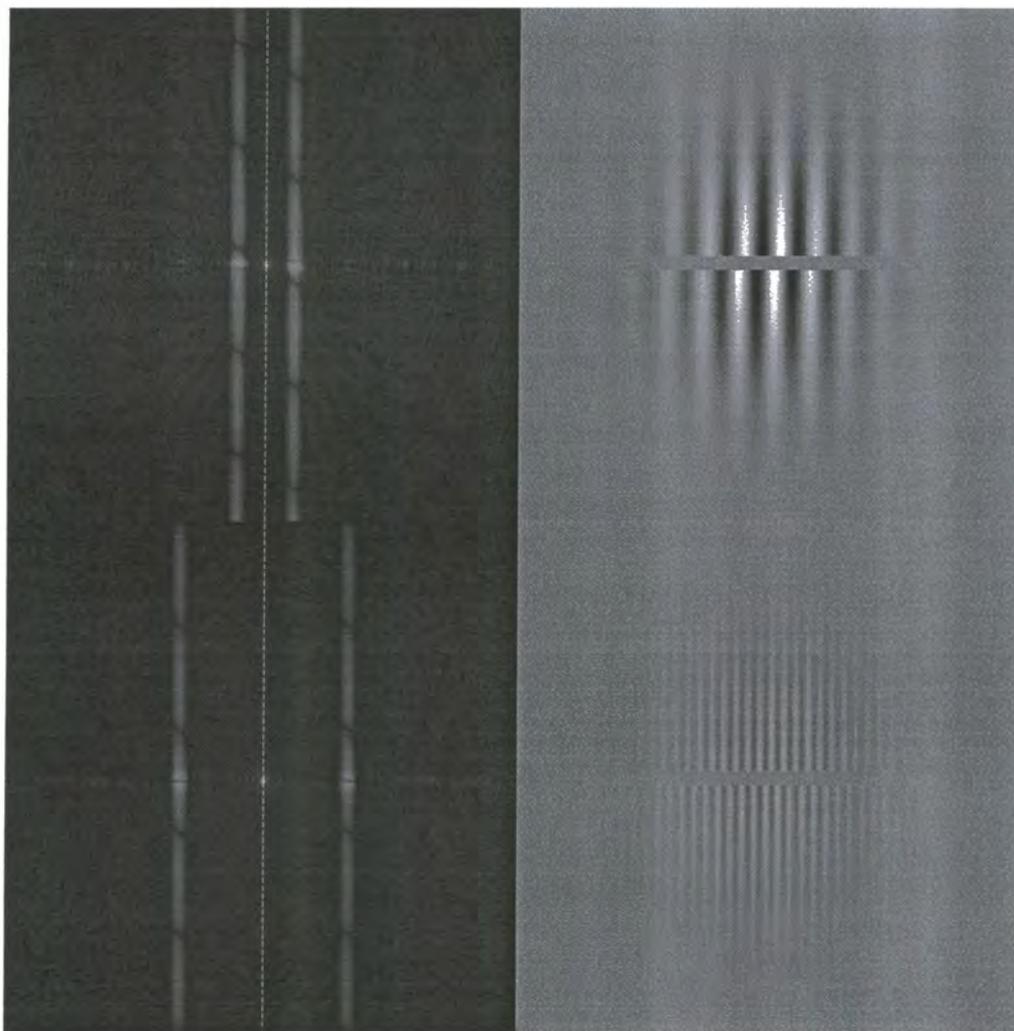


Figure 5.4: Vernier stimuli and their power spectra. The left hand column shows a detail of the power spectra of the Vernier stimuli shown in the right hand column. The images shown at the top are of a 5 cpd spatial frequency, while below this is shown the 15 cpd spatial frequency. Only a section of these power spectra is shown as the details were obscured in the full power spectra images due to their large scale. It can be seen that the predominant energy in the Vernier stimuli are at their respective spatial frequencies, however energy at all spatial frequencies is present in the images because of the sharp edges introduced by the gap between the two grating halves. Note that the low spatial frequency contrast envelope in the stimuli has been introduced by the printing process.

the F would be significantly lower than that for the 3F as the relational information would be very poor for the 3F. It was also predicted that the Vernier acuity thresholds for the 3F / F would be significantly lower than either the F, 3F or the F / 3F Vernier acuity thresholds, while the F / 3F would not be significantly different from the F as this provided the better relational information. This was examined in a series of planned comparisons.

5.3.1 *Method*

All methods are as experiment 11a. Higher spatial frequencies of 5 and 15 cpd were employed and are shown in figure 5.4.

Subjects. Six subjects, three female and three male, with an age range of 21 to 55 took part in this experiment. One subject, the author, was fully aware as the purpose of the experiment.

5.3.2 *Results*

Table 5.3 shows the Vernier thresholds found as a function of the spatial frequencies shown. The values are in seconds. It can be seen that the 3F spatial frequency when presented alone produces very large thresholds with 4 subjects showing thresholds which are greater than the largest horizontal displacement tested. For this reason the 3F condition was excluded from further analysis. The important comparison involves the single spatial frequency which produced the lowest threshold, clearly this is the F alone condition.

A one way repeated measures ANOVA examining image condition: F alone; F / 3F ; 3F / F was found to be significant ($F(2, 10) = 6.9$; $M.S.e = 70.37$; $p < 0.05$).

Planned comparisons were carried out examining the temporal orders 3F / F and F / 3F with each other and respectively with the single spatial frequency conditions of F. It was found that the Vernier threshold for the 3F / F pair was significantly lower than that found with the F / 3F pair ($F_1 = 10.2$; $p < 0.01$) and importantly the 3F / F threshold was lower than the F alone ($F_1 = 10.49$; $p < 0.01$). The Vernier threshold for the F / 3F pair was found not to be significantly different from the F alone ($F_1 < 1$).

Subject	F	3F	F / 3F	3F / F
1	64.33	99.90	61.74	50.75
2	74.52	173.02	69.81	64.59
3	54.76	117.24	80.89	38.38
4	55.16	1899.57	49.40	39.47
5	47.73	136.54	31.57	29.03
6	69.75	129.46	68.53	49.91
mean	52.32	365.53	51.71	39.3

Table 5.3: shows the Vernier acuity thresholds for each of the 6 subjects. The values are in seconds. It can be seen that the Vernier acuity thresholds determined for the 5 cpd are slightly higher than those determined for the 6 cpd spatial frequency grating used in experiment 11a, this is likely to be due to subject variability.

5.3.3 Discussion

It was predicted that the maximisation of positional information from the two spatial frequencies would improve the Vernier acuity thresholds found when both spatial frequencies were presented from fine to coarse when compared to the Vernier acuity

thresholds for the same spatial frequencies when presented alone or from coarse to fine. Furthermore it was predicted that the coarse to fine presentation of these spatial frequencies would not significantly differ from the F spatial frequency when presented alone. The results clearly show that the both of these predictions have been borne out. The Vernier acuity threshold for the 3F / F pair was significantly lower than the F alone and the F / 3F pair, while the threshold for the F / 3F was found not to be significantly lower than the F alone. The proposed explanation was that the Vernier acuity thresholds are dependent upon two limiting forms of positional information which can be maximised so that the combination of the two positional cues produces a threshold significantly lower than would be predicted from the threshold of either alone. It is interesting to note that while the relational information provided by the 3F was very unreliable (as suggested from the extremely large Vernier acuity thresholds determined from that condition), the locational information which the 3F provided was integrated with the low spatial frequency relational information, provided by the F, to lower Vernier acuity thresholds when they were both presented from fine to coarse. Furthermore the results show that the integration of these positional cues only occurs when the spatial frequencies are presented from fine to coarse.

This suggests that the lower thresholds reflect an integration of spatial frequencies which occurs when spatial frequencies are presented from fine to coarse, not as a consequence of a masking effect between the F and 3F spatial frequencies. This supports the findings of the experiments presented in chapter 3 and 4 (experiments 6, 8, 9a, and 9b) showing that a fine to coarse sequence of spatial frequencies is integrated while a coarse to fine sequence is not integrated at all. In terms of the results of experiment 10, examining the contrast thresholds of the 3F when presented preceding or succeeding the F, the lower Vernier acuity thresholds found in the 3F / F pair of spatial frequencies suggests that the higher contrast threshold for the 3F when succeeded by an

F is due to the integration between spatial frequencies when presented in this temporal order rather than masking. Thus the results of the masking experiment reflect the operation of a process which integrates spatial frequencies from fine to coarse not coarse to fine. Furthermore this in turn suggests that the fine to coarse preference found in the square-wave likeness experiments (6, 8, 9a and 9b) also reflect an integration process.

5.4 General Discussion

Overall the results suggest that the Vernier acuity task, as utilised here, is limited by two positional uncertainties: locational uncertainty and relational uncertainty.

Locational uncertainty is concerned with the physical positioning of the grating, while relational uncertainty is concerned with the positioning of the two halves of the grating with respect to each other. Locational uncertainty is expected to improve as spatial frequency increases, however relational information is limited by gap size and matching ambiguity and this becomes degraded as spatial frequency increases. The higher spatial frequency channels provide the best locational information as they are responsive to a smaller amount of space while the lower spatial frequency channels provide the best relational information as they span the gap (or there are fewer low spatial frequency channels spanning the gap) between the two halves of the gratings thereby reducing matching ambiguities. The results of experiment 11a utilising spatial frequencies of 2 and 6 cpd show no significant difference between the Vernier acuity thresholds for the F , $3F$, $F/3F$ and the $3F/F$. This can be explained when it is considered that the positional information provided by both spatial frequencies is similar and therefore their combination does not allow an increase in positional certainty, assuming the noise associated with both signals was uncorrelated. This is also indicated from the results of experiment 11b where the Vernier acuity thresholds for a compound of the F and $3F$

show no improvement. Furthermore, the results of experiment 12 also suggest this dual role of positional information as the spatial frequencies of 5 cpd and 15 cpd give different forms of positional information with the 5 cpd providing the best relational information and the 15 cpd providing better locational information. Their combination then allows for an increase in positional certainty which is reflected in the lowering of the Vernier acuity threshold when both spatial frequencies are presented from fine to coarse.

The important point here is that once the nature of the task and the limits (the two types of positional uncertainty) which it imposes upon the integration of spatial frequencies are considered then the evidence clearly shows that the integration of spatial frequencies does take place and it does so more effectively from fine to coarse. The evidence found in this chapter suggests that the results found in experiments employing the square-wave likeness paradigm (experiments 6, 8, 9a and 9b) and the masking paradigm (experiment 10) are also due to the integration of spatial frequencies.

Thus the findings suggest that spatial frequencies are more effectively integrated when they are presented from fine to coarse, not from coarse to fine as previously suggested. Indeed the results show that with the Vernier acuity paradigm a coarse to fine sequence of spatial frequencies is not integrated at all: experiment 14 showed that the Vernier acuity thresholds found for a coarse to fine temporal order of spatial frequencies was not significantly different from the spatial frequency which produced the lowest threshold when presented alone. A fine to coarse temporal order of spatial frequencies is integrated whereas a coarse to fine temporal order is not.

The results found in experiment 12 and as a consequence the results found in experiment 10 and experiments 6, 8, 9a and 9b (masking and square-wave likeness experiments) all suggest that spatial frequencies are integrated more effectively when they are available from fine to coarse. However, experiments 1, 5 and 7 all show

evidence suggesting a coarse to fine temporal order improves the efficacy with which spatial frequencies are integrated. This suggests that either the evidence for coarse to fine integration may be a result of some other cue or that the integration process is flexible, that is the temporal order in which spatial frequencies are more effectively integrated depends upon the task and the stimuli employed. That the latter is the case is suggested by the results of Parker et al (1996), Schyns & Oliva (1997; Oliva & Schyns, 1997) and Watt (1987) where evidence for a coarse to fine and a fine to coarse integration of spatial frequencies was also found (See section 1.3).

The results show then that the integration of spatial frequencies is not a temporally anisotropic one where spatial frequencies are more effectively integrated from coarse to fine because this mirrors the natural progression of spatial frequency availability in the visual system. In terms of the generic processing model of filter – non-linearity – integrate, outlined in section 1.3.2 this suggests that the integration of spatial frequencies is not limited to occur from coarse to fine as a result of the first stage spatial frequency processing delays but rather it indicates that the integration of spatial frequencies takes place in whichever manner is most optimal for the successful completion of the task at hand.

However, a consideration of the differential delays involved in the processing of spatial frequencies (see section 1.3.2) suggests that the integration stage of spatial frequency processing is limited to occur from coarse to fine. Thus it may be the case that the evidence for the fine to coarse integration of spatial frequencies found here is a consequence of the methodology of presenting spatial frequency information sequentially, that is, the finding for a fine to coarse integration may never occur in normal visual processing because of the spatial frequency processing delays imposing a coarse to fine structure on further processing. Therefore, while a fine to coarse integration has been shown to take place it never actually occurs in normal visual

processing because spatial frequencies are constrained to be made available from coarse to fine. The finding that spatial frequency integration can take place more effectively from fine to coarse is then of no functional importance for visual processing as a coarse to fine integration of spatial frequencies is compelled to take place.

Therefore the evidence for a flexibility in the integration of spatial frequencies shows only that the integration mechanism is flexible and can integrate spatial frequency information from coarse to fine or fine to coarse, however, this is restricted at the earlier stage of spatial frequency processing so that the fine to coarse integration of spatial frequencies never takes place. The critical aspect of this argument then is that the temporal delays in the processing of spatial frequencies restricts the integration of spatial frequencies to occur from coarse to fine. This will be addressed in the following chapter.

Spatial frequency processing delays

Summary

Despite the evidence for flexibility in the spatiotemporal integration of spatial frequencies it is restricted to operate only from coarse to fine because of early spatial frequency processing delays. In this chapter the nature of these spatial frequency processing delays was examined using temporal order judgements, that is which of two visual events presented at different times was shown first. First a replication of Parker & Dutch (1987) was carried out where subjects judged the temporal order of a LED and a variety of spatial frequencies (Experiment 13). No clear evidence was found for a consistent pattern of spatial frequency processing delays. Second, it was found that this lack of consistent pattern was not due to the mismatches between the onset and offset of the LED and the spatial frequencies (Experiment 14). It was argued that the lack of support for spatial frequency processing delays, that is the difference from the Parker & Dutch finding, may be due to the use of a broadband LED as a match. The remaining four experiments examined the temporal order judgements of the onset of gratings with two widely different spatial frequencies (Experiment 15, 16, 17 and 18). Overall the results provide no clear evidence for spatial frequency processing delays. It is suggested then that temporal delays involved in the processing of spatial frequencies may be compensated for, negligible or variable across subjects. It is concluded that the

flexibility found when examining the integration of spatial frequencies could operate in normal visual processing.

6.1 Introduction

It has been shown (see chapter 3, 4 and 5) that the integration of spatial frequencies does not rely on a particular temporal order of spatial frequencies, integration can occur effectively from coarse to fine or from fine to coarse. This has been suggested to be a function of the flexible nature of the integration process, that is spatial frequencies are integrated from coarse to fine or fine to coarse according to the most efficient way in which to successfully complete the task demands of the experiment. However, it has been pointed out that because of the differential delays involved in the processing of spatial frequencies (Breitmeyer, 1975; Gish et al, 1986; Parker & Dutch, 1987; see section 1.3.2) the integration of spatial frequencies is restricted to take place from coarse to fine regardless of the optimal strategy which could be invoked if spatial frequencies were available simultaneously.

Therefore, given the findings for differential delays in the processing of spatial frequencies, it is difficult to understand how flexibility in the integration of spatial frequencies could have any role to play in normal visual processing, that is when spatial frequencies are presented concurrently not artificially stretched to appear progressively from coarse to fine or fine to coarse. The temporal lags introduced by the different processing delays for different spatial frequencies show that spatial frequencies must be initially integrated from coarse to fine.

The results showing a fine to coarse integration of spatial frequencies (see chapter 3, 4 and 5) suggest that the integration of spatial frequencies does not depend upon this coarse to fine availability of spatial frequencies. Spatial frequencies are only

integrated from coarse to fine because of the temporal lags introduced in the initial processing of spatial frequencies. This has been pictured as a two stage model (see section 1.3.2), where the output of a first stage of spatial frequency filters, such as those described by Wilson & Gelb (1984), are subject to a non-linearity, such as halfwave rectification (Watt & Morgan, 1985), and then passed through a second stage of spatial frequency integration.

The results presented in this thesis, for example experiment 5 and 6, suggest that the second stage of spatial frequency integration is a malleable mechanism in which spatial frequencies are integrated in the manner which is optimally suited to the efficient completion of the task or efficient processing of the stimuli. The findings of temporal anisotropies in the integration of spatial frequencies show that the mechanism is flexible and shifting in its preference as spatial frequencies are not integrated isotropically, that is they are not necessarily integrated as and when they become available. It is suggested that the order in which they are more effectively integrated depends upon the nature of the information to be derived from the stimuli in order to carry out the task involved in the experiment. For example, when performing an image detection experiment (experiment 5) a coarse to fine presentation of spatial frequencies is more effectively integrated, but when performing a Vernier acuity task (experiment 12) a fine to coarse presentation order is more effectively integrated.

However when the first stage of spatial frequency filtering is considered it appears that the findings showing temporal delays in the initial processing of those spatial frequencies is critical to the restriction of the second stage. The processing delays in the first stage force the second stage of the model to mirror the first stage and act to all intents and purposes as a coarse to fine integration process.

Many studies have reported results showing differential temporal delays with spatial frequencies (Breitmeyer, 1975; Gish et al, 1986; Parker & Dutch, 1987; see

section 1.3.2). The actual delays found in the temporal processing of spatial frequencies has been found to vary quite substantially with the method employed (See section 1.3.2). Breitmeyer (1975) found that reaction times increased as spatial frequencies became higher. The actual value of this delay was found to be approximately 21 msec per octave. Parker & Dutch (1987) found, using a temporal order judgement experiment that the delay was as little as 5 msec per octave. They argued that this reflected the delays involved in spatial frequency processing more accurately as it removed any added delays which may be a result of the manual response involved in making reaction times (see Ejima & Ohtani, 1987). The temporal order judgement experiments of Parker & Dutch (1987) involved setting the onset of a spatial frequency grating so that it coincided with the onset of an LED. As an LED contains a broadband of spatial frequencies it is difficult to assess whether the temporal delays they found reflect the actual processing delays per octave. Barr (1986) carried out temporal order judgement experiments with two spatial frequency gratings presented at the top and bottom of the screen separated by a strip of mean luminance and containing a fixation point. He found that most of his subjects judged the onset of the gratings of different spatial frequencies to be simultaneous when they were in fact physically simultaneous. That is, he found little evidence supporting spatial frequency delays occurring when the simultaneity of spatial frequency gratings is judged. This suggests that when the necessity of making manual responses are removed the differential delays involved in the processing of spatial frequencies disappears. This gains support from the VEP studies which find no evidence for differential delays in the processing of spatial frequencies (Musselwhite & Jeffreys, 1985; Kulikowski, 1977). Therefore it may be the case that the first stage of processing delays restricting spatial frequency integration to take place from coarse to fine may not be totally restrictive on the processing carried subsequently, this would allow the

integration of spatial frequencies to occur in whichever manner was optimal to the processing of visual information.

6.1.1 Preview of Chapter

In this chapter the evidence for the differential delays in the processing of spatial frequencies will be examined. The first two experiments to be presented were based upon the temporal order judgement experiment of Parker & Dutch (1987) (Experiment 13 and 14). In these experiments a spatial frequency grating and a LED were cyclically presented and the subject matched the onset of the peripherally presented LED with the centrally presented grating. In the second set of experiments the judged simultaneity of two very disparate spatial frequencies was examined in a single presentation method and with a cycled presentation method (Experiment 15 and 16). A third set of experiments assessed the judged simultaneity of two equally detectable spatial frequencies in both single and cycled presentation methods (Experiment 17 and 18).

6.2 Experiment 13: Parker & Dutch (1987) Replication

This experiment was carried out as a reassessment of the Parker & Dutch (1987) processing delay estimate. Their estimate was found to be very small at 5 msec per octave and was suggested to reflect only the perceptual processing delay without involving any manual response delay.

6.2.1 Method

Subjects: Four subjects took part in this experiment, 3 female and 1 male, with an age range of 27 to 34 years. All subjects had normal or corrected to normal vision and were naive as to the purpose of the experiment.

Apparatus & Materials: The stimuli used in these experiments were vertically oriented spatial frequency patches of 0.5, 1.5, 3, 4.5 and 9 cpd shown at 17.7 % contrast subtending 6 degrees of visual angle. Purpose written software was used to create the stimuli employing a VSG display system (Cambridge Research Systems Ltd). These were presented for 40 msec on a 21 inch(1024 x 768 pixels) monochrome Eizo flexi-scan 6600-M monitor with a refresh rate of 69 Hz and a mean luminance of 12.9 cd/m².

A 5 mm diameter general purpose red LED was attached to the top centre of the monitor approximately 2.5 cm away from the screen. The onset and duration of the LED was controlled by a purpose built device. A trigger was passed to the LED via software (this took three msec), at which point the delay between this trigger signal and the actual onset of the LED could be varied over a period of 100 msec using a potentiometer. The duration of the LED was 40 msec. These readings were verified using an oscilloscope (Trio 100 Mhz Oscilloscope, CS-2100A).

Stimuli had a square-wave temporal waveform and were viewed binocularly at 1m. A fixation mark was presented for 1500 msec followed by the onset of the grating and the LED for 40 msec. The stimulus onset asynchrony (SOA) of the LED could be varied by the subject using the potentiometer from a setting where the onset occurred 50 msec prior to the grating to one where it occurred 50 msec after. The subjects head was placed in a chin rest throughout the duration of the experiments.

As with Parker & Dutch (1987) a second series of readings was taken for the 0.5 and the 9 cpd grating with its duration increased to 400 msec. The LED duration remained at 40 msec. This was to ensure that the subjects' LED onset judgements were carried out with regard to the onset and not the some other aspect of the grating timing such as the offset.

Please note that there are differences between this experiment and that carried out by Parker & Dutch (1987). First the timing in the Parker & Dutch (1987) experiment 1 was controlled by a series of relays and thus subjects wore headphones with white noise played to mask the noise made by the relays. Second the luminance of the screen in the Parker & Dutch experiment was set at 5.0 cd / m^2 .

Design: The method used was one of adjustment where the subject adjusted the onset of the peripheral LED so that it appeared simultaneously with the onset of a grating of a variety of spatial frequencies. Six readings of the five spatial frequencies were determined in a randomised order. Thus there were 30 trials overall.

Procedure: The subject's task was to fixate centrally and manipulate the peripherally viewed LED onset so it occurred simultaneously with the onset of the grating. The sequence of LED / grating onset was cycled with an inter - stimulus interval of 1500 msec, at which point the fixation mark was presented. Once the subject was satisfied with the LED onset with respect to the grating, the difference in their onsets was noted from the screen of the oscilloscope, where both the onsets and durations of the LED and grating were represented. The next trial was then presented.

6.2.2 Results

Figure 6.1 shows the LED latency estimate as a function of spatial frequency for the four subjects. It can be seen that at a spatial frequency grating duration of 40 msec there is no definite pattern supporting a consistent processing delay. Subjects latency estimates showed a variety of patterns with the LED being judged to have appeared simultaneously with the spatial frequency grating when it had in fact been presented both preceding and succeeding the onset of the spatial frequency gratings. Two subjects

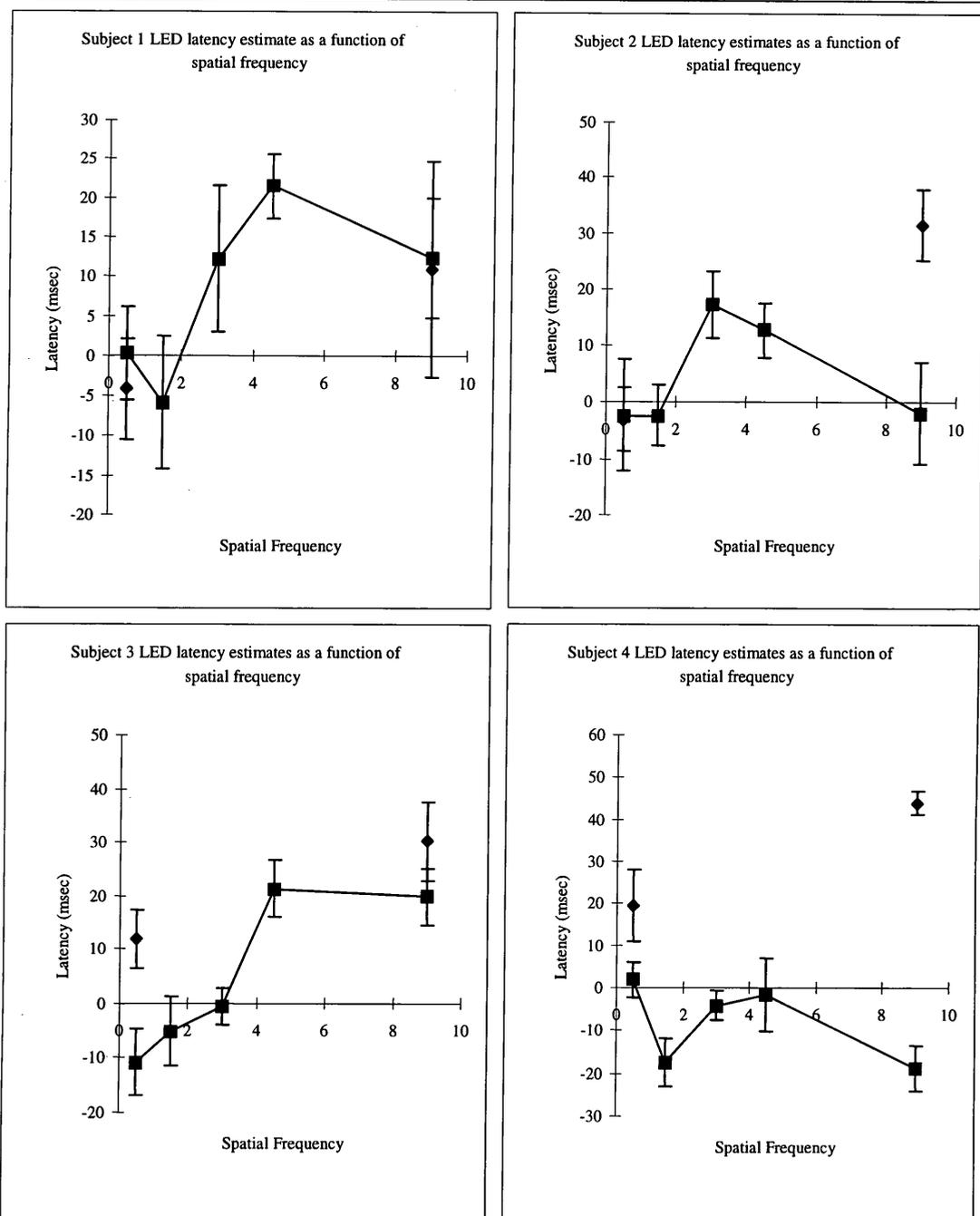


Figure 6.1: shows the estimated onset of a grating of various spatial frequencies in msec, using a LED as a matching stimulus. The actual onset of the gratings is represented as occurring at zero. The data points denoted by black squares represent those conditions where the LED and spatial frequency gratings had an equal duration of 40 msec. The black diamonds denote conditions where the LED was shown for 40 msec and the spatial frequency grating was shown for 400 msec.

showed an increase in spatial frequency processing time as spatial frequency increased, but even this latency effect seemed to plateau at 4.5 cpd. A more consistent pattern was found when the spatial frequency grating was shown for 400 msec. These data points are denoted by the black diamonds in figure 6.1 at 0.5 and 9 cpd. For three of the four subjects there was an increase in the estimated point of simultaneity as spatial frequency increased when the duration of the two stimuli differ.

6.2.3 Discussion

The results show no consistent support for spatial frequency temporal delays involved in early spatial frequency processing when the duration of the LED and spatial frequency grating were the same. However there was found to be an increase in the judged simultaneity of an LED with a duration of 40 msec and a spatial frequency grating with a duration of 400 msec. This suggests that there is a increase in the processing time of spatial frequencies as they increase. The average increase over all subjects was found to be 5.67 msec per octave which is very similar to that found by Parker & Dutch (1987).

The difference in the results when the LED and grating were of the same duration could be due to a problem with matching the onset of the LED and spatial frequency grating when both were shown for a short duration, that is the subject may have on some occasions matched the offset not the onset. When the duration of the grating is increased to 400 msec this mismatch of onset and offset is more difficult to make. This suggests then that to assess the temporal delays involved in spatial frequency processing using this form of temporal order judgement experiment it is necessary to dissociate the durations of the two components.

In order to assess the delays involved in early spatial frequency processing a second experiment was carried out using an LED at a duration of 40 msec and spatial frequencies at durations of 400 msec.

6.3 Experiment 14: temporal order judgement of a LED and spatial frequencies with different durations

In experiment 13 little evidence was found for delays involved in the processing of spatial frequencies such that the low spatial frequencies are processed faster than the higher ones. It was suggested that this may be due to a mismatch between the onset and offset of the LED and gratings. In order to investigate the delays involved in the processing of spatial frequencies the following experiment used the same design as experiment 13, however the duration of the gratings was extended to 400 msec while the LED duration was held at 40 msec. This would ensure that the match between the LED and grating would be made on the basis of their onsets.

6.3.1 Method

All methods are identical to experiment 13 except were noted.

Subjects: Two subjects took part in this experiment, both female, with an age range of 28 to 31 years. Both subjects were naive as to the purpose of the experiment and neither participated in experiment 13.

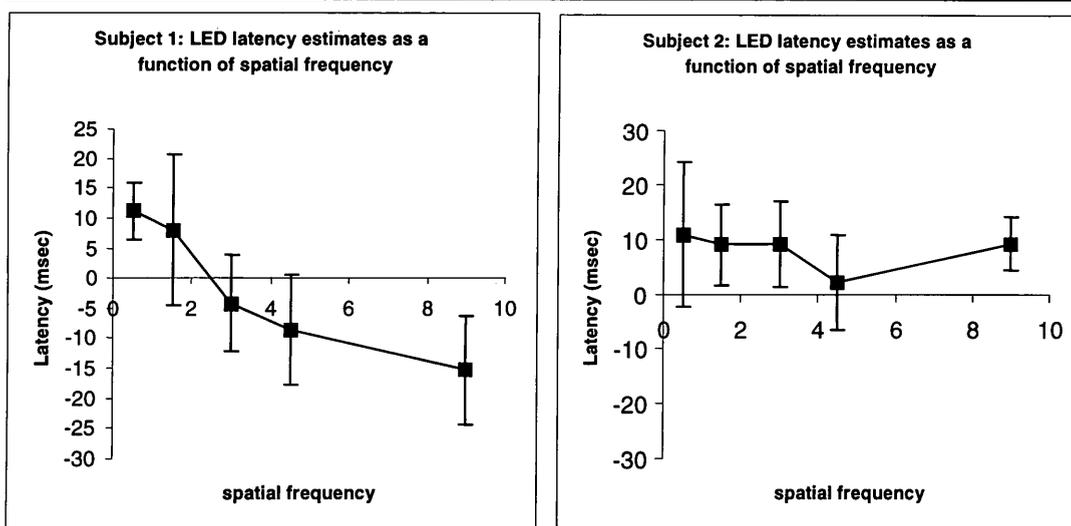


Figure 6.2: shows the estimated onset of a grating of various spatial frequencies in msec, using a LED as a matching stimulus. The actual onset of the gratings is represented as occurring at zero.

Apparatus & Materials: The duration of the gratings were increased to 400 msec. The LED was kept at a duration of 40 msec. Exactly the same spatial frequencies of 0.5, 1, 3, 4.5 and 9 cpd were employed.

6.3.2 Results

Figure 6.2 shows the results for the latency estimates for the simultaneity judgements of the LED as a function of spatial frequency. It can be seen that there is no consistent pattern of delay found. Subjects showed no support for an increase in processing delays as spatial frequency increased.

6.3.3 Discussion

The results show no clear support for any pattern of delay structure imposed on the simultaneity of an LED with increasing spatial frequency. Thus the estimated onset of spatial frequencies does not increase as spatial frequencies become higher. A mismatch of onset and offset, suggested to be the cause of the inconsistency found in

experiment 13, cannot account for the inconsistent pattern found here. The results may be due to two reasons: there are no temporal processing delays in spatial frequency processing; or there is a problem with the methodology. Given the large amount of support for spatial frequency processing delays it appears that using this method of temporal order judgements of a LED and gratings of various spatial frequencies does not give any clear indication of the existence of differential delays in the processing of spatial frequencies. It may be the case that this methodological problem is due to the use of an LED as a match as it is broadband in its spectrum so it is difficult to know what spatial frequencies exactly the subject is matching.

One way to address this problem of being certain which spatial frequencies are being matched is to employ matching stimuli which contain a narrowband of spatial frequencies. It has already been pointed out that Barr (1986), using two spatial frequencies, found little evidence for delays involved in the processing of spatial frequencies. He found different spatial frequencies were judged to have appeared simultaneously when they did in fact appear simultaneously. However there are some problems with the paradigm employed by Barr. When judging the temporal order of two gratings they appeared in two halves at the top and bottom of the screen. The edges of the gratings will have introduced spurious spatial frequencies into the display. These may have been used in the temporal order judgement. Thus, even when the spatial frequencies were widely different, it may be the case that the reason that the gratings were judged as having appeared simultaneously is because the spatial frequencies they were judged upon were the same and were introduced by the edges of the gratings where the judgement was most likely made. A further problem with Barr's experiments is that the temporal order judgement of spatial frequencies was assessed when they were presented in the periphery. It is not clear what effect this would have on the estimated processing delays as it is generally assessed for foveal presentation. For these reasons it

was decided to explore this temporal order judgement of two foveally presented spatial frequency Gabor patches which contain a narrow band of spatial frequency information.

6.4 Experiment 15: Examining the Temporal Order Judgements of spatial frequencies with stimuli equally above contrast threshold

Experiments 13 and 14, using a temporal order judgement experiment, showed little supporting evidence for a consistent pattern of spatial frequency processing delays. However evidence for such delays has been found when measured by reaction times (Lupp et al, 1976) and some VEP studies (Parker & Salzen, 1977). It was argued that this discrepancy may be due the use of a LED for matching the onset of a spatial frequency grating, as an LED will contain a broadband of spatial frequencies. In order to examine spatial frequency processing delays using a temporal order judgement methodology it would be better to match the temporal order of spatial frequencies amongst themselves. This was carried out by Barr (1986) and he found little evidence for spatial frequency processing delays. However, these results may have been affected by the use of gratings the edges of which will have introduced a number of spatial frequencies into the display and the peripheral presentation of these gratings.

In the following experiment the temporal order judgement of two widely different spatial frequencies were assessed using Gabor patches. These contained only a narrowband of spatial frequencies and thus no spurious spatial frequencies were introduced into the display. Furthermore these Gabor patches were presented at the same central location and the temporal stimulus onset asynchrony of the two was varied.

In this experiment the spatial frequencies of 0.5 and 16 cpd were employed. If indeed processing delays do occur then the onset at which both spatial frequencies will

appear as being simultaneous will differ from the physically simultaneous stimulus onset asynchrony of zero, that is the two spatial frequencies will appear as being simultaneous when the 16 cpd was presented prior to the 0.5 cpd spatial frequency. If no temporal delays are involved in the processing of spatial frequencies then the point at which they are estimated to be simultaneous will be when they have a stimulus onset asynchrony of zero.

It is possible, from the experiments examining the temporal order judgement of spatial frequencies and reaction times to spatial frequencies, to predict the stimulus onset asynchrony at which the two spatial frequencies will be judged to have appeared simultaneously. If a conservative estimate is taken such as that found by Parker & Dutch (1987) of 5 msec per octave of spatial frequency then as the 0.5 cpd and 16 cpd spatial frequencies are 5 octaves apart the processing of the 16 cpd will take 25 msec longer than the 0.5 cpd. Therefore the 16 cpd must be presented 25 msec prior to the 0.5 cpd for it to be judged as having appeared simultaneously with the 0.5 cpd spatial frequency. If a temporal delay more commonly estimated is taken, such as that found by Breitmeyer (1975) or Gish et al (1986) of 21 msec per octave then the 16 cpd must be presented 105 msec prior to the onset of the 0.5 cpd spatial frequency. Thus in order for the two spatial frequencies to be judged as having appeared simultaneously it would be predicted that the 16 cpd should be shown 25 msec or 105 msec prior to the onset of the low spatial frequency.

The two estimates of spatial frequency processing delays are vastly different. The results of Parker & Dutch (1987) gives figures which are much lower than that predicted from the reaction time data. While Parker & Dutch (1987) argue that the delay estimated from their data reflects the perceptual processing delays of spatial frequencies without the added delays incurred when employing a reaction time measure, their results were not replicated in experiment 13 or 14. Thus there is little supporting evidence for

such a low estimate of spatial frequency processing delays, therefore the longer estimate will be taken as the predicted values expected in the following experiments.

In this experiment then the evidence for spatial frequency processing delays is examined. The contrast thresholds are determined for a 0.5 and a 16 cpd Gabor patch presented full screen for 14 msec. This threshold is then doubled to maintain equal activation at both spatial frequencies (see section 1.3.2 and 6.5.3) and used in the second experiment of temporal order judgements. The SOA between the 0.5 and the 16 cpd were varied from -224 msec to +224 msec. If the onset of the 0.5 cpd spatial frequency is considered as occurring at time $t = 0$ it would be expected that, if the delays proposed do exist then, in order for the stimuli to be judged as occurring simultaneously the 16 cpd should be shown substantially prior to the 0.5 cpd patch, on the order of 105 msec, as its associated processing time is increased.

6.4.1 *General Methods*

Subjects: There were six subjects, four males and two females, with an age range of 27 to 31 all had normal or corrected to normal vision. One of the subjects was fully aware of the purpose of the experiments the remaining five were naive.

Apparatus and Materials: The stimuli used in these experiments were vertically oriented spatial frequency Gabor patches of 0.5 and 16 cpd and a spatial envelope of standard deviation 2.48 degrees. Purpose written software was used to create the stimuli employing a VSG display system (Cambridge Research Systems Ltd). These were presented on a 21 inch(1024 x 768 pixels) monochrome Eizo flexi-scan 6600-M monitor with a refresh rate of 69 Hz and a mean luminance of 12.9 cd/m^2 . Stimuli were viewed binocularly at 1.4m. for 14 msec and had a square-wave temporal waveform. The subjects head was placed in a chin rest throughout the duration of the experiments.

Procedure: Subjects were fully informed about the nature of the experiments prior to their participation, including examples of the stimuli set and the temporal form the experiments were to take. All experiments were carried out in a blacked out room.

More detailed methods for the contrast threshold determination and temporal order judgements are given below.

Experiment 15a: Contrast threshold determination

The contrast thresholds for the two spatial frequencies of 0.5 and 16 cpd were estimated in separate blocks each utilising a temporal two alternative forced choice, two up, one down, double staircase as described by Cornsweet (1962). The two intervals were denoted by an auditory signal, following the second of which the subject made a decision as to which interval contained the briefly presented Gabor patch. Following the contrast sensitivity function (Schade, 1956) it was expected that the 16 cpd patch would require a much higher contrast level than the 0.5 cpd. The starting value for the staircases reflected this. The 0.5 cpd staircases started at zero and 29.3% and initially changed the contrast values in steps of 1.83% to allow a rapid convergence to threshold after ten reversals this step size was changed to 0.37% to allow a fine tuning of the threshold value. The 16 cpd staircases started at zero and 93.75 %, initially decreasing by 6.25% and thereafter by 1.25%. The threshold values were computed by averaging the last five reversal pairs beyond the initial ten reversal points. Prior to proceeding with experiment 15a subjects adapted to the luminance of the screen for two minutes.

6.4.1.2 *Experiment 15b* temporal order judgements

This experiment used a method of constant stimuli paradigm with a temporal order judgement. The two contrast thresholds determined for the 0.5 and 16 cpd Gabor patches from experiment 15a were doubled and these were presented briefly to the subject with a varying stimulus onset asynchrony (SOA). The subject's task was to identify which stimulus was shown first, either the high spatial frequency or the low. If the low spatial frequency patch is taken as representing a standard of onset time $t = 0$ then the stimulus onset asynchronies of the 16 cpd spatial frequency were $t = -224; -112; -56; -14; 14; 56; 112; \text{ and } 224$ msec, where a minus figure indicates the high spatial frequency patch preceding the low spatial frequency standard and positive numbers indicating the high spatial frequency patch succeeding the standard. The experiment was a within subjects design as all subjects completed all conditions concurrently. The subject was required to make 100 decisions per SOA condition. Due to the lengthy nature of this experiment the 800 decisions were blocked into smaller packages of 4 blocks of 200 trials. Within these 200 trials an equal number of decisions per SOA condition were made, that is 25 decisions per condition per block. The subjects were given the freedom to take as long a break between blocks as necessary without actually leaving the experimental room.

6.4.2 Results

6.4.2.1 Experiment 15a: Contrast threshold determination

Table 6.1 shows the contrast thresholds for the 0.5 and the 16 cpd patch for the six subjects. It can be seen that as expected from the contrast sensitivity function and the reduced duration of the stimuli that the contrast threshold for the 0.5 cpd Gabor patch is much lower than that found for the 16 cpd Gabor patch.

Subject	0.5 cpd	16 cpd
1	3.1	29.9
2	1.95	47.34
3	1.57	43.77
4	2.1	50
5	2.95	44.4
6	1.7	34.5

Table 6.1: shows the contrast threshold (in percent) of six subjects for a 0.5 and a 16 cpd spatial frequency Gabor patch in percent.

6.4.2.2 *Experiment 15b* temporal order judgements

Figure 6.3 shows the temporal order judgements for the 0.5 and 16 cpd Gabor patch for all subjects as a function of the 0.5 cpd spatial frequency onset. A data point offset to the left of zero indicates that the subject judged that the 16 cpd was shown prior to the 0.5 cpd spatial frequency, and an offset to the right of zero indicates that the subject judged the 16 cpd as having an onset after that of the 0.5 cpd spatial frequency. The curve through the data points shows the probit fit to the data. Subjects 4 and 6 showed no deviation from chance at any stimulus onset asynchrony and so were discounted from further analysis. The point of subjective equality, where the 16 cpd spatial frequency is judged to have an onset equal to the onset of the 0.5 cpd spatial frequency, was taken as the 50% mark of 0.5 cpd presented first responses. These are shown in table 6.2.

It can be seen from Table 6.2 that all subjects judged the two spatial frequencies as occurring simultaneously when in fact the 16 cpd spatial frequency had been presented prior to

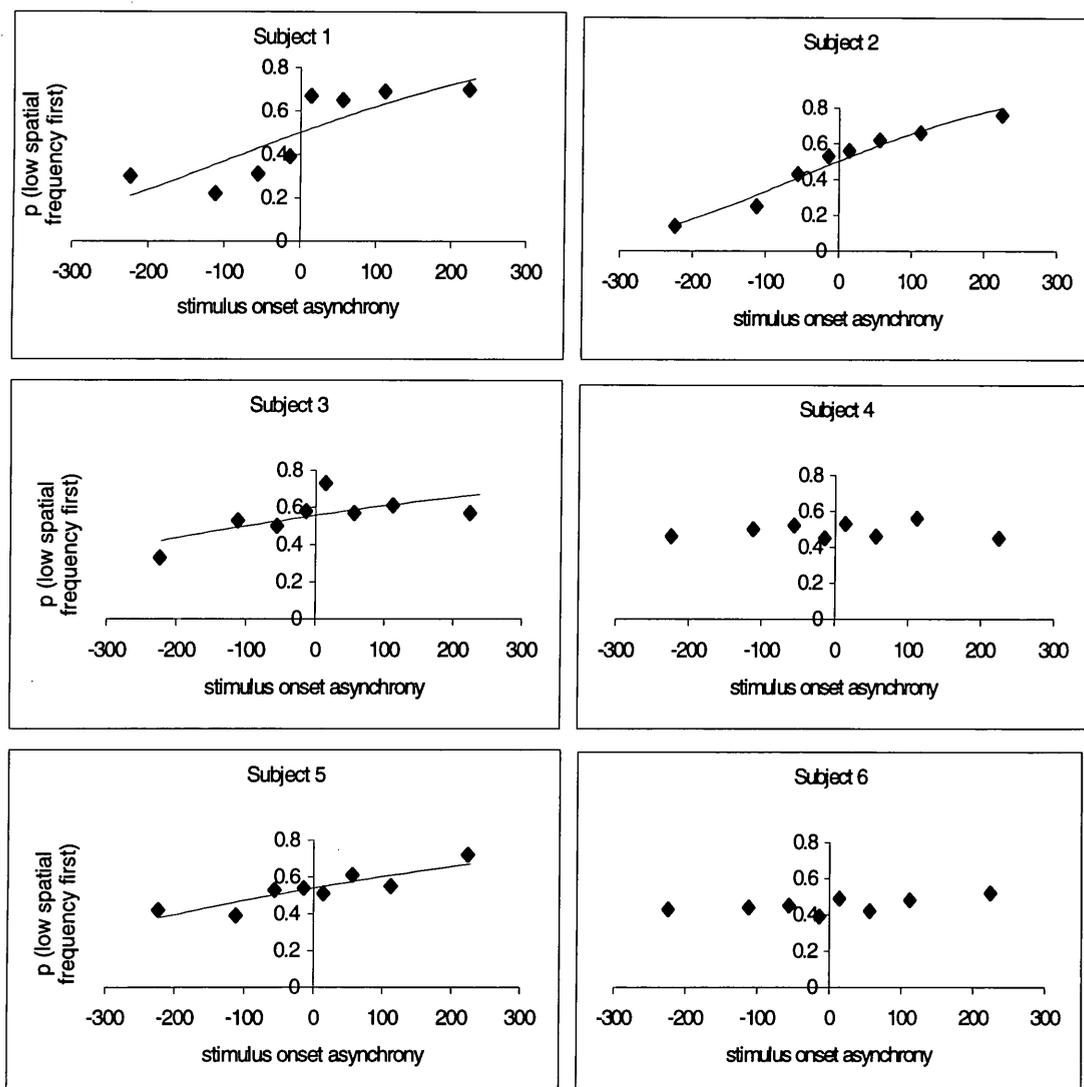


Figure 6.3: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject.

the 0.5 cpd spatial frequency. Two subjects show a point of simultaneity which is quite close to zero but the remaining two show a strong deviation from zero showing

that in order for the high spatial frequency to be simultaneous with the low spatial frequency the high spatial frequency must be shown 60 to 100 msec prior to the low spatial frequency. However, the confidence intervals were found to be very large and in all cases but one (subject 5) included zero, the point of physical simultaneity. Therefore three of the latency estimates were not significantly different from zero. In one case, subject 2, the predicted point of simultaneity was found to significantly different from the predicted stimulus onset asynchrony of 105 msec.

	P.S.E.	Negative Conf	Positive Conf
1	-1.05	-115.99	130.53
2	-1.33	-23.44	21.21
3	-100.57	-473.86	274.37
5	-60.18	-127.47	-8.33

Table 6.2: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval).

6.4.3 Discussion

The results show that the point of simultaneity for the two spatial frequencies of 0.5 and 16 cpd was found to vary largely across subjects. The results show that the point of simultaneity was judged to be as little as -1.05 msec and as large as -100.57 msec. All estimates of simultaneity were to the left hand of zero thus supporting a increased in

processing time associated with the higher spatial frequency. However two of these estimates were smaller than that predicted (105 msec) and two of these estimates were larger than that predicted. The larger two estimates were not consistent showing approximately 40 msec difference between each other. The confidence intervals for these points of simultaneity were found to be very large. In all cases, except one, the confidence intervals associated with each P.S.E. included the physically simultaneous point of zero. One of the subjects showed a pattern of processing delays which was not significantly different from the predicted and was significantly different from zero. However the remaining three subjects showed processing delay patterns which did not significantly differ from a stimulus onset asynchrony of zero.

Therefore it seems that these results show no clear support for the temporal delays involved in spatial frequency processing, nor do they support there being no delays, as found by Barr (1986). The fact that two subjects showed little deviation from chance performance in the temporal order judgement task and the large confidence intervals suggest that the subjects may have found the task difficult to complete. This may be due to the very high contrast thresholds necessary for the detection of the 16 cpd spatial frequency.

Parker & Dutch (1987; and experiments 13 and 14) presented the spatial frequencies on cyclical basis where the temporal order judgement could be assessed over a larger period of time. It was decided that in order to reduce the difficulty of this temporal order judgement task the experiment presented here, assessing the temporal order judgement of two disparate spatial frequencies, would be repeated with a cycled presentation method.

6.5 Experiment 16: Examining the Temporal Order Judgements of spatial frequencies with cycled stimuli equally above threshold

In this experiment the evidence for spatial frequency processing delays was again examined. This experiment is an extension of the temporal order judgement experiment where the stimuli were shown at double the contrast threshold and shown in a single presentation interval. In this experiment the stimuli are again shown at double contrast threshold but the presentation interval was cycled. As with experiment 15a the contrast thresholds are first determined for a 0.5 and a 16 cpd Gabor patch presented full screen for 14 msec. This threshold is then used in the second experiment to assess temporal order judgements in experiment 16b which is based on experiment 15b but the presentation of the two gratings is cycled. The stimuli are presented briefly for 14 msec while the SOA between the 0.5 and the 16 cpd were varied from -224 msec to +224 msec if the 0.5 cpd onset is considered as occurring at time $t = 0$. The presentation period cycles every 1500 msec. It would be expected that if the delays proposed do exist then in order for the stimuli to be judged as occurring simultaneously the 16 cpd should be shown substantially, 105 msec, prior to the 0.5 cpd patch as its processing time is increased. Subjects can examine the display for as long as they wish.

6.5.1 Method

All methods were as experiment 15 except where noted.

Subjects: There were six subjects, four males and two females, with an age range of 18 to 28. One of the subjects was fully aware of the purpose of the experiments the remaining five was naive.

6.5.1.1 *Experiment 16b*: Temporal order judgements

The stimuli used in these experiments were exactly the same as in experiment 15 but the contrast of the spatial frequencies was determined from experiment 16a. The presentation of the two stimuli in the temporal order determination of onset was cycled every 1500 msec. The subject was required to make 60 decisions per SOA condition. Due to the lengthy nature of this experiment the 480 decisions were blocked into smaller packages of 3 blocks of 160 trials. Within these 160 trials an equal amount of decisions per SOA condition were made, that is 20 decisions per condition per block. The subjects completed the blocks on successive days.

6.5.2 *Results*

6.5.2.1 *Experiment 16a*: Contrast threshold determination

Table 6.3 shows the contrast thresholds for the 0.5 and the 16 cpd patch for the three subjects. As with experiment 15a the contrast thresholds were found to be much higher for the 16 cpd spatial frequency than those for the 0.5 cpd spatial frequency.

	0.5 cpd	16 cpd
1	3.1	29.9
2	1.90	37.2
3	1.86	24.9
4	1.98	46.76
5	2.05	42.77
6	1.56	47.6

Table 6.3: Contrast thresholds (in percent) for the 0.5 and 16 cpd Gabor patches for six subjects.

6.5.2.2 *Experiment 16b*: temporal order judgements

Figure 6.4 shows the temporal order judgements for the 0.5 and 16 cpd Gabor patches for all subjects as a function of the 0.5 cpd spatial frequency onset. A data point offset to the left of zero indicates that the subject judged that the 16 cpd was shown prior to the 0.5 cpd spatial frequency, and an offset to the right of zero indicates that the subject judged the 16 cpd as having an onset after that of the 0.5 cpd spatial frequency. The curve through the data points shows the probit fit to the data. The results of both subject 5 and 6 show little deviation from chance and are thus not included. The point of subjective equality, where the 16 cpd spatial frequency is judged to have an onset equal to the onset of the 0.5 cpd spatial frequency, was taken as the 50% mark of 0.5 cpd presented first responses. These are shown in table 6.4.

The results show that the point of simultaneity was found to vary with subject. Both subjects 1 and 2 show a point of simultaneity when the 0.5 cpd spatial frequency was shown prior to the 16 cpd spatial frequency, while subjects 3 and 4 show a point of simultaneity when the 16 cpd spatial frequency was presented prior to the 0.5 cpd spatial frequency. The confidence intervals of three of the subjects (subject 1,2 and 3) show that the points of simultaneity are not significantly different from zero and are significantly different from the predicted stimulus onset asynchrony value of -105 msec. The fourth subject shows a point of simultaneity which is significantly different from the predicted point of 105 msec but which is also significantly different from the physical point of simultaneity.

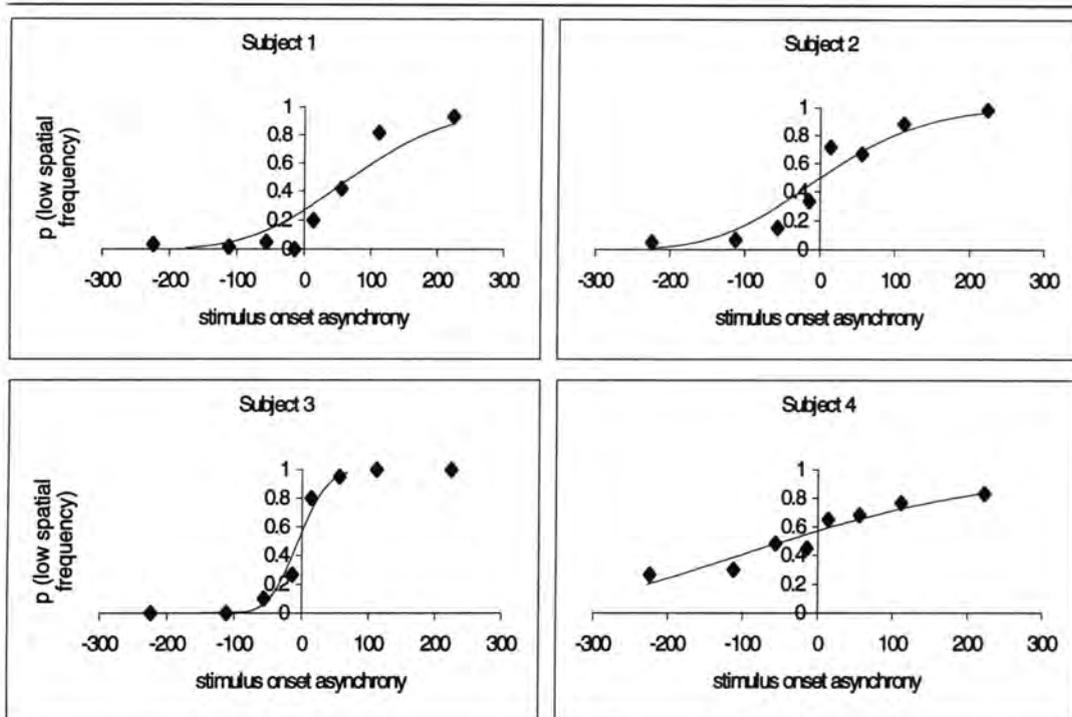


Figure 6.4: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject.

Subject	P.S.E.	Negative Conf	Positive Conf
1	71.8	-49.7	216.11
2	1.2	-81.79	94.02
3	-5.31	-11.92	1.35
4	-44.85	-76.18	-16.163

Table 6.4: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval).

6.5.3 Discussion

The results show little evidence for the expected latency differences in the processing of spatial frequencies. Indeed they show that the processing delays estimated are not significantly different from zero for three of the subjects, while the fourth did not approach the size of the delays associated with spatial frequency processing that have been shown with reaction time and VEP experiments. Thus the results show little consistent evidence for temporal delays involved in the processing of spatial frequencies. The results are, however, not inconsistent with a much smaller estimate of processing delay such as that found by Parker & Dutch (1987). In line with this, as with experiment 15, there was found to be a large variation in the point simultaneity judged by the subjects, shown by the large differences between subjects' points of stimulus simultaneity, despite the use of an extended presentation paradigm. The difficulty of the task is again underlined by the lack of deviation from chance shown by two of the subjects.

The lack of consistent performance found in this experiment mirrors that found in experiment 15. A cycled form of spatial frequency presentation does not appear to greatly increase the ease with which this task can be completed. The result may be due to the method of using multiple contrast threshold levels as representing an equally level of detectability. The method of increasing contrast thresholds multiplicatively reflects the assumption that at contrast threshold levels the level of activation which the contrast thresholds produce is equal at each spatial frequency. Thus increasing the contrast threshold equally across all spatial frequencies maintains this equal level of activation in the visual system. However this assumes that the rate at which activation increases is the same across all spatial frequencies for contrast which are equally above contrast threshold. While there is evidence for this from work examining perceived or apparent contrast (Kulikowski, 1976) other work does not show this. Georgeson & Sullivan

(1975) found evidence that the apparent contrast of high spatial frequencies increased much faster than the apparent contrast of lower spatial frequencies (see also Cannon & Fullenkamp, 1991). If it is the case that the activation caused by different spatial frequencies is not the same when they are shown at contrasts equally above contrast threshold then the 0.5 cpd and 16 cpd spatial frequency would not be equally detectable. It was decided therefore that a temporal order judgement experiment be carried out using equally detectable spatial frequencies.

6.6 Experiment 17: Temporal order judgement as function of two equally detectable stimuli

The results of experiment 15 and 16 were found to be highly variable. This may have been due to an unequal detectability of the two spatial frequencies. In this experiment the contrast thresholds of a 0.5 and 16 cpd spatial frequency was determined. Four contrast values were chosen including and above this contrast threshold value and the detectability of the spatial frequencies was assessed at each contrast level. The point at which the two spatial frequencies were detectable 90% of the time was chosen as the point at which spatial frequencies were equally detectable. The temporal order of these two equally detectable spatial frequencies was then assessed.

6.6.1 Method

All methods are as experiment 15 except were noted.

Subjects: There were three subjects, two male and one female, with an age range of 25 to 34. Two of the subjects were fully aware of the purpose of the experiments the remaining one was naive.

6.6.1.1 *Experiment 17a: Contrast threshold*

The contrast thresholds for a 0.5 and 16 cpd spatial frequency was determined. All methods are as experiment 16a.

6.6.1.2 *Experiment 17b: Equally detectable*

The design of this experiment was a temporal two alternative forced choice experiment utilising a method of constant stimuli paradigm. The task of the subject was to detect the presence of either the high or low spatial frequency patch depending upon block in two auditory signalled intervals. The detectability of four different stimuli contrasts were measured. The individual thresholds as measured in experiment 17a and log steps above this value. In the case of the 0.5 cpd Gabor patch the log steps were 0.1, while the log step difference in the 16 cpd patch case was 0.15. This enabled the detectability of a broad range of contrasts to be assessed. The detectability of the two spatial frequencies were determined in separate blocks. The subject is presented with two temporal intervals both denoted by an auditory signal. Subjects indicated in which interval a Gabor patch was presented. Each contrast level was presented 100 times, requiring 400 hundred decisions per experimental block.

6.6.1.3 *Experiment 17c: temporal order judgements*

The subject was required to make 100 decisions per SOA condition. Due to the lengthy nature of this experiment the 800 decisions were blocked into smaller packages of 4 blocks of 200 trials. Within these 200 trials an equal amount of decisions per SOA condition were made, that is 25 decisions per condition per block. The subjects were given the freedom to take as long a break as necessary without actually leaving the experimental room. All other methods are the same as experiment 15b.

6.6.2 Results

6.6.2.1 Experiment 17a: contrast thresholds determination

Subject	0.5 cpd	16 cpd
1	3.1	29.9
2	2.3	26.2
3	1.6	20.7

Table 6.5: the contrast thresholds (in percent) for the 0.5 and the 16 cpd Gabor patch

Table 6.5 shows the contrast thresholds for the 0.5 and the 16 cpd patch for the three subjects. These values were used to set four levels of contrast to be employed in the determination of the detectability of the stimuli. For the low spatial frequencies, three other contrast levels 0.1 log steps above the contrast threshold were used. For the 16 cpd spatial frequency log steps of 0.15 were used to set three contrast levels above contrast threshold. The actual contrast values for each subject are shown in table 6.6.

6.6.2.2 Experiment 17b: Equally detectable

The detectability ratings for the 0.5 and 16 cpd Gabor patches when presented at contrast threshold and log steps above are shown in table 6.7. It can be seen, as would be expected, that as contrast is raised so the detectability ratings increase. From this data the contrast values for ensuring that the spatial frequencies of 0.5 cpd and 16

Subject	0.5 cpd				16 cpd			
	Threshold	0.1	0.2	0.3	Threshold	0.15	0.3	0.45
1	3.1	3.9	4.9	6.2	29.9	42.2	59.7	84.3
2	2.3	2.9	3.7	4.7	26.2	37	52.3	73.8
3	1.6	2	2.5	3.2	26.7	37.7	53.3	75.3

Table 6.6: Four contrast values (in percent) for the two spatial frequencies to be used in the temporal order judgements of experiment 17c. The contrast thresholds for both spatial frequencies were used and from this contrast value three other contrast levels were chosen. The contrast increased in log steps of 0.1 or 0.15 for the 0.5 cpd spatial frequency and the 16 cpd spatial frequency respectively. This enabled the detectability of a good spread of contrast values to be assessed.

cpd were equally detectable stimuli were estimated using a probit curve fitting procedure (Finney, 1952). From these probit estimates the contrast which produces stimulus detectability of 90% was chosen to represent equally detectable stimuli for use in experiment 19c. The 90% detectability level was chosen because of the increase in the error of equal detectability estimation with the increase in contrast level which reflects the nature of the psychometric curve. For example if a contrast level which reflected 99% detectability was chosen then the error that this contrast level was incorrect would be greater due to the flattening of the psychometric curve as it reaches 1.0. This is illustrated in figure 6.5.

Subject	0.5 cpd				16 cpd			
	Threshold	0.1	0.2	0.3	Threshold	0.15	0.3	0.45
1	81	89	98	99	60	74	86	96
2	75	78	92	95	61	81	90	98
3	82	94	99	100	64	66	83	96

Table 6.7: the detectability (in percent) of the 0.5 cpd and 16 cpd spatial frequencies as a function of their contrast in percent.

It can be seen then that it is advantageous to select a lower detectability rating to minimise the possible error associated with the choice of contrast level. It is for this reason that the detectability level 90% was chosen. These are shown in table 6.8.

6.6.2.3 *Experiment 17c*: temporal order judgements

The results of the temporal order judgements of the individual subjects are shown in Figure 6.6. The solid curves show the probit fit to the individual subjects data. It can be seen, overall, that temporal order judgements are relatively poor, with stimulus onset asynchronies as high 224 msec producing judgements far short of perfect. However, it is still clear that the expected bias to one side of the stimulus onset asynchronies was not found. The probit estimates show that all subjects that stimulus onset asynchronies of zero are close to the chance levels of 50%. The actual points of subjective equality are shown in table 6.9.

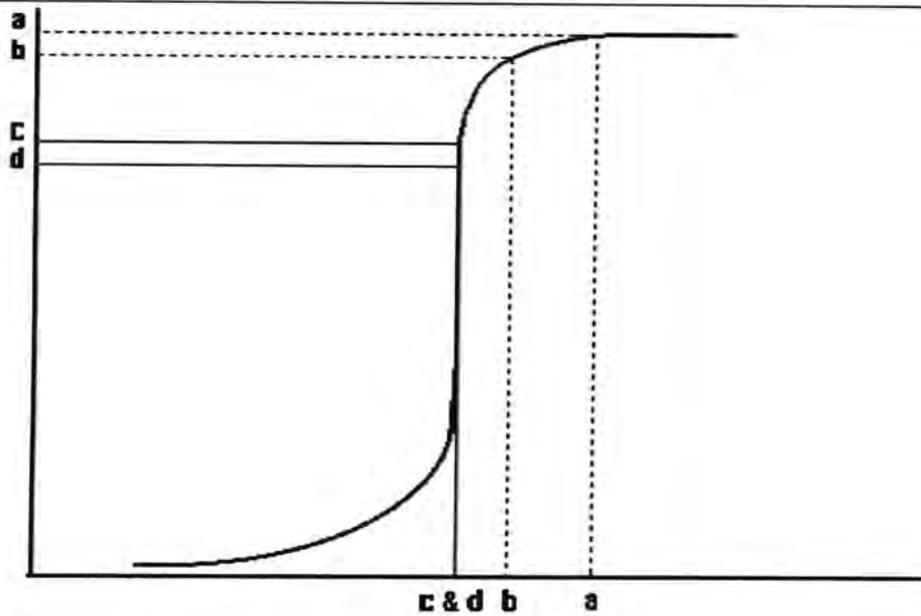


Figure 6.5 : Showing detectability level on the ordinate and contrast level on the abscissa. If detectability level a (ordinate) is chosen as the level for stimulus equality then contrast for the stimulus in this case equals a (abscissa). However, a short shift down the ordinate produces a much larger shift on the abscissa thus suggesting that any small variability in the estimation of the detectability level produces a much larger error in the choice of contrast level. A point further down the psychometric function detectability level c produces contrast level c an equal amount of variability as in the case of a produces no change in the choice of contrast level.

	0.5 cpd	16 cpd
100	4.6	79.1
90	4.8	61.38
50	2.25	73.17

Table 6.8: shows the 90% detectability contrasts (in percent) for the 0.5 and 16 cpd spatial frequencies.

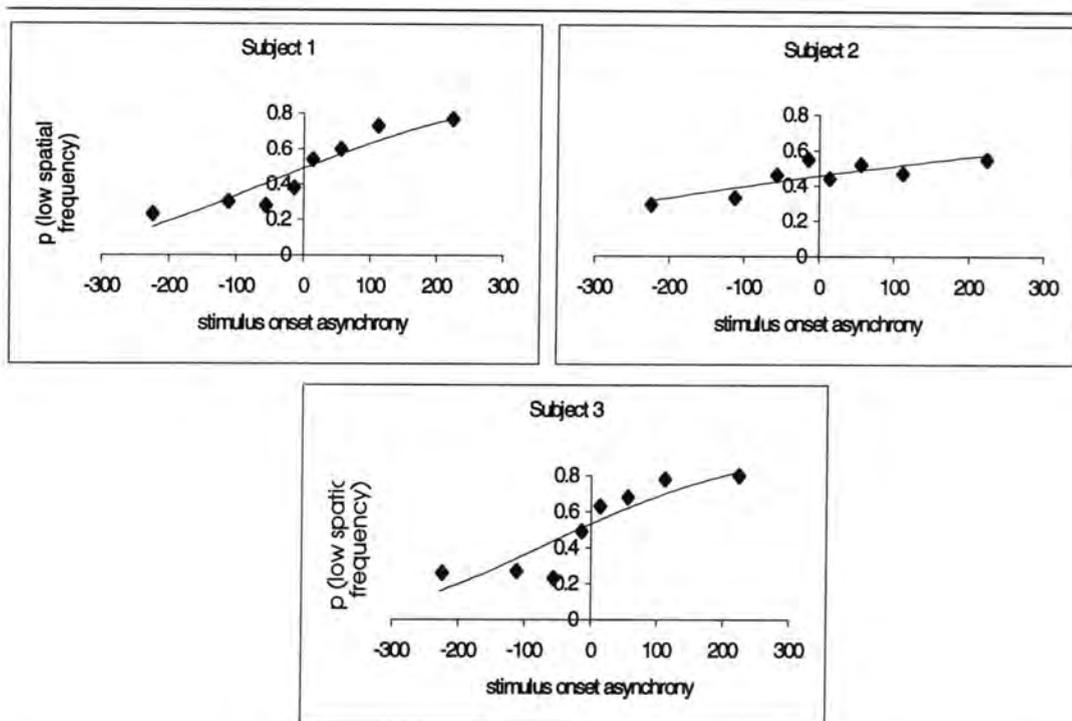


Figure 6.6: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject.

Subject	P.S.E.	Negative Conf	Positive Conf
1	6.69	-74.17	97.18
2	76.2	14.28	189.45
3	-19.37	-114.92	85.56

Table 6.9: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval).

It would be expected from the predicted values of spatial frequency processing delays that the point of perceptual simultaneity would occur on the minus SOA side of zero, at -105msec, that is when the high spatial frequency patch was presented about 105 msec before the low spatial frequency patch. However it is clear that only one subject (subject 3) showed a point of simultaneity on this side of the graph and this was far less than the 105 msec estimate and much less than the conservative estimate of a 25 msec processing lag difference. The remaining two subjects both show a bias to the positive side of the stimulus onset asynchronies, that is when the high spatial frequency was presented after the low spatial frequency. However when the confidence intervals are examined it can be seen that two subjects show no evidence for the expected processing delays, both being significantly different from the expected -105 msec value. The remaining subject's point of simultaneity was found not to be significantly different from the predicted point. Furthermore two of the subjects show that the point at which the two spatial frequencies were simultaneously presented was significantly different from zero. However both were in the opposite direction. Thus one subject's estimate of simultaneity occurred when the high spatial frequency was presented prior to the onset of the low spatial frequency, while the second point of simultaneity was when the 0.5 spatial frequency was presented prior to the 16 cpd spatial frequency.

6.6.3 Discussion

The results found in experiment 17c show again that there is little consistent support for increasing delays in the processing of higher spatial frequencies. Two of the subjects show that the point at which the two spatial frequencies are judged as being simultaneous was significantly different from that expected if the a 21 msec per octave delay is assumed. Only one subject showed a point of simultaneity which included this point. The results found here and those presented in experiment 16 show no consistent

evidence for processing delays in the processing of spatial frequencies. Indeed one subject in this experiment showed a significant difference from zero on the opposite side to that expected.

Overall then the results found here argue against the neural processing lag in the processing of different spatial frequencies. However, the variability amongst subjects is still a concern. In order to attempt to reduce this the form of presentation was changed to a cyclical one which reduced the variability found in experiment 16.

6.7 Experiment 18: Examining the Temporal Order Judgements of equally detectable cycled spatial frequencies

In this experiment the evidence for spatial frequency processing delays was again examined. This experiment was an extension of the temporal order judgement experiment where the stimuli were shown at double the contrast threshold, in this experiment the stimuli were cycled when the temporal order judgement were made. The contrast thresholds were determined for a 0.5 and a 16 cpd Gabor patch presented full screen for 14 msec. This threshold was then used in the second experiment to determine the detectability of the patches. As with experiment 17a contrast level at which both the 0.5 and the 16 cpd Gabor patches were 90% detectable was determined. These contrast levels were then used in the temporal order judgement experiment. In this experiment the stimuli were presented briefly for 14 msec while the SOA between the 0.5 and the 16 cpd were varied from -224 msec to +224 msec. The presentation period cycled every 1500 msec. It was expected that if the spatial frequency delays proposed do exist then in order for the stimuli to be judged as occurring simultaneously the 16 cpd should be

shown substantially prior to the 0.5 cpd patch as its processing time is increased.

Subjects examined the display for as long as they wished.

6.7.1 *General Methods*

In this experiment the temporal order judgements of two equally detectable spatial frequencies was assessed. The contrast thresholds of 0.5 and 16 cpd spatial frequencies were first determined. These thresholds and three other values of contrast were employed in a detectability task where subjects indicated in which of two temporal intervals a Gabor patch was shown. The contrast values for the 0.5 and 16 cpd Gabor patches which were 90% detectable were then employed in the temporal order judgement experiment where either spatial frequency was presented first. The presentation of the two spatial frequencies was cycled every 1500 msec so that the subject could examine the presentation order for as long as they wished. Once the subjects were satisfied they indicated which spatial frequency appeared first. All methods are as experiment 17 except where noted.

Subjects: There were three male subjects, with an age range of 25 to 26. One of the subjects was fully aware of the purpose of the experiments, the remaining two were naive.

6.7.2 *Results*

6.7.2.1 *Experiment 18a: Contrast threshold determination*

Table 6.10 shows the contrast thresholds for the 0.5 and the 16 cpd patch for the three subjects. These values were used to set four levels of contrast to be employed in the determination of the detectability of the stimuli. For the low spatial frequencies, three other contrast levels 0.1 log steps above the contrast threshold were used. For the

16 cpd spatial frequency log steps of 0.15 were used to set three contrast levels above contrast threshold. The actual contrast values employed for each subject are shown in table 6.11.

6.7.2.2 Experiment 18b: Equally detectable

The detectability ratings for 0.5 and 16 cpd when presented at contrast threshold and log steps above are shown in table 6.12. It can be seen, as would be expected, that as contrast is raised so the detectability ratings increase. From this data the contrast values for ensuring that the spatial frequencies of 0.5 cpd and 16 cpd were equally detectable stimuli were estimated using a probit curve fitting procedure (Finney, 1952). From these probit estimates the contrast which produces stimuli detectability of 90% was chosen to represent equally detectable stimuli for use in experiment 18c. These values are shown in table 6.13.

Subject	0.5 cpd	16 cpd
1	3.1	29.9
2	1.90	17.2
3	1.86	24.9

Table 6.10: Contrast thresholds (in percent) of the 0.5 cpd and 16 cpd spatial frequencies for three subjects.

Subject	0.5 cpd				16 cpd			
	Threshold	0.1	0.2	0.3	Threshold	0.15	0.3	0.45
1	3.1	3.9	4.9	6.2	29.9	42.2	59.7	84.3
2	2.3	2.9	3.7	4.7	26.2	37	52.3	73.8
3	1.6	2	2.5	3.2	26.7	37.7	53.3	75.3

Table 6.11: Four contrast values (in percent) for the two spatial frequencies to be used in the temporal order judgements of experiment 18c. The contrast thresholds for both spatial frequencies were used and from this contrast value three other contrast levels were chosen. The contrast increased in log steps of 0.1 or 0.15 for the 0.5 cpd spatial frequency and the 16 cpd spatial frequency respectively. This enabled the detectability of a good spread of contrast values to be assessed.

Subject	0.5 cpd				16 cpd			
	Threshold	0.1	0.2	0.3	Threshold	0.15	0.3	0.45
1	81	89	98	99	60	74	86	96
2	75	78	92	95	61	81	90	98
3	82	94	99	100	64	66	83	96

Table 6.12: the detectability (in percent) of the 0.5 cpd and 16 cpd spatial frequencies as a function of their contrast.

Subject	0.5 cpd	16 cpd
1	3.39	14.83
2	3.99	71.4
3	3.36	96.59

Table 6.13: shows the 90% detectability contrasts (in percent) for the 0.5 and 16 cpd spatial frequencies.

6.7.2.3 *Experiment 18c*: temporal order judgements

The results of the temporal order judgements of the individual subjects are shown in figure 6.7 and table 6.14. The table shows the estimated point at which number of low spatial frequency first responses was 50%. This was taken as representing the point at which the two spatial frequencies of 0.5 and 16 cpd were presented simultaneously. This was estimated by probit curve fitting procedure (Finney, 1952). All subjects show a point of simultaneity which is positive, that is the 0.5 cpd was presented prior to the 16 cpd for the two to be judged as having appeared simultaneously. This is the opposite to the result expected if spatial frequency processing delays played a role. The confidence intervals show that the points of simultaneity were not significantly different from zero but were significantly different from the 105 msec processing delay predicted from reaction time experiments.

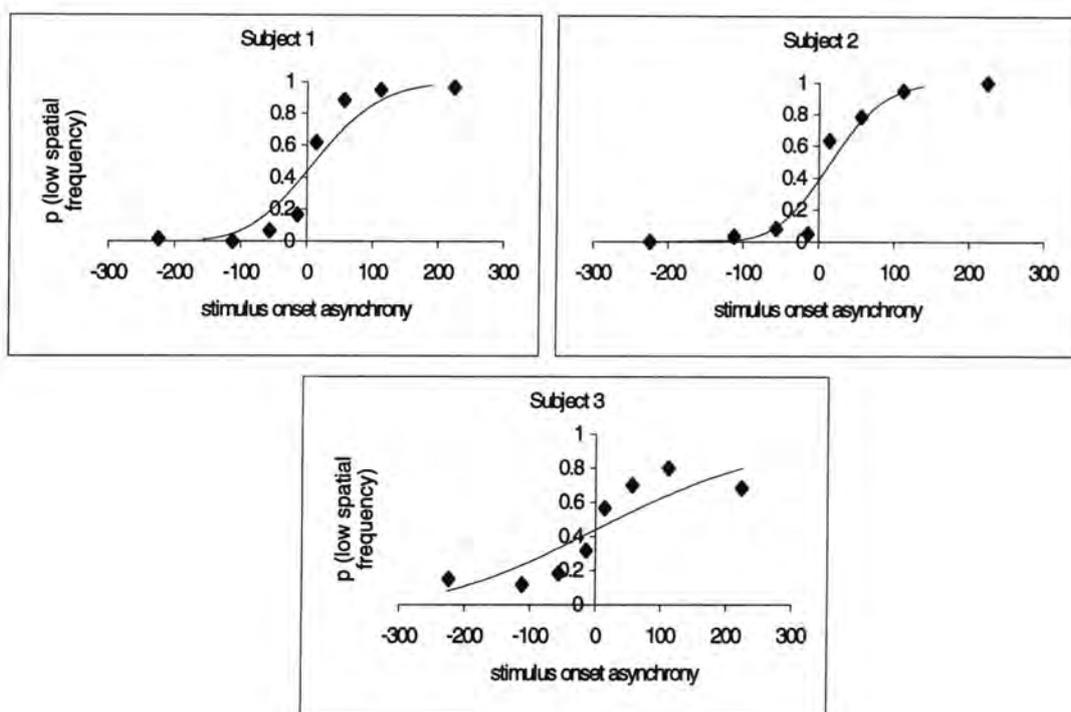


Figure 6.7: shows the proportion of low spatial frequency first responses as a function of stimulus onset asynchrony. The onset of the 0.5 cpd spatial frequency was normalised to zero. Thus a minus value stimulus onset asynchrony indicates that the 16 cpd spatial frequency was shown prior to the lower spatial frequency and a positive value indicates that the 16 cpd spatial frequency was shown after the low spatial frequency. The data points show the subject data while the solid line shows the probit fit for each subject.

Subject	P.S.E.	Negative Conf	Positive Conf
1	12.22	-49.82	78.66
2	15.93	-36	72.12
3	31.27	-78.89	160.02

Table 6.14: shows the stimulus onset asynchrony of the point of subjective equality (P.S.E.), that is the stimulus onset asynchrony at which the two spatial frequencies are estimated as having been shown simultaneously. The remaining two columns show the bounding 95% confidence intervals (negative conf = negative confidence interval; Positive conf = positive confidence interval).

6.7.3 Discussion

Overall then the results show a point of perceptual simultaneity which suggests that the low spatial frequency should be shown prior to the high spatial frequency in order for the two to appear as having been presented simultaneously. This is the opposite pattern of results that would be expected given the delays involved in processing of spatial frequencies. It was predicted that the high spatial frequency would have to be shown 105 msec prior to the low spatial frequency in order for the two to appear as having been presented simultaneously. The confidence intervals associated with each estimated point of simultaneity are significantly different from this 105 msec point and importantly they are not significantly different from zero, the point of physical simultaneity. Thus it can be suggested that the results show no evidence for spatial frequency processing delays of the magnitude predicted from reaction time and VEP experiments. Indeed the results show no support for spatial frequency processing delays at all, as the points of simultaneity found do not significantly differ from the point of physical simultaneity. Thus this experiment finds little evidence for the operation of increasing temporal lags associated with spatial frequencies as they become higher.

6.8 General Discussion

Overall then, the results of experiment 13 and 14 lend no support for the spatial frequency processing delays found by Parker & Dutch (1987). It was suggested that this lack of spatial frequency processing delays may be due to the use of a broadband LED as a matching stimulus. In the remaining experiments the temporal order judgements of two widely spatial frequencies was assessed. In experiment 15 no clear evidence was found for spatial frequency processing delays although some evidence did suggest that

the variability in subjects' performance may be masking more subtle spatial frequency processing delays. In experiment 16 the presentation of the temporal order trials was cycled so that the order of the spatial frequencies could be thoroughly assessed. The results showed no evidence for spatial frequency delays. This pattern was also found in experiment 17 and 18 where the low and high spatial frequencies were made equally detectable.

Therefore the results show little evidence in support of the suggestion that spatial frequency processing delays increase as spatial frequencies become higher. Indeed the results show no support that spatial frequency processing delays occur at all as the estimated point of simultaneity of the presentation of two widely different spatial frequencies was found not to differ from zero. Thus spatial frequencies appeared simultaneously "perceptually" when they appeared simultaneously physically. It could be argued that the variability in the judgement of temporal order are consistent with delays which are not as large in magnitude as that found in reaction time and VEP experiments. However, it remains the case that the overall pattern of the results of experiment 16, 17 and 18 show that the delays are not different from cases when the two spatial frequencies are physically presented at the same time. Thus while the results show variation they do not show any evidence for a delay in the processing of spatial frequencies.

This lack of support for spatial frequency processing delays and the variation amongst subjects suggest that the spatial frequency delays may not be important in restricting any further processing to occur from coarse to fine. That subjects can show variations indicates that delays in spatial frequency processing may be compensated for or are so small that they are negligible. For example, in experiment 18c the point of simultaneity was estimated to occur when the 0.5 cpd spatial frequency was presented prior to the 16 cpd spatial frequency, suggesting that the low spatial frequency takes

longer to process than the high spatial frequency. While it is the case that these estimates were not significantly from a zero stimulus onset asynchrony, they still raise the possibility that spatial frequency processing is not rigidly processing spatial frequency information from coarse to fine.

The results show then that there is no evidence for spatial frequency processing delays. The variability in spatial frequency delays suggests that spatial frequency processing delays are either mutable, negligible or compensated for. When considered in terms of the integration of spatial frequencies and the two stage model outlined in the introduction to this chapter (see also section 1.3.2 *conclusions*) that the integration stage is not restricted to occur in a coarse to fine manner. The mutable, negligible or compensated delays thus allow the integration of spatial frequencies to occur from both coarse to fine or fine to coarse as may be necessary. Thus the findings supporting a fine to coarse integration of spatial frequencies show not only that the integration of spatial frequencies is flexible but also that it is not limited to occur from coarse to fine because of differential delays involved in spatial frequency processing.

Discussion: The temporal nature of the integration of spatial frequencies

There were two main aims underpinning the work described in this thesis. The first aim was to examine the form which the integration of spatial frequencies takes and the second was to examine the delays involved in initial spatial frequency processing. The purpose of this chapter is to examine whether these aims have been fulfilled and to further discuss the implications of the results obtained. In the first section (7.1) an overview of the results of the thesis is given followed by a brief evaluation of the significance of this work in relation to aims just outlined. In the second section future research is discussed. Finally the main conclusions from the thesis are restated.

7.1 Overview of thesis findings

This section will outline the results from the experiments reported in this thesis. First the experiments specifically examining the temporal nature of the integration of spatial frequencies will be reported. Second, a discussion of the results pertaining to spatial frequency processing delays will be presented. Finally a brief overview of a simple model is suggested.

7.1.1 *The temporal nature of spatial frequency integration*

The first twelve experiments reported in this thesis have examined the form and nature of the temporal integration of spatial frequencies. It was suggested in section 1.2.3 that the integration of spatial frequencies may be temporally anisotropic, that is spatial frequencies are most efficiently integrated when they are processed from coarse to fine, mirroring the normal availability of spatial frequencies suggested by the increase in processing times found with higher spatial frequencies (see section 1.3.2). An alternative account was also proposed where the integration of spatial frequencies was flexible, that is spatial frequencies are integrated effectively when they are available in task dependent temporal orders. This difference between temporally anisotropic integration and flexible spatial frequency integration was examined with artificially extended presentation times of spatial frequencies: spatial frequencies were presented sequentially from coarse to fine or fine to coarse. This enforced a particular form of spatial frequency availability on to the visual system from which the interactions of those spatial frequencies could be examined.

Both direct and indirect measures were used to assess the integration of spatial frequencies. The direct measures involved a subjective assessment of the quality of the sequence of spatial frequencies, such as a quality rating of the sequences (experiment 1) or a forced choice, between the two temporal orders, of which appeared to most resemble the fullbandwidth image which would be formed by the summation of the spatial frequencies presented (experiments 6, 7, 8, 9a and 9b). The indirect measures involved either measuring errors of detection (experiments 2, 3 and 4) contrast thresholds (experiment 10) or Vernier acuity thresholds (experiments 11 and 12). These were indirect in that they did not involve a direct judgement of the phenomenal appearance of the two spatial frequency temporal orders.

Summary of Results

In line with previous work examining the integration of spatial frequencies (Parker et al, 1992; 1997) a coarse to fine temporal order of spatial frequencies was found to improve their subsequent integration suggesting that the integration of spatial frequencies is temporally anisotropic. In experiment 1 a coarse to fine temporal order of filtered natural images was rated as being of better quality than a fine to coarse sequence. The same pattern was found in experiment 2, 3 and 4 where a fullbandwidth image was more likely to be mistaken as having been presented in a sequence of filtered natural images if that sequence progressed from coarse to fine, although this was found not to be as strong as had previously been found (Parker, Lishman & Hughes, 1992).

However, the results from these experiments employing natural images were found not to be as clear cut as those found in previous experiments (Parker et al, 1992). It was suggested that this may be due to the use of natural images which may not have isolated the spatial frequency integration process. This was indicated from the results of experiment 4, where it was found that using a scene image and a face image produced a different pattern of performance. It was pointed out that the use of filtered images essentially involved the use of large bandwidths of spatial frequencies within each filtered image which would, due to the intrinsic processing delays associated with spatial frequencies, be processed from coarse to fine¹. Therefore even a fine to coarse presentation of filtered images would involve coarse to fine processing within each of the filtered images. Thus even though the overall spatial frequency content of the images are shifting from fine to coarse, the initial image presented is integrated from coarse to fine and the additional presentation of further filtered images introduces the problem of how the spatial frequency content of these interact with the already integrated spatial frequency content of the first image. Therefore this pattern of processing which occurs for each filtered image suggests that the method of sequentially presenting triplets of

filtered images might examine only a “second order” interaction between the filtered images not the integration of spatial frequencies as such. Therefore it was proposed that the use of filtered images was not ideally suited to the isolation and examination of the temporal nature of the integration of spatial frequencies.

It was further suggested that the results from the two tasks of image detection and image quality judgement differ because the detection task examined spatial frequency integration indirectly while the quality task examined spatial frequency integration more directly. The detection task required that a target was to be discriminated from other distracting filtered images; while the quality task required a judgement based on the overall image triplets. On the basis of this consideration it was suggested that the quality task encouraged the second order interaction between the filtered images to be one of integration; while the detection task actively discouraged this. Thus although the filtered images when considered separately are processed in spatial frequency content from coarse to fine their overall spectral content does vary from coarse to fine or fine to coarse allowing a test of whether the interaction between the filtered images is one of a coarse to fine or a fine to coarse integration of spatial frequency content. The results from the quality judgement task suggests that the interaction across filtered images is one of coarse to fine temporally anisotropic spatial frequency integration. Therefore it was concluded that while the problems involved when employing large bandwidths of spatial frequencies resulted in the task introducing a second order interaction across the bandwidths, the quality judgement task encouraged this interaction to take the form of integration. The results from this experiment (experiment 1) suggest that spatial frequencies are more effectively integrated when they are available from coarse to fine thus showing evidence for a temporal anisotropy. However it was also found that the data were well accounted for by a temporally isotropic integration of spatial frequencies where the contribution made by each filtered

image is a weighted average and the weights depend upon the filtered images' position in the temporal sequence (see section 2.2.2.3).

In order to address these concerns about the use of natural images the integration of spatial frequencies was examined with narrowband spatial frequency stimuli, thus circumventing the problems of second order interactions. The narrowband spatial frequency stimuli took the form of the fundamental and harmonics of a square-wave series. In experiment 5, using the indirect measure of fullbandwidth image detection, a coarse to fine preference was found, that is a coarse to fine sequence of spatial frequencies was more likely to be integrated to form a square-wave than a fine to coarse sequence thus leading to more incorrect "square-wave present" responses being made when the spatial frequencies were presented from coarse to fine. This suggestion that the coarse to fine sequence was mistaken for the fullbandwidth target because it integrated more effectively so that it resembled the square-wave was tested directly in a forced choice experiment (experiment 6) in which a coarse to fine and fine to coarse sequence were presented one after the other. Subjects judged which sequence appeared to resemble the fullbandwidth image more. Contrary to expectation the fine to coarse sequence was judged to be more like the fullbandwidth image than the coarse to fine sequence. This judgement was found not to be based on the final spatial frequency presented alone (experiment 9a and 9b). However the result could be explained by masking between the spatial frequencies themselves (experiment 10). The contrast threshold for the detection of the 3F was measured when it was preceded or succeeded by the F. It was found that the threshold was higher when the 3F was succeeded by the F, that is when the spatial frequencies were shown from fine to coarse. It could be the case then that the fine to coarse sequence of 3F / F was chosen to resemble a square-wave more because the 3F produced a lower activation in this instance than when compared to the 3F presented after the F. This suggests that the square-wave likeness

decision is based on the relative visibility of the 3F. However it was pointed out that the higher contrast threshold of the 3F when presented prior to the F would also be predicted from an integration hypothesis: If the F and 3F are integrated when they are presented from fine to coarse then the 3F would require more contrast in order for it to appear as a separate component. Thus the same pattern of results would be predicted from a masking or integration hypothesis.

In order to assess whether the results indicating a fine to coarse integration of spatial frequencies were due to spatial frequency masking or integration, a Vernier acuity task was introduced (Experiment 11a, 11b and 12). In this task the position of a grating was judged relative to another. The spatial frequencies of these were either F or 3F and were presented alone, from coarse to fine or from fine to coarse. When considered in terms of positional information conveyed by the spatial frequencies Vernier acuity thresholds were found to improve when spatial frequencies were presented from fine to coarse in comparison with the coarse to fine presentation order or with either spatial frequency presented alone (experiment 12). Thus a second situation has been found in which the fine to coarse sequence of spatial frequencies does improve the efficacy of the integration of spatial frequencies. This supports an integration explanation, rather than a masking explanation, for the masking result of experiment 10 and also suggests that the square-wave likeness results of experiments 6, 8, 9a and 9b are also a result of spatial frequency integration.

Discussion of results

Summarising the results so far, both a coarse to fine and a fine to coarse temporal order of spatial frequencies improves the efficacy of spatial frequency integration. The results of experiment 5 suggested that a coarse to fine sequence of spatial frequencies is mistaken for a fullbandwidth target. If the integration of spatial

frequencies is considered as occurring in a fixed fashion this finding for a coarse to fine integration of spatial frequencies cannot be due to that sequence integrating to form a fullbandwidth like representation as the fine to coarse sequence was judged to be more target like in a forced choice paradigm (experiment 6). However, the results of experiments assessing quality judgements employing a natural image (experiment 1 and 7) both show evidence for a coarse to fine integration of spatial frequencies. In particular the results of experiment 7 show that a coarse to fine sequence of filtered natural images is judged to resemble a fullbandwidth natural image more than a fine to coarse sequence in a forced choice paradigm. This coupled with the results of experiment 6 using individual spatial frequencies suggests that the integration of spatial frequencies is a flexible one which can take place from coarse to fine or fine to coarse. There are a number of other experiments which also suggest that spatial frequency information is integrated flexibly as evidence has been found for a spatial frequency integration occurring from coarse to fine (Parker et al, 1992; 1997; Schyns & Oliva, 1994; Watt, 1987) and from fine to coarse (Schyns & Oliva, 1997; Oliva & Schyns, 1997). Therefore both of these temporal forms of spatial frequency integration occur in the visual system. This may reflect a degree of flexibility within the visual system which allows the integration of spatial frequencies to take place in a manner which will best enable the task at hand to be completed successfully.

It is interesting to note that while the flexibility explanation can account for the results of this thesis a model where a weighted average temporally isotropic integration of spatial frequencies takes place can account for a great deal of the data. In this model information presented in each frame shown is weighted in its contribution to the task decision depending upon when it was presented in a temporal sequence (see section 2.2.2.3). Thus the first spatial frequency shown is given a low weighting and the final spatial frequency presented is given a much higher weighting. This weighted average

model was found to account well for the natural image quality data of experiment 1 and can also account for the results found in experiments 7, 8, 9a, 9b and 10. For example, in the square-wave likeness in which two of the square-wave harmonics are used (experiment 8) the model would weight the 3F such that its contribution would be much less when shown prior to the F when compared to when it was shown after the F. Thus the 3F and F's appearance would be judged to be more square-wave like when the contribution made by the 3F was less. In this case that is the 3F / F pair, which indeed was consistently judged to be more square-wave like. However the weighted average temporally isotropic model has difficulty explaining the square-wave detection data and the Vernier acuity results. For example, in the square-wave detection task (experiment 5) it would not predict a coarse to fine integration of spatial frequencies, but would rather predict that the fine to coarse sequence would elicit more false square-wave reports for precisely the same reasons why it predicts a fine to coarse pattern of results in the square-wave likeness experiments. It is the flexible integration of spatial frequencies which best accounts for the overall pattern of results found in this thesis.

The nature of the flexibility involved in the integration of spatial frequencies suggested from the results in this thesis is different to that suggested by Oliva & Schyns (1997). They suggest that the flexible nature of the integration of spatial frequencies found in their experiments is one which involves a high level constraint or compensation on the coarse to fine delay structure imposed on spatial frequencies by the different temporal lags involved in their processing. This high level constraint operates in such a way that the effectiveness of different spatial frequencies for the successful completion of a recognition task compensates for the effect of the initial spatial frequency processing delays. Furthermore, Oliva & Schyns (1997) also provide evidence for a shift in the temporal order of the integration of spatial frequencies using the same task. They showed that a fine to coarse integration of spatial frequencies can take place

despite the presence of a coarse to fine sequence of spatial frequencies, if on the previous trial the natural image had been presented from fine to coarse. This shows a shift in the integration preference using a single task but changing the diagnosticity of information across trials. This suggests that the integration of spatial frequencies is flexible, but only if the information it has previously encountered demand it to be (See section 1.3.3 for a discussion of some problems with the methodology of Oliva & Schyns, 1995; 1997).

However, given the discussion of different models of early visual processing (section 1.2.1) it can be suggested that the results of Oliva & Schyns (1995; 1997) do not reflect the operation of higher level influences on low level visual processing but rather the results may reflect the operation of different temporal forms of spatial frequency integration depending upon the task. As suggested by Wilson & Wilkinson (1997) “in sum, current research suggests that early-visual-channel responses are almost always combined, but in a variety of task-dependent manners.” (p 956). This suggestion is also supported by the results of this thesis.

In the experiments presented in this thesis, the integration of spatial frequencies has been shown to operate effectively with a coarse to fine or fine to coarse sequence of spatial frequencies depending upon the task and the stimuli employed. Therefore the mechanism involved in the integration of spatial frequencies can be suggested to be flexible in response to the demands of the processing necessary to derive different forms of information. Thus the integration of spatial frequencies is not a unitary process taking a common form, but rather the temporal nature of spatial frequency integration differs depending upon the information to be derived from the image. For example, it was pointed out in section 1.2.1 that the models which underlie the processing of motion and texture differ in some aspects of their derivation from the image these differences may extend to the temporal form of their spatial frequency integration. As an illustration of

this, it can be imagined that a coarse to fine integration of spatial frequencies takes place in order to derive a particular motion signal while a fine to coarse integration of spatial frequencies takes place in order to derive particular texture information.

The results of the experiments employing natural images presented in chapter 2 also suggest that there may be a further flexible dimension to the integration of spatial frequencies. It was found in experiment 4 that the type of natural image used had an effect on the errors elicited for the detection of a fullbandwidth image. This suggests that different types of natural images may affect the form that spatial frequency integration takes. Thus the type of stimuli used may also reflect the flexible integration of spatial frequencies. It was argued that these results may be simply due to the invocation of other factors such as recognition mechanisms which may mask the operation of an early spatial frequency integration process. However, when natural image quality is assessed in a rating or forced choice paradigm (experiments 1 and 7) the coarse to fine integration of spatial frequencies is found; but when narrowband spatial frequencies are utilised in this forced choice quality task a fine to coarse integration is found (experiment 6). This suggests that the stimuli employed may affect the form spatial frequency integration takes. This again indicates that the integration of spatial frequencies is dependent upon the information which is to be derived from the image rather than any higher level influence as proposed by Oliva & Schyns (1995; 1997).

Overall the experiments reported in this thesis have found that spatial frequency integration can take place effectively from coarse to fine or fine to coarse. This pattern of integration is not isotropic such that spatial frequencies are integrated as and when they become available, rather the integration of spatial frequencies is primarily determined by the information to be derived from the image.

It is difficult to understand why the integration of spatial frequencies would take different temporal forms when the end product is the same. A possible reason why concerns the nature of incremental representations discussed in Chapter 1 section 1.3.1. In that section the computational advantages for the coarse to fine and the fine to coarse integration of spatial frequencies was outlined. These advantages were time and positional certainty respectively. Thus a coarse to fine integration enabled an incremental representation of positional information to be available very quickly but very inaccurately (Watt, 1987). However, a fine to coarse integration of spatial frequencies allowed greater positional accuracy and also enabled edge detection to take place with greater efficiency (Marr, 1982; Canny, 1986). It may be the case then that the flexible temporal nature of spatial frequency integration reflects trade-offs like these between different types of information. A further possibility is that of Ramachandran's (1990) Utilitarian theory of perception in which it is suggested that rather than using a single algorithm to solve a perceptual processing problem a whole host of strategies and heuristics may be employed. This could indeed be the case for the results found here. Thus the differences in the temporal forms of spatial frequency integration may be due to the existence of multiple strategies which have evolved to deal with different visual processing problems as and when they have occurred.

7.1.2 Spatial frequency processing delays

Whether the integration of spatial frequencies could take place flexibly when spatial frequencies are not presented sequentially as in the previous experiments discussed was addressed in chapter 6. In section 1.3.2 a simple two stage model of spatial frequency integration was outlined where the output of a spatial frequency filtering stage was subjected to a non-linearity. This was followed by an integration processing stage. It was suggested that a coarse to fine integration is supported by the

evidence for delays involved in the processing of spatial frequencies which are introduced in the first stage of filtering and these constrain the second stage of integration to integrate spatial frequencies from coarse to fine. It was pointed out that this did not necessarily mean that the integration stage itself was biased to integrate spatial frequencies more effectively if they are available from coarse to fine. Integration may take this form solely because of the spatial frequency processing delays introduced in the first stage. Therefore it may be the case that spatial frequencies could be integrated from coarse to fine and fine to coarse just as well if the processing delays were countered. The results presented in this thesis (see section 7.1.1) suggest that this is indeed the case: the integration of spatial frequencies does not take place more effectively if presented from coarse to fine but is flexible and dependent upon processing requirements. However, it follows from the consideration of the two stage model that while the integration stage is indeed flexible it has no ecological importance as the first stage processing delays still exist and therefore constrain the second stage of spatial frequency integration to take place from coarse to fine.

This consideration was addressed in Chapter 6. Six experiments were carried out examining the existence of spatial frequency processing delays using temporal order judgement tasks. In experiments 13 and 14, based on the experiments carried out by Parker & Dutch (1987) no evidence for a consistent pattern of processing delays was found. This lack of evidence was suggested to be due to the use of a broadband LED as a matching stimulus to the onset of spatial frequency gratings. In the remaining experiments temporal order judgements to two widely different spatial frequencies was assessed, following Barr (1986). Although experiment 15 showed no consistent pattern of results, experiments 16, 17 and 18 showed no evidence for spatial frequency processing delays when assessed using temporal order judgements of two spatial frequencies. This was found to be the case when spatial frequencies were equally above

contrast threshold or made equally detectable. The results show no support for processing delays increasing when spatial frequency become higher, indeed the delays were found not to be significantly different from the physical point of simultaneity.

These results suggest that spatial frequency processing delays may not restrict the integration of spatial frequencies to occur from coarse to fine and indicate that the integration of spatial frequencies may take place from coarse to fine or from fine to coarse in normal visual processing. In order for this to occur there must be a tolerance in the processing of spatial frequencies such that spatial frequencies are able to be stored for a short period of time regardless of the order in which they become available. They would then be able to be differentially integrated from coarse to fine or fine to coarse. Thus the important finding from these set of results is that the integration of spatial frequencies is not limited to occur from coarse to fine.

7.2 Future avenues of research

Future experimentation is required to fully explore the details of the flexible integration of spatial frequencies. This involves addressing why spatial frequencies would be integrated flexibly and the nature of the integration process. While the results show that the spatial frequencies can be integrated flexibly they do not address why it should be flexible. It has been suggested that this may be a result of an optimisation process where the nature of spatial frequency integration depends upon the task and stimuli involved or this may be a reflection of the utilitarian theory of perception where the results reflect the use of several different integration algorithms rather than the operation of a single mechanism. Further work examining the nature of the integration of spatial frequencies will help form a more concrete explanation. The nature of the integration process may also be involved in this determination of why spatial

frequencies are integrated flexibly. The results do not show whether there are any principles behind the flexible integration of spatial frequencies. The determination of these principles would not only provide further information about the operation of spatial frequency integration but may also provide answers to the question of why spatial frequencies are integrated flexibly.

A further interesting and more specific aspect of the integration of spatial frequencies is the determination of the bandwidths of these effects, both phase and spatial frequency. The results of experiment 9b, where subjects decided that a 3F / F pair in peaks subtract phase appeared to be more square-wave like than a 3F / F pair in peaks add phase, shows a phase sensitivity to spatial frequency integration. Thus it would be interesting to examine the phase bandwidth. It may be the case that, as with other experiments that phase information is not coded at contrast threshold levels but once sufficiently above threshold sensitivity to phase is found (Nachmias & Weber, 1975). The Vernier acuity task would allow the determination of phase bandwidths and also spatial frequency bandwidths.

It is also important that other explanations for the data found be examined. For example, an alternative interpretation of the detection results is one of masking, where the F component, although easily discriminable from a square-wave, when present prior to a 3F and 5F becomes sufficiently square-wave like, due to the backward masking properties of the higher components, to result in a larger number of square-wave present responses than that found in the fine to coarse sequence. Given the results of Parker et al (1992; 1997) and Schyns & Oliva (1994; 1997; Oliva & Schyns, 1997) this interpretation seems unlikely. However, if it were the case that the results are a result of masking this would lead to the questioning of the methodology employed in the detection tasks for examining the coarse to fine integration of spatial frequencies as the results are a function of spatial frequency presentation and thus of the nature of the

stimuli alone not the underlying mechanisms of the visual system. This would also lend support to the weighted average temporally isotropic model of spatial frequency integration. If the results of the square-wave detection experiment are not due to the temporally anisotropic integration of spatial frequencies then it removes one of the biggest objections to the weighted average temporally isotropic model of spatial frequency integration.

Throughout the thesis a number of problems with the Parker, Lishman & Hughes (1992) detection paradigm have been identified. First, there are problems with interpreting the findings when a fullbandwidth image was present in the triplet sequence and there are problems with interpreting the sequences when a fullbandwidth image was not present because of the overlaps in spatial frequency content between the filtered images which may result in secondary interactions between the overlapping spatial frequency content. Even when employing bandlimited stimuli in the square-wave detection experiment (experiment 5) there was still the possibility that the results were due to the bias of the subjects. Second, it was not entirely clear what the subject was actually doing when performing these detection tasks: were they attempting to detect the presence of the target within the triplet sequences or were they attempting to identify sequences where the target was present as well as the other spatial frequency content of either the filtered images or the other harmonics of the square-wave. For these reasons this area of the thesis requires further empirical research. A detection paradigm is required which allows for a more concrete interpretation of results to be made while using stimuli which do not overlap in spatial frequency content and one in which the task carried out by the subject matches what the subject is actually doing. One such possibility takes the form of a two interval forced choice signal detection task in which only the target and low pass and high pass versions of the target are employed. In this task the subject identifies which interval contained the fullbandwidth target. In one

interval the low pass and high pass images are presented from coarse to fine or fine to coarse while in the other interval low pass and high pass images are presented from coarse to fine or fine to coarse and a fullbandwidth image target accompanies these. The contrast energy of the fullbandwidth image can be manipulated so that its presence is either very weak or very strong. An example of a trial would be as follows: the low pass and high pass images are shown in the first interval from coarse to fine, in the second interval the low pass image and a low contrast energy version of the fullbandwidth image is presented followed by a high pass image with the same low contrast energy fullbandwidth image present. By using this paradigm it is possible to determine which of the coarse to fine and fine to coarse sequences is more likely to be mistaken for the target regardless of bias. This design addresses all of the problems associated with the Parker, Lishman & Hughes (1992) paradigm and allows any temporal anisotropy in the integration of spatial frequencies to be identified.

It was shown in chapter 2 experiment 4 that the use of different stimuli such as faces and scenes produced different results in an image detection paradigm. The reasons why this was so is not clear and deserves further examination. For example it would be informative to assess whether the differences exist using an image quality paradigm varying the type of stimuli and the power spectra separately.

A further avenue of interest is the examination of the idea that the fine to coarse integration of spatial frequencies could be explained in terms of the overlap which occurs neurally between the components if the spatial frequency processing delays are in place (See section 3.3.3 and 1.3.2 discussion of spatial frequency processing delays). Although the results of experiments 1, 5 and 7 suggest that this is not the case the role of neural overlaps could be examined using the square-wave likeness paradigm and different onsets of the F, 3F and 5F thus shifting the amount by which they overlapped and also the direction of the availability from coarse to fine, fine to coarse and mixed

conditions. This would enable the assessment of whether neural overlaps, directionality or both are important in a fine to coarse integration of spatial frequencies.

A great deal of research could also be carried out further examining the temporal delays involved in the processing of spatial frequencies. A large amount of variability was found in the experiments of chapter 6 which may reflect the difficulty of the temporal order judgement task. Thus a determination of spatial frequency delays across a number of spatial frequencies is necessary in temporal order judgement experiments, VEP recordings or fMRI studies. These measures would provide a measure of the perceptual delays involved in spatial frequency processing.

7.3 Conclusion

In conclusion then the results of this thesis are two-fold. They provide evidence which suggests that the integration of spatial frequencies is flexible and dependent upon the information to be derived by any particular processing, that is the goal dictates the temporal form of spatial frequency integration, and possibly the stimuli involved. This goes beyond previous evidence in that this has been found with narrowband Gabor patches and with natural images. It is suggested that the variability found across models of early visual processing, in which the nature of the processing varies with the information which that model derives, is also reflected in the different temporal forms of the integration of spatial frequencies. The experiments in this thesis also indicated that this flexibility may not just operate in situations where the presentation of the high spatial frequency information is manipulated so that it appears prior to the low spatial frequencies. The lack of evidence for spatial frequency processing delays suggest that a flexible integration of spatial frequencies may be possible when spatial frequencies are presented simultaneously.

Notes

Chapter 1:

1. The terms coarse and fine are used purely from convention and to link with common usage in the literature. Their meaning is synonymous with low and high spatial frequencies and are generally used to indicate the spatial domain. They are used here to indicate low and high spatial frequencies as the experiments described here are carried out with stimuli defined in the frequency domain
2. As an aside it should be noted that the concept of a coarse to fine matching strategy or one where information at one scale constrains the processing carried out at another scale is different to the that involved in the integration of spatial frequencies. It was suggested in section 1.2.2 that when the integration of spatial frequencies was discussed this involved the combining of spatial frequency information across scale, it did not presuppose the mathematical operation involved in this combination. The conception of spatial frequency integration which is implied in the approach of Canny (1986) is one where spatial frequency information held at one scale constrains the spatial frequency processing carried out at other spatial frequencies. This suggests that the information held in one scale constrains where in the next scale processing will first occur. It can be conceived as a weighting function where, in the case of edge detection, the finding of an edge in one scale weights where in the next scale processing will occur to check for a corresponding edge. It can be seen that this constraining process does not necessarily

involve an integration of spatial frequency processing. Thus evidence for a constraining process does not necessarily support the integration of spatial frequencies.

3. While these results are suggestive of a coarse to fine integration of spatial frequencies they can be seen to be an example of the coarse to fine constraining of spatial frequency processing as pointed out in section 1.2.2. Thus the results do not directly address the nature of spatial frequency integration but they only highlight possible advantages for processing spatial frequency information from coarse to fine and are included only as an illustration of this.

4. The shift in decision process may reflect the subjects' use of different strategies dictated by the first phase which showed a filtered image in either the low or high spatial frequencies. This strategy may, for example, involve using the low pass images only i.e. "large blobs are sufficient to complete the task", and there is no need to modify this strategy for the second phase as it is still true. The argument could be applied to the use of the high pass image only. This would be possible if the two images were represented separately if, for example, the integration across scale collapsed at the 4 cpd point due to lack of zero-crossing matches (Marr, 1982; Canny, 1986). For example, Marr (1982) suggests that the image tokens in the low spatial frequencies are superseded by those in the higher spatial frequencies and if the low and high spatial frequency information do not match then they are represented at both levels individually. That the images are separately represented is suggested from the results of Oliva & Schyns (1997).

Chapter 2:

1. The natural image used in experiments 1, 2 and 3 presented in this chapter were taken from the PICS image database at the University of Stirling. This is available as an internet resource at the following address: <http://pics.psych.stir.ac.uk/index.html>.

2. It has been suggested that when discussing form perception the correct scale on which to describe image spatial frequency content is not in terms of cycles per degree but rather in terms of the image itself, that is the distal scale. In terms of face perception this is generally cycles per face width (Riley & Costall, 1980) and in object recognition this is specified in cycles per object width (Norman & Ehrlich, 1987; Parish & Sperling, 1991). Although there is strong evidence for this in the face processing literature it is not clear whether this is justified when all objects are considered. Norman & Ehrlich (1987) examined this question in an object identification task by varying the viewing distance for one group of subjects. Thus if the distal scale is the correct one then no difference in the performance would be found. They showed that changes in performance found when viewing distance was changed was not entirely consistent with a proximal scale or a distal scale. They concluded that both scales played a role in object identification. As there is no conclusive evidence has been found either way the proximal scale has been used throughout this chapter.

3. This size was not actually reported in Parker et al (1992). The size was ascertained from Parker (1997) through personal communication

4. The natural images used in experiment 4 presented in this paper were taken from the PICS image database at the University of Stirling (See note 1 for this chapter). The face

stimuli are taken from a large set of faces compiled by the University of Aberdeen. They are also available from the PICS image database at the University of Stirling.

Chapter 7:

1. Although the results of chapter 6 examining spatial frequency processing delays suggests that this pattern of processing need not take place, the point that spatial frequencies are initially integrated within each filtered image prior to integration across filtered images still stands. Therefore, although the example of coarse to fine processing is used in the following discussion of this point, it is not a integral to the argument but rather is used to make the point clearer.

References

Adelson, E. H. & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2, 284-299.

Allport, A. (1993). Attention and Control: Have we been asking the wrong questions - A critical review of the last twenty five years. In D. E. Meyer & S. Kornblum (Eds.), *Attention and Performance XIV: A silver Jubilee* (Vol. XIV,). Cambridge, Mass.: M.I.T.

Anderson, S. J. (1993). Visual processing delays alter the perceived spatial form of moving gratings. *Vision Research*, 33, 2733-2746.

Asada, H. & Brady, M. (1986). The curvature primal sketch. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8, 2 - 14.

Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing. *Network*, 3, 213 - 251.

Badcock, J. C., Whitworth, F. A., Badcock, D. R. & Lovegrove, W. J. (1990). Low frequency filtering and the processing of local - global stimuli. *Perception*, 19, 617 - 629.

Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 217-234).

Cambridge, M. A.: M.I.T. Press and John Wiley & Sons, Inc.

Barlow, H. B. (1989). Unsupervised learning. *Neural Computation*, 1, 295 - 311.

Barr, J. M. (1986). *A study of visual detection latency using spatial frequency stimuli*.

Unpublished D. Phil., University of Oxford.

Barrett, B. T. & Whitaker, D. (1998). Vernier acuity measured with compound gratings.

Investigative Ophthalmology and Visual Science Supplement, 39, S622.

Bell, A. J. & Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, 37, 3327 - 3338.

Bergen, J. R. & Landy, M. S. (1991). Computational modelling of visual texture segregation. In M. Landy & J. A. Movshon (Eds.), *Computational models of Visual Processing* (pp. 253 - 271). Cambridge, M. A.: MIT Press.

Blakemore, C. & Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology (London)*, 203, 237-260.

-
- Blakemore, C., Nachmias, J. & Sutton, P. (1970). The perceived spatial frequency shift: evidence for frequency-selective neurones in the human brain. *Journal of Physiology (London)*, 210, 727-750.
- Blakemore, C. & Sutton, P. (1969). Size adaptation: A new after-effect. *Science*, 166, 245-247.
- Boer, L. C. & Keuss, P. J. G. (1982). Global precedence as a postperceptual effect: An analysis of speed-accuracy tradeoff functions. *Perception & Psychophysics*, 31, 358 - 366.
- Bowne, S. F. (1990). Contrast discrimination cannot explain spatial frequency, orientation or temporal frequency discrimination. *Vision Research*, 30, 449-462.
- Bracewell, R. N. (1965). *The Fourier transform and its applications*. New York: McGraw-Hill.
- Bradley, A. & Freeman, R. D. (1985). Is reduced Vernier acuity in amblyopia due to position, contrast or fixation deficits? *Vision Research*, 25, 55-66.
- Bradley, A. & Skottun, B. C. (1987). Effects of contrast and spatial frequency on Vernier acuity. *Vision Research*, 27, 1817-1824.
- Breitmeyer, B. (1984). *Visual Masking: An integrative approach*. Oxford: Clarendon.

Breitmeyer, B., Levi, D. M. & Harwerth, R., S. (1981). Flicker masking in spatial vision. *Vision Research*, 21, 1377-1386.

Breitmeyer, B. G. (1975). Simple reaction time as a measure of the temporal response properties of transient and sustained channels. *Vision Research*, 15, 1411-1412.

Breitmeyer, B. G. & Ganz, L. (1977b). Temporal studies with flashed gratings: Inference about human transient and sustained channels. *Vision Research*, 17, 861-865.

Breitmeyer, B. G. & Ganz, L. (1977a). Temporal studies with flashed gratings: Inference about human transient and sustained channels. *Vision Research*, 17, 861-865.

Breitmeyer, B. G. & Julesz, B. (1975). The role of on and off transients in determining the psychophysical spatial frequency response. *Vision Research*, 15, 411-415.

Burbeck, C. A. (1987). Position and spatial frequency in large-scale localization judgements. *Vision Research*, 27, 417-428.

Burbeck, C. A. & Yap, Y. L. (1990a). Spatial Filter selection in large scale spatial interval discrimination. *Vision Research*, 30, 263 - 272.

Burbeck, C. A. & Yap, Y. L. (1990b). Spatial-filter selection in large-scale spatial-interval discrimination. *Vision Research*, 30, 263-272.

Burr, D. C. (1980). Sensitivity to spatial phase. *Vision Research*, 20, 391-396.

-
- Burt, P. J. & Adelson, E. H. (1983). The laplacian pyramid as a compact code. *IEEE Transactions on Communications*, 31, 532 - 540.
- Burton, G. J. (1973). Evidence for non-linear response processes in the human visual system from measurements on the thresholds of spatial beat frequencies. *Vision Research*, 13, 1211-1225.
- Burton, G. J. & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, 26, 157 - 170.
- Campbell, F. W. & Gubisch, R. W. (1966). Optical quality of the human eye. *Journal of Physiology (London)*, 186, 558-578.
- Campbell, F. W. & Kulikowski, J. J. (1966). Orientational selectivity of the human visual system. *Journal of Physiology (London)*, 187, 437-446.
- Campbell, F. W. & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology (London)*, 197, 551-566.
- Cannon, M. W. & Fullenkamp, S. C. (1991). A transducer model for contrast perception. *Vision Research*, 31, 983-989.
- Canny, J. (1986). A computational approach to edge detection. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 8, 679-698.

Carter, B. E. & Henning, G. B. (1971). The detection of gratings in narrow-band visual noise. *Journal of Physiology (London)*, 219, 355-365.

Cavanagh, P. & Mather, G. (1989). Motion: the long and short of it. *Spatial Vision*, 4, 103-129.

Christman, S. D. (1993). Local-global processing in the upper versus lower visual fields. *Bulletin of the Psychonomic Society*, 31, 275 - 278.

Chubb, C. & Sperling, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America A*, 5, 1986-2006.

Chubb, C. & Sperling, G. (1989). Two motion perception mechanisms revealed through distance driven reversal of apparent motion. *Proceedings of the National Academy of Sciences U.S.A.*, 86, 2985-2989.

Chubb, C. & Sperling, G. (1991). Texture quilts: Basic tools for studying texture from motion. *Journal of Mathematical Psychology*, 35, 411-442.

Cleland, B. G., Dubin, M. W. & Levick, W. R. (1971). Sustained and transient neurones in the cat's retina and lateral geniculate nucleus. *Journal of Physiology (London)*, 217, 473-496.

Cornsweet, T. N. (1962). The staircase-method in psychophysics. *American Journal of Psychology*, 75, 485-491.

-
- Daugman, J. D. (1980). Two-dimensional spectral analysis of cortical receptive field profiles. *Vision Research*, 20, 847-856.
- Daugman, J. G. (1993). Quadrature-phase simple-cell pairs are appropriately described in complex analytic form. *Journal of the Optical Society of America A*, 10, 375 - 377.
- DeValois, K. K. (1977). Spatial frequency adaptation can enhance contrast sensitivity. *Vision Research*, 17, 1057-1065.
- DeValois, R. L., Albrecht, D. G. & Thorell, L. G. (1982). Spatial frequency selectivity of cells in macaque visual cortex. *Vision Research*, 22, 545-560.
- DeValois, R. L. & DeValois, K. K. (1980). Spatial Vision. *Annual Review of Psychology*, 31, 309-341.
- DeValois, R. L. & DeValois, K. K. (1988). *Spatial vision*. New York: Oxford University Press.
- Ejima, Y. & Ohtani, Y. (1987). Simple reaction time to sinusoidal grating and perceptual integration time: contributions of perceptual and response processes. *Vision Research*, 27, 269-276.
- Enroth-Cugell, C. & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *Journal of Physiology (London)*, 187, 517-552.

-
- Felipe, A., Buades, M. J. & Artigas, J. M. (1993). Influence of the contrast sensitivity function on the reaction time. *Vision Research*, 33, 2461 - 2466.
- Felleman, D. J. & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex*, 1, 1-47.
- Ferrero, M. & Foster, D. H. (1986). Discrete and continuous modes of curved-line discrimination controlled by effective stimulus duration. *Spatial Vision*, 1, 219 - 230.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4, 2379-2394.
- Field, D. J. (1993). Scale invariance and self similar "wavelet" transforms: an analysis of natural scenes and mammalian visual systems. In M. Farge, J. C. R. Hunt & J. C. Vassilicos (Eds.), *Wavelets, fractals and fourier transforms* (pp. 151 - 193). Oxford: Clarendon Press.
- Field, D. J. (1994). What is the goal of sensory processing. *Neural Computation*, 6, 559 - 601.
- Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local association field. *Vision Research*, 33, 173 - 193.
- Findlay, J. M. (1973). Feature detectors and visual acuity. *Nature*, 241, 135-137.

Fink, G. R., Halligan, P. W., Marshall, J. C., Frith, C. D., Frackowiak, R. S. J. & Dolan, R. J. (1996). Where in the brain does visual attention select the forest and the trees. *Nature*, 382, 626 - 628.

Finney, D. J. (1952). *Probit analysis*. (3rd ed.). Cambridge: Cambridge University Press.

Fiorentini, A., Maffei, L. & Sandini, G. (1983). The role of high spatial frequencies in face perception. *Perception*, 12, 195 - 201.

Flavell, J. H. & Draguns, J. (1957). A microgenetic approach to perception and thought. *Psychological Bulletin*, 54, 197 - 217.

Foley, J. M. (1994). Human luminance pattern-vision mechanisms - masking experiments require a new model. *Journal of the Optical Society of America A*, 11, 1710-1719.

Foley, J. M. & Yang, Y. (1991). Forward pattern masking: effects of spatial frequency and contrast. *Journal of the Optical Society of America A*, 8, 2026 - 2037.

Gabor, D. (1946). Theory of Communication. *Journal of the Institute of Electrical Engineers (London)*, 93, 429 - 457.

Georgeson, M. (1991). Over the limit: Encoding contrast above threshold in human vision. In J. J. Kulikowski, V. Walsh & I. J. Murray (Eds.), *Limits of Vision* (Vol. 5,). Boca Raton, Fl.: CRC Press.

-
- Georgeson, M. A. (1987). Temporal properties of spatial contrast vision. *Vision Research*, 27, 765 - 780.
- Georgeson, M. A. (1992). Human Vision Combines Oriented Filters to Compute Edges. *Proceedings of the Royal Society of London B* 249, 235 - 245.
- Georgeson, M. A. (1994). From filters to features: location, orientation, contrast and blur, *Higher-order processing in the visual system* . Chichester: John Wiley.
- Georgeson, M. A. & Georgeson, J. M. (1987). Facilitation and masking of briefly presented gratings: time course and contrast dependence. *Vision Research*, 27, 369-380.
- Georgeson, M. A. & Meese, T. S. (1997). Perception of stationary plaids: The role of spatial filters in edge analysis. *Vision Research*, 23, 3255 - 3271.
- Georgeson, M. A., Schofield, A. J., Guest, S. J. & Anderson, R. (1998). Detection and representation of variations in luminance and contrast: separate streams in spatial vision? *Perception*, 27, 38.
- Georgeson, M. A. & Shackleton, T. M. (1994). Perceived contrast of gratings and plaids: non - linear summation across oriented filters. *Vision Research*, 34, 1061 - 1075.
- Georgeson, M. A. & Sullivan, G. D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. *Journal of Physiology (London)*, 253, 627-656.

-
- Gibson, J. J. (1950). *The perception of the visual world*. Westport, CT: Greenwood Press.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton - Mifflin.
- Gilinsky, A. S. (1968). Orientation-specific effects of patterns of adapting light on visual acuity. *Journal of the Optical Society of America*, 58, 13-18.
- Ginsburg, A. P., Cannon, M. W. & Nelson, M. A. (1980). Suprathreshold processing of complex visual stimuli: Evidence for linearity in contrast perception. *Science*, 208, 619-621.
- Gish, K., Shulman, G. L., Sheehy, J. B. & Leibowitz, H. W. (1986). Reaction times to different spatial frequencies as a function of detectability. *Vision Research*, 26, 745-748.
- Glass, L. (1969). Moire effect from random dots. *Nature*, 223, 578 – 580.
- Grabowski, A. & Nowicka, A. (1996). Visual-spatial-frequency model of cerebral asymmetry: A critical survey of behavioural and electrophysiological studies. *Psychological Bulletin*, 3, 434 - 449.
- Graham, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Graham, N. & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channel models. *Vision Research*, 11, 251-259.

Graham, N. V. S. (1980). Spatial frequency channels in human vision: Detecting edges without edge detectors. In C. S. Harris (Ed.), *Visual coding and adaptability* (pp. 215 - 262). Hillsdale, New Jersey: Lawrence Erlbaum Associates Ltd.

Graham, N. V. S. & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (non - Fourier) texture channels. *Vision Research*, 38, 231 - 257.

Green, M. (1981a). Psychophysical relationships among mechanisms sensitive to pattern, motion, and flicker. *Vision Research*, 21, 971-983.

Green, M. (1981b). Spatial frequency effects in masking by light. *Vision Research*, 21, 861-866.

Greenlee, M. W. & Magnussen, S. (1988). Spatial waveform discrimination following higher-harmonic adaptation. *Journal of the Optical Society of America A*, 5, 1744-1749.

Grice, G. R., Canham, L. & Burroughs, J. M. (1983). Forest before trees? it depends where you look. *Perception & Psychophysics*, 33, 121 - 128.

Hancock, P. J. B., Baddeley, R. J. & Smith, L. S. (1992). The principle components of natural images. *Network*, 3, 61 - 70.

-
- Harmon, L. & Julesz, B. (1973). Masking in two-dimensional recognition: effects of two-dimensional filtered noise. *Science*, *180*, 1194-1197.
- Harwerth, R. S. & Levi, D. M. (1978). Reaction time as a measure of suprathreshold grating detection. *Vision Research*, *18*, 1579-1586.
- Henning, G. B., Hertz, B. G. & Broadbent, D. E. (1975). Some experiments bearing on the hypothesis that the visual system analyzes patterns in independent bands of spatial frequency. *Vision Research*, *15*, 887-899.
- Hering, E. (1920). *Outline of a Theory of the Light Sense* [English translation by L. Hurvich and D. Jameson, 1964]. Cambridge, Massachusetts: Harvard University Press.
- Hess, R. F. & Dakin, S. C. (1997). Absence of contour linking in peripheral vision. *Nature*, *390*, 602 - 604.
- Hoffman, J. E. (1980). Interaction between global and local levels of a form. *Journal of Experimental Psychology: Human Perception and Performance*, *6*, 222 - 234.
- Hughes, H. C. (1986). Asymmetric interference between components of Suprathreshold compound gratings. *Perception & Psychophysics*, *40*, 241 - 250.
- Hughes, H. C., Fendrich, R. & Reuter-Lorenz, P. A. (1990). Global versus local processing in the absence of low spatial frequencies. *Journal of Cognitive Neuroscience*, *2*, 272 - 282.

Hughes, H. C., Nozawa, G. & Kitterle, F. (1996). Global precedence, Spatial frequency channels and the statistics of natural images. *Journal of Cognitive Neuroscience*, 8, 197 - 230.

Jamar, J. H. T. & Koenderink, J. J. (1985). Contrast detection and detection of contrast modulation for noise gratings. *Vision Research*, 25, 511-522.

Jones, R. & Keck, M. J. (1978). Visual evoked response as a function of grating spatial frequency. *Investigative Ophthalmology and Visual Science*, 17, 652 - 659.

Julesz, B. (1960). Binocular depth perception of computer generated patterns. *Bell System Technical Journal*, 39, 1125-1162.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

Julesz, B. & Miller, J. E. (1975). Independent spatial-frequency-tuned channels in binocular fusion and rivalry. *Perception*, 4, 125-143.

Kaplan, E., Lee, B. B. & Shapley, R. M. (1990). New views of primate retinal function. *Progress in Retinal Research*, 9, 273-336.

Kinchla, R. A. (1974). Detecting target elements in multi-element arrays: A confusability model. *Perception & Psychophysics*, 15, 149 - 158.

-
- Kinchla, R. A. & Wolfe, J. M. (1979). The order of visual processing: Top - down, bottom - up or middle - out. *Perception & Psychophysics*, 25, 225 - 231.
- King-Smith, P. E. & Kulikowski, J. J. (1975). The detection of gratings by independent activation of line detectors. *Journal of Physiology (London)*, 247, 237- 271.
- Kitterle, F. L., Christman, S. & Conesa, J. (1993). Hemispheric differences in the interference among components of compound gratings. *Perception & Psychophysics*, 54, 785 - 793.
- Klein, S. & Stromeyer, C. F., III. (1980). On inhibition between spatial frequency channels: adaptation to complex gratings. *Vision Research*, 20, 459-466.
- Klein, S. A. & Buetter, B. (1992). Minimizing and maximizing the joint space - spatial frequency uncertainty of Gabor - like functions: comment. *Journal of the Optical Society of America A*, 9, 337 - 340.
- Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics*, 50, 363 - 370.
- Kulikowski, J. J. (1976). Effective contrast constancy and linearity of contrast sensation. *Vision Research*, 16, 1419 - 1431.
- Kulikowski, J. J. (1977). Visual evoked potentials as a measure of visibility. In J. E. Desmedt (Ed.), *Visual Evoked Potentials in Man: New Developments* (pp. 168 - 183). Oxford: Clarendon Press.

-
- Kulikowski, J. J. (1991). On the nature of visual evoked potentials, unit responses and psychophysics. In A. Valberg & B. B. Lee (Eds.), *From Pigments to Perception* (pp. 197 - 208). New York: Plenum Press.
- Kulikowski, J. J. & King-Smith, P. E. (1973). Spatial arrangement of line, edge and grating detectors revealed by subthreshold summation. *Vision Research*, *13*, 1455-1478.
- Kulikowski, J. J. & Tolhurst, D. J. (1973). Psychophysical evidence for sustained and transient detectors in human vision. *Journal of Physiology (London)*, *232*, 149-162.
- LaGasse, L. L. (1993). Effect of good form and spatial frequency on global precedence. *Perception & Psychophysics*, *53*, 89 - 105.
- Lamb, M. R. & Robertson, L. C. (1988). The processing of hierarchical stimuli: Effects of retinal locus, locational uncertainty and stimulus identity. *Perception & Psychophysics*, *44*, 172 - 181.
- Lamb, M. R. & Robertson, L. C. (1989). Do response time advantage and interference reflect the order of processing of global and local level information? *Perception & Psychophysics*, *46*, 254 - 258.
- Lamb, M. R. & Robertson, L. C. (1990). The effect of visual angle on global and local reaction times depends on the set of visual angles presented. *Perception & Psychophysics*, *47*, 489 - 496.

-
- Lamb, M. R. & Yund, E. W. (1993). The role of spatial frequency in the processing of hierarchically organized stimuli. *Perception & Psychophysics*, *54*, 773 - 784.
- Lamb, M. R. & Yund, E. W. (1996a). Spatial frequency and attention: Effects of level-, target-, and location - repetition on the processing of global and local forms. *Perception & Psychophysics*, *58*, 363 - 373.
- Lamb, M. R. & Yund, E. W. (1996b). Spatial frequency and interference between global and local levels of structure. *Visual Cognition*, *3*, 193 - 219.
- Lawden, M. C. (1983). An investigation of the ability of the human visual system to encode spatial phase relationships. *Vision Research*, *23*, 1451-1464.
- Legge, G. E. (1978). Sustained and transient mechanisms in human vision: temporal and sustained properties. *Vision Research*, *18*, 69 - 81.
- Legge, G. E. & Foley, J. M. (1980). Contrast masking in human vision. *Journal of the Optical Society of America*, *70*, 1458-1471.
- Lennie, P. (1993). Roles of M and P pathways. In R. M. Shapley & D. K.-L. Lam (Eds.), *Contrast Sensitivity* (pp. 201 - 213). Cambridge, M. A.: MIT Press.
- Lennie, P., Tervarthen, C., Van Essen, D. & Wässle, H. (1990). Parallel processing of visual information. In L. Spillmann & J. Werner (Eds.), *Visual Perception: The Neurophysiological Foundations* (pp. 103-128). New York: Academic Press.

Lindeberg, T. (1993). Detecting salient blob-like image structures and their scales with a scale-space primal sketch: A method for focus-of-attention. *International Journal of Computer Vision*, *11*, 283 - 318.

Lovegrove, W. J., Lehmkuhle, S., Baro, J. A. & Garzia, R. (1991). The effects of uniform field flicker and blurring on the global precedence effect. *Bulletin of the Psychonomic Society*, *29*, 289 - 291.

Lueschow, A., Miller, E. K. & Desimone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex*, *4*, 523 - 531.

Lupp, U., Hauske, G. & Wolf, W. (1976). Perceptual latencies to sinusoidal gratings. *Vision Research*, *16*, 969 - 972.

Malik, J. (1994). General Discussion. In J. Goode (Ed.), *Higher-order processing in the visual system* (pp. 121). New York: John Wiley.

Malik, J. & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, *7*, 923 - 932.

Marcelja, S. (1980). Mathematical description of the responses of simple cortical cells. *Journal of the Optical Society of America A*, *70*, 1297-1300.

Marr, D. (1982). *Vision*. San Francisco: Freeman.

Marr, D. & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London B*, 207, 187-217.

Marr, D. & Poggio, T. (1976). The co-operative computation of stereo disparity. *Science*, 194, 283-287.

Marr, D. & Poggio, T. (1979). A Computational theory of human stereo vision. *Proceedings of the Royal Society of London B*, 204, 301-328.

Martin, M. (1979). Local and global processing: The role of sparsity. *Memory and Cognition*, 7, 476 - 484.

Maunsell, J. H. R., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E. & Noerager, B. D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual Neuroscience*, 16, 1 - 14.

Mayhew, J. E. W. & Frisby, J. P. (1976). Rivalrous texture stereograms. *Nature*, 264, 53-56.

Merigan, W. H., Byrne, C. & Maunsell, J. H. R. (1991). Does primate motion perception depend on the magnocellular pathway? *Journal of Neuroscience*, 11, 3422-3429.

Merigan, W. H. & Eskin, T. A. (1986). Spatio-temporal vision of macaques with severe loss of P β retinal ganglion cells. *Vision Research*, 26, 1751-1761.

-
- Merigan, W. H., Katz, L. M. & Maunsell, J. H. R. (1991). The effects of parvocellular lateral geniculate lesions on the acuity and contrast sensitivity of macaque monkeys. *Journal of Neuroscience*, *11*, 994-1001.
- Miller, J. (1981). Global Precedence in Attention and Decision. *Journal of Experimental Psychology: Human Perception and Performance*, *7*, 1161 - 1185.
- Mitov, D., Vassilev, A. & Manahilov, V. (1981). Transient and Sustained Making. *Perception & Psychophysics*, *30*, 205 - 210.
- Miyashita, Y. & Chang, H. S. (1988). Neuronal correlate of pictorial short term memory in primate temporal cortex. *Nature*, *331*, 68 - 70.
- Morgan, M. J. (1991). Hyperacuity. In D. Regan (Ed.), *Spatial Vision* (Vol. 10, pp. 87-113). Boca Raton, Fl.: CRC Press.
- Morgan, M. J. & Hotopf, W. H. N. (1989). Perceived diagonals in grids and lattices. *Vision Research*, *29*, 1005 - 1015.
- Morgan, M. J. & Ward, R. M. (1985). Spatial and spatial-frequency primitives in spatial- interval discrimination. *Journal of the Optical Society of America A*, *2*, 1205-1210.
- Morgan, M. J., Ward, R. M. & Hole, G. J. (1990). Evidence for positional coding in hyperacuity. *Journal of the Optical Society of America A*, *7*, 297 - 304.

-
- Morgan, M. J. & Watt, R. J. (1984). Spatial frequency interference effects and interpolation in Vernier acuity. *Vision Research*, 24, 1911-1919.
- Morgan, M. J. & Watt, R. J. (1997). The combination of filters in early spatial vision: a retrospective analysis of the MIRAGE model. *Perception*, 26, 1073 - 1088.
- Morrone, M. C., Burr, D. C. & Ross, J. (1983). Added noise restores recognizability of coarse quantized images. *Nature*, 305, 226 - 228.
- Morrone, M. C. & Burr, D. C. (1988). Feature detection in human vision: A phase – dependent energy model. *Proceedings of the Royal Society of London, B*, 235, 221 - 245
- Moulden, B. (1994). Collator units: Second-stage orientation filters, *Higher-order processing in the visual system* (pp. 170 - 192). Chicester: John Wiley.
- Mussap, A. J. & Levi, D. M. (1996). Spatial properties of filters underlying Vernier acuity revealed by masking: Evidence for collator mechanisms. *Vision Research*, 36, 2459 - 2473.
- Musselwhite, M. J. & Jeffreys, D. A. (1985). The influence of spatial frequency on reaction times and evoked potentials recorded to grating pattern stimuli. *Vision Research*, 25, 1545 - 1555.
- Nachmias, J. (1967). Effect of exposure duration on visual contrast sensitivity with square-wave gratings. *Journal of the Optical Society of America*, 57, 421-427.

-
- Nachmias, J. & Sansbury, R. (1974). Grating contrast: discrimination may be better than detection. *Vision Research*, *14*, 1039-1042.
- Nachmias, J., Sansbury, R., Vassilev, A. & Weber, A. (1973). Adaptation to square-wave gratings: in search of the elusive third harmonic. *Vision Research*, *13*, 1335-1342.
- Nachmias, J. & Weber, A. (1975). Discrimination of simple and complex gratings. *Vision Research*, *15*, 217-223.
- Navon, D. (1977). Forest before trees: The Precedence of Global features in Visual Perception. *Cognitive Psychology*, *9*, 353 - 383.
- Navon, D. (1981). Reply to Miller, J. 1981 article. *Journal of Experimental Psychology: Human Perception and Performance*, *7*,.
- Norman, J. & Ehrlich, S. (1987). Spatial frequency filtering and target identification. *Vision Research*, *27*, 87-96.
- Nowak, L. G., Munk, M. H. J., Girard, P. & Bullier, J. (1995). Visual latencies in areas V1 and V2 of the macaque monkey. *Visual Neuroscience*, *12*, 371 - 384.
- Olhausen, B. A. & Field, D. J. (1996). Emergence of simple cell receptive field properties by learning a sparse code for natural images. *Nature*, *381*, 607 - 609.

Oliva, A. & Schyns, P. G. (1995,). *Mandatory scale perception promotes flexible scene categorisation*. Paper presented at the Proceedings of the XVII Meeting of the Cognitive Science Society.

Oliva, A. & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognitive Psychology*, 24, 72 - 107.

Olzak, L. A. & Thomas, J. P. (1981). Gratings: Why frequency discrimination is sometimes better than detection. *Journal of the Optical Society of America*, 71, 64-70.

Olzak, L. A. & Thomas, J. P. (1992). Configural effects constrain Fourier models of pattern discrimination. *Vision Research*, 32, 1885 - 1898.

Olzak, L. A. & Thomas, J. P. (1999). Neural recoding in human pattern vision: model and mechanisms. *Vision Research*, 39, 231 - 236.

Olzak, L. A. & Wickens, T. D. (1997). Discrimination of complex patterns: orientation information is integrated across spatial scale; spatial-frequency and contrast information are not. *Perception*, 26, 1101 - 1120.

Palagi, P. M. & Guerin-Dugue, A. (1995). An architecture for texture segmentation: From energy features to region detection., *Lecture Notes in Computer Science* (pp. 956 - 962).

-
- Pantle, A. (1974). Motion aftereffect magnitude as a measure of the spatiotemporal response properties of direction-sensitive analyzers. *Vision Research*, 14, 1229-1236.
- Pantle, A. & Sekuler, R. W. (1968). Velocity sensitive elements in human vision: Initial psychophysical evidence. *Vision Research*, 8, 445-450.
- Paquet, L. & Merikle, P. M. (1984). Global precedence: The effect of exposure duration. *Canadian Journal of Psychology*, 38, 45 - 53.
- Parish, D. H. & Sperling, G. (1991). Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination. *Vision Research*, 31, 1399 - 1415.
- Parker, A. J. & Hawken, M. J. (1985). Capabilities of monkey cortical cells in spatial-resolution tasks. *Journal of the Optical Society of America A*, 2, 1101-1114.
- Parker, D. M. (1980). Simple reaction times to the onset, offset and contrast reversal of sinusoidal grating stimuli. *Perception & Psychophysics*, 28, 365 - 368.
- Parker, D. M. & Dutch, S. (1987). Perceptual latency and spatial frequency. *Vision Research*, 27, 1279-1284.
- Parker, D. M., Lishman, J. R. & Hughes, J. (1992). Temporal integration of spatially filtered visual images. *Perception*, 21, 147 - 160.

-
- Parker, D. M., Lishman, J. R. & Hughes, J. (1996). Role of coarse and fine spatial information in face and object processing. *Journal of Experimental Psychology: Human Perception and Performance*, 22, 1448 - 1466.
- Parker, D. M., Lishman, J. R. & Hughes, J. (1997). Evidence for the view that temporospatial integration in vision is temporally anisotropic. *Perception*, 26, 1169 - 1180.
- Parker, D. M. & Salzen, E. A. (1977). Latency changes in the human visual evoked response to sinusoidal gratings. *Vision Research*, 17, 1201 - 1204.
- Parker, D. M. & Salzen, E. A. (1982). Evoked potentials and reaction times to the offset and contrast reversal of sinusoidal gratings. *Vision Research*, 22, 205 - 207.
- Perrett, D. I., Mistlin, A. J. & Chitty, A. J. (1987). Visual neurones responsive to faces. *Trends in NeuroScience*, 10, 358-364.
- Phillips, G. C. & Wilson, H. R. (1984). Orientation bandwidths of spatial mechanisms measured by masking. *Journal of the Optical Society of America A*, 1, 226- 232.
- Phillips, L. & Framo, J. L. (1954). Developmental theory applied to normal and psychopathological perception. *Journal of Personality*, 22, 465 - 474.
- Plant, G. T., Zimmern, R. L. & Durden, K. (1983). Transient visually evoked potentials to the pattern reversal and onset of sinusoidal gratings. *Electroencephalography and Clinical Neurophysiology*, 56, 147 - 158.

Polat, U. & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33, 993-999.

Polat, U. & Sagi, D. (1994). The architecture of perceptual spatial interactions. *Vision Research*, 34, 73 - 78.

Pollen, D. A., Lee, J. R. & Taylor, J. H. (1971). How does the striate cortex begin the reconstruction of the visual world. *Science*, 173, 74 - 77.

Pomerantz, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General*, 112, 516 - 540.

Previc, F. H. (1993). Functional specialisation in the lower and upper visual fields in humans: Its ecological origins and neurophysiological implications. *Behavioral and Brain Sciences*, 13, 519 - 575.

Ramachandran, V. S. (1990). Interactions between motion, depth, color and form: the utilitarian theory of perception. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 346 - 359). Cambridge: Cambridge University Press.

Regan, D. (1989). *Human Brain Electrophysiology*. New York: Elsevier.

Riley, D. & Costall, A. (1980). Comment on "Recognition of faces in the presence of two - dimensional sinusoidal masks" by Tieger & Ganz. *Perception & Psychophysics*, 27, 373 - 374.

Rizzolatti, G., Riggio, L. & Sheliga, B. M. (1994). Space and Selective Attention. In C. Umiltà & M. Moscovitch (Eds.), *Attention and Performance XV* (Vol. XV, pp. 231 - 265). Cambridge, Mass.: M.I.T.

Robertson, L. C., Lamb, M. R. & Knight, R. T. (1988). Effects of lesions of temporal-parietal junction on perceptual and attentional processing in humans. *Journal of Neuroscience*, 8, 3757 - 3769.

Robson, J. G. (1966). Spatial and temporal contrast sensitivity functions of the visual system. *Journal of the Optical Society of America*, 56, 1141-1142.

Robson, J. G. (1975). *Receptive fields: Neural representation of the spatial and intensive attributes of the visual image*. (Vol. V). New York San Francisco London: Academic Press, INC.

Rogowitz, B. E. (1983). Spatial/temporal interactions: backward and forward metacontrast masking with sine-wave gratings. *Vision Research*, 23, 1057-1074.

Ross, J. & Speed, H. D. (1991). Contrast adaptation and contrast masking in human vision. *Proceedings of the Royal Society B*, 264, 61-69.

Ruderman, D. L. (1994). The statistics of natural images. *Network*, 5, 517 - 548.

Ruderman, D. L. (1997). Origins of scaling in natural images. *Vision Research*, 37, 3385 - 3398.

- Russell, M. H. A., Kulikowski, J. J. & Murray, I. J. (1987). Spatial frequency dependence of the human visual evoked potential. In C. Barber & T. Blum (Eds.), *Evoked Potentials III* (pp. 231 - 239). London: Butterworths.
- Sachs, M. B., Nachmias, J. & Robson, J. G. (1971). Spatial-frequency channels in human vision. *Journal of the Optical Society of America*, *61*, 1176-1186.
- Sakata, H., Taira, M., Kusunoki, M., Murata, A. & Tanaka, Y. (1997). The parietal association cortex in depth perception and visual control of hand action. *Trends in Neurosciences*, *20*, 350 - 357.
- Schade, O. H. (1956). Optical and photoelectric analog of the eye. *Journal of the Optical Society of America*, *46*, 721-739.
- Schober, H. A. W. & Hilz, R. (1965). Contrast sensitivity of the human eye for square-wave gratings. *Journal of the Optical Society of America*, *55*, 1086-1091.
- Schyns, P. G. & Oliva, A. (1994). From blobs to boundary edges: Evidence for time and spatial scale dependent scene recognition. *Psychological Science*, *5*, 195 - 200.
- Schyns, P. G. & Oliva, A. (1997). Flexible, diagnosticity - driven rather than fixed, perceptually determined scale selection in scene and face recognition. *Perception*, *26*, 1027 - 1038.
- Sekuler, R. & Levinson, E. (1974). Mechanisms of motion perception. *Psychologia*, *17*, 38-49.

Sergent, J. A. (1982). The cerebral balance of power: Confrontation or co-operation. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 253 - 272.

Shulman, G. L., Sullivan, M. A., Gish, K. & Sakoda, W. J. (1986). The role of spatial frequency channels in the perception of local and global structure. *Perception*, 15, 259 - 273.

Shulman, G. L. & Wilson, J. (1987). Spatial Frequency and selective attention to local and global information. *Perception*, 16, 89 - 101.

Smallman, H. S. (1995). Fine-to-Coarse scale disambiguation in stereopsis. *Vision Research*, 35, 1047-1060.

Solomon, J. A., Watson, A. B. & Morgan, M. J. (1999). Transducer model produces facilitation from opposite-sign flanks. *Vision Research*, 39, 987 - 992.

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision*, 4, 183 - 207.

Srinivasan, M. V., Laughlin, S. B. & Dubs, A. (1982). Predictive coding: A fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B*, 218, 427 - 459.

Stork, D. G. & Wilson, H. R. (1990). Do Gabor functions provide appropriate descriptions of visual cortical receptive fields. *Journal of the Optical Society of America A*, 7, 1362 - 1373.

Stromeyer, C. F., III & Julesz, B. (1972). Spatial frequency masking in vision: critical bands and spread of masking. *Journal of the Optical Society of America*, 62, 1221-1232.

Stromeyer, C. F., III & Klein, S. (1975). Evidence against narrow-band spatial frequency channels in human vision: the detectability of frequency modulated gratings. *Vision Research*, 15, 899-910.

Swanson, W. H. & Wilson, H. R. (1985). Eccentricity dependence of contrast matching and oblique masking. *Vision Research*, 25, 1285-1296.

Tartaglione, A., Goff, D. & Benton, A. (1975). Reaction time to square-wave gratings as a function of spatial frequency, complexity and contrast. *Brain Research*, 100, 111 - 120.

Teo, P. C. & Heeger, J. (1994). Perceptual image distortion. In B. E. Rogowitz & J. P. Alebach (Eds.), *Human Vision, Visual Processing and Digital Display V* (pp. 127 - 139).

Thomas, J. P. (1970). Model of the function of receptive fields in human vision. *Psychological Review*, 77, 121 - 134.

-
- Thomas, J. P., Fagerholm, P. & Bonnet, C. (1999). One spatial filter limits speed of detecting low and middle frequency gratings. *Vision Research*, *39*, 1683 - 1693.
- Tolhurst, D. J. (1972). On the possible existence of edge detector neurones in the human visual system. *Vision Research*, *12*, 797-804.
- Tolhurst, D. J. (1973). Separate channels for the analysis of the shape and the movement of a moving visual stimulus. *Journal of Physiology (London)*, *231*, 385-402.
- Tolhurst, D. J. (1975a). Reaction times in the detection of gratings by human observers: a probabilistic mechanism. *Vision Research*, *15*, 1143-1149.
- Tolhurst, D. J. (1975b). Sustained and transient channels in human vision. *Vision Res.*, *15*, 1151-1155.
- Tolhurst, D. J. & Barfield, L. P. (1978). Interactions between spatial frequency channels. *Vision Research*, *18*, 951-958.
- Tolhurst, D. J., Tadmor, Y. & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic & Physiological Optics*, *12*, 229 - 232.
- Turvey, M. T. (1973). On peripheral and central processes in vision: Inferences from an information processing analysis of masking with patterned stimuli. *Psychological Review*, *80*, 1 - 52.

-
- van der Schaaf, A. & van Hateren, J. H. (1996). Modelling the power spectra of natural images: Statistics and information. *Vision Research*, 36, 2759 - 2770.
- van Hateren, J. H. & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London B*, 265, 359 - 366.
- Vassilev, A. & Mitov, D. (1976). Perception time and spatial frequency. *Vision Research*, 16, 89-92.
- Vassilev, A., Stomonyakov, V. & Manahilov, V. (1994). Spatial frequency specific contrast gain and flicker masking of human transient VEP. *Vision Research*, 34, 863 - 872.
- Vassilev, A. & Strashmirov, D. (1979). On the latency of human visually evoked response to sinusoidal gratings. *Vision Research*, 19, 843 - 845.
- Ward, L. M. (1982). Determinants of attention to local and global features of visual forms. *Journal of Experimental Psychology: Human Perception & Performance*, 8, 562 - 581.
- Watson, A. B. (1986). Temporal Sensitivity. In K. R. Boff, L. Kaufman & J. P. Thomas. (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 6.1 - 6.43). New York: Wiley.

-
- Watson, A. B. & Nachmias, J. (1977). Patterns of temporal interaction in the detection of gratings. *Vision Research*, 17, 893-902.
- Watson, A. B. & Solomon, J. A. (1997). Model of visual contrast gain control and pattern masking. *Journal of the Optical Society of America A*, 14, 2379 - 2391.
- Watt, R. J. (1984a). Further evidence concerning the analysis of curvature in human foveal vision. *Vision Research*, 24, 251-254.
- Watt, R. J. (1984b). Towards a general theory of the visual acuities for shape and spatial arrangement. *Vision Research*, 24, 1377-1386.
- Watt, R. J. (1987). Scanning from coarse to fine spatial scales in the human visual system after the onset of a stimulus. *Journal of the Optical Society of America, A*, 4, 2006-2021.
- Watt, R. J. (1988). *Visual processing: Computational, psychophysical and cognitive research*. Hove: Lawrence Erlbaum Associates.
- Watt, R. J. & Andrews, D. P. (1982). Contour curvature analysis: hyperacuities in the discrimination of detailed shape. *Vision Research*, 22, 449-460.
- Watt, R. J. & Morgan, M. J. (1983). The recognition and representation of edge blur: evidence for spatial primitives in human vision. *Vision Research*, 23, 1465-1478.

-
- Watt, R. J. & Morgan, M. J. (1984). Spatial filters and the localization of luminance changes in human vision. *Vision Research*, 24, 1387-1398.
- Watt, R. J. & Morgan, M. J. (1985). A theory of the primitive spatial code in human vision. *Vision Research*, 25, 1661-1674.
- Waugh, S. J. & Levi, D. M. (1993). Visibility, timing and Vernier acuity. *Vision Research*, 33, 505 - 526.
- Waugh, S. J. & Levi, D. M. (1995). Spatial alignment across gaps: Contributions of orientation and spatial scale. *Journal of the Optical Society of America A*, 12, 2305 - 2317.
- Werner, H. (1956). Microgenesis and aphasia. *Journal of abnormal and social psychology*, 52, 347 - 353.
- Westheimer, G. (1975). Visual acuity and hyperacuity. *Investigative Ophthalmology and Visual Sciences*, 14, 570-571.
- Westheimer, G. (1979). The spatial sense of the eye. *Investigative Ophthalmology and Visual Sciences*, 18, 893-912.
- Westheimer, G. (1981). Visual hyperacuity. *Progress in Sensory Physiology*, 1, 1-30.
- Westheimer, G. & McKee, S. P. (1977). Perception of temporal order in adjacent visual stimuli. *Vision Research*, 17, 887-893.

-
- Whitaker, D. (1993). What part of a Vernier stimulus determines performance. *Vision Research*, 33, 27 - 32.
- Whitaker, D. & MacVeigh, D. (1991). Interaction of spatial frequency and separation in Vernier acuity. *Vision Research*, 31, 1205-1212.
- Williamson, S. J., Kaufman, L. & Brenner, D. (1978). Latency of the neuromagnetic response of the human visual cortex. *Vision Research*, 18, 107 - 110.
- Wilson, H. R. (1986). Response of spatial mechanisms can explain hyperacuity. *Vision Research*, 26, 453-469.
- Wilson, H. R. (1991). Psychophysical models of spatial vision and hyperacuity. In D. Regan (Ed.), *Spatial Vision* (Vol. 10, pp. 64-86). Boca Raton, FL: CRC Press.
- Wilson, H. R., Blake, R. & Halpern, D. L. (1991). Coarse spatial scales constrain the range of binocular fusion on fine scales. *Journal of the Optical Society of America A*, 8, 229-236.
- Wilson, H. R. & Gelb, D. J. (1984). Modified line-element theory for spatial-frequency and width discrimination. *Journal of the Optical Society of America A*, 1, 124-131.
- Wilson, H. R., McFarlane, D. K. & Phillips, G. C. (1983). Spatial frequency tuning of orientation selective units estimated by oblique masking. *Vision Research*, 23, 873-882.

Wilson, H. R. & Wilkinson, F. (1997). Evolving concepts of spatial channels in vision: From independence to non-linear interactions. *Perception*, 26, 939 - 960.

Wilson, H. R., Wilkinson, F. & Assad, W. (1997). Concentric orientation summation in human form vision. *Vision Research*, 37, 2325 - 2330.

Yang, J. (1992). Do Gabor functions provide appropriate descriptions of visual cortical receptive fields: comment. *Journal of the Optical Society of America A*, 9, 334 - 336.

Young, R. A. (1987). The Gaussian derivative model for spatial vision: I. Retinal mechanisms. *Spatial Vision*, 2, 273-293.

