Quality Measurements on Quantised Meshes

AESHAH ALI S ALMUTAIRI

Quality Measurements on Quantised Meshes

Aeshah Ali Almutairi

A Thesis presented for the degree of
Doctor of Philosophy

Department of Computer Sciences
University of Durham
England
April 2021

Dedicated to

I dedicate this thesis to my beloved parents and family.

Quality measurements on quantised meshes

Aeshah Almutairi

Submitted for the degree of Doctor of Philosophy
April 2021

Abstract

In computer graphics, triangle mesh has emerged as the ubiquitous shape representation for 3D modelling and visualisation applications. Triangle meshes, often undergo compression by specialised algorithms for the purposes of storage and transmission. During the compression processes, the coordinates of the vertices of the triangle meshes are quantised using fixed-point arithmetic. Potentially, that can alter the visual quality of the 3D model. Indeed, if the number of bits per vertex coordinate is too low, the mesh will be deemed by the user as visually too coarse as quantisation artifacts will become perceptible. Therefore, there is the need for the development of quality metrics that will enable us to predict the visual appearance of a triangle mesh at a given level of vertex coordinate quantisation.

In this thesis, we present the results of four psychophysical experimental studies to assess the visual quality of quantised meshes. To achieve that, we used triangles meshes which varied in some important geometric characteristics, such as the number of triangles, the average shape of the triangles, and the level of detail in their global shape. The meshes in the experiments were quantised using different quantisation methods, rendered with the use of various textures and lighting environments. We employed various experimental designs, such as 2-AFC with or without staircases, and MLDS, and in all cases lightly trained participants were invited to compare the visual qualities of the models.

The main findings of our experiments cab be summarised as follows. The discrimination threshold, that is, the level of quantisation below which the viewer can perceive quantisation artifacts, is lower when dithered quantisation is used instead

of the most commonly used rounding. A large number of triangles in the mesh, and rendering methods with high specular components, increase the discrimination thresholds, that is, they require more bits per vertex coordinate to make the quantisation artifacts imperceptible. In, perhaps, the strongest result in the thesis, we established a strong correlation between the discrimination threshold and the amount of information carried by the mesh, as measured by the file size of the mesh when compressed with a state-of-the-art method. A fourth experiment, based on a more complex design and the MLDS method, was not conclusive, but enabled a preliminary investigation on the challenges facing these types of complex experiments with lightly trained participants.

Declaration

The work in this thesis is based on research carried out at the Department of Computer Sciences, Durham University, England. No part of this thesis has been submitted elsewhere for any other degree or qualification and it is all my own work unless referenced to the contrary in the text.

Acknowledgements

Throughout my research journey, I have received a great deal of support and assistance either directly or indirectly, which I sincerely appreciated it.

I would like first to thank my thesis supervisor Dr.Ioannis Ivrissimtzis for his guidance and continues challenges that sharpened my thoughts and brought the experiments to higher level.

I would also like to thank Dr.Toni Saarela from University of Helsinki for agreeing to take a part in the research and appreciate his remarks, his questions, and discussions. This allowed me to take a step back and think clearly about the next step.

I would also like to thank all the participant who took part in these experiments. I will not forget their contribution and patience especially through COVID19, which were essential to obtain a reliable and precise results.

Certainly, I must not forget to thank Dr.Hanan and Altaf for all the support and love.

Finally, I don't even have the words to thank you enough, my parent, my love Faisal, Lama and Saif. I feel blessed to have such a great family.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Summary

A measure for evaluating the 3D mesh quality is essential for the purpose of determining whether a particular operation on the mesh, including compression and quantisation, impacts its perceived quality. The number of studies done in the area of 3D meshes is limited when it is compared with the studies for 2D images, even though large databases of 3D meshes, which would benefit from such insights, are in use for many years now [11]. In this thesis, we are addressing this gap in the literature, by presenting the results of a series of psycho-physical experiments aimed principally at analysing the effect of the mesh vertex quantisation on the perceived mesh quality.

1.1 Context and Motivation

Language, imagination, perception, and planning reveals the psychological cognitive processes of how a human observer experiences and understands the world around. The ability to receive and process information with the eyes defines the visual perception, which is a part of the human cognition. This makes the human subject not only the final receiver, but an important part of measuring the perceived quality of graphics.

Today, graphical data are used in different applications including engineering

design, video gaming, virtual reality, architectural walk-through and e-commerce. Since the demands for efficiency and quality have never ceased to increase, 3D geometric models are often undergoes processing by shape optimization techniques such as simplification, remeshing and fairing, as well as compression for efficient storage and transmission, before being actually used in a practical application. In most cases, these manipulations will lead to artifacts and noises, which may alert their visual quality of the 3D models.

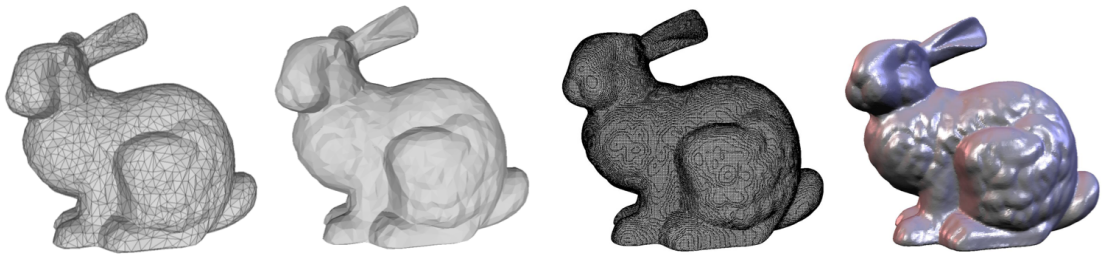


Figure 1.1: Example of different quality of Bunny model. **Left**: Two images of Bunny mesh at low resolution and vertex number of 2500; **Right**: Two images at high level of resolution and vertex number of 139990.

Realism in computer graphics is achieved through the use of complex models. These models are obtained from different sources, such as 3D scanning and modelling software. These usually demand a huge number of computer resources in terms of storage space and transmission bandwidth in the raw data format. More computer resources are needed as the number and complexity of the 3D images increase [81], as shown in Figure 1.1 where we start with simple Bunny with few triangles to more complex and almost replicated model of the real one. Thus, leads to the generation of 3D meshes which is acknowledged to be the most prevalent discrete virtual surface and volume representation with complex storage, processing, and visualization. Followed by growing number of 3D meshes and its complexity the demands for high resource including power, storage space, and bandwidth have increased subsequently. This calls for efficient compression mechanisms to allow for real-time interactivity of high quality meshes.

Among the various visualisation applications, triangular meshes are one of the main representation formats in use, that have been developed for representation of

3D models. Triangular meshes can be characterized using different features including geometry, connectivity, and property data also called attributes. Connectivity describes the relationship between the vertices while geometry specify location of the vertices [81]. The data represent different features, such as texture coordinates, normal vector and material reflectance. Geometry and property data are generally called vertex data, as they are often attached to vertices. Consequently, most algorithms for the compression of triangular meshes handle geometry and property data in the same fashion.

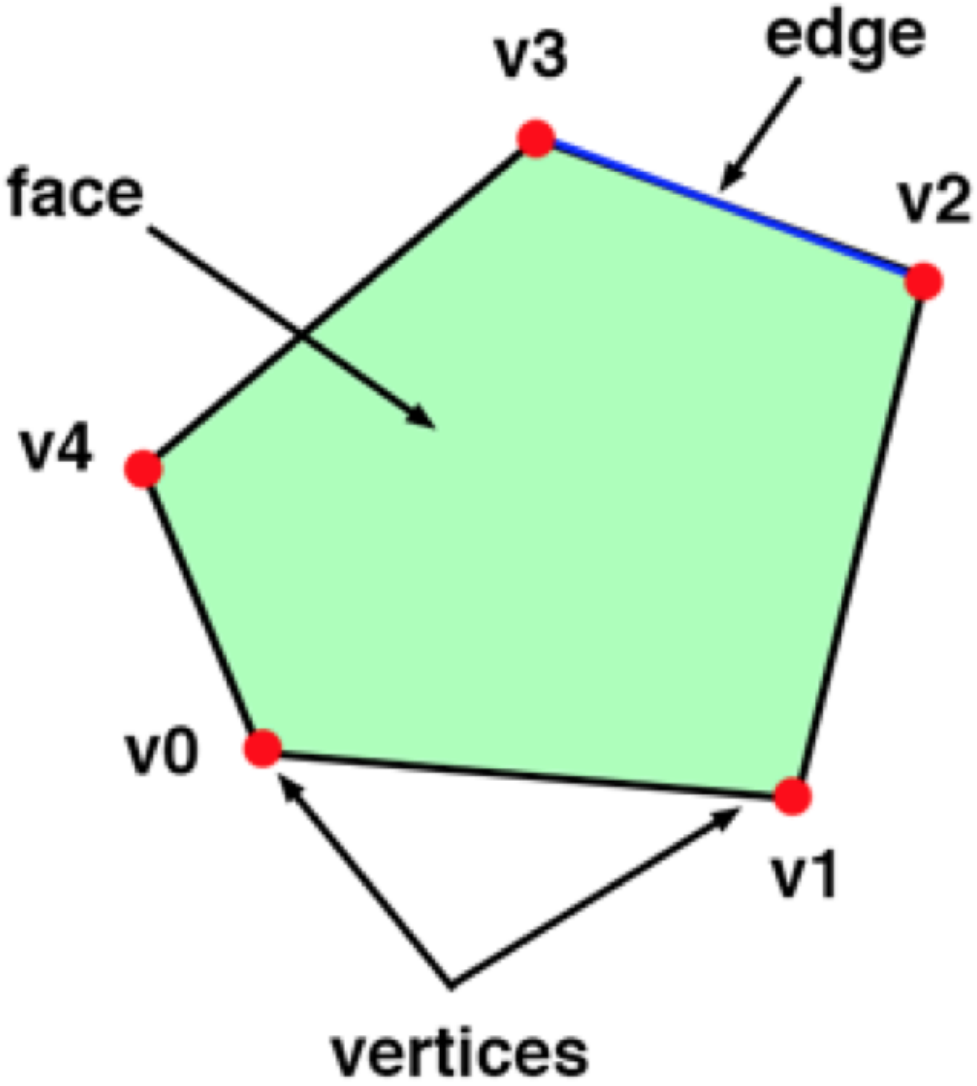


Figure 1.2: Polygon components.

Figure 1.2 shows the polygon components. According to `http://www.Scratchapixel.`

com, polygon meshes are considered as the oldest form of geometric representations that have been used in computer graphics. The individual points are the vertices. This can be represented in x and y coordinate system in 2D representation. The vertices can be connected in either clockwise or counterclockwise direction to form the faces. The ordering of the faces is important as it helps in computing. The edge is the line connecting the faces. The minimum number of vertices in a face is 3; such a face is called triangle. When there are four vertices, then this is called, quad while when there are more than four vertices, then this is called a general polygon.

Triangular meshes are preferred because of their ability to maintain mesh quality. Mesh quality is defined as the manifoldness, faithfulness and uniformness of the mesh [70]. This is different from the case of polygon meshes that are not able to maintain the same level of quality. Faithfulness is the ability of the mesh to preserve topology and geometry, such as volume and surface area. Manifoldness means that each point in the surface mesh has a neighborhood that is topologically similar to a disk [70]. Beyond manifoldness, certain types of meshes, such as general polygon meshes may cause certain difficulties in generating surface elements that conform to the finite element computations. Uniformness is associated with the triangle shape, complexity and regularity. The quality of the triangle in surface mesh is important in determining its robustness in certain applications [70].

Mesh compression is often being used by the researchers to decrease the size of the data without severe loss on quality. The early studies on the mechanisms of 3D mesh compression focused mainly on single-rate compression. The objective of this was to save the bandwidth between the graphics card and the Central Processing Unit (CPU). In a single-rate 3D compression, all the geometry and connectivity data are compressed and decompressed as a whole. This means that the graphics card is not able to render the original signal or mesh until the whole signal stream has been received. However, with the advent of the internet, there has been development of progressive compression and transmission mechanisms. One of the benefits of progressive compression is that it enhances interactivity. This is because compression can be interrupted when a user realizes that the mesh, they are downloading is not in the resolution that they expected [81]. The early mesh compression algorithms

are obtained based on the fundamental connectivity coding. Though the geometry data ask for a greater number of bits rather than the topology data, various methods have been suggested for effective compression of geometry data devoid of topology data reference. Further research has resulted in the use of 3D mesh compression in audio and video technology leading to different developments, such as MP3 and MPEG-4 among others [81].

Quantisation is a compression mechanism which focuses on compressing continuous infinite values to smaller finite values. The purpose of quantization is to reduce the number of discrete symbols within a given stream. When this is achieved, then it becomes easy to compress the stream [41]. For example, reducing the number of vertices for a given shape makes it easy to reduce the size of the image. Quantization introduces different errors, such as rounding errors and computational noise [50].

The motivation for our a study is derived from the visualisation impact of the quantization technique on the mesh quality. Currently, visual applications use float 32 bits or double 64 bits without knowing the ideal size of precision needed in these applications that will give results that are visually accepted. Hence, using a small quantisation level l may lead to a significant loss of information and a large l may lead to redundancy in the mesh, with unnecessary large files as shown in Figure 1.3.

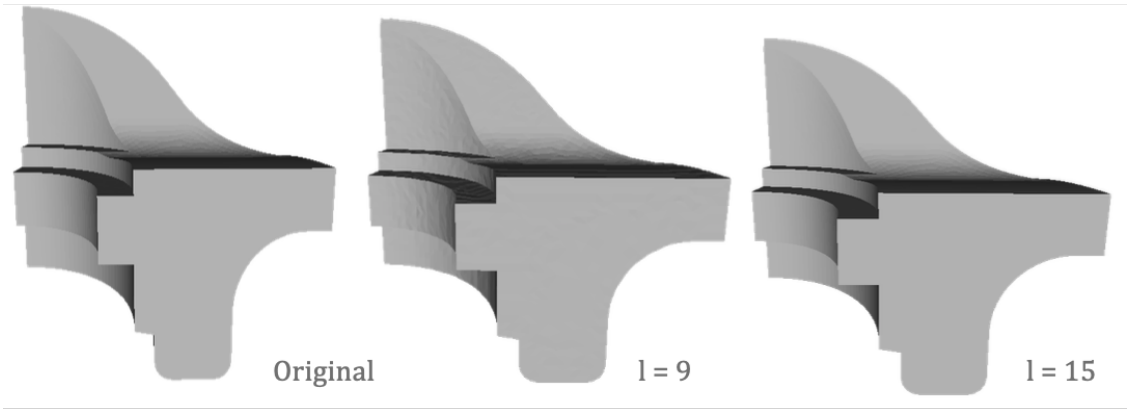


Figure 1.3: Visual mesh quality at various level of quantisation

3D triangle meshes are always quantised. While the standard formats use a floating point arithmetic of high precision, they quite often are transformed to fixed point arithmetic, of relatively low resolution. Quantisation can reduce the size of

3D model without any identifiable quality loss. The triangles x, y, z can be seen as real numbers quantized at default level l, typically at $l = 32$ bits (floats) or at $l = 64$ bits (doubles), for more efficient storage and modeling. Yet, this simple modelling can be upgraded in complexity by adding more and more triangles.

The use of a float should be reasonable for most of application since it is so widely used, thus this has to be also checked. Moreover, the compression algorithms do quantisation to less than 32 bits and we do not know the visual impact of this procedure on the quality of the mesh. If we know the visual significance of each bit, we can still use these 32 bits and use the least significant bits for carrying other information such as in steganography.

1.2 Objectives and Methodology

The research topic of this thesis is to study experimentally the effect of quantisation on 3D polygonal models and to evaluate the impact of the choice of the level of quantisation of a mesh on its quality and define methods to assess that without the need for any user input.

1.2.1 Main objective:

How can we choose the appropriate level of quantisation for a high quality, yet memory efficient representation of a triangle mesh?

While the issue of finding the appropriate level of mesh quantisation has been encountered in literature, especially the one related to mesh compression, to the best of our knowledge, there are no systematic experiments studying the impact of mesh quantisation, and trying to estimate an optimal quantisation level for triangle meshes. In our context, an optimal quantisation level is the one that uses the least number of bits per vertex coordinate, without creating any visible artifacts to the mesh.

There are many factors to take into consideration that affect the visual quality of the mesh and the optimal quantisation level. The rendering conditions play a significant role as they could amplify noise and make it easier for the observer's eye

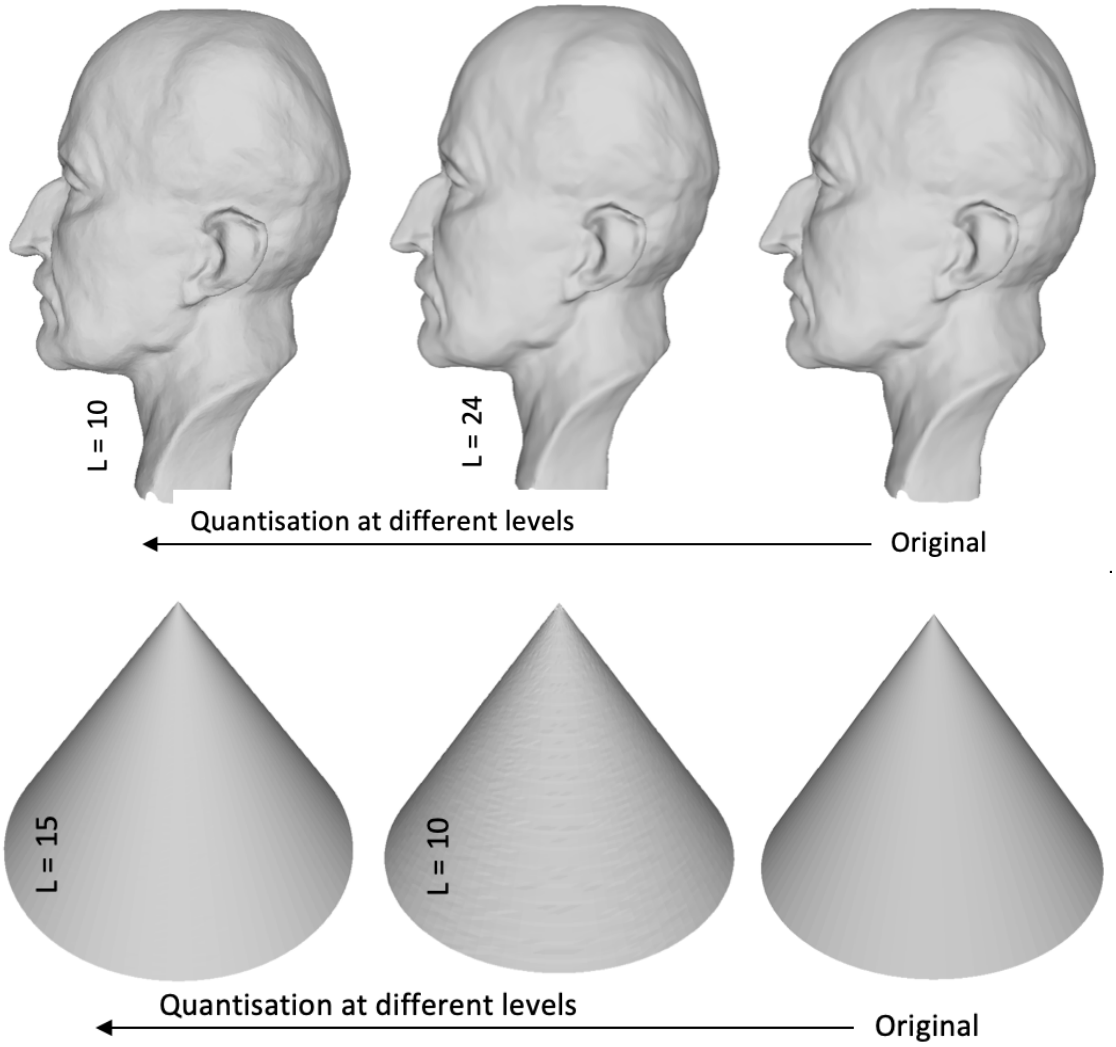


Figure 1.4: Two different Models quantised.**Top:** The Max-Planck Model. **Bottom:** The Cone Model.

to detect the artifacts. For example, specular rendering, where the reflection of light is strong, makes it easier to detect any noise by human eyes as will be discussed in some of our experiments. The size and the characteristics of the triangles influence the surface quality as well. Complex models contains a lot of details with different types of triangles, make it an easy target to be visually disturb by the choice of level of quantisation. It is evident that different factors have different visual quality impact on the models, in Figure 1.4, the size and the complexity of the Max-planck model will increase the optimal level of quantisation, while the characteristics of the triangles have an effect on the level of quantisation in the Cone model comprising

of skinny thin triangles.

1.2.2 Methodology

To accomplish the objective of studying the perceptual impact of mesh quantisation, we designed four psycho-physical experiments with unquantised and quantised 3D meshes as stimuli. In three of them we used a Two-Alternative Forced choice (2-AFC) design in which observers were shown a quantised alongside an unquantised mesh and were asked to choose the mesh of higher quality. We also designed one psycho-physical experiment, described in Chapter 6, where the participants were presented two pairs of quantised meshes and had to choose the pair with the highest perceived difference between these meshes.

The collected data were used to evaluate and scale the visual quality of the human performance. In each experiment, we introduce quantised models with both geometric and texture distortions, conduct a paired-comparison subjective assessment, and invite university subjects to evaluate the visual qualities of the models under different rendering condition and resolutions. Finally, we evaluated the results based on the human visual perception and quantisation level.

1.3 Outline

The remainder of this thesis is organised as follow:

Chapter 2 provides a background knowledge on mesh quantisation in terms of choosing the appropriate level of quantisation in order to process a real-life image and compress it without having a significant impact on mesh quality. This is because to present an identical image with respect to the real world [30] it should be quantised at proper level l, which identifies the factors of mesh quality as it is shown in Figure 1.3. It covers the elements of rendering, especially surface rendering, where the surface is the major part affected from quantisation procedure; further discussion includes the rendering algorithms and volume rendering where some inspiration for studying how different mesh characteristics affect the perceived quality of the quantised mesh came from David Roberts's work on volume rendering and quality perception [87].

We discuss in detail the measurements of quality on meshes and how these will influence the visual quality of the renderings of the 3D models. Finally, although algorithms in computer graphics mostly function in a three-dimensional setting, the 3D model is then mapped for visualisation into a two-dimension image at a the stage of the overall process [113] called rendering. For that reason, it is very important to review the literature on the measures of the visual quality of two-dimensional images as well.

Chapters 3-6 describe the results of four psycho physical experiments.

Chapter 3 Does the quantisation method have any effect on the values of the optimal quantisation level?

We introduce the first experiment where we differentiate between two common used methods of quantisation, dithered quantisation and non-dithered (rounding). The non-dithered method sets the bits above the quantisation level to zero which could lead to blocky artifacts and alter the visual quality of the 3D object. On the other hand, in the dithered method, all bits above the quantisation level have random values and that could cause high frequency noise which again can easily be detected by the human eyes. We present the results of a user study on estimating a quantisation threshold above which the quantised triangle mesh is perceived as indistinguishable from its unquantised original. The experimental design follows a 2-AFC process. That is, in each trial of the experiment the subject is forced to choose between two stimuli. The results show that dithering has higher quantisation threshold and while the difference between the two methods is small, around one bit per vertex coordinate, it is nevertheless statistically significant.

Chapter 4 What are the factors that play significant role on the perceived quality of the quantised triangle meshes?

Our study is based on a psychophysical experiment following again a 2-AFC design. Two versions of the same model are presented to the participant, the original and the quantised one at a certain level, and they chooses the one with the higher visual quality. We used three models in total, two of them are the Max-Planck model at two resolutions, 100K and 5K triangles, respectively, while the third is a spherical model with 5K triangles. The aim is to establish whether the geometric

complexity of the model, as manifested by its number of triangles, as well as its regularity as manifested by various measures of triangle shape quality, affect the quantisation threshold. We also used two rendering methods one of which had a much higher specular component than the other.

Our initial results indicate that factors such as the number of triangles in the mesh, and the strength of the specular component in the reflection model of the rendering algorithm, do affect the quantisation thresholds.

Chapter 5 Does the geometry of the mesh, that is, the shape of the 3D model and the properties of the underlying mesh are related to the discrimination threshold beyond which the quantised and unquantised meshes are perceived as identical?

Our study is based again on a 2-AFC ,psycho physical experiment, where two stimuli of one model are presented — the original and one which is quantised at a certain level — and the participant chooses the one with the higher visual quality. We used four different 3D models, the Max-Planck, the Cone, the Sphere, and the Human-Head, fixing all the experimental parameters we had studied previously. In particular, we chose rounding as the quantisation method, we used only meshes with a large number of triangles, between 200K and 315K, and we used a single rendering method, which had a high specular component. The results show a strong and statistically significant correlation between the discrimination threshold and the amount of geometric information carried by the mesh, as measured by the filesizes of the compressed meshes.

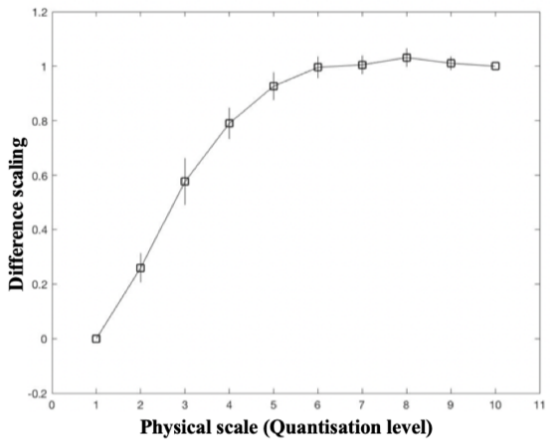


Figure 1.5: Psychometric Function.

Chapter 6 In this experiment we want to go beyond the estimation of discrimination thresholds and understand the scaling of the perceived visual quality. We designed a psycho physical, 2AFC experiment and used Maximum Likelihood Difference Scaling (MLDS) method for modeling the perceived visual differences when the vertices of a triangle mesh are quantized at various levels. The aim is to obtain a scale of the perceived quality in terms of the quantisation level instead of a a single threshold value as shown in the figure example of difference scaling function, Figure 1.5. We presented two pairs of images to the observer and asked to choose which pair consists of more different images than the other. The results were inconclusive, but indicate that there could be a relationship between quantisation thresholds, geometric characteristics of the mesh and properties of the rendering style which could be further studied with larger experiment based on the MLDS method.

Chapter 7 summarizes the contributions of the thesis, and proposes future work on the assessment of visual quality for 3D objects.

Chapter 2

Background

Over the last two decades, various mesh compression methods have been proposed to compress 3D models, usually in the form of triangle meshes, in order to limit the bandwidth usage and reduce the data transfer time. These operations may introduce geometric distortions in form of perturbation of vertex coordinates, which might be visible to a human observer. This is key issue for human-centered applications, as the visibility of these geometric distortions can directly impact the quality of experience of the user. It is therefore important to be able to predict or control the visibility of such geometric distortions.

This chapter is divided into two parts. The first part gives a brief overview about the 3D models (Section 2.1). In particular, we focus on the rendering methods and understanding the visual information that the human eyes perceive (Section 2.2). The second part focuses on the major characteristics of the triangle meshes that are of relevance to our perceptual studies (Section 2.3), and we discuss the quality measures, for both 2D images and meshes (Section 2.4). Finally, we discuss the types and the design of the experiments that we implemented throughout this research, as well as the software we used.

2.1 3D models (meshes)

3D geometry studies the mathematics of shapes in three-dimensional space, the points of which are described by a triple consisting of the 3 coordinates x, y and

z, see Figure 2.1. In 3D geometry, the three coordinates determine uniquely the location of the point in space.

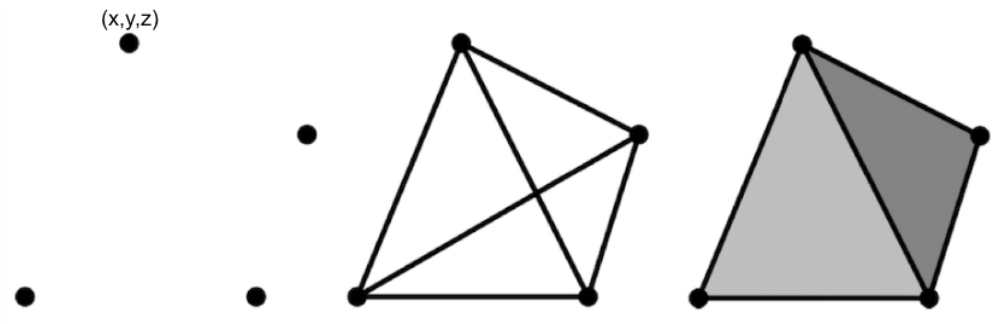


Figure 2.1: x,y,z of Triangle Mesh.

Polygonal meshes consist of three different kinds of mesh primitives: vertices, edges, faces. The vertex is a point positioned in 3D space, which might in addition be appended with other information, such as color, normal vector, and texture coordinates. The edges are connection between two vertices. The faces are ordered sets of vertices, and consist of three vertices in the faces of a triangle mesh, or three or more vertices in the faces of general polygonal meshes. The set of faces describes the topology of the triangle meshes. The information described by the vertex coordinates is usually referred to as the geometry of the mesh, while the information described by the faces, that is, the incidence relation between the mesh vertices is called the connectivity. The incidence relations can be used to specify for each face the vertices and edges on the bounding loop, for each edge the end vertices and the faces to which the edge is incident, and for each vertex the incident edge and face. Two vertices or two faces are called adjacent if there exists an edge incident to both.

2.2 Rendering

Rendering creates 2D images from 3D models through computational methods. The geometry of these 3D models may be described in various formats and data structures, while lighting environments and texture details should also be described. This mentioned data is transmitted to the rendering program and then the digital

2D image is produced as an output of this rendering process. In most visual applications, 3D geometry is described by surfaces, however, the field of volumetric data rendering has also received considerable attention from the early days of graphics development [52], where several rendering techniques for 3D volumetric data of scalar or vector quantities are discussed.

In a study regarding surface and volume rendering [105], several surface rendering schemes are presented in terms of 3D displaying of data obtained by imaging devices. Figure 2.2 illustrates a particular data set, which has been used extensively in the literature. The volume renderings of it were not created at the highest possible resolution because of the constraints in terms of computational power and the memory requirements associated with volume rendering. Images A and B show surface renderings, and images C and D show volume renderings.

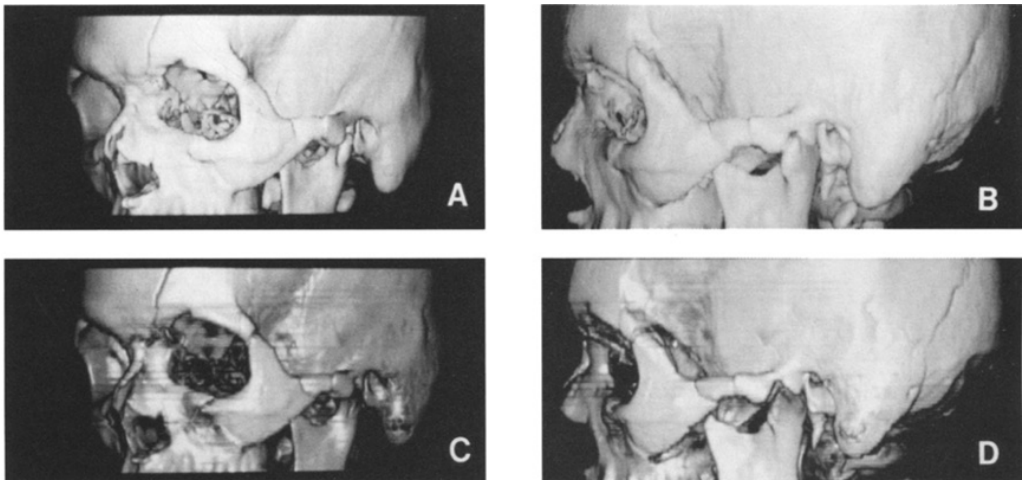


Figure 2.2: Image A and B are surface renderings and image C and D are volume renderings [105].

2.2.1 Rendering Algorithms

Rendering processes may make use of real-world illumination captured by special equipment, or a set of synthetic point or area light sources, or a combination of the above [30]. Generally, the the generation of a high quality image require the use of appropriate, efficient rendering processes [53]. The physics behind the various rendering algorithms are usually described in some basic mathematical form, and

the most efficient algorithms will split the problem into a large numbers of smaller problems that will be solved one piece at a time, rather than attempting to solve the particular instance of the problem for the surface as a whole [76].

2.2.2 Volume Rendering

Johnson and Hansen [55] indicate that various methods have been developed that approach the rendering problem by focusing on visualising certain geometric primitives; and that most surface rendering techniques essentially rely on surface approximations in terms of how the surface geometric primitives are treated. Moreover, a whole dimension of the information is significantly underutilised when visualising volumetric data with the use of surface rendering techniques [15]. Based on this, volume-rendering techniques have been developed in order to process directly the volumetric 3D data and produce a 2D image.

In the study by Boucheny et al [7], it was identified that volume rendering techniques can be used to visualise information related to the spatial layout of the model, in terms of the supported primitives, and within a dynamic framework. This study aimed to measure scenarios for the volume rendering techniques for presenting clear depth indications in a dynamic context. Jänicke and Chen [51] also studied volume rendering techniques for processing 3D data sets in terms of 2D projections that demonstrate layered depictions of complicated 3D structures. The main problem encountered was in analysing the suitability of the various transfer functions that would lead to suitable visualisations of the various highly complex and overlapping structures in a 3D volume [6].

The impact of the choice of rendering techniques can be observed in Figure 2.3, which is based on various volume renderings of a cylinder [7]. Perlin noise was added to the cylinder, the data was stored as 3D textures and the rendering process had been conducted through 100 planes, with their accumulation being from back to front.

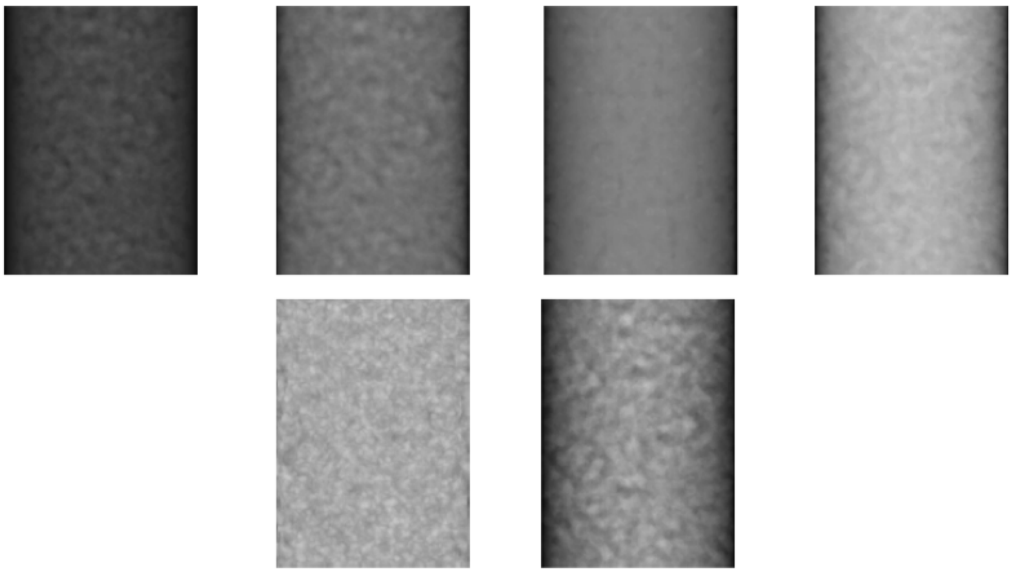


Figure 2.3: The difference images of the renderings of a cylinder produced with various volume rendering techniques [7].

2.3 Quantisation

Data in form of images, audio files, or videos are often processed for improving their suitability for a specific application, or for extracting information. In network communication, data transmit in the form of signals from a source to a destination. The received signals are processed, and then analysed to extract the information they carry. As signals can be of different types (analog or digital) their processing often involves their conversion from one type to another. An analog signal is a continuous type of signal in which data are represented by time varying quantities. Whereas, a digital signal is a discrete signal in which data are represented by a sequence of separate values at any given point of time.

Sometimes data is represented in the digital form but needs to be converted into analog, i.e. when a modem converts the digital data to analog signals which then are transmitted over telephone lines. In such situations Digital-to-Analog converters are required. Generally, in the real world, most of the time, signals are presented in an analog form i.e. temperature, light, and sound. However, in computing, information usually needs to be in digital form (1's or 0's) because digital computer systems require discrete binary information. For such purposes, Analog-to-Digital

converters are being used to convert an analog signal into a digital one, and making it thus usable by a computing machine [38]. On the other hand, if a signal is already presented in a discrete form (0s and 1s generated by the computer and then translated into voltage pulses that can be broadcasted through a wire), then the extraction of digital information from it is known as Digital-to-Digital encoding [32].

These conversion from Analog-to-Digital, Digital-to-Analog, and Digital-to-Digital are done based on application requirements and the suitability of the signal type for specific signal processing operations [38]. These conversion processes are based on sampling and quantisation. The spatial resolution of an image is defined by the sampling rate, while, on the other hand, the number of grey shades in an image representation are determined by quantisation level of the intensities. Quantisation is a mapping of infinite values (input) to finite and values (output). This way input values are digitised, that is, the input is transformed from continuous to discrete [41]. Gray and Neuhoff [41], states that the quantisation levels of the intensity values need to be high in order for the human visual system to be able to extract detailed information from the image.

Quantisation can also be seen as discretisation by the division of the allowable range into small parts. The level of the quantisation, that is, the number of parts into which we split the allowable range, is critical in a wide range of applications in image and signal processing. For example, in images, high quantisation levels are required for applications that require fine shading. Even in very fine quantisations, there will still be a difference between the input and quantised output, which is known as the quantisation error. A device or function that is used for quantisation is known as a quantiser. The simplest and most common form of quantiser is the zero-memory quantiser. Another class of quantisers is the sequential quantisers, including techniques such as the delta modulation, of the Differential Pulse Code Modulation (DPCM) systems, and their adaptive versions. Quantisation errors can introduce several knock-on errors in an algorithm, including underflow or overflow, computational noise, and rounding errors, which, as a consequence, lead to dissimilarities between the expected ideal and the computed numerical behaviour of a system.

Quantisation can be applied on any type of information. For example, audio quantisation takes an analog signal (i.e. a sound wave) and converts it into a sequence of discrete values, each one corresponding to the amplitude of each sample [86]. The range of the values of the amplitude levels depends on the number of bits, that is, the level of the quantisation. For example, 8-bit quantisation means 256 possible values. One of the classic examples of quantisation is CDs with music tracks, encoded in formats such as MP3. Quantisation has also become vital in the optimisation of deep learning models, as it increases their power of inference upon deployment in embedded systems.

Adaptive image quantisation involves the determination of various segments in the image, a method for assessing which parts of the image are less important and therefore, can be quantised more aggressively with minimal only loss of visual quality [18]. It should be noted however, that adaptive image quantisation is still a lossy process, which is thus reducing the quality of the image. Finally, regarding colour quantisation, there are various techniques which help to represent an image with the minimum possible number of different colours. In a typical application, the image would be converted to the GIF format, reducing the number of different colours to 256.

2.3.1 Mesh quantisation in compression algorithms

Even though still much less prevalent than the ubiquitous 2D images, 3D geometry in various formats, but essentially encoded by three dimensional coordinates, the x-coordinate, the y-coordinate and the z-coordinate, is now widely used in animations and 3D modelling applications. It is thus quite important to develop techniques for the quantisation and compression of such 3D geometry models, which will reduce the processing time and would require less memory [115].

Aiming at exploiting the high spatial correlation of adjacent mesh vertices, the current geometry compression techniques are based on the pre-quantisation of the vertex positions, the prediction of the quantised positions through various schemes utilising connectivity information, and the entropy coding of the prediction residuals. We note that the IEEE 32-bit floating-point numbers that are used to encode

uncompressed geometry data, that is, vertex coordinates, support a level of detail which is beyond the limits of the human perception and way more than the requirements of the vast majority of the common applications. Therefore, in such situations, quantisation can be safely conducted, without causing any impairment to the visual quality of the mesh.

Torkhani et al [103], studied experimentally the mesh distortions that are expected under realistic application scenarios, including the distortions caused from lossy compression algorithms and network transmission errors. The paper also studied distortions in 3D mesh animations, which are also used in several applications and which, often, are also undergoing lossy compression operations that would affect quality. This issue was also been discussed in [41]. Similar techniques can also be applied for the compression of general polygonal, rather than triangular, meshes.

As discussed by Peng et al. [81], compression techniques for single-rate mesh require the connectivity to be encoded in a lossless way, as the connectivity is regarded a discrete mesh characteristic which must be preserved. In contrast, encoding of geometry data is usually conducted in a lossy manner. The current mesh compression algorithms are able to encode the connectivity of the mesh using less than four bits per vertex, and their performance in terms of the achieved compression ratios, is close to the optimal theoretical lower bounds. In contrast, the encoding of the mesh geometry, that is the quantisation of the vertex coordinates, has not been the focus of previous studies, which almost invariably pick a certain quantisation level empirically. However, in a compressed mesh, the size of the data representing geometry is significantly larger than that of the connectivity, thus, more recent studies focus more on geometry coding.

2.4 Quality Measures on Images and Meshes

Quality measures for images are generally based on two main methods: the subjective method and the objective method [26]. The subjective method refers to human involvement for the quality evaluation of the images [119]. The objective method includes a computational process for calculating a measure of the image quality in an

automatic manner. In terms of the image quality metrics (IQM) within the objective methods, these can predict the perceived image quality in an automatic and fast manner. The objective IQMs have been classified into three main categories: those of full reference, reduced-reference (RF) and no-reference. In terms of the subjective IQMs, they are based on the judgement of human beings, which make them more reliable with respect to the quality of images. However, this procedure has been often criticised as irrelevant, slow, and expensive, when it comes to implementing it [101].

Despite their obvious drawbacks, the distinctive advantage of the subjective methods is again that they involve human users for quality evaluation of images, while the objective comprise of computational methods for image quality. Human beings are often the end users, especially in mesh animations. Thus, analysing the perceptual quality of 3D dynamic meshes is being the focus of research studies, and is consider a critical issue as discussed in [102]. This paper studied the perceptual quality of distorted dynamic meshes by collecting human opinion scores in a large-scale subjective experiment [12].

Quality assessment of an image is essential in many applications [12]. That study focused on automatic methods for computing quality scores that were validated by correlating them with the scores that are provided by human observers. [81] developed image quality assessment processes by studying user perception of structural data. As described by Weiskopf and Erlebacher [62], in many contexts, the generation of high quality images is a major challenge, especially when it comes to visualisation. Indeed, the suitability of image quality metrics for visualisation depends on many factors, including domain-specific requirements, the needs and expectations of the user, the source data and the acquisition and processing techniques that have been used [51].

The simplification method presented by Lindstorm and Turk [69] is based on the minimisation of the root-mean-squared difference with respect to the generated images from several views regarding the object that has been simplified in comparison to the original. In [94], a similar approach is employed, their process comprising of comparing the structural data of the distorted image against the original image.

In similar studies, the distinction between full reference, reduced reference, and no reference is crucial [90].

Image quality metrics compute a value for the image quality in an almost immediate manner. In a survey of image quality measures [101], the authors identified that the assessment of image quality is still a challenging task in terms of digital image processing systems, and Peak Signal-to-Noise Ratio (PSNR)

$$PSNR = 10 \log 10 \frac{255^2}{MSE} \tag{2.4.1}$$

is commonly used in terms of objective IQM, but does not always correlate well with validating results from respective subjective IQM. Here, 255 is the maximum grey level of a 8bits/pixel monochrome image, and MSE is the Mean Square Error. Several other objective image quality metrics have been developed as substitutes for the PSNR.

In order to control quality when various image processing operations are applied to an image, IQM could be inserted into the image-processing pipeline for the optimisation of the algorithmic and parametric settings. A study regarding the performance of image quality measures [26] conducted evaluations of greyscale images under compression. Figure 2.4 shows different versions of a Lenna image, starting from the original, and then encoded in the compressed JPEG, EPIC, and RLPQ formats. The difference in the quality of the image that can be visually observed in this context, can be quantifying by the various objective image quality metrics, which can thus be used to control the impact of the image compression algorithms through parameter selection.

In another related study [108], image quality measures had been studied in relation to the human visual system. Their study included metrics extracting and utilising structural information with respect to the viewing field. Such measures are particularly suitable in the analysis of image distortion. This is illustrated in Figure 2.5, which depicts the visual impact of various types of distortion on the same image.

In Figure 2.6, this aspect is further illustrated. The figure shows the impact of changing the values of various image processing algorithmic parameters, affecting brightness, contrast and other related factors.

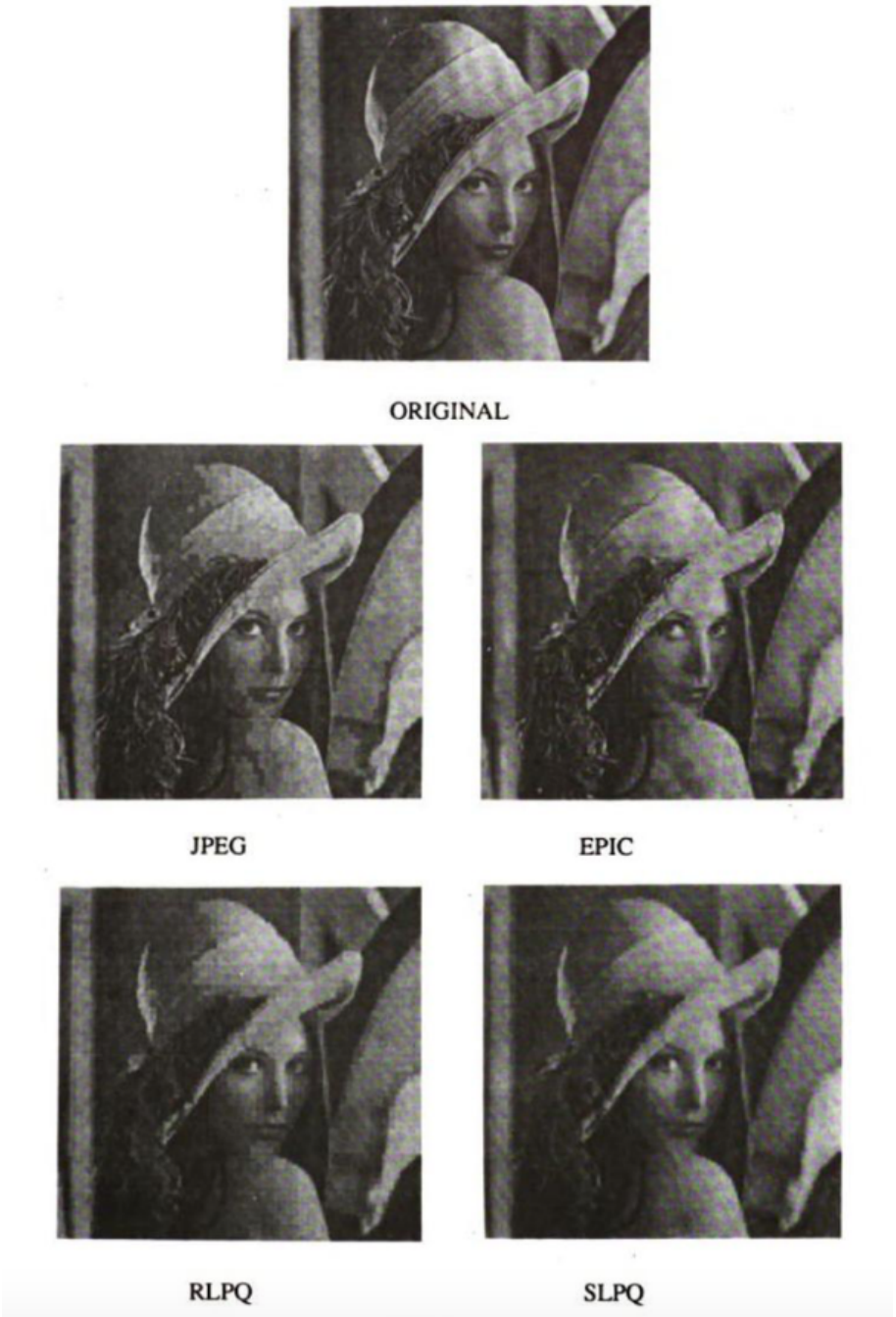


Figure 2.4: The Lena image encoded in various compressed image formats, [26].

2.4.1 Quality measures on meshes

In the context of mesh quantisation, there are again various approaches to the measuring of the quality of meshes. As per the image quality metrics, in meshes too, we distinguish between objective and subjective measures. Objective measurements often involve factors that are closely related to the ability of the mesh to produce

Figure 2.5: Comparing Boat images which have undergone various types of distortion: (top left) original image; (top middle) contrast stretched image; (top right) mean-shifted image; (bottom left) JPEG compressed image; (bottom middle) blurred image; (bottom right) salt-pepper impulsive noise contaminated image. See, [108].

images of measurably high quality through rendering processes. Standard objective mesh quality metrics are based on measures of smoothness, curvature, as well as the average shape of the mesh triangles. A second class of objective mesh quality measures are defined through objective measurements of the deviation of a mesh from an ideal mesh models. Subjective mesh quality measures are also used, usually based on colour renderings of the mesh.

When studying how noise affects the quality of a mesh, we investigate the relationship between the strength and type of the noise and the information extracted by the viewer, rather than the the noise itself. In order to improve the quality of a noisy mesh, mesh smoothing algorithms can be applied, creating smooth approximations of the original mesh data, which can then be used to obtain the essential patterns of the underlying data. Smoothing, by eliminating excessive mesh noise, can be a useful pre-processing step before quantisation. Figure **??** shows the effect

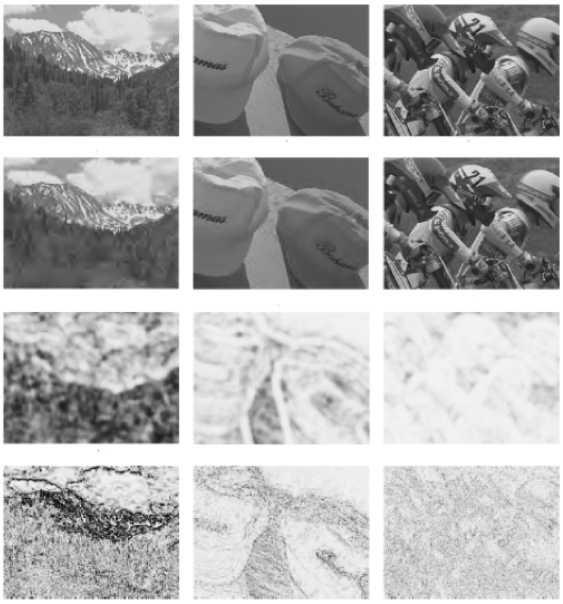


Figure 2.6: The impact of the parameter selection for various image processing algorithms can be quantified by Image Quality Metrics. See, [108].

of a smoothing algorithm on a noise mesh.

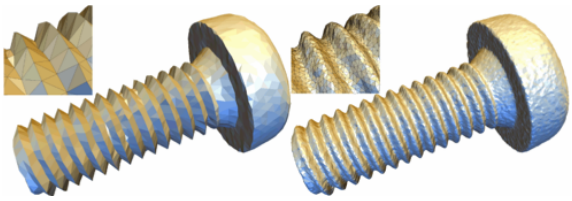


Figure 2.7: Smoothing a noisy mesh triangle, see [116].

In terms of quality measures on dynamic meshes, they are categorised by distortions that have the same strength on a given vertex over all frames, or the same strength over all vertices for a given frame, [102]. Figure 2.8 shows the distortions of the first type. Part a shows the original image, in part b noise was added weighted by a measure of local mesh roughness, and part in part c the added noise was inversely weighted by the same measure of roughness.

Several denoising methods, in order to maintain the quality of the mesh and its fidelity to the original data while reducing the noise, utilise mesh fairness. A related study [47] studied the effect of fairness through the implementation of several techniques mesh processing techniques on meshes that were created from a Kinect's

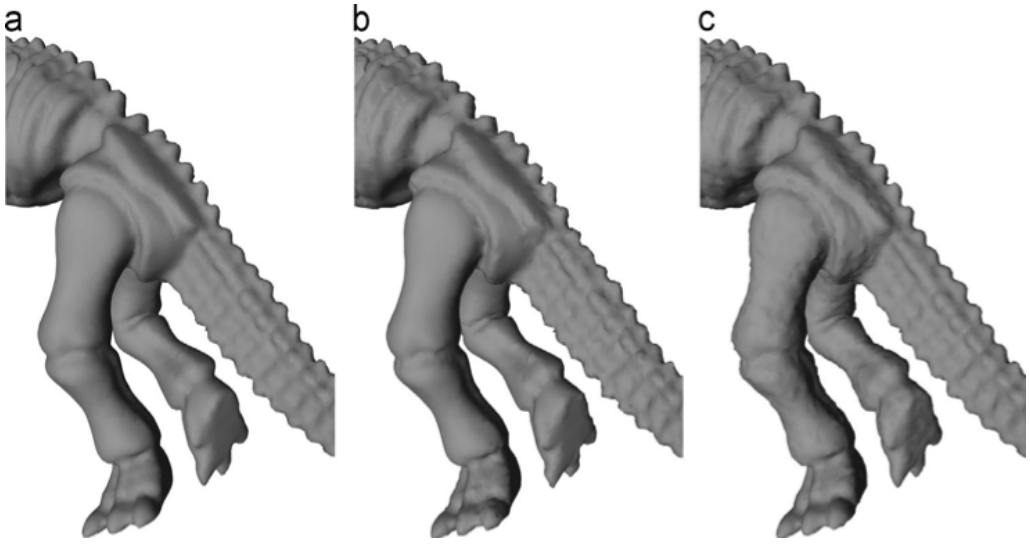


Figure 2.8: Impact of distortions by added noise on the quality of a mesh, see [102].

depth map of a person, see Figure 2.9. The figure clearly depicts the improvements in the fairness of the person's face, through the methods implemented in the study by Haque and Govindu, [47].

The quality measures that are related to curvature continuity, facilitate the improvement of the mesh by curvature-driven techniques. Such techniques work have the advantage that the new mesh remains close to the original surface and preserves quite well features such as normals and curvatures. Figure 2.10 depicts the differences between the outputs of two remeshing algorithms, one driven by surface curvature computations, and one where curvature is ignored.

In comparison with the quality measures that have been discussed above, the quality of a surface can be also be evaluated with the use of visualisation tools. Visualisation techniques for assessing surface quality include isophotes, reflection lines, and iso-curve. In terms of assessing curvature continuity and fairness, [108] recommends two major techniques; iso-curvature lines and lines of curvature. In a volumetric mesh, unlike the quality measures that have been discussed above, the use of visualisation tools for assessing mesh quality is restricted to analysing one layer at a time.

Reflection lines can also be used to evaluate the fairness of a surface. In CAD applications in particular, where reflection lines are widely used, they can be used to

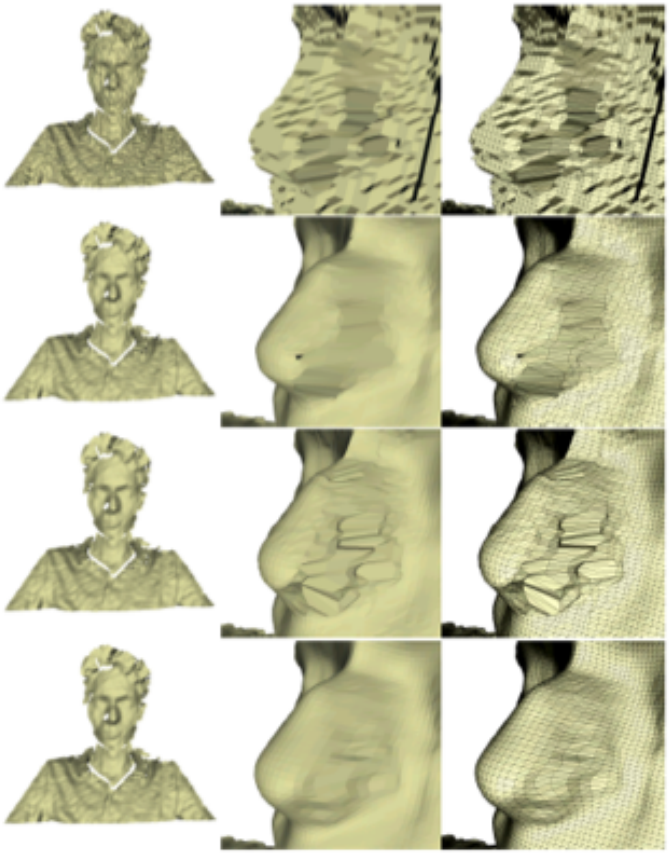


Figure 2.9: Mesh fairing by various methods, see [47].

infer an assessment on the quality of the shape design. Figure 2.11, from [103], shows an example of computation of reflection lines and their use for mesh optimisation. They also indicate the changes in the directions of the features of the mesh that are required in order to achieve the desired quality of the designed shape.

Regarding the method of isophotes as a visualisation tool [68], the technique is examining surfaces through lines of the same light intensity. Isophotes are considered a useful tool for visualising and highlighting local irregularities and defect, whose small size within the surface makes them difficult to spot under a wire frame or a shading based rendering of the surface. The first image in Figure 2.12, from [46], shows the original surface, while the other two are isophote visualisations.

The curvature based visualisation methods for assessing mesh fairness and mesh quality are based on two major techniques, see Wang et al. [107]. Iso-curvature lines and lines of curvature. The Iso-curvature lines are the lines of constant curvature on a surface. In [45], curvature lines and directions are utilised in order to compute

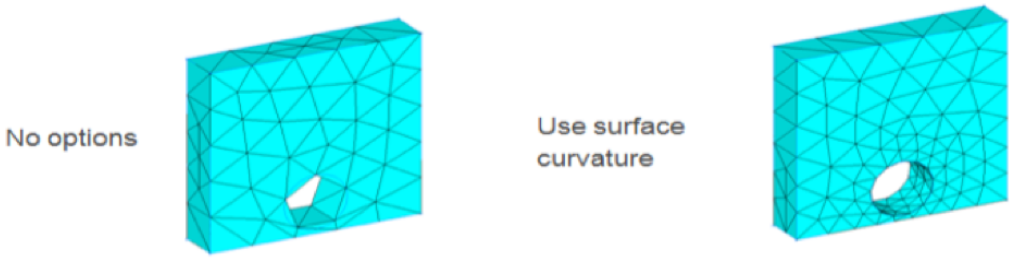


Figure 2.10: Outputs of two remeshing algorithms. Left: no curvature computations. Right: the remeshing process is driven by curvature computations. See [108].

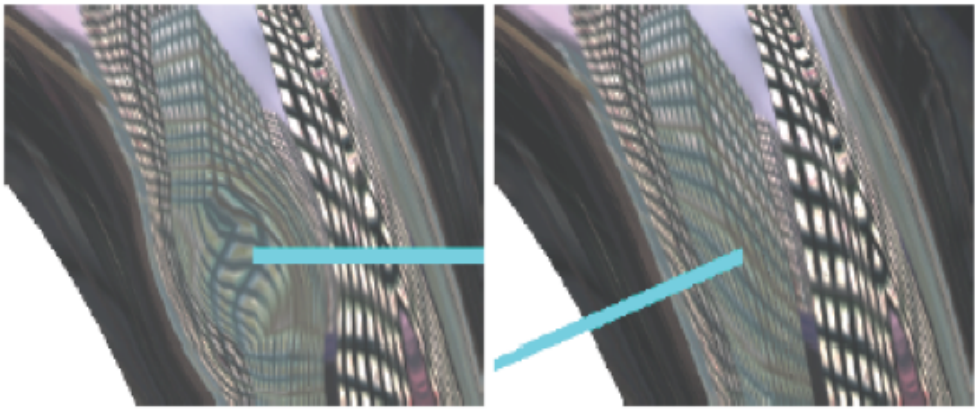


Figure 2.11: Reflection lines optimisation, see [103].

efficiently and then draw the surface's iso-curvature lines.

In [37], quality measures for optimal meshing of parameterised CAD surfaces were proposed. Their technique was based on the extension of distortion (quality) measures for planar meshes to the mentioned parameterised surfaces. The proposed distortion measures are computed with respect to the parametric surface nodes and their coordinates. In order to asses validity and the quality of a particular meshing of the parametric surface, the proposed distortion measure was utilised. The study also utilised this measure by minimising it on concurrent smooth and untangled surface meshes [44].

Various evaluation methods for 3D meshes directly utilise metrics for 2D images, including VDP [20], SSIM [109], RMS error and the Sarnoff VDM [54]. According to Lavoué et al [65], there are eight key attributes needed for the evaluating the visual qualities of a mesh and constructing a mesh visual quality metric. Namely,

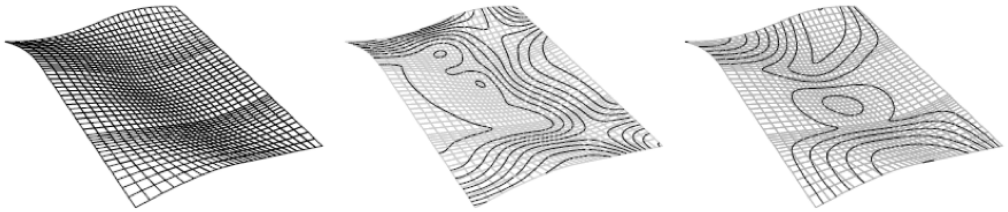


Figure 2.12: Isophote visualisations, see [46].

minimum curvatures, maximum curvatures, mean curvatures, Gaussian curvatures, dihedral angles, geometric Laplacians, Laplacian of the Gaussian curvatures, and 3D geometric position of the vertices. A combination of these geometric attributes can help in constructing suitable formulas for evaluating 3D mesh quality. These geometric attributes help in predicting apparent distortions that occur in processed 3D mesh models. It should be noted that such mesh visual quality metrics do not apply to textured meshes, and compute each of the above individual factors separately. It was deduced that a combination of the above eight mesh quality attributes can be quite effective in producing application appropriate mesh quality measures. According to the studies by Lavoué and Corsini [64] and [17], the definite factors that contribute most in accurate predictions of the mesh visual quality are the mean and maximum curvature, as well as dihedral angles. The least performing predictor of visual mesh quality is the geometric differences, with a correlation factor of 34%. Moreover, generally, the attributes that are based on Laplacians seem to give poor outcomes.

The high quality of the mesh triangles guarantees good behaviour of the mesh, especially when analytic properties of the mesh are studied, [66]. The key metrics that relate with the quality of mesh triangles are aspect ratio, skewness, orthogonality and smoothness. When a 3D mesh is being created, the average quality of its triangles can be enhanced by selecting the right design topology and utilising smoothing algorithms for evenly distributing all the elements and the features of the model. Perhaps the most important factor that should be maintained is the aspect ratio, that is, the ratio between the longest and the shortest edge length of a triangle. That would be always equal to 1 in the case of an ideal mesh consisting exclusively of equilateral triangles.

2.5 Types of experiments

2.5.1 Psychophysical experimental designs

In our context, the term experiment refers to process by which the responses of a set of users is collected and analysed, and the effect on dependent variables is measured with respect to changes in the independent variables. The term design refers to the specification of certain procedures with the ultimate aim to assess a hypothesis. Therefore, a suitable experimental design needs thorough understanding of the environment in which it will take place. First, the experimental design considers the variables and their relations within the environment in which takes place, followed by formulation of the hypotheses, and finally their testing. There are three main types of experimental design in the literature, based on different strategies: time based, choice based and standard scale based.

The task based experimental design is a systematic procedure performed under controlled conditions. The user is assigned a task, and the time to finish this task, under various factors, is measured. The design of such an experiment should identify dependent variables, outcome, usability goal, the independent variables, as well as any other factors that can be under our and manipulated directly [61, 114]. For describing the task, we need to identify what we think the users should be asked to do while using the system. Examples include, find the smallest object, or the object of a specific colour. The user is asked to perform the task and the time of completion, as well as the success or failure in the completion of the task are recorded [89]. A drawback of such designs is that an unsupervised user could deviate completely from the task, making simple statistical methods, such as the average time spent on the task, unsuitable for the analysis of the results. To give an example, the might needs to find a particular 3D object, but instead they are spending their time surfing on the site. Addressing such issues may lead to more complicated designs. However, overall, this type of design provides high level of control, it is quite straightforward to replicate, and the type of data that are usually collected are suitable for deep statistical analysis.

Staircase is one of several methods used for controlling the sequence of trials

presented to the participants of an experiment. It was first introduced by Dixon and Mood [22]. The method is quite straightforward, and will be used in our first experiment to compute discrimination thresholds, that is the lowest possible level of quantisation that give acceptable mesh quality. The reason for discussing studies that developed the staircase approach is to analyse this technique in order to properly understand its implementation regarding mesh quantisation.

In [67], they introduced a novel up-down procedure for the sequential choice of observations that will presented to the participants, which offered several benefits. Based on [67], which has in the context of psychoacoustic testing, several variations of the simple up-down procedure were introduced in the past decades. The main objective is efficient parameter estimation, that is, as few user trials as possible. The benefits associated with the up-down method include robustness, simplicity, high efficiency, relative freedom in the design, and small sample requirement.

Another research paper [83], evaluates the statistics of fatigue strength distribution, employing an up-down staircase method. In order to determine the fatigue strength and its standard deviation with as much accuracy as possible, the system's parameters and their effect were evaluated using the traditional staircase statistic with a large scale numerical simulation, which enabled the quantification of the standard deviation and the bias as functions of the staircase step size and the sample size. Emulating these designs, and employing a yes/no procedure, our first experiment will aim a identifying the difference, if any, amongst the stimuli that have been produced from 3D models quantised at various levels.

A two-alternative forced choice (2-AFC) is an experimental design for analysing the participants' responses through their selection of exactly one of two available choices, in a time constraint environment. In our case, a 2-AFC staircase will be based on the responses of the participants to the stimuli by a yes or no. In response to a yes answer, the intensity of the stimulus will be decreased by one, and the procedure be repeated iteratively, until the answer is no. In other words, for each yes or no, an increment of one or, respectively, an increment of one will be done, with the stimuli hovering back and forth around the threshold. In our experiment, a trial would be to present to the participant two images and ask them if they are

identical or not, expecting an either yes or no answer. In this staircase method, the values of the stimuli concentrate in the region around the threshold, thus making its estimation an efficient process. However, there are also a drawbacks related to the expected high variation in a subject's response to similar stimuli around the threshold [24].

As far as benefits of 2-AFC are considered, it is the simplest design based on the choice between two given options. In contrast to scaling methods discussed next, it offers a threshold measured in physical units. Furthermore, it can be better than the task based designs due to the unambiguity of the binary decision making process [36]. This binary nature also makes it an unbiased model, as it will be the case with our quality discrimination experiments based on the choice between a left and a right image. Moreover, 2-AFC is based on consistent behavioral outcomes hence, making it helpful in decision making modeling. Another advantage of using 2-AFC is that it can benefit from user performance enhancement, due to the fact that on each trial the subjects are provided with more factual information [36]. On the other hand, its performance mainly depends on the question presented to the users and their relevant assumptions during the procedure. That is, it can still be affected by the subject's personal preferences during decision making.

Generally, one of the most common types of experimental design, one that is widely being used in practice, is the scale based, where the response of the user is measured on a standard scale. In scale base designs, the responses correspond to the user's subjective attitudes, and they are not considered as absolute measures, but more of an instant estimation of the user's emotional state and intention. Moreover, the approach is uni-dimensional and with limited options for corrective analysis when, due to various misunderstandings, or lack of experience from the part of the user, it fails to calculate the correct user attitudes and intentions. Indeed, in most settings, the development of a universally accepted scale, one which will be easy to understand and implement, seems to be an elusive goal. on the other hand, scale based designs can evaluate the data in ordinal, nominal, interval, and ratio formats, offering thus some design flexibility [28].

The maximum-likelihood methods are designed to estimate the threshold as the

middle point of the psychometric curve, identified as the 50% level in a yes/no task, or as the 75% of presentations in 2-AFC. For more details we refer to chapter 5, where such methods were used. While the simplicity and flexibility of the staircase method make it the preferred choice in many instances, on the other hand, the maximum-likelihood estimation procedures might offer some benefits over staircases. In particular, the duration of the trial is usually longer with a staircase as the stimuli change by a fixed step size. As a result, robust estimation of thresholds is possible with just few trials in maximum-likelihood methods, while the staircases require a number of trials to produce the turnarounds required to calculate the threshold [95].

2.6 Software

various software tools were used in the implementation of the experiments and the analysis of the results. In particular, Meshlab and Mitsuba were used for producing and rendering the 3D meshes of the experiments, the interfaces of the experiments and the collection and analysis of the results was done in Matlab, while Palamedes was used in the analysis of the results of one of the second experiment.

Figure 2.13: Matlab, Meshlab and Mitsuba software.

Meshlab: The Visual Computing Lab of ISTI - CNR [13] developed a graphical front-end, which called Meshlab. The software provides a friendly interface which make it easy for the user to experiment with its tools interactively. It is an open source, portable, and extensible system, for editing and modifying 3D triangle meshes which can be downloaded from this link `https://www.meshlab.net/#download`. It provides tools for processing, cleaning and rendering meshes. It is

widely used in various technical fields that require 3D model development and 3D data handling.

We used the MeshLab software in our experiments for various purposes. From cleaning the 3D models, remeshing them, for simplification to reduce the number of vertices, subdivision to increase the number of vertices, and map colouring. The original models were downloaded from the Visualization Virtual Services (VVS) workbench.

Matlab: It is a programming language and numeric computing environment developed by Mathworks. It provide an environment where the programmer could write and run programs, plot functions, implement algorithms and create user interfaces.

Our experiments place a strong emphasis on the participant's behaviour and interaction. Therefore, we needed responsive, intuitive interfaces which would have been easy for the user to interact with. Thus, we used the Matlab to design each experiment's interface along with functions for collecting the data from the participants' trials, and in most of the experiments we used Matlab to write the code for analysing the results.

Palamedes: We used the Palamedes software to fit Psychometric Functions (PFs) to the data generated by the psychophysical experiment, using the Maximum Likelihood Estimation (MLE) method. For more details we refer to the chapter 6, section 6.3.1.

Mitsuba: It is a free open source, research oriented rendering system developed by the Realistic Graphics Lab at EPFL. It is a physical simulation based rendering system written in portable C++ and consists of a small set of libraries and more than 100 various plugins that implement functionality from support of various light sources and environments to the customisation of rendering algorithms. It comes with a graphical user interface, which allows the user to interactively explore the scene, and after choosing a viewpoint, various rendering techniques can be used to generate a high quality image of the scene. It also supports advanced rendering techniques such as volumetric rendering.

We used the Mitsuba software for rendering our 3D models with imported com-

plex complex environments, making the resulted images that were shown to the users closer to the images that one would expect to see in modern real-life applications.

Chapter 3

A user study on the effect of quantisation methods on thresholds of triangle meshes

3.1 Introduction

Triangle mesh is the ubiquitous shape representation for 3D graphics and visualisation applications. In their simplest form, they consist of a set of vertices, which are points in $\mathbf{R}^3$ connected between them by triangular faces. The encoding of the vertex coordinates most often makes use of 32-bit floats, however, the use of fixed-point with less than 32 bits per vertex coordinate is also common, especially when we want the triangle mesh in a compressed form. While strictly speaking geometry encoded at any finite precision, including 32-bit floats, is quantised, here following a widely accepted convention we refer to the process of transformation from 32-bit floats to fixed-point arithmetic as *quantisation*, to the resulted mesh as *quantised* and to the original mesh as *unquantised*.

The effect of the quantisation on the visual quality of the mesh naturally depends on the *quantisation level*, that is, the number of bits per vertex coordinate. While it is well-known that coarse quantisations often result to meshes of low visual quality, to the best of our knowledge there is no systematic study aiming at finding the minimum number of bits per vertex coordinate that are required for a quantised

mesh that will be visually indistinguishable from the unquantised. While there could be several possible explanations for the lack of study of this *quantisation threshold*, we note as a prominent one that the threshold seems to depend on several of the mesh characteristics in conjunction with the rendering algorithm used and that, generally, it should be considered as application dependent. A classic example where a quantisation level must be chosen outside the context of a specific visual applications the testing and evaluation of mesh compression algorithms. In early seminal papers such as [104], the quantisation levels range from 8 to 10 bits per vertex, while in some of the more recent approaches surveyed in [71], the standard quantisation level seems to be 16 bits per vertex coordinate. In [49], general, not necessarily triangle meshes were tested at quantisation levels ranging from 12 to 16 bits.

The experimental study of quantisation thresholds in this chapter focuses on the comparison between two different quantisation methods. The first is *rounding*, which sets all the bits above the quantisation level to zero. The second method is *dithering*, where all bits above the quantisation level are considered as having a random value. While the simplicity of rounding makes it the most commonly used quantisation method, dithering has the advantage that the randomised bits could represent encoded information in applications such as high capacity steganography [118]. Figure 3.1 shows an example of rounding and dithering at 8 bits per vertex coordinate.

To the best of our knowledge, in the literature there is no experimental comparison of the visual properties of the two quantisation methods. It is well-known that truncation creates blocky artifacts, which one could argue are easily detected by the eye while, on the other hand, dithering causes high frequancy noise which again the eye is sensitive to it.

The findings of the experiment are summarised as follows:

- Dithering has a higher threshold than rounding, that is, with dithering we need more bits per vertex coordinate to make the quantised model indistinguishable from the unquantised. The increase is small, around one bit per vertex coordinate in average in our experiments, but nevertheless statistically

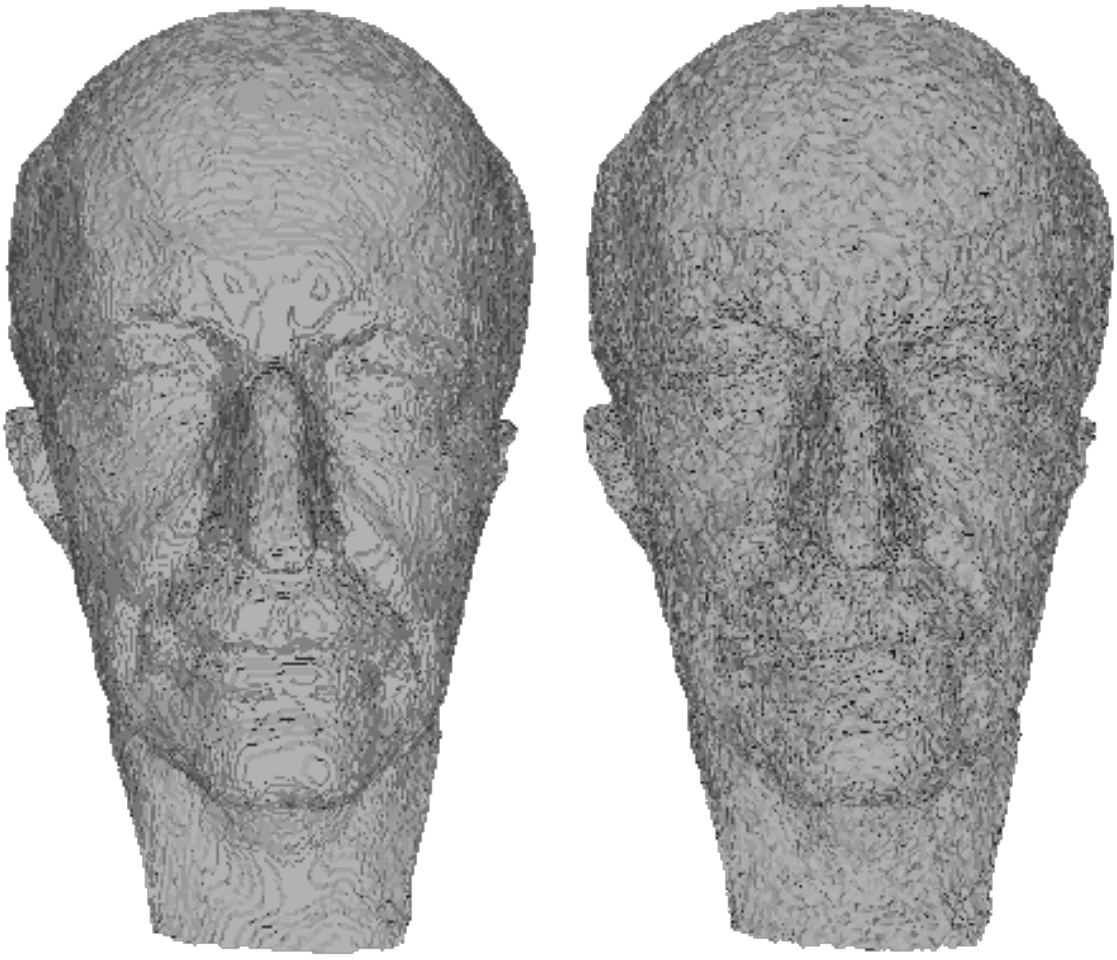


Figure 3.1: **Left:** Rounding at 8 bits per vertex coordinate. **Right:** Dithering at 8 bits per vertex coordinate.

significant. To the best of our knowledge, this is the first study establishing such a result.

- As expected, the characteristics of the mesh model affect the quantisation threshold. Regarding the type of characteristics affecting most the quantisation threshold, the first indications we have from our experiment suggest that the size of the model is more important than smoothness. In particular, larger models with many triangles and thus more detail require, as expected, more bits per vertex coordinate.

The main limitation of our approach is that we use only one rendering method. Moreover, by opting for the interface of experiment to be interactive the renderings

presented to the participants were of low quality, while on the other hand though it should be noted that our rendering setting, essentially Phong shading, is often met in real-world applications. The second limitation is that the set of models we used was limited to three models in total, even though their characteristics were very diverse. Overall, while we think that the comparison of the two quantisations methods was fair and broad enough to have limited only threats to the validity of the main finding that dithering has higher quantisation thresholds, the results regarding the effect of the mesh characteristics on the thresholds should be considered preliminary, and of course, the actual values of the thresholds computed in each case should be treated with caution as application depended.

3.2 Background

Quantisation techniques are most often studied in the context of signal theory [39]. According to an extensive survey of the technique in [41], rounding, which is historically the oldest example of quantisation and was first analysed in [96] for estimating densities by histograms. Dithering was introduced in [88] for improving the visual quality of a digitally encoded image by removing the visual artifacts caused by coarse quantisations of the grayscale range.

3.2.1 Perception

Subjective experiments have been employed by various researchers studying 3D model visual quality degradation under common mesh manipulation processes such as lossy compression [111], or watermarking [16]. More recent work utilises large databases containing meshes that have undergone a variety of distortions including compression, lossy tranmission and noise addition [63, 98], while in [102] dynamic meshes are considered. The types of mesh distortions studied in those papers are not as simple and natural as the vertex coordinate quantisations of our case, and the ultimate aim there is not just a comparison between two specific distortions, but rather the development and validation of metrics of visual mesh quality which can then be computed automatically.

We are not aware of any systematic experimental comparison between the quantisation thresholds of rounding and dithering. It is of course well-known that rounding creates blocky artifacts, which could be easily detected by the human visual system. On the other hand, dithering causes high frequency noise which human observers are also sensitive to. With 2D images, blockiness in the form of averaging of pixel values over a given area, is known to decrease visual recognition performance [48]. Similarly to the 3D model case, such blocky artifacts may be the result of certain lossy image compression algorithms. Although dithering in the form of added noise also degrades performance in many visual tasks [80], it can actually improve performance when added to a blocky 2D image: the added noise disrupts the high-frequency edge structure of the blocky image, making it easier to recognise [73].

The perceptual effect of quantisation of a 3D model is, of course, more complex. The stimulus for the human observer is not the quantised model itself, but a 2D image that is a result of a rendering process. As such, the perceptual effects of quantisation depend on the rendering algorithm and, eventually, how blocky the result appears depends at least partly on how good a job the rendering algorithm does in smoothing out the quantisation effects. On the other hand, the noise introduced in dithering might itself be highly visible to the observer, possibly making the quantised version perceptually even more dissimilar from the original.

Given that blockiness resulting from vertex coordinate rounding and high frequency noise introduced by dithering are both causes of visual degradation, it was difficult formulate a firm hypothesis prior to the execution of the experiment on how the quantisation thresholds of the two methods compare. Instead, we expected statistically non-significant differences as the most probable outcome of the experiment and lower dithering thresholds as the second most probable outcome, given the cues we had from the literature on possible visual improvement of images through dithering. While the eventual outcome of the experiment was the opposite, i.e.,lower thresholds for rounding, it should be noted that we did not compare the general visual quality of the two quantisation methods but something rather more specific, i.e. the indistinguishability thresholds.

3.3 Experimental Design

In each trial of the experiment the participant was presented with two meshes, one unquantised at the left hand side of the screen and a quantised one at the right hand side of the screen. The participant had to decide if the two meshes were different or not by giving a Yes/No answer to the question *Do the two meshes look the same?*. The interface of the experiment was interactive, allowing the user to use the mouse to grab any of the two meshes and rotate them, or zoom in and out of them. All implementation was done in Matlab and a screenshot of the interface is shown in Figure 3.2.

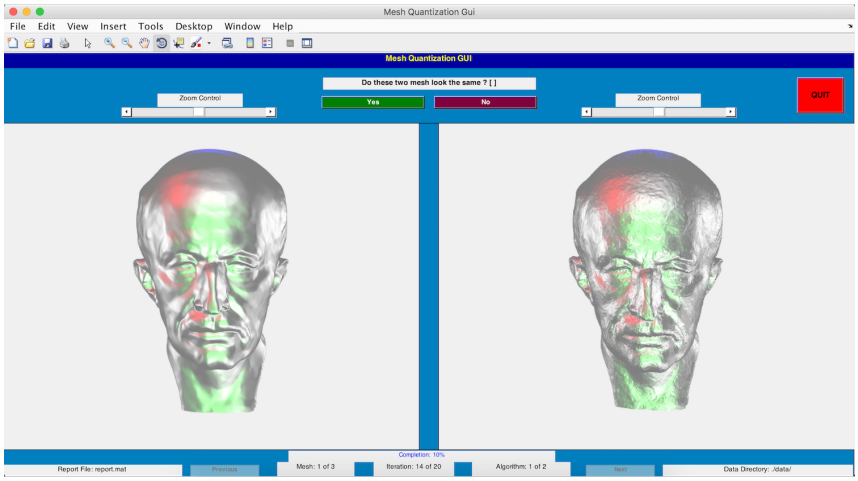


Figure 3.2: The interactive interface of the experiment.

The three meshes, chosen primarily for their large variation in size, are shown in Figure 3.3. The smallest was the *Cube* with 766 vertices, the *Eight* with 15K vertices was chosen as mid-sized and the *Max-Plank* model with 100K vertices as large. We also note that there is significant variation in the natural characteristics of the models: the Cube is a CAD model with sharp features, the Eight is an analytic model that is very smooth and has non-trivial topology, while the Max-Planck is a natural model which contains both smooth areas and sharp features.

The three models and the two quantisation methods created a 2-dimensional space of six in total conditions. For each condition the participant was presented with 20 trials meaning 120 trials in total. The order in which meshes were presented

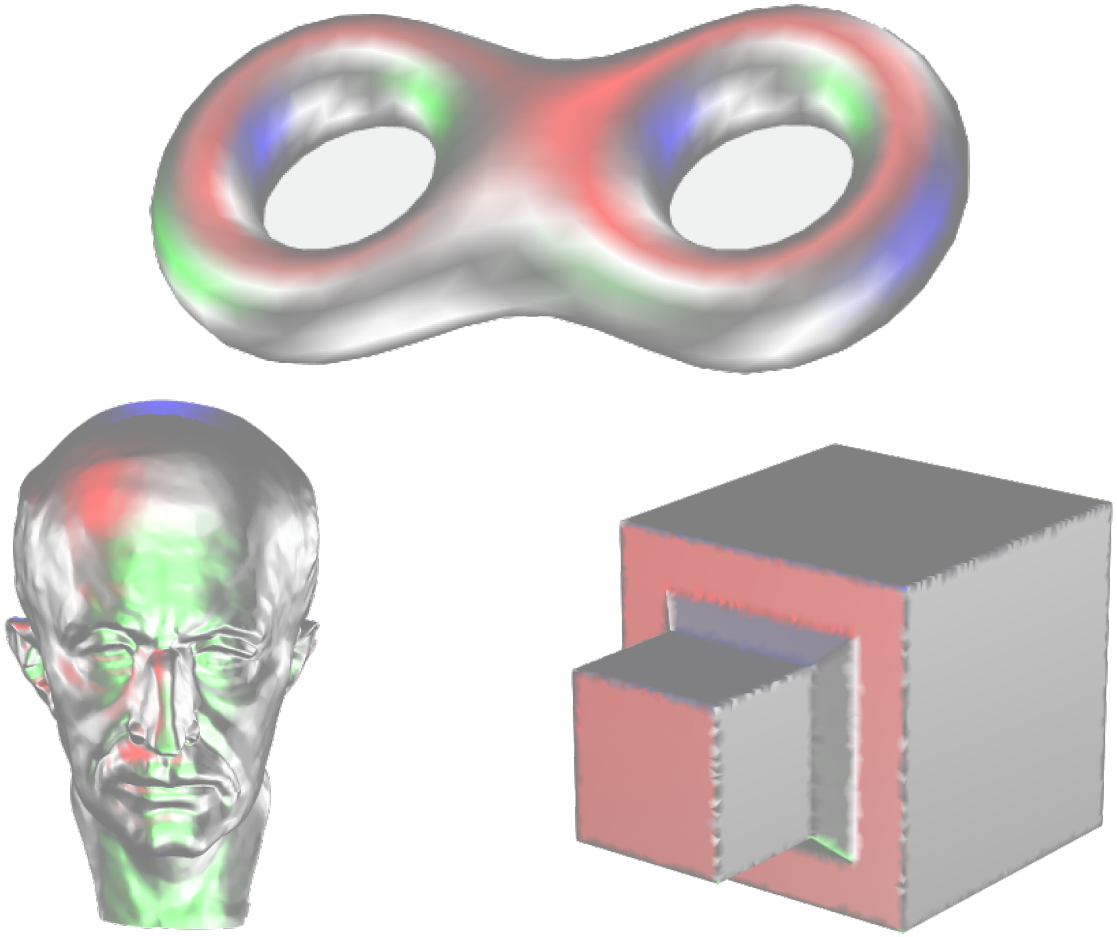


Figure 3.3: The models used in the experiments were the Eight, the Max-Planck and the Cube.

was fixed as Cube, Eight and Max-Planck, while the order in which the quantisation algorithms were presented was random. After a Yes answer, meaning that the participant was perceiving the two models as looking the same, meaning in turn that the quantisation level was on or above the threshold, quantisation level of the next trial was decremented by one. After a No answer the level of quantisation was incremented by one. As it has been established in the literature [14,92], in this type of experiments it is useful to start a staircase series of trials as near to the actual threshold as possible. Therefore, we established rough estimates of the thresholds by running a pilot and then the set of 20 trials for each condition was starting at these estimated thresholds. For example, for the Maxc-Planck model the initial threshold estimated by the pilot was 12 bits per vertex coordinate for either of the

two quantisation methods. Figure 3.4 shows a series of renderings for the dithered quantised Max-Planck model around the initially estimated threshold.

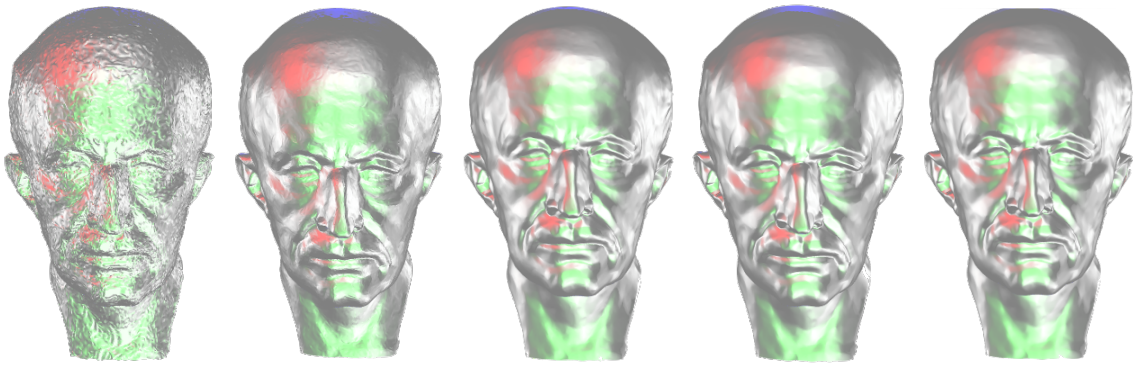


Figure 3.4: **From left to right:** The Max-Planck model at dithered quantisation levels of 8,10,12,14 and 16 bit per vertex coordinate.

The pilot was conducted in November 2016 at Durham University while the main experiment was conducted in January 2017 with participants students from Qassim University, Saudi Arabia. Ethical clearance for the experiment was obtained from Durham University.

The experiment conducted in a computer lab with 19-inch screen and a resolution of 1280 x 1024 pixels. The participants viewed the screen from a distance of 50 cm facing a desktop. The lab has a natural light and quiet so no distraction could affect the process. All participant involved in the experiment were computer science students. At the beginning of the experiment the participants were signing consent forms and were given a brief oral introduction to the purpose of the experiment. Next, they were presented with a pre-trial using a mesh that was different from the three meshes of the main experiment before, finally, being presented with the main experiment. There were no time limits for any single trial, or for the whole experiment, and no timings were recorded, however, all participants completed the experiment in around 30 minutes. Data from twenty one participants in total were collected and analysed, but as we discuss in Section 3.4.2, data from one participant were excluded as outliers.

3.4 Results

For each participant and for each of the six conditions of the experiment we compute a point estimate of the quantisation threshold, which is not necessarily an integer number, as follows. From the corresponding set of 20 Yes/No trials we exclude the first five. The exclusion of a number of initial trials is for allowing the staircase to reach the threshold and is recommended in [14], if we did not exclude them, the average would depend on the arbitrary choice of the starting value. The estimate of the quantisation threshold is then computed as the average of the first two peaks and the first two valleys.

Next, we screened the results for possible exclusions of outliers. This step is highly recommended, not only in subjective but also in physical experiments [19]. In a user study, screening for outliers can lead to the exclusion of participants from all or parts of the analysis, or to the exclusion of results associated with parts of the experimental dataset [82]. In our case, participant number 16 was found to be above the average quantisation threshold by more than two standard deviations for four out of the six conditions and was excluded from any further analysis. We believe that this participant systematically overestimated the threshold by a high margin due to a misunderstanding of the instructions. There were three more participants that were outside the ± 2 standard deviation zone in one of the six conditions, but they were not excluded. We note that here we did not follow the empirical recommendations of ITU [1] protocol for participant exclusion, firstly because their recommendation does not explicitly cover the format of our experiment, i.e. a Yes/No staircase, and secondly because it seems to be very strict when the data are not deemed normally distributed in which case the outlier zone is $\pm\sqrt{20}$ standard deviations.

3.4.1 Normality tests

Table 3.1 shows the results of Shapiro-Wilks normality test for each condition. We notice that in four out of the six cases the data are classified as non-normal and the non-normality can be the result of either positive or negative skewness.

Figure 3.5 shows frequency histograms for the Truncated Cube and the Dithered

	S-W p-value	skewness
Cube Trunc.	.006	1.299
Cube Dith.	.003	1.389
Eight Trunc.	.376	.032
Eight Dith.	.018	-1.182
Max Planck Trunc.	.539	-.222
Max Planck Dith.	.001	1.914

Table 3.1: The results of the Shapiro-Wilks normality test and the skewness of the distributions

Max-Planck models. In the case of the Cube, which has a low number of vertices and thus low quantisation threshold, the non-normality can be attributed to a naturally one-sided distribution of the observed thresholds. That is, the left tail of the distribution is very short because it was quite unlikely that a participant would underestimate considerably the threshold. On the other hand, for higher quantisation thresholds as in the case of the Max-Planck model, the high skewness value seems to be the result of outliers.

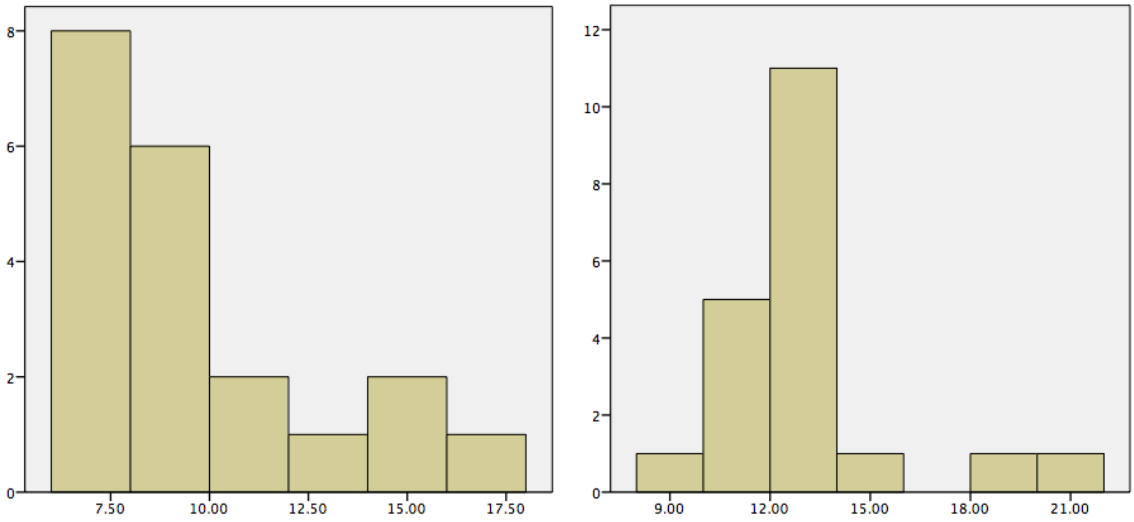


Figure 3.5: **Left:** The frequency histogram of the estimated thresholds for the Cube with truncation. **Right:** The frequency histogram of the estimated thresholds for the Max-Planck with dithering.

3.4.2 ANOVA test and post-hoc analysis

Since ANOVA tests are considered robust under non-normality conditions, we proceeded with a 2-way ANOVA test. The quantisation method is significant with $p = 0.045$ and $F = 4.094$, while the mesh model is significant with $p < 0.001$ and $F = 11.248$. Figure 3.6 shows the averages for each condition of the experiment and we notice that there is a small but consistent across the three meshes difference between the average thresholds of the two quantisation methods.

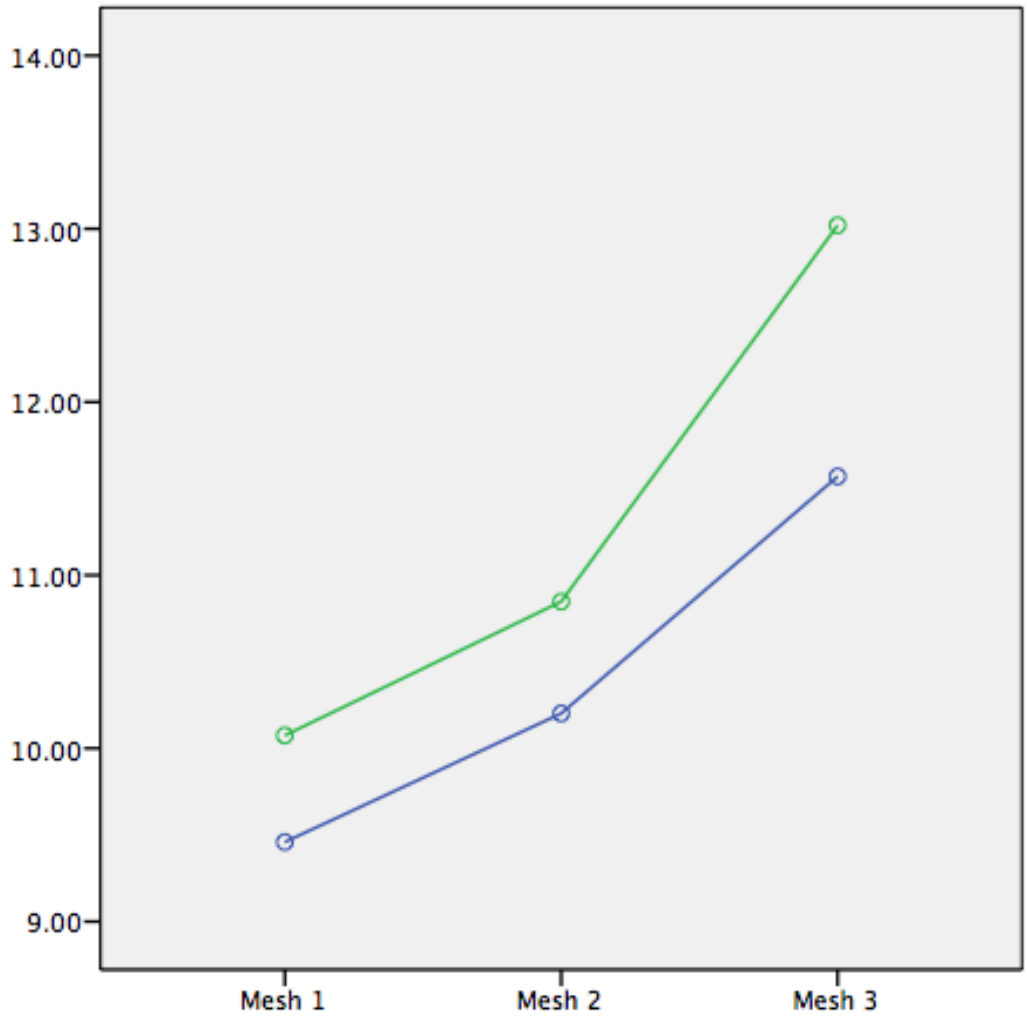


Figure 3.6: The means for each mesh for truncation as Eight, Cube and Max-Planck (top line) and dithering (bottom line).

Finally, in a post-hoc analysis of the results we performed pair-wise comparisons between the three meshes after collapsing the quantisation method variable. Figure 3.7 shows boxplots for the three meshes. The difference between mesh 1 and

mesh 2 was not statistically significant with $p = .06$ for the Bonferroni correction test. On the other hand, mesh 3 was significantly different from mesh 1 and mesh 2 with $p < 0.001$ and $p = 0.005$ for the corresponding Bonferroni correction tests.

While the focus of the experiment was on the comparison between the two quantisation methods and thus, it was not designed to answer questions regarding the effect of mesh characteristics on the quantisation threshold, we note that the results indicate that the size of the mesh is the most important factor in determining the quantisation threshold.

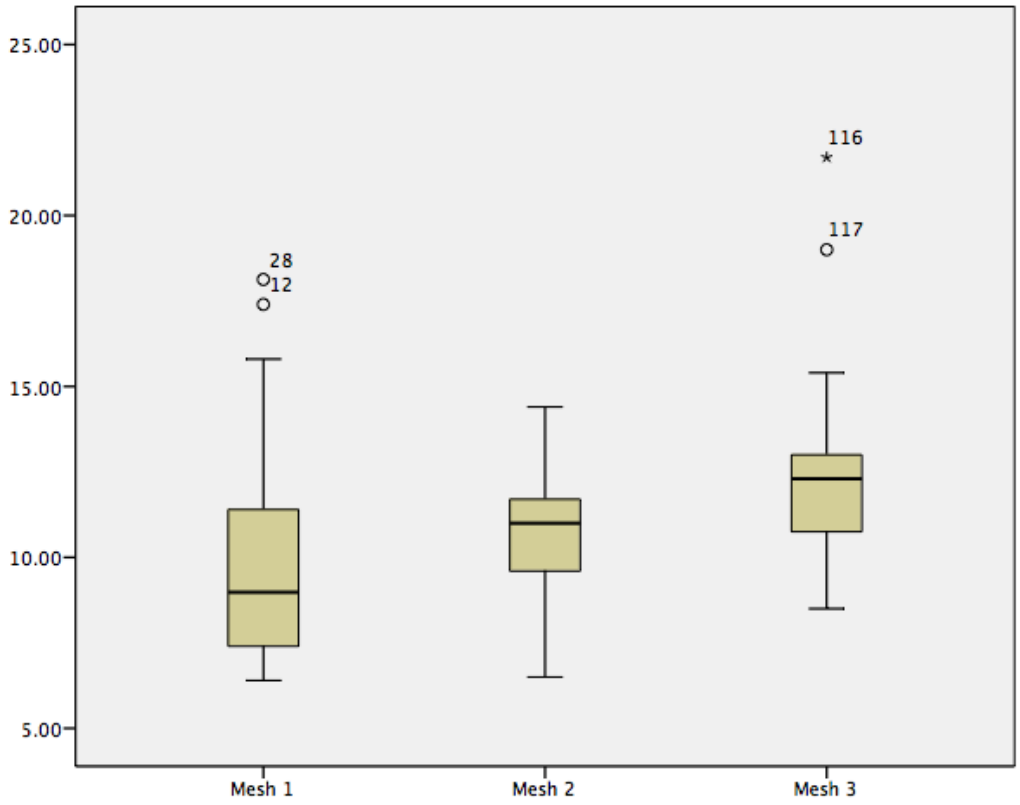


Figure 3.7: The boxplots of the meshes after collapsing the quantisation method variable.

The results for the other conditions are shown in the Appendix A.

3.5 Discussion and Conclusion

We presented an experimental study of the quantisation threshold of triangle mesh vertices, above which a quantised mesh becomes visually indistinguishable from the

original unquantised. The focus of our study was the comparison between two quantisation methods, rounding and dithering, and our main finding was that dithering has a higher quantisation threshold than rounding. While that result does not contradict any prior findings of the existing literature, we note that it could not have been easily predicted before the execution of actual experiment since, in the particular setting of 3D model quantisation, it was not known a priori whether blockiness or high frequency noise would prove to be perceptually stronger.

This experiment focused on the comparing between two well know quantisation algorithms and thus didn't considered any rendering style. Next, we plan to use the results of this experiment to inform the design of a larger experiment, aiming at confirming the correlation between the amount of geometric information in a mesh, as measured by the number of triangles, the discrimination thresholds, and properties of the rendering algorithms.

Chapter 4

A user study on the relationship between the quantisation threshold and the characteristics of the triangle meshes

4.1 Introduction

The advances in $3D$ mesh representation have widely developed to the point that they are employed in several mass-market applications, including networked $3D$ games, and $3D$ visualization applications. Triangles meshes, due to reasons such as, ease of acquisition, manipulation and visualisation, have become the ubiquitous standard in 3D geometry representation.

3D triangle meshes are always quantised. While the standard formats use a floating point arithmetic of high precision, they quite often are transformed to fixed point arithmetic, of relatively low resolution. This can happen through mesh manipulation algorithms. The most characteristic example is mesh compression, where the entropy of the vertex position information is reduced using predictive encoding based on the parallelogram rule [49], which however works at its best after the vertex positions have been quantised to levels between 8 to 24 bits per coordinate. After decompression, the vertex coordinates are still essentially encoded with fixed-point

arithmetic, even though, technically, the format they appear in might be that of floating-point arithmetic.

The level of the quantisation, that is, the number of bits per vertex coordinate, is obviously critical for the visual quality of the mesh. By choosing a sufficiently high number of bits per vertex coordinate, the quantised mesh will be visually indistinguishable from the original mesh. On the other hand, much higher levels of quantisation, well beyond the discrimination threshold, will increase unnecessarily the memory and computational overheads for storing, manipulating and visualising the triangle mesh, without any tangible benefit in terms of visual quality.

In this chapter we present an initial study of the relationship between the quantisation level threshold, beyond which the quantised mesh is visually indistinguishable from the original, and some fundamental characteristics of the mesh, such as its number of triangles, as well as the choice of rendering method for visualising the 3D model.

Figure 4.1 shows one of the meshes used in the experiment at nine different quantisation levels, from 6 bits per vertex coordinate to 14 bits per vertex coordinate.

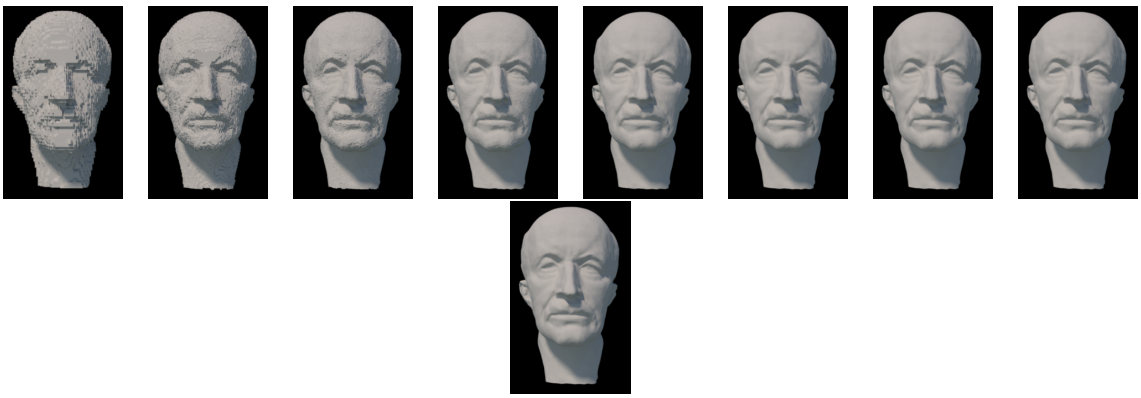


Figure 4.1: The Max-Planck model at resolution 100K diffuse rendering. **Left to Right:** from 6 bits per vertex coordinate to 14 bits per vertex coordinate.

Our study is based on a simple two-alternative, forced-choice psychophysical experiment, where two versions of the same model are presented — the original and one which is quantised at a certain level — and the user chooses the one with the

higher visual quality. We used three models in total, two of them are the Max-Planck model at two resolutions, 100K and 5K triangles, respectively, while the third is a spherical model with 5K triangles. The aim is to establish whether the geometric complexity of the model, as manifested by its number of triangles, as well as its regularity as manifested by various measures of triangle shape quality, affect the quantisation threshold. We also used two rendering methods one of which had a much higher specular component than the other.

Contribution: The experiment established that meshes with larger number of triangles require higher level of quantisation. As an intuitive explanation for this, we note that higher resolution models have more detail, which require higher quantisation levels to be represented with fidelity. Moreover, larger number of triangles resolutions means smaller triangles, the normals of which are more sensitive to the spatial perturbations caused by the quantisation and thus, the rendering process, which is based on normal information, is affected more severely. Moreover, as expected, renderings based on reflectance models with higher specular component require higher quantisation level, the reason again being that normal perturbations are easier to perceive in glossy rendering with a strong specular component.

Limitations: The main limitation of the study is its relatively limited scope. While we were able to demonstrate the existence of statistically significant relationships between quantisation thresholds, geometric characteristics of the mesh and rendering settings, we could not quantify them, that is, we did not have enough data to produce a formula relating, for example, the number of triangles in the mesh with the quantisation threshold.

4.2 Background

Vertex coordinate quantisation is the first step of all mesh compression algorithms [71]. Indeed, geometric information encoded in the least significant bits is visually redundant and moreover, has very high entropy since essentially it is noise, and thus, it is incompressible from an information-theoretic point of view. The use of 16 bits per vertex coordinate seems to be emerging as the defacto standard, as far as mesh

compression is concerned [71].

Quantisation techniques are most often studied in the context of signal theory [39]. According to an extensive survey of the technique in [42], rounding is the most commonly used and historically the oldest example of quantisation. It was first analysed back in 1897 [97] for density estimations via histograms. Here, we use rounding for the quantisation of the spatial coordinates of the mesh vertices. We note that quantisation techniques applied on the various frequency domain representations of the mesh geometry, see for example [99], might have significant theoretical interest, as well as significant applications, but are nevertheless less relevant to the everyday real-life use of meshes.

A previous study by the authors for determining the visual effect of quantisation, used a simple yes/no task experiment and was aimed at determining a discrimination threshold beyond which the quantized mesh is not perceived to differ from the original [2]. However, the focus there was on understanding the effect the choice of quantisation method has on the threshold, focusing in particular on the effect of dither.

The main alternative methodological approach to the subjective psychophysical experiment we have chosen here, would have been the use of an objective mesh quality metric. Various such metrics have been proposed in the literature, measuring mesh quality based on criteria, such as size, shape, and smoothness, [5], [33], [34], [60], [58]. There are also various mesh quality metrics which are computed as averages over the whole mesh of a single triangle quality metric. Examples of such metrics, which are often used in practice for mesh optimization, include: edge length ratio [59], area [107], edge length root mean square [59], inverse mean ratio [74], and aspect ratio [23].

We note that the use of objective metrics, such as those mentioned above, seems to be a more appropriate methodology in the context of CAD and Finite Element Method applications, while in the context of visual applications they are mostly employed as a cheap alternative to the systematic user studies. For example, as Vanhoey et al. [106] stated, for two main reasons, only a few subjective studies have been performed in the field of interactive visualisation: firstly, it is a relatively

new field with less than twenty years history; secondly, perceptual experiments are expensive and time-consuming processes. In some of the first studies based on psychophysical experiments, Rushmeier et al. [93] study the effect of geometry and texture resolution on the perceived quality, however, all their models are unquantised and the geometric resolution of the model is its number of triangles. In Rogowitz and Rushmeier [91], a psychophysical experiment is conducted to establish perceptual differences between animated 3D models and 2D still renderings of them. Away from triangle meshes, [77] conduct a psychophysical experiment to study the effect of the wireframe and the texture resolution on the perceived visual quality of wireframe models.

The experiment presented in this chapter is based on a two-alternative forced-choice design. The simplicity of the method makes for experiments that are relatively simple to design and run, and fare favourably compared to other more complex designs, especially when the number of participants is small [112]. In our context, the main benefit from the simplicity of the experimental design is that it minimises the risks to the validity of the results from any misunderstandings, or subjective interpretations of the tasks by the non-expert participants. Indeed, we have conducted analogous experiments based on the more complex *Maximum Likelihood Difference Scaling* design, [72], [9], and we found that although specific users would return meaningful results, the statistical aggregation of all users was not possible because of the large proportion of participants who either did not understand the task, or interpreted it in their own subjective way. Another possibility regarding the type of the psycho-physical experiment would have been the use of a task based experiment, as for example in [110], where the users are asked to recognise, as fast as possible, 3D models presented to them at various resolutions. We note that such designs are rarely used in the assessment of perceptual quality of 3D models, as they cannot detect very fine grain differences and, moreover, they suffer from high variance between participants.

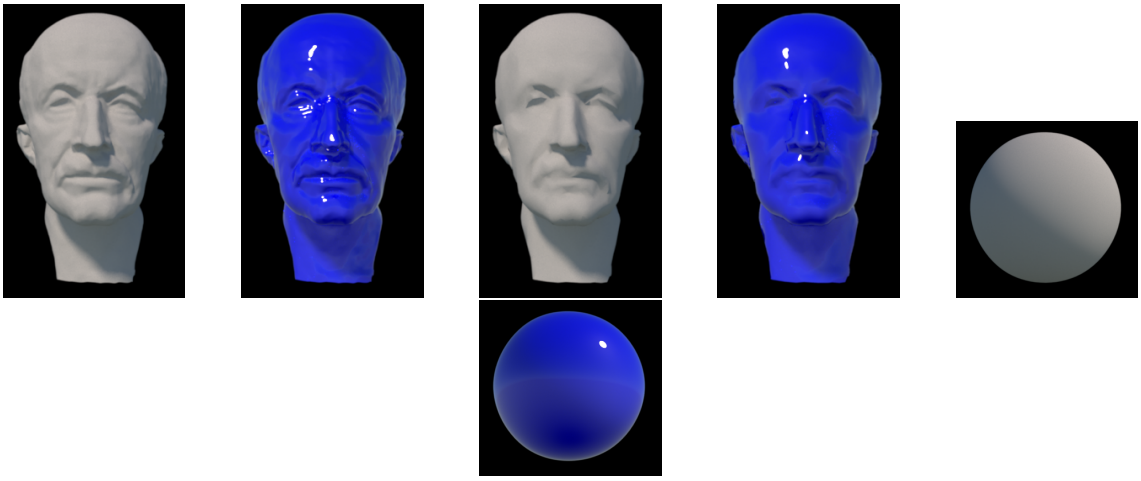


Figure 4.2: The six conditions of the experiment. From left to right: the Max-Planck at resolution 100K and 5K triangles and the sphere. Diffuse and specular renderings.

4.3 Experimental Design

The experiment is based on three different triangle mesh models, each one rendered with two different methods. For each of the six model/rendering method combinations shown in Figure 4.2, nine different stimuli were produced, each one corresponding to a different quantisation level of the triangle mesh, going from 6 bits per vertex coordinate to 14.

The first two models were the Max-Planck model with 100K triangles, and a coarser version of it, decimated down to 5K triangles. The aim is the comparison of the discrimination thresholds of these two models, which will show the relation between quantisation thresholds and model detail as expressed by the number of triangles in the mesh. The third model is a sphere with the same number of triangles as the coarse Max-Planck model. The sphere model is geometrically more regular than the other two, and consists of triangles that on average are very close to equilateral as shown by the average edge ratios, which are 0.7 for the Max-100K model, 0.6 for the Max-5K, and 0.9 for the Sphere-5K. The comparison of the thresholds between the Sphere-5K and the coarse Max-Planck model Max-5K will give an indication of how objective mesh quality metric relate to the quantisation

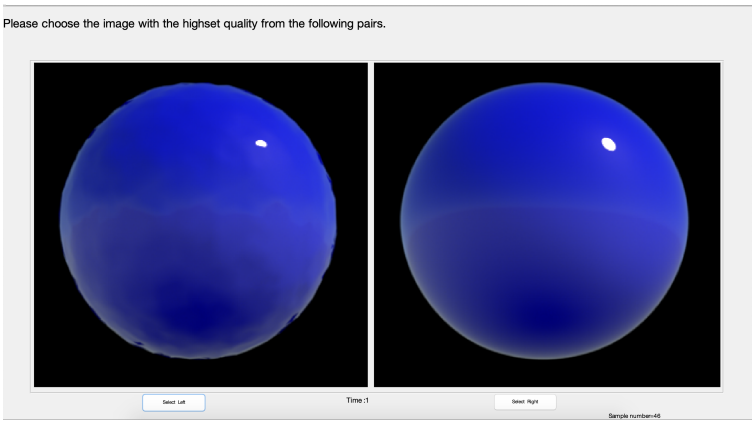


Figure 4.3: The interface of the experiment.

thresholds.

The first rendering method has a high diffuse component, therefore, it is less sensitive to small perturbations of the model's normals that are caused by the quantisation. The second rendering method has a high specular component, resulting into a rendered surface with the characteristic glossy plastic appearance caused by the mirror-like reflections of the specular component. Since normal perturbations are the main source of visual mesh degradation [117], and since a glossy rendering method is even more sensitive to normal perturbations than a diffuse one, we expect that it would make quantisation artifacts easier to detect.

Each of the 54 stimuli were presented on a 1280×725 pixels screen and a resolution of 2560×1600 pixels. The observer was presented with quantised alongside unquantised stimulus from a distance of 50 cm, in a two-alternative forced-choice experimental design, and in each trial, each participant was asked to choose the model with highest quality. Figure 4.3 show the interface of the experiment. The experiment conducted in a room with natural light and there is no distraction could impact the procedure.

Each of the 54 stimuli, was presented to the participants alongside the corresponding unquantised stimulus, in a two-alternative forced-choice experimental design, and in each trial, each participant was asked to choose the highest-quality model. See Figure 4.3 for the interface of the experiment.

Participant number	Max-100K		Max-5K		Sphere-5K	
	Diffuse	Plastic	Diffuse	Plastic	Diffuse	Plastic
1	0.7167	0.8222	0.5222	0.7556	0.6222	0.6556
2	0.6667	0.7889	0.5556	0.6944	0.6222	0.6389
3	0.6500	0.7944	0.5778	0.7222	0.5833	0.6556
4	0.6611	0.8167	0.5611	0.7000	0.5833	0.6389
5	0.6389	0.8056	0.5667	0.7167	0.5889	0.6500
6	0.6833	0.8167	0.5222	0.7222	0.5833	0.6500
7	0.6889	0.8111	0.5500	0.7389	0.6000	0.6667
8	0.6778	0.8056	0.5389	0.7111	0.6111	0.6444
9	0.6722	0.8000	0.5611	0.7278	0.5944	0.6556
10	0.7111	0.8278	0.5500	0.7444	0.5722	0.6778
AVG	0.67667	0.8089	0.5506	0.7233	0.5961	0.6534

Table 4.1: Probability of correct answer per condition of the experiment.

Before the experiment the 10 users were given instructions on how to conduct the experiment, and were given the opportunity to test the interface with models different from those of the main experiment. For each of the six model/rendering method combinations the nine quantisation levels were presented to the user in a random order and the place of the quantised model in each trial, that is, at the left or at the right of the screen, was again chosen randomly. For each trial the user had 5 seconds to respond, after which the screen would go blank while awaiting the user's response. Each participant repeated the entire 54-trial experiment 10 times.

4.4 Results

The first statistic we analysed is the probability of a correct answer, computed as the ratio of correct answers to the total number of trials. Table 4.1 shows these probabilities for each of the 6 conditions of our 2×3 experimental design.

We observe that:

- On each of the three models, the specular rendering gives a higher probability

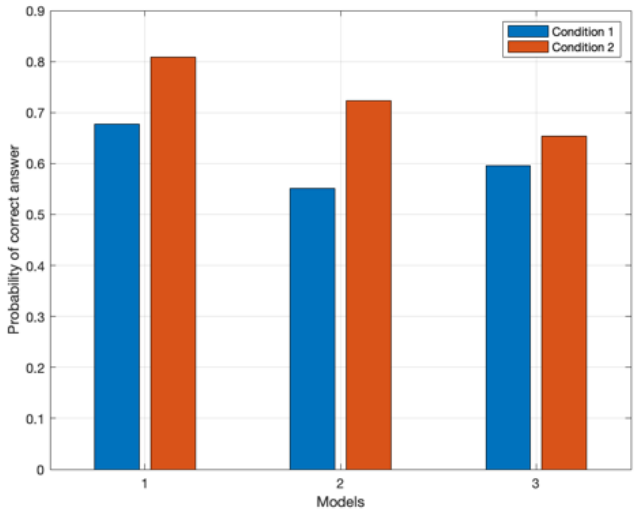


Figure 4.4: Mean values of probability of correct answer for each independent comparison group.

of correct answer, that is, a higher probability of detection of quantisation artifacts, as expected.

- With either rendering method, the detailed Max-Planck model gives a higher probability of correct answer than the coarse one, indicating that, as expected, more detailed geometry is more sensitive to quantisation artifacts.

- The comparison between the two models with the same number of triangles is not so conclusive. We note that, on the one hand, one would expect the higher quality of the triangles of the Sphere-5K model to make it more resilient against quantisation noise, however, on the other hand, its very smooth and regular global shape might function as a very uniform background, on top of which quantisation artifacts become easier to detect.

in Figure 4.4 the mean values for each comparison group are shown and the no-interaction between the two factors is evident, because the trend among conditions is the same for each model (plastic condition has higher means than diffuse condition in all cases), and for both conditions, the model 1 has the highest mean.

We performed a 2×3 two-way ANOVA analysis on the probabilities for correct

answer of the 10 participants of the experiment. The effect of the rendering method was significant, with p-value 0.0028 (F-statistic 25.896). The effect of the geometric model was also significant with p-value 0.0088 (F-statistic 9.981). The interaction between rendering method and geometric model was not statistically significant, with p-value 0.1290 (F-statistic 0.1290).

4.4.1 Psychometric function fitting

The above statistical analysis allows statistically significant inferences on how model geometry and rendering method affect the quantisation imperceptibility threshold, however, it does not produce estimates of those thresholds. To obtain such estimates we fitted to the data the psychometric function

$$f(x;\alpha,m,s) = 1 - \alpha \cdot \Phi_{m,s}(x) \tag{4.4.1}$$

where x denotes quantisation level, $\Phi_{m,s}$ is the cumulative Gaussian distribution with mean m and standard deviation s, and α is a third free variable of the model, representing the asymptotic probability of a wrong answer for high quantisation levels. A Matlab program was implemented to fit $f(x;\alpha,m,s)$ to a given set of observations by a maximum likelihood estimation (MLE) of its three free variables α, m, s.

Notice that we would normally expect the value of α to be equal to 0.5, reducing the number of free variables to two. That is, for increasingly higher levels of quantisation, and as quantised and unquantised models become indistinguishable, we would expect the probability of correct answer to tend to 0.5. Nevertheless, we treat α as a free variable to be estimated along m and s, allowing for the possibility of a systematic bias in favour, or against, the quantised model, something which often was indeed the case.

Fig. 4.5 shows the MLE fitted curves for the six conditions of the experiment. The estimates of the variables α, m, s are reported in Table 4.2.

Since a Gaussian probability distribution function has its maximum at m, the corresponding cumulative probability distribution has at m its inflection point, which is also the point where the maximum of its derivative is obtained. Thus,

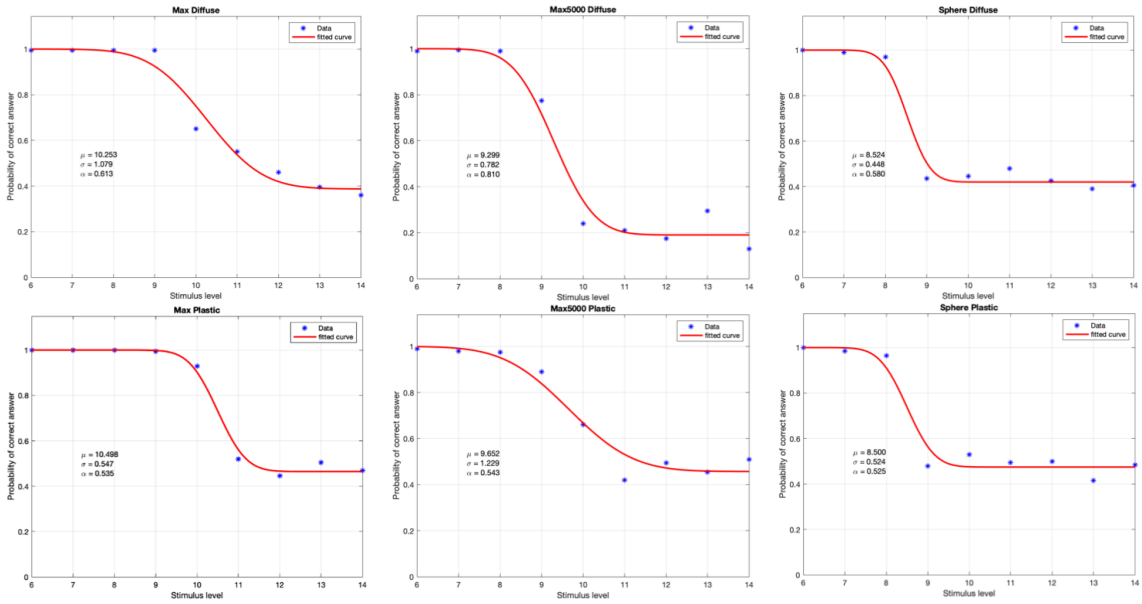


Figure 4.5: **Top:** Diffuse rendering. **Bottom:** Specular rendering. **Left to Right:** Max-Planck with 100K triangles, Max-Planck with 5K triangles, and Sphere.

	Max-100K		Max-5K		Sphere-5K	
	Diffuse	Plastic	Diffuse	Plastic	Diffuse	Plastic
m	10.253	10.498	9.299	9.652	8.524	8.500
s	1.079	0.547	0.782	1.229	0.448	0.524
α	0.613	0.535	0.810	0.543	0.580	0.525

Table 4.2: The free parameters of the psychometric function $f(x; \alpha, m, s)$ estimated by the MLE method.

m corresponds to the level of quantisation where the probability of a correct answer takes a value exactly at the middle between its absolute maximum of 1 and its asymptotic minimum $(1 - \alpha)$. Therefore, m is the best, in the maximum likelihood sense, estimate of the imperceptibility threshold.

The values of m in Table 4.2 verify our previous observations that were based on the raw probabilities of correct answer. For example, on the more detailed Max-100k model, the user needs a lower increase of the stimulus level to go from a correct to a wrong answer, as compared against the coarser Max-5K model, hence the discminination threshold is higher. Perhaps most importantly, we note that the

differences between the discrimination thresholds of different models and rendering methods are rather small, indicating that in certain application domains it could be possible to find universal discrimination thresholds that will also be efficient in terms of memory usage.

The parameter s, that is, the standard deviation of the Gaussian, controls the steepness of the transition from the highest to the lowest value of the psychometric function. We observe that the smooth and regular Sphere-5K has smaller s values than the more irregular Max-5K model, indicating that the latter degrades more gradually than the former as the quantisation level decreases.

The parameter α represents the asymptotic minimum of the probability of correct answer, approached as the value of quantisation level becomes large. The value of α is expected to be 0.5, that is, as the model quality increases, the user should eventually be at change level of picking the higher-quality mesh. Surprisingly, in many cases α differed from 0.5. We note that the diffuse rendering of the Max-5K model gave the highest value of α, followed by those of Max-100K and the Sphere-5K. A possible explanation for the much higher α value of the diffuse Max-5K model is that, under the diffuse rendering, the coarseness of the model produces some sharp rendering features, which by some participants were interpreted as faithful rendering of fine surface detail, indicating a higher mesh quality rather than an artifact.

Finally, we performed a two-way ANOVA on the m values computed on each of the 10 participants separately, that is, on our estimates of the discrimination thresholds of those individuals shown in table 4.3 organized by model and condition; subsequently, this were employed for the ANOVA test calculations. As it was also the case with the raw probability of correct answer, the geometric model had a significant effect with a p-value 0.0034 (F-statistic 16.554). However, unlike in the case of raw probability of correct answer, the effect of the rendering method on the imperceptibility thresholds was not significant, corresponding to a p-value of 0.4078 (F-statistic 0.468). The interaction between geometric model and rendering method, again, was not significant, corresponding to a p-value of 0.2010 (F-statistic 0.054) as shown in table 4.4.

Participant number	Max-100K		Max-5K		Sphere-5K	
	Diffuse	Plastic	Diffuse	Plastic	Diffuse	Plastic
1	9.977	10.441	9.273	9.903	8.683	8.735
2	10.297	10.572	10.034	10.891	8.694	8.681
3	10.601	10.582	8.938	8.463	8.235	8.546
4	10.551	10.147	9.093	9.716	8.445	8.750
5	9.817	10.381	9.376	9.847	8.670	8.540
6	9.971	10.645	9.213	10.062	8.605	9.442
7	9.400	10.769	9.241	9.584	8.672	8.117
8	10.938	9.950	9.378	9.054	8.647	8.614
9	9.933	10.501	9.266	8.630	8.392	8.352
10	10.951	10.486	9.289	9.462	8.519	8.360

Table 4.3: Results of fitted parameter m for each user by model and rendering condition.

Variables	Sums of Squares	Degree of Freedom	Mean Squares	F	P-Values
Models	15.501	2	7.751	16.554	0.0034
Conditions	0.219	1	0.219	0.468	0.4078
Models/Conditions	0.051	2	0.025	0.054	0.2010
Error	25.283	54	0.468		
Total	41.054	59			

Table 4.4: ANOVA test results for experiment.

4.5 Conclusions

We presented the results of a two-alternative forced-choice psycho-physical experiment, aiming at studying the quantisation thresholds below which the visual quality of the mesh can be considered degraded as quantisation artifacts become more and more perceptible. This preliminary study established that meshes with larger number of triangles should be quantised at higher levels, while the choice of rendering method should also be taken into account since a high specular component can reveal artifacts that are not perceivable under renderings with high diffuse compo-

nents. There is also some tentative evidence that mesh quality metrics are related to the quantisation thresholds, however further studies are needed to establish the nature of that link.

We used the results of this experiment to inform the design of a larger experiment, the scope of which would go beyond establishing the existence of links between quantisation thresholds and characteristics of the mesh or the rendering method. Instead, the aim of that extended experiment would be to quantify such links by establishing simple formulas relating the quantisation threshold with numerical information derived from the mesh, such as the number of its triangles, or the value of a mesh quality metric.

Chapter 5

A user study on the impact of the geometry of the quantised triangle meshes on the quality perception and discrimination thresholds

5.1 Introduction

Triangle meshes have emerged as the ubiquitous standard for 3D content representation for most graphics applications. Being, essentially, piece-wise linear representations of surfaces, triangle meshes are simple and scalable, and benefit from the existence of specialised sophisticated algorithms covering the whole graphics pipeline, from mesh generation, to processing and rendering, to transmission and storage.

The vertex coordinates of a triangle mesh are usually represented by 32-bit floats. However, at various stages of its life cycle, and most notably during compression for transmission or storage, the vertex coordinates may be quantised and represented in a fixed-point arithmetic, typically, by 12, 16 or 24 bits per vertex coordinate. In this chapter we present the results of a psycho-physical experiment, part of a series of similar experiments we conducted [2, 3], aiming at studying the visual effect of such quantisations. In particular, given that vertex coordinate quantisation is an

irreversible process, we want to study the discrimination threshold, above which an observer is not able to perceive quantisation artifacts.

In [2], the main focus of the experiment was on the impact of the quantisation method. We compared rounding, where the least significant bits are put to zero, and dithering where the least significant bits are randomised, and found that, in general, dithering has a slightly higher discrimination threshold. In [3], the focus of the experiment was on the impact of the number of triangles in the mesh, and that of the rendering method. We found that, generally, larger meshes have higher discrimination thresholds, and also that renderings with a higher specular component have higher discrimination thresholds.

In this chapter we study the impact of the geometry of the mesh, that is, how the shape of the 3D model and the properties of the underlying mesh are related to the discrimination threshold. Our study is based on a two Alternative Forced Choice, psycho-physical experiment, where two stimuli of one model are presented — the original and one which is quantised at a certain level — and the participant chooses the one with the higher visual quality. We used four different 3D models, the Max-Planck, the Cone, the Sphere, and the Human-Head, fixing all the experimental parameters we had studied previously. In particular, we chose rounding as the quantisation method, we used only meshes with a large number of triangles, between 200K and 315K, and we used a single rendering method, which had a high specular component.

Our first hypothesis was that there is an inverse correlation between the amount of geometric information carried by a 3D model and its discrimination threshold. In particular, we hypothesised that the simpler in shape synthetic models, the Sphere and the Cone, will have higher discrimination thresholds, because it will be easier for a participant to detect quantisation artifacts on them, while, in contrast, quantisation artifacts will be more difficult to detect on the more complex surfaces of the natural models. This hypothesis was verified and moreover, in a post-hoc analysis, we computed a strong, statistically significant correlation between the discrimination threshold and the filesize of the compressed mesh, which was used as measure of the amount of geometric information carried by the mesh.

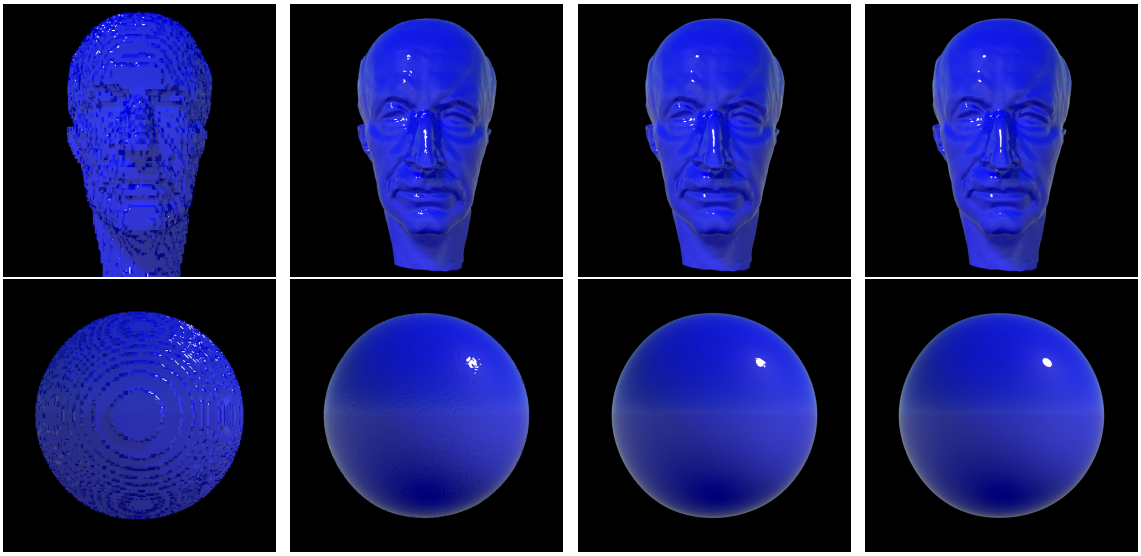


Figure 5.1: From left to right: The Max-Planck model (top) and the Sphere model (bottom) quantised at 8, 12 and 14 bits per vertex coordinate, and the original unquantised model.

For an illustration of our hypothesis, Figure 5.1 shows, at various levels of quantisation, the Max-Planck model, which has the largest filesize when compressed, and the Sphere models which has the smallest compressed filesize. We notice that when we use 12 bits per vertex coordinate, quantisation artifacts in the form of surface texture are still visible on the Sphere, while one cannot detect this type of artifacts on the Max-Planck model.

Our second hypothesis was that the average quality of the triangles of the mesh has an impact on the discrimination threshold. In particular we hypothesised that meshes comprising long thin triangles will have higher discrimination thresholds compared to meshes comprising well-rounded, near equilateral triangles. That was also a quite intuitive hypothesis, given that the same amount of spatial perturbation will most likely cause a larger perturbation of the normals of the thin triangles, which in turn will be easier to detect in a rendering of high specularity [117]. As a measure of thinness of a triangle we used the aspect ratio, that is, the ratio of the smallest edge-length by the largest. Our experiment did not verify that second hypothesis, as we there was no significant correlation between the mean aspect ratio of the mesh triangles and the discrimination thresholds.

5.2 Background

Quantise the triangle mesh is the first step of all mesh compression algorithms [71]. In order to encode the vertex coordinate, most of application frequently utilise the 32-bit floats. On the other hand, when a triangle mesh is needed in the compressed form, the use of 16 bits per vertex coordinate seems to be the usual practice in mesh compression [71].

Quantisation techniques are mainly studied from the perspective of signal theory [40]. According to an extensive survey of the technique in [42], rounding is the most commonly used and traditional example of quantisation for density estimations via histograms back in 1897 [97]. Here, we use rounding for the quantisation of the spatial coordinates of the mesh vertices. We note that quantisation techniques applied on the various frequency domain representations of the mesh geometry, see for example [99], might have significant theoretical interest, as well as significant applications, but are nevertheless less relevant to the everyday real-life use of meshes.

A previous study by the authors for determining the visual effect of quantisation, used a simple yes/no task experiment and was aimed at determining a discrimination threshold beyond which the quantized mesh is not perceived to differ from the original [2]. However, the focus there was on understanding the effect the choice of quantisation method has on the threshold, focusing in particular on the effect of dither.

5.3 Experimental Design

In the experiment we used 4 triangle meshes, each consisting of about 100K vertices. The Max-Planck and the Human Head models are both natural models acquired through laser scanning of physical objects, and between them the Max-Planck model has more geometric information. The Sphere and the Cone are synthetic models created by CAD software. The Sphere model consists of almost equilateral triangles, while the Cone mostly comprises long skinny triangles. Overall, the choice of the models of the experiment aimed at establishing the relationship between the discrimination threshold on the one hand, and two shape related factors on the other,

that is, the amount of geometric information carried by the mesh, and the average shape of its individual triangles.

For each of the 4 original unquantised meshes, 10 different quantised meshes were produced, one for each integer quantisation level, from 8 bits per vertex coordinate to 16 bits per vertex, while the quantisation level of 20 bits per vertex was also used.

Rendered images were produced from these meshes and were presented to the participants as the stimuli of the experiment, see Figure 5.3. The high specular component of the rendering method resulted into surfaces with a characteristic glossy appearance. As it was shown in our previous experiment [3], quantisation artifacts are easier to detect on high specularity renderings, because they reveal better the normal perturbations of the underlying mesh, which normal perturbations are considered the main source of visual degradation in a mesh [117].

The experiment was conducted in Saudi Arabia, in October 2020, over a period of 25 days, with 20 participants in total.

The stimuli were presented on a computer screen MacBook Pro with a resolution of 2560 $\times$ 1600 pixels. The screen width and height were 30.41 and 21.24 centimeters, respectively. The observer viewed the screen from a distance of 50 cm. The room, where the experiment took place, has a natural light and quiet so no distraction could affect the process. The stimulus size on the screen was 1280 $\times$ 725 pixels.

Discrimination thresholds were measured using a two-alternative forced-choice method. On each trial, the original, unquantised, stimulus and the quantised one were presented side by side on the screen for 4000 ms. The observer then used the computer keyboard or mouse to indicate which of the two stimuli has the highset quality. See Figure 5.3 for the interface of the experiment. The next trial started after the response.

Each quantisation level was repeated 10 times during the experiment. As there were four different models and 10 quantisation levels for each model, there were a total of 4*10*10=400 trials in the experiment for each observer. The order of models and quantisation levels was randomised across trials. Due to a programming error, the left/right order of the original and the quantised stimulus was not randomised, but switched after each trial. That is, on every other trial the original was on the

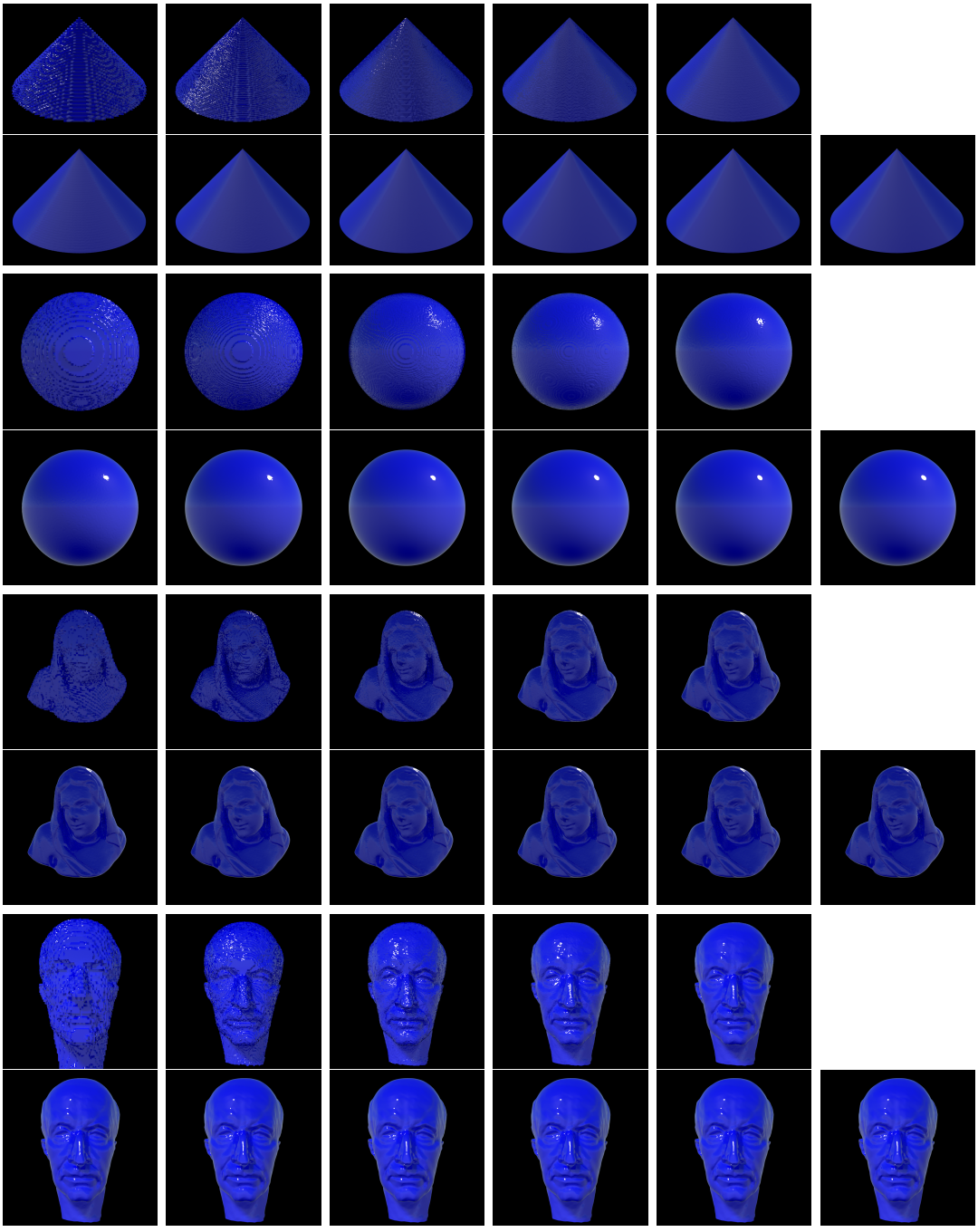


Figure 5.2: For each model, the top row shows meshes quantised at levels 8-12. The bottom row shows meshes quantised at levels 13-16, 20 and the original unquantised model.

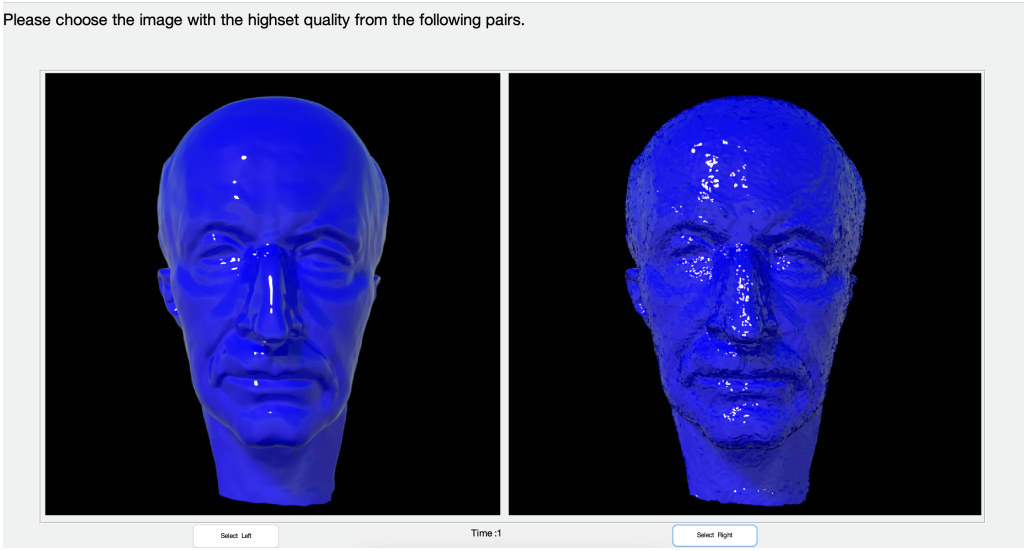


Figure 5.3: The interface of the experiment.

left, and on every other, on the right.

Before the experiment the participants were given instructions on how to conduct the experiment. They were then allowed to test the interface and practice the task. The models in this practice experiment were different from those of the main experiment.

Following the experiment in [3], we obtain discrimination thresholds by fitting to the data the psychometric function

$$f(x;\alpha,m,s) = 1 - \alpha \cdot \Phi_{m,s}(x) \tag{5.3.1}$$

where x denotes quantisation level, $\Phi_{m,s}$ is the cumulative Gaussian distribution with mean m and standard deviation s, and α is a third free variable of the model, representing the asymptotic probability of a wrong answer for high quantisation levels. A Matlab program was implemented to fit $f(x;\alpha,m,s)$ to a given set of observations by a maximum likelihood estimation (MLE) of its three free variables α, m, s.

Notice that we would normally expect the value of α to be equal to 0.5 in a two-alternative forced-choice experiment, reducing the number of free variables to two. That is, for increasingly higher levels of quantisation, and as quantised and unquantised models become indistinguishable, we would expect the probability of

correct answer to tend to 0.5. We add α to the model as a free variable to be estimated along m and s. This is because the left-right order of the two stimuli was not randomised in the experiment (see above), and thus it is possible that any bias in the observer's responses leads to a value of α different from 0.5. If we forced α to 0.5, we would then get biased estimates of the other parameters, which are of main interest to us. Thus, in order to get a good estimate of the threshold, we also fit α.

5.4 Results

Figure 5.4 shows the MLE fitted curves for the four models of the experiment, and Table 5.1 summarises the estimates of the variables α, m, s. Since a Gaussian probability distribution function has its maximum at m, the corresponding cumulative probability distribution has at m its inflection point, which is also the point where the maximum of its derivative is obtained. Thus, m corresponds to the level of quantisation where the probability of a correct answer takes a value exactly at the middle between its absolute maximum of 1 and its asymptotic minimum $(1 - \alpha)$. Therefore, m is the best, in the maximum likelihood sense, estimate of the discrimination threshold. As expected, the values of m for the four models follow the inverse order of the values of the mean probability of choosing the unquantised model.

As seen in Figure 5.4, the psychometric functions do not asymptote at 0.5. The fit values for α were systematically smaller than 0.5 indicating that, counter-intuitively, when there was very little difference between the two stimuli, the observer chose the quantised one as having a better quality. It is very unlikely that this is a true perceptual effect, however, as this holds also for the quantisation level 20, which was practically identical with the original. It is more likely to result from non-independence of the observer's responses across trials. Several types of sequential effects are known to exist between trials in a psychophysical experiment [29, 35]. We do not have enough data here to distinguish between them and we focus on the other parameters m and s.

As a measure of the amount of geometric information carried by a 3D model we use the filesize, after applying a state-of-the-art mesh compression algorithm, here

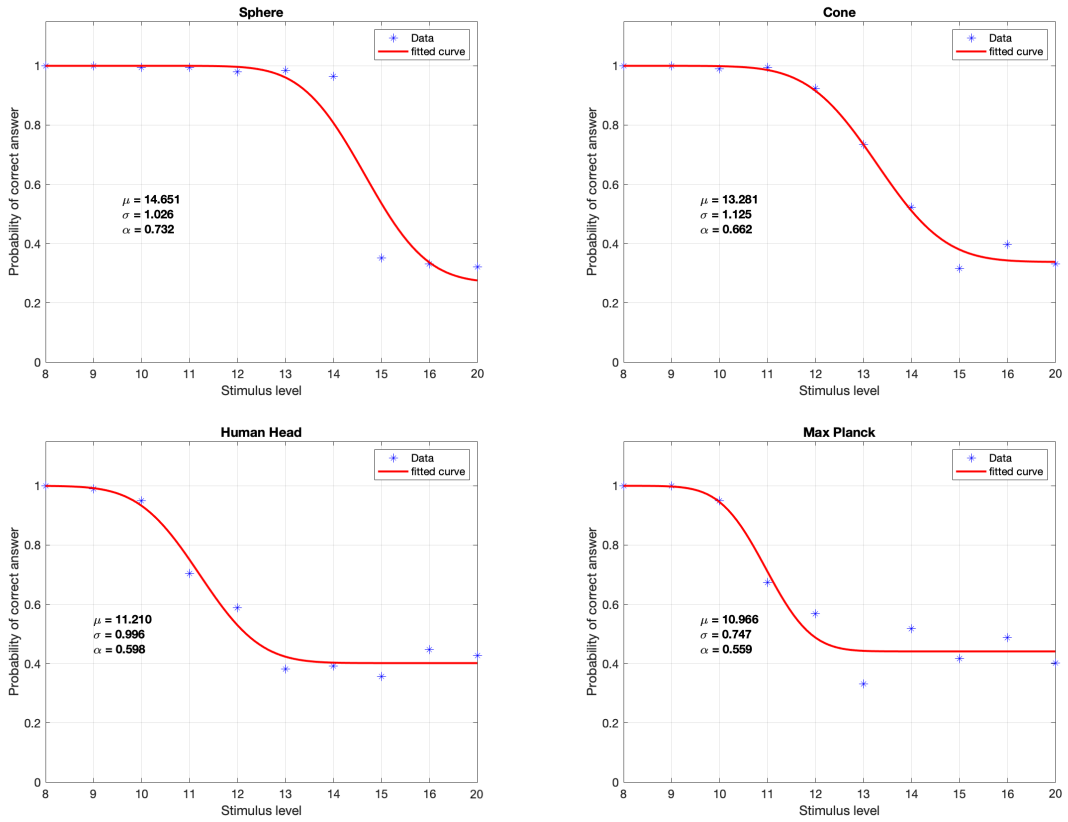


Figure 5.4: The Maximum Likelihood Estimated psychometric function for each of the models of the experiment.

the publicly available Draco software `https://google.github.io/draco/`. The use of absolute filesizes it is justified by the fact that the uncompressed meshes have all about 100K vertices and a filesize of about 22MB in uncompressed .obj format. Most importantly we are interested in an absolute measure of the amount of geometric information carried by the mesh, which is provided by its filesize when compressed with a state-of-the-art compression algorithm, which will try and remove information redundancies.

As a measure of the average quality of the individual triangles of the mesh, we computed the average over the triangles of the mesh aspect ratio, that is the length of the shortest edge of the triangle, divided by the length of the longer. The aspect ratio of an equilateral triangle is 1, while for long skinny triangles, the aspect ratio tends to 0. The aspect ratio is one of the various element quality metrics described in [78]. We note that all the various other metrics also favour equilateral triangles against thin ones, and we do not expect any different results from their use. We also note that averaging over triangle quality metrics can be used to derive quality metrics that would apply to whole meshes [65].

Table 5.1 summarises the variables from the analysis of the experimental results and the analysis of the models themselves.

	Sphere	Cone	Head	Max
m	14.651	13.281	11.210	10.996
s	1.026	1.125	0.996	0.747
α	0.732	0.662	0.598	0.559
slope $= \alpha/s$	0.713	0.588	0.600	0.748
# Mesh triangles	307.520	314.400	216.928	199.996
Compressed filesize	310.848	342.020	402.932	423.724
Mean aspect ratio	0.9635	0.3913	0.8043	0.6401

Table 5.1: The parameters of the fitted psychometric function (top three rows), the slope at the inflection point computed as α/s (fourth row), and geometric characteristics of the original meshes (bottom three rows).

The values of m in Table 5.1 verify our previous observations that were based

on the raw probabilities of correct answer. For example, on the Sphere model, the user needs a lower increase of the stimulus level (here, the quantisation artifacts) to go from a correct to a wrong answer, as compared against the Sphere, hence the discrimination threshold is higher. We also note that the differences between the discrimination thresholds of different models are rather small, indicating that in certain application domains it could be possible to find universal discrimination thresholds that will also be efficient in terms of memory usage. That is, we can use universal discrimination thresholds that are not, unnecessarily high, e.g. 16 or 24, rather than 32 bits per vertex.

The steepness of the transition from the highest to the lowest value of the psychometric function depends on its slope at the inflection point. For the type of psychometric function we employed, this slope is proportional to α/s, see [100], the values of which are reported in the fourth row of Table 5.1. We observe that the Sphere and the Max-Planck models have higher slopes, and thus sharper discrimination thresholds than the Cone and the Head models. However, it is clear that the experiment does not provide enough evidence, not even for a qualitative study of the issue.

The values of α show a systematic bias in favour of the models quantised at a high quantisation level, and against the unquantised model. As we mentioned in Section 5.3, we suspect that this could in part be explained by a flaw in the implementation of the experiment, that is, using alternation rather than the randomisation in the relative positions of the two stimuli within the interface.

5.4.1 Correlations between discrimination thresholds and mesh geometry

The results show a clear correlation between the discrimination threshold and the filesizes of the compressed meshes. Quantitatively, we computed a correlation coefficient of $r = -0.9915$, with $p = 0.0085$, indicating a strong inverse correlation, with a high statistical significance. The result is intuitive, as one would expect that given the quantisation level, an observers' ability to detect quantisation artifacts at that particular scale would depend on the amount of geometric information the artifacts

are embedded in. That is, at a given level of quantisation, artifacts should easier to detect over meshes that carry little geometric information, such as the sphere, rather than on meshes with more geometric information such as the Max-Planck.

In the experiment we used both natural meshes, acquired by laser scanners and carrying large amounts of information, and synthetic meshes that carried less information. We note that the correlation between discrimination thresholds and filesize of the compressed mesh is evident both across the two mesh types, as well as within each one of them. Regarding the comparison across mesh types, the synthetic meshes, carrying less information, had higher discrimination thresholds than the natural ones. Within the synthetic mesh type, the Sphere, carrying the least geometric information, had higher threshold than the Cone. Within the natural mesh type, the Max-Planck, having more prominent features and carrying more geometric information, had as expected a larger compressed filesize than the Head, and eventually a lower discrimination threshold.

We consider this correlation as the most significant result of our experiment. We note that, generally, and especially in signal theoretic studies, the ratio of the carrier strength to the amount of noise is considered an important measure of the expected performance of a system. However, to the best of out knowledge, it is the first time that a similar observation is made in such a setting, that is, about the visual perception of a mesh as established by a psycho-physical experiment on the one hand, and the amount of geometric information carried by that mesh, as measured by its compressed filesize, on the other.

The analysis of the results does not show any statistically significant correlation between the discrimination thresholds and the average aspect ratio of the triangles of the mesh. Specifically, we computed a correlation coefficient of just $r = 0.2198$, with $p = 0.7802$.

The lack of a statistically significant correlation could be interpreted as an indication of the unsuitability of that mesh quality metric to predict mesh discrimination thresholds. Indeed, the metric averages aspects ratios over all the triangles of the mesh, including triangles in the non-visible part of the mesh. On the other hand, it could be the case that some users were evaluating some meshes by focusing their

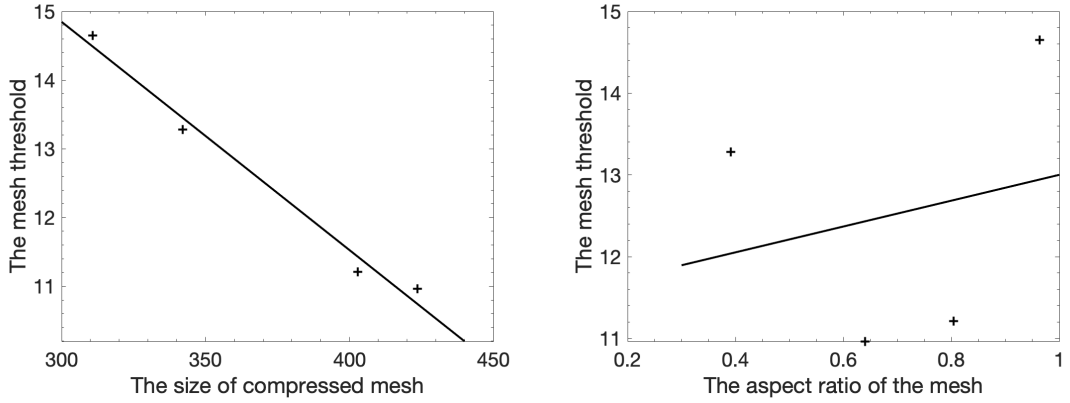


Figure 5.5: Scatter plots of the compressed mesh filesize (left) and the mean aspect ratio (right), against the discrimination thresholds. The best fitting lines are shown too.

attention on specific parts of them. Especially, the parts of the mesh which, depending on the mesh position and orientation, and the lighting conditions, reflect most of the incident light directly on to the camera.

Alternatively, we cannot altogether exclude the possibility that there is such a correlation, which however must be weaker than that between the compressed filesize and the discrimination threshold. This possibility could be investigated by a follow-up experiment with meshes that have similar compressed filesizes and different mean aspect ratios.

Figure 5.5 shows the scatter-plots of the compressed filesizes and the mean aspect aspect ratios, respectively, against the discrimination thresholds. The tightness of the best fitting line in the first scatter-plot illustrates the high correlation between the compressed filesizes and the discrimination thresholds.

5.5 Conclusions

We presented the results of a two-alternative forced-choice psychophysical experiment, aiming at studying the quantisation thresholds, below which the degradation of the visual quality of the mesh by the quantisation artifacts can be detected. Our main finding is that there is a strong inverse correlation between the discrimination

threshold of a mesh and its filesize after compression. We also studied a possible link between discrimination thresholds and the quality of the mesh triangles, as measured by the mean aspect ratio, but we did not detect any significant correlation.

The main limitations of the experiment presented in this chapter stem from its relatively small size. Next, we plan to investigate the relationship between quantisation thresholds, geometric properties of the mesh and properties of the rendering algorithms influence the visibility of models. Such a study would require higher dimensional experiment and perhaps more subtle experimental designs too. In particular, we plan to use the maximum likelihood difference scaling method which has been proven to be a powerful approach to similar problems [10, 72].

Chapter 6

A perceptual difference scaling study on quantised 3D models

6.1 Introduction

Vertex coordinate quantisation is the first step of all mesh compression algorithms [71]. Indeed, geometric information encoded in the least significant bits is visually redundant and moreover, it has high entropy and thus, it is incompressible from an information theoretic point of view. The use of 16 bits per vertex coordinate seems to emerge as a defacto standard as far as mesh compression is concerned [71].

Previous studies for determining the visual effect of quantisation are based on simple yes/no task experiments, aiming at determining an undetectability threshold beyond which the quantised mesh is perceived as identical to the original [2]. Here we describe a Maximum Likelihood Difference Scaling (MLDS) experiment [72], aiming at modeling perceived differences as a function of the quantisation level.

In this chapter we present MLDS method to estimate the perceptual scale of differences in stimulus based on the participant judgment. [72] . The simplest way to explain MLDS method is to describe the kind of task that it is intended to scale. Consider the upper and lower pair of images, the arrangement in Figure 6.2 is an example of a typical trial from a difference scaling experiment. The observer is asked to check the two pairs of stimuli (‘a quadruple’) then select the pair with the larger perceptual difference (’up or down’). All four of the image samples selected from

Figure 6.1: The two Conditions of the MAx-Planck. **Left:** Plastic. **Right:** Diffuse.

data set where models are quantised and rendered from Max-Planck model shown in Figure 6.1. During one trial of an experiment, the subject is asked to make this decision for a large number of quadruples of image samples, all randomly drawn from the data set.

6.2 Maximum Likelihood Difference Scaling

The Maximum Likelihood Difference Scaling (MLDS) estimate perceptual scales based on Maximum Likelihood Estimations. The maximum-likelihood estimation (MLE) based adaptive psycho-physical procedures have been used nowadays for the minimization of testing time. In 1982, Shelton et al. [95] compared the MLE with other techniques and find out that it can converge on the threshold in less time that means it is more efficient in terms of speed. Therefore, it can be said that the applications that are time critical should uses MLE algorithm. Also, Green in 1990 [43], worked on the stimulus selection and analyze that how it affects the threshold estimation for psychometric functions. He concluded that the effectiveness of maximum likelihood estimation improves as the number of trials that are used to threshold estimation value increases [4].

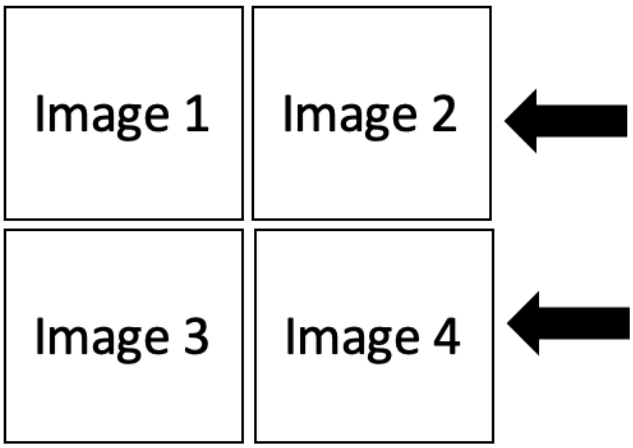


Figure 6.2: Difference scaling experiment.

Various characteristics of MLE including accuracy and speed were checked in comparison to a transformed up-down algorithm (an adaptive procedure for the estimation of points on the psychometric function known as the traditional Levitt method procedure) in a two-interval forced choice task. It has been witnessed from the results that a MLE procedure can provide high performance (90%) level and significant speed advantages than previous transformed up-down approaches. Thus, this advantage has increased variability, and a potential for the estimated thresholds to follow a skewed distribution. These skewing appeared when relatively low performance is estimated using MLE technique. However, by increasing the number of trials or restricting the criteria to stop the trials at the cost of speed can be an option to overcome these difficulties [4]. Despite this, there are some scenarios where in short span of time, repeated measurements are required. In such cases, MLE procedures are better than the traditional Levitt method due to their speed efficiency with many turnarounds [4].

There have been many examples of successful use of MLE for quantisation, in areas such as digital image processing and JPEG image compression. Thus, if an image (JPEG) is previously coded then the information of the quantization table is sometimes needed during the compression process. This quantisation table information helps in minimizing the error of quantisation. In [27], a maximum likelihood based algorithm has been proposed to detect the JPEG compression history of an

image. The author has proposed detection method MLE for finding the quantisation history of image (JPEG). It is important to notice here that only information related to bitmap of the coded image was available previously, which in combination with the other method has been used to find out compression history. A detection algorithm has been proposed which is based on the difference between the neighboring pixels in the histogram form as shown in 6.3, where these differences can either be 1D or 2D. The comparisons of these histograms helped in analysing the history of compression that either compression has done before or not. In this study the histograms are normalised and the absolute difference between these histograms and between pixels across the boundaries of image are compared. Thus providing information about the degree of compression of an image.

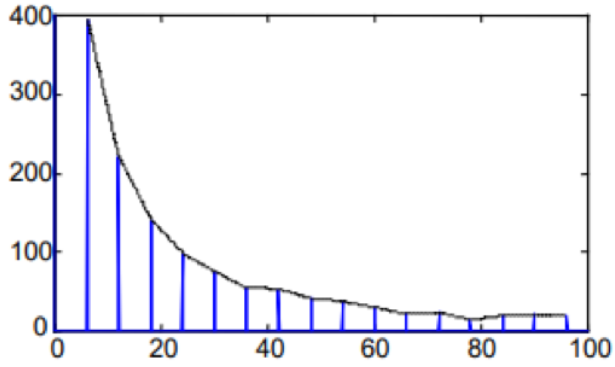


Figure 6.3: Histrogram for image [27].

In another article, the author used maximum likelihood difference scaling (MLDS), to estimate the supra threshold differences [10]. This method can be useful in image compression specifically in measuring the perceptual effects using vector quantization. Different rates of vector quantization have also been illustrated, however, this compression is at the cost of image quality.

In this respect, an ideal compression method is one that maximized the compression while minimized the perceptual distortion. Furthermore, this method has been implemented over a wide range of images and then results were analyzed. This approach by-passes the limitations of using rating methods and focused on using multiscale structural similarity (MS-SSIM) for diverse categories of image. This re-

search also demonstrated data collection through MLD and in this way enhance the performance of the method [10].

Charrier et al. [8] proposed an IQA method, based on Maximum Likelihood Difference Scaling (MLDS), a psychophysical method. The proposed MLDS technique is forced-choice task and require a limited trials to obtain measureable effects of the estimation of distortion. The effectiveness of the proposed algorithm is measured using MLDS in terms of compression quality trade off. MS-SSIM index has been used as a trial image quality assessment (IQA) algorithm due to its significantly high correspondence with human ratings.

In non-metric multidimensional scaling (NMDS), MLDS was applied to determine the scaling problem [56], [72], [57]. It is evident in the literature that there are mainly two assumptions, firstly, perceptual scale in a scaler quantity and secondly, this scale propagates on accordance with stimulus. More specifically, it undertakes that the order of two physical space stimuli and the perceptual scale stimuli denote the same order. However, advantages of the MLDS method include the unbiasedness of MLE and for algorithmic convergence, a small subset of quadruplets. As a conclusion, correlation between input and output has also been observed.

Devinck and Knoblauch [21] worked on the MLDS for the threshold estimation. In their research they noticed that a signal detection model (SDM) determines the discrimination performance. They have used, MLDS in terms of luminance ratio between the two components, for the measurement of the perceptual strength. MLDS is centered over a Gaussian, equal-variance, signal detection model and thus resulting into a perceptual scale.

Moreover, it has also been said that it is a psycho-physical method that specifies the efficient description of perceptual scales [57], [72]. Maloney and Yang [72] proposed the model of supra threshold perceptual differences. A maximum likelihood difference scaling technique has been used for parametric estimation and for the evaluation of the method reliability. An approach for testing the efficacy of the method in various context has also been discussed. It can be noticed from the literature that threshold stimuli had always been under the observation of researches.

In this respect, in 2010, Emrith [25] presented a research where virtual differ-

ences for the threshold stimuli were judged by the observer, further based on these differences a scale is estimated. The method has been used to study variety of visual domains such as texture properties. In this research, the author studied the human perception related to the changes in the texture of an image accrued due to statistical changes. In the proposed technique, first and second order statistic were kept constant while analysing the 3rd and high order stats through randomization of phase spectra. Stimuli include natural and synthetically produced images, where for each observer, pixel wise comparison was done using synthetically produced images. However, derivation of perceptual scale has been done through difference scaling. In addition to this, a biologically plausible model was used to calculate the changes of local measurements of phase congruency.

Furthermore, colour differences and glossiness of an image are the essential attributes of a visual appearance. In 2004, Obein [75] has used maximum likelihood difference scaling to find out glossiness of surface over an extended range. In their study, MLD had proven to be a robust method for perceptual scale estimation. They have gathered the observers' judgments for 10 coated samples (black color), in binocular vision and monocular vision. The results demonstrated the relationship of gloss percept with the specular gloss value and found to be a nonlinear one. It has been noticed that the sensitivity value is more at edges (extreme) as compare to the middle. However, in binocular vision, gloss sensitivity is more as compare to the monocular one. At the end, it has been noticed that gloss difference scales change in accordance to illumination. Thus, it can be said that gloss scaling is not dependent on the geometrical variants of the luminous flux over surface of a sample.

When talked about different material, an observer can discriminate the material (silk, granite, etc.) under viewing conditions and this is considered to be an achievement of a visual system, however, it is a challenging situation. Early researchers have focused on the flat but thin filters for the observation of transparency. However, in this research [10], focus is particularly on the irregular shapes of thick transparent objects with fluctuated refractive index (for example, ice cube). A vital part of the visual evidence is distortion that indicates the existence of previously mentioned objects that has been noticed in the observed shape of other objects. Therefore,

new category of visual cues has been proposed based on induced distortion. Furthermore, experimental evaluation has also been done through prediction of failure of success in refractive indices judgement [31].

Along with all the above mentioned properties, physical properties are also important in image processing that can also be examined through MLD. It has been observed that most of the literature is centered towards the optical properties however, physical property like shape is also as important as the visualization of an image. Researchers have used maximum likelihood difference scaling for the reconstruction of perceptual scales for perceived viscosity of fluid [79].

Additionally, MLD has also used for the compression of image however, it resulted to relatively poor image quality. In a research by Charrier [10], nine images have been taken and their quality has been examined with the help of maximum likelihood difference scaling. Each image is compressed with the help of VQ (vector quantisation) to 2 distinct colour spaces and up to 10 distinct levels. It has been seen that an RGB image can be compressed up to 32% on average that leads to changes in perception and loss in quality. Therefore, the proposed technique provides a fast and direct way to measure the compression consequences [10]. Effects of different compression rates as shown in Figure 6.5, using vector quantization compression on an image is represented in the Figure 6.4. A trade-off between compression rate and quality can be seen in the figure.

Figure 6.4: (Color online) Effects of VQ compression. The original image (0% compression) is shown after VQ compression using a codebook based on the LBG algorithm applied to the image. Larger compressions lead to evident decreases in image quality [10].

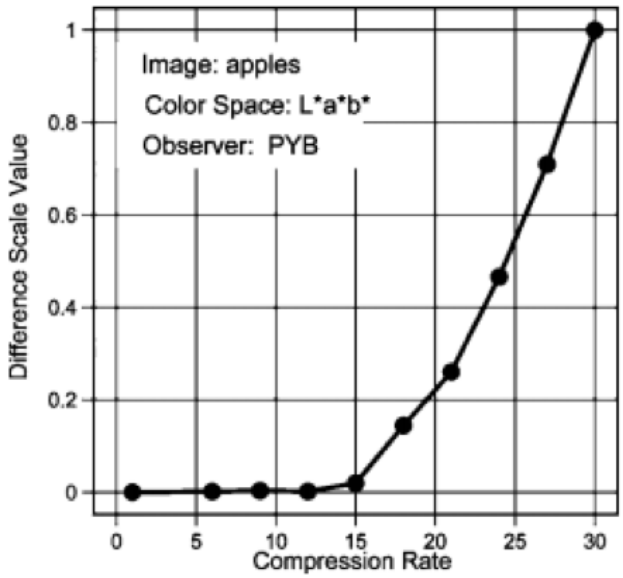


Figure 6.5: The difference scale values are estimates based on the observer's judgments of superthreshold perceptual differences between the images portrayed in 6.4 [10].

6.3 Background

The maximum likelihood difference scaling (MLDS) method [72] was applied in this experiment to determine the perception threshold among a series of visual stimuli. The experiment was performed with $N = 10$ quantization levels $L_1, L_2, ..., L_N$. According to the MDLS method, there is a real number ψ_k associated with each stimulus level L_k, and such levels are numbered in a way that $\psi_1 < \psi_2 < ... < \psi_N$. In this particular case, these numbers are the image quality levels from 8 to 12.5 with a precision of 0.5, as follows:

$$\psi_k; k = 1, 2, ..., 10 = 8, 8.5, 9, 9.5, 10, 10.5, 11, 11.5, 12, 12.5 \tag{6.3.1}$$

The experiment is designed to present the observer with quadruples $(L_a, L_b; L_c, L_d)$, which means, two pairs of images a, b and c, d on each trial. The observer is instructed to select which pair a, b and c, d is perceived as more different. During the experiment, the participant observes pairs of images that are considerably different; that is, image a is evidently different from image b, and image c in evidently different from image d. Nevertheless, the observer is not asked to distinguish the two

images in each pair, but to order the perceived magnitude of threshold perceptual differences. Using the MDLS method, the objective is to determine a set of difference scale values $\psi_1, \psi_2, ..., \psi_N$ that correspond to the stimuli, $L_1, L_2, ..., L_N$, in such way that given a quadruple $(L_a, L_b; L_c, L_d)$, the observer judges L_a, L_b to have more difference than L_c, L_d precisely when:

$$\psi_b - \psi_a > \psi_d - \psi_c \tag{6.3.2}$$

Thus, these scaling values predict judgments of perceptual difference. It is unlikely that human observers are 100% reliable in judgment and satisfy the given criterion, particularly if the differences $\psi_b - \psi_a$ and $\psi_d - \psi_c$ were very close. Then, a model that allows the observer to exhibit stochastic variation in judgment is applied [72], [10], which is based on an equal-variance Gaussian signal detection model. For this model, the signal Δ is the difference in the length $L_{ij} = |\psi_j - \psi_i|$ of the intervals, as follows;

$$\delta(a, b; c, d) = L_{cd} - L_{ab} = \psi_d - \psi_c - \psi_b - \psi_a \tag{6.3.3}$$

If the observer chooses the second interval (L_c, L_d) as larger, then the signal is positive, and if the observer chooses the first interval as larger, δ is negative. Then, it is assumed that the decision variable employed by the observer is;

$$\Delta(a, b; c, d) = \delta(a, b; c, d) + \varepsilon = L_{cd} - L_{ab} + \varepsilon \tag{6.3.4}$$

where ε is a Gaussian random variable with mean zero and standard deviation $\alpha > 0$.

In the experiment, the observer completes P trials, each based on a quadruple q^m with $m = 1, 2, ..., P$. Then, the observer's response is coded as $R^m = 0$ if the difference of the first pair is judged larger or $R^m = 1$ if the second pair's difference is judged larger.

The difference scale values $\Psi = \psi_1, \psi_2, ..., \psi_N$ and the standard deviation α are

the parameters to be obtained by maximizing the likelihood function:

$$L(\psi, \alpha) = \prod_{m=1}^{p} \phi(\frac{\delta(q^m)}{\alpha})^{1-R^m}(1 - \phi(\frac{\delta(q^m)}{\alpha}))^{R^m} \tag{6.3.5}$$

where $\phi(x)$ denotes the Gaussian cumulative distribution function, and $\delta(q^m)$ was defined in the equation 6.3.3. Without loss of generality, we set $\psi_1 = 0$ and $\psi_N = 1$, leaving $N - 1$ parameters to estimate: $\psi_2, ..., \psi_N - 1$ and α.

6.3.1 Palamedes toolbox

Palamedes routine has the brain of MLDS method for deriving the perceptual scale. It was developed by Prins and Kingdom 2009 [85]. It is a free MATLAB toolbox and does not require high level programming skills but basic statistical understanding. The routine can be downloaded from `http://www.palamedestoolbox.org/download.html`.

All functions in Palamedes are prefixed with PAL, to avoid any confusion with existing functions in Matlab software. After PAL, are the names of function they implement. Table 6.1 lists the acronyms we used for MLDS Palamedes routine.

Function Name	Meaning
PF	Psychometric function
PFML	Psychometric function: Maximum likelihood
MLDS	Maximum likelihood difference scaling

Table 6.1: Acronyms used in Palamedes software.

The routine draw out parameter estimates describing the perceptual scale based on the judgment of perceived differences between stimuli from the observer [72], [84]. The stimuli was presented as two pairs of images and the standard errors can be determined by using a bootstrap procedure. The task is 2AFC, as a pair of stimuli will be observed as to which appears to have the greatest magnitude.

Palamedes can be used to find out the relative merits in order to establish a perceptual scale. As found in fitting procedures, free parameters must be estimated in the beginning of the routine. Specify the type of methods (the pairs, triads, or

quadruples), the amount of stimulus levels, the number of repeats for each stimulus combination, and an estimated internal noise level. Therefore, the arguments should be as 2,10,40 respectively and the best SDnoise for each user is initially set to 0.3 but in our case we did it manually based in each model. The routine will display a graph similar to what can be found in Figure 6.5 and an output of the number of trials that occurred in the simulation. In this study, the number was found to be around 4200.

To understand the Palamedes, MLDS routine uses simulated data set for the representation of their operation. So, the first step is the generation of dataset using the $PAL\_MLDS\_GeneratestimList$ routine. This routine is being used for MLDS demonstration as well as for stimulus generation. After this, we need to simulate the response of the observer for that perceptual scale to measure the response by using $PAL\_MLDS\_SimulateObserver$ then fitting the data with MLDS using the MLDS fitting routine PAL_MLDS_Fit. A function PAL_PFML_Fit is used to perform a simplex search. The simplex search will chase after using a likelihood function and finds a maximum after certain iterations. Precision of parameters can also increase with the help of available options to change the tolerance. For example, low tolerance means high precision. In order to decrease the tolerance, the maximum number of iterations needs to increase using PAL_PFML_Fit function. The output of this function is a vector containing all fitted parameters known. $PAL\_MLDS\_Bootstrap$ function is used for bootstrap analysis for error estimation. At the end, results are demonstrated using demonstration routine PAL_MLDS_Demo.

Therefore, it can be said that, Palamedes use a maximum likelihood criterion to fit the psychometric functions among all possible psychometric functions like all combinations of possible values for the free parameters, Palamedes finds that psychometric function with a responses that maximum matches to the observer response.

Palamedes can fit psychometric functions to many different conditions at the same time, although provide flexibility to the user in defining a model. Statistical comparisons between models can be done using Palamedes. For examples a test to check the difference between at least two conditions, like to check the slopes equality

between the conditions or to check the lapse rate difference from a certain value. Finally, Palamedes permits to check the goodness-of-fit of a model defined by the user that defines the efficacy of experiment.

6.4 The Experiment

We followed the experimental design used in [9] for compressed images. We used Max-Planck model rendered with two different methods, diffuse and plastic, with ten quantisation levels of the mesh going from $l = 8$ bits per vertex coordinate (the Level 1 of our Figures) and going up to to $l = 12.5$ bits per vertex coordinate (the Level 10 of our Figures), in intervals of half as shown in Figure 6.6. The mesh vertices were lying on a regular $(2^l) \times (2^l) \times (2^l)$ grid.

The aim is to use MLDS estimation method to scale thresholds of this model, which will show the relation between quantisation thresholds and rendering algorithms and evalute the human performance. Diffuse rendering method is know to be less sensitive to small perturbations of the model's normals caused by different level of quantisation. On the other hand, specular rendering condition has a glossy plastic surface which makes it more sensitive to normal perturbations than a diffuse one, we expect that it would make quantisation artifacts easier to detect.

Each participant was presented with four images of quantised meshes of the Max-Planck model arranged in a 2×2 design, see Figure 6.7 for the experiment's interface. The participants were asked whether the top or the bottom pair of images have larger difference. We use all 210 possible combinations of choosing 4 levels out of the 10, and use one matte and one glossy rendering for a total of 420 trials per participant.

Prior to the actual experiment the 14 candidates were given instructions about how they will run and choose the stimuli. For each of the two rendering method combinations of the ten quantisation levels were presented to the user in random order. For each trial the user had 6 seconds to respond, after which the screen would go blank while awaiting the user's response.

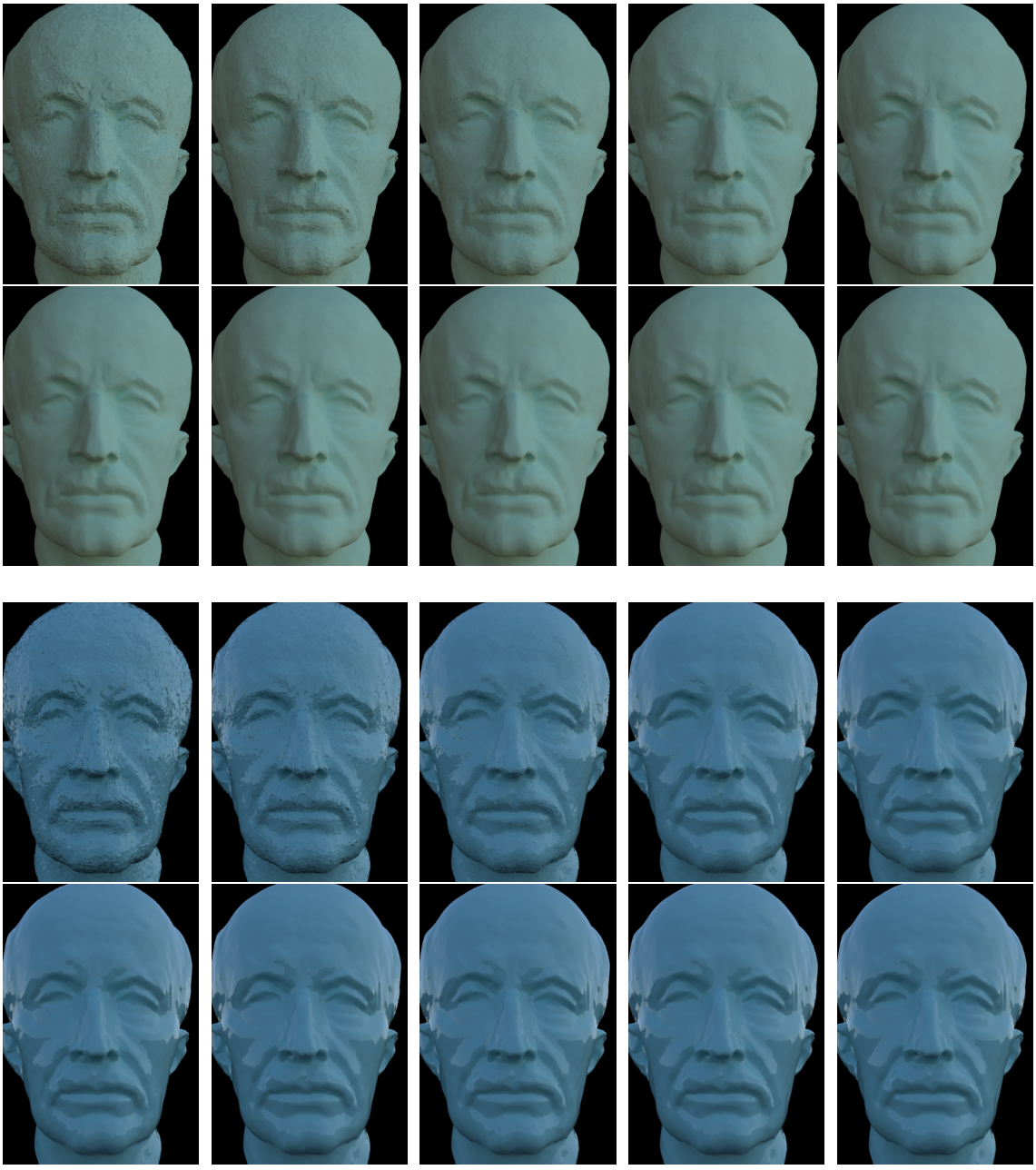


Figure 6.6: Quantised Max-Planck Models from Level 8 to 12.5. **left to right** Diffuse rendering then Plastic rendering.

Figure 6.7: The design of the MLDS experiment.

6.5 Results and Discussion

We fitted a difference scaling curve on the data collected from each participant using the Palamedes software. Notice that MLDS is not a statistical aggregation method requiring large number of participants for valid results and, for example, the experiment in [9] was based on only two participants. Since the convergence of the iterative fitting scheme depends on the initial Noise value, we run the optimization for a range of initial Noise values and chose the result with the highest log-likelihood. In many of the 14 cases we could not fit a good curve, that is, the error estimates exceeded the range of the values.

The results of applying a MLDS analysis on each of the 14 participants are shown below. The horizontal axis is the quantisation level. The vertical axis is the estimated scaling difference. After analysing the results, we can group them together in 3 different groups, based on a qualitative analysis of the resulting curves. Group I comprises of participants 1, 3 and 13. Group II comprises of participants 2, 4, 5,

6, 7, 8, 9, 10, 11 , and 14. Lastly, group III includes participant 12 only.

Figure 6.8 shows a typical difference scaling curve with low error estimates. We notice that the curve rises sharply, despite the low 0.5 bit per vertex coordinate granularity, and levels off at 10 bits, perhaps when a perceptual limit imposed by screen resolution is reached. In this particular participant, as well in some but not all of the others participants, the difference between the two coarser quantisations is not the largest one, indicating that some participants were distinguishing between different coarse quantisations, while others would only distinguish between coarse and fine quantisations.

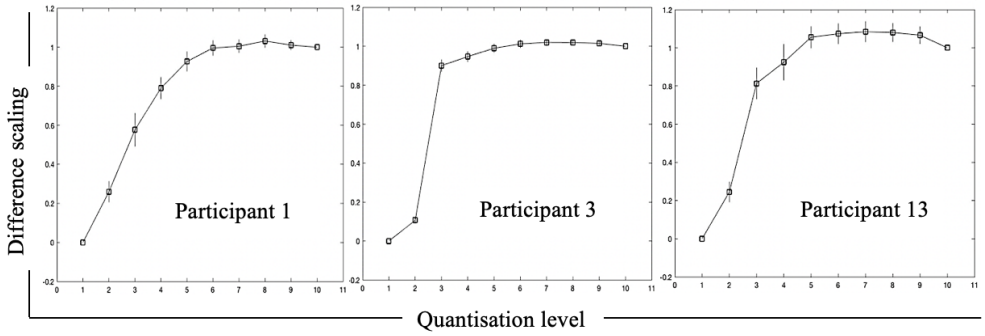


Figure 6.8: Difference scaling curves for participant 1, 3 and 13.

It is evident from Figure 6.8, particularly participant 1, the image quantisation level has significantly high effect on the image till the quantisation up to Level 6, but, above that point, small changes in result in least perceived differences in the image. Similarly, in the participant 3 and 13 the image quantisation level has significantly high and sharp effect on the image between the Levels 2 to 3. Where it took the scaling difference suddenly from approximately 0.1 to 0.9, that means a clear degradation in the image quality. However, above that point, a plateau can been observed. Thus, it can be inferred that beyond this threshold image quality will remain stable even with high level of quantisation. Thus, the results of the difference scaling do suggest that above the quantisation level, the benefits of image compression come with relatively little change in image quality. That clearly depicting it as a stable graph with an ideal curve that does not further effecting the quality as the cost of compression.

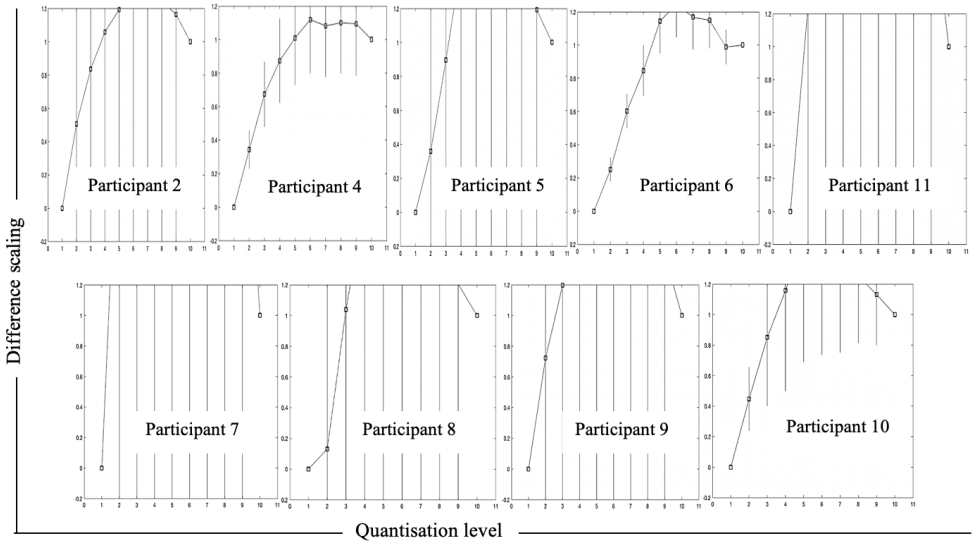


Figure 6.9: Difference scaling curves.

The scaling differences shown in Figure 6.9 exhibit the results from the group II participants. Irregular curves with abrupt changes can be observed indicating the low reliability of our statistical modelling method in this case. By taking the curves at face value, they would be indicating that even small changes at the image quantization level have consistently high effect on the image quality, throughout the curve. The more you will compress the image, the more quality will be degraded due to large scaling differences and above a certain threshold this difference continue to be high. However, given the wide confidence intervals, shown by the thin vertical curves, it is really difficult to infer any meaningful information or significant trends from these type of graphs.

A unique but irregular dangling situation in scaling difference is evident by the participant 12, shown in Figure 6.10, where more scaling difference has been observed from 0 to 0.8 which is quite significantly high difference resultant into low quality of image for the quantization level 1 to 2. However, after level 2, scaling difference is observed between 0.8 and 1.2 for the rest of all quantization levels. High scaling difference at one quantization level and then low at other moment, in this way a zigzag symmetry has been observed.

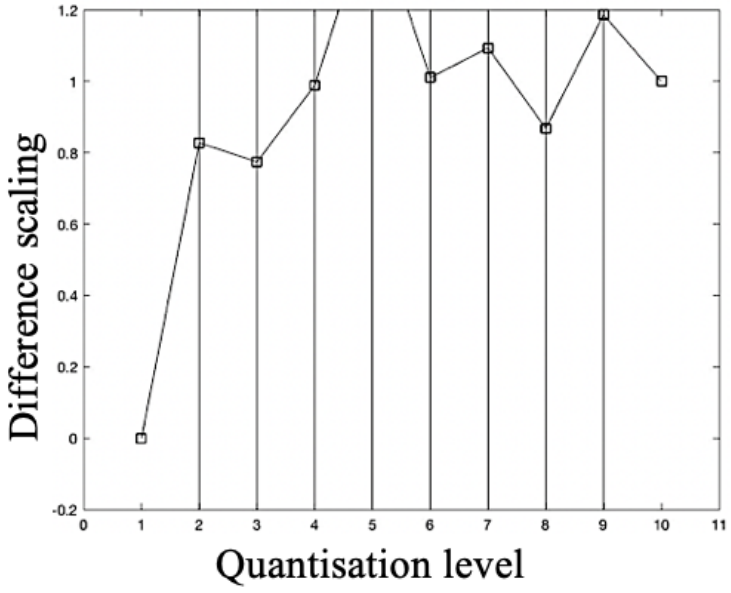


Figure 6.10: Difference scaling curves for participant 12.

6.5.1 Quality difference vs differences in visual appearance

From the results, it seems that there was some sort of confusion in participants regarding the meaning of the trial's question, and some of them were comparing image quality, as instructed, while some others were comparing visual differences between image pairs, irrespective of their quality.

Indeed, there seems to be an inherent difficulty when using the MLDS design, namely, the non-uniformity of the quality scale when parametrised by the amount of differences in visual appearance between two models. To demonstrate this by a schematic example, two different quantisations of the Max-Planck model, both of them at the high-end of the quality scale, let say at around 90 out of 100, will look very similar between them, and indeed, very similar to the original. In contrast, two different quantisations, both of them at the low-end of the quality scale, let say at around 10 out of 100, could be very different between them in appearance.

To demonstrate this phenomenon, in Figure 6.11 we show the number of times each pair of quantisation levels, which could either have been (a, b) or (c, d), was chosen by a certain participant. We notice that we were able to construct good MLE curves, in the cases of participants 1 and 3. There, the number of times

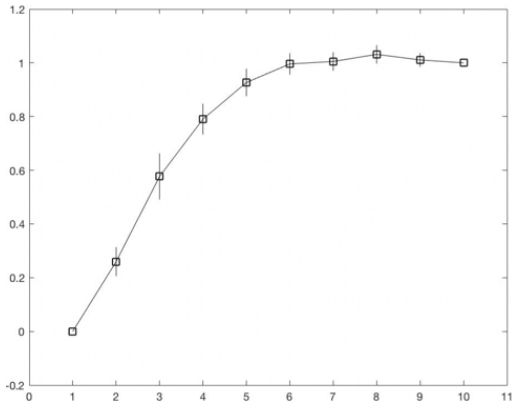
that pairs of two coarse quantisations were chosen, for example (8,8.5), (8,9), or (8.5,9), is relatively low. That is, these participants were able to understand that two very coarse quantisations, such as for example 8 and 8.5, even though they were visually very different, they had nevertheless similar very low quality. On the other hand, participants 5 and 12, they were choosing coarse pairs such as (8,8.5), (8,9), and (8.5,9) almost every time they were presented to them. That means that they interpreted the large difference in the visual appearance between two coarse quantisations as difference in quality. We notice that the corresponding MLE curves for such participants were bad.

6.6 Conclusion

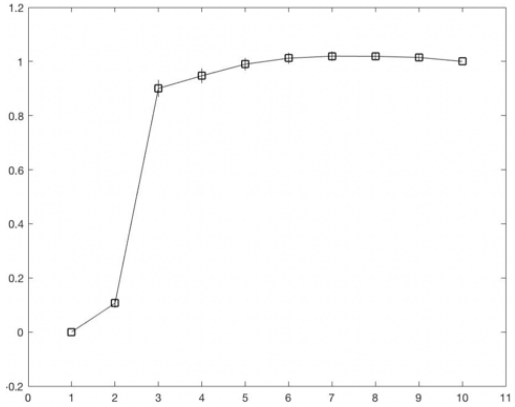
We presented an Maximum Likelihood difference scaling (MLDS) method to estimate the perceptual scale of differences in stimulus based on the participant judgment [72]. The main goal was to use the MLDS to question the relationship between quantisation thresholds, geometric characteristics of the mesh and properties of the rendering style which has been proven to be a powerful approach to similar problems [10, 72].

In all, it was an ambitious undertaking to make use of scaling method for estimating the perceived quality of mesh thresholds, even though we have shown at least that the quantisation had the expected significant impact on discrimination thresholds, in some cases. The main drawback in our implementation of the experiment was that we did not train the participants sufficiently well for what was a more complex task than a 2-AFC experiment. In the future, we would like to repeat this experiment, with better trained participants. In particular, in a follow-up experiment, the participants should be clearly instructed, and trained, to judge differences in perceived visual quality, not differences in visual appearance.

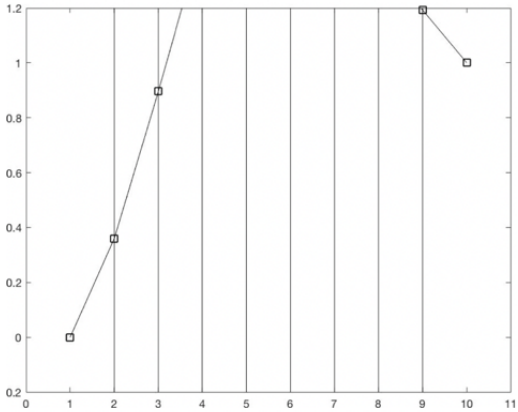
Quantisation level of a or c	Quantisation level of b or d									
	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5
8	0	34	34	28	17	11	4	1	0	0
8.5	0	0	30	26	16	9	5	1	0	0
9	0	0	0	25	15	8	6	3	2	0
9.5	0	0	0	0	18	8	7	2	3	1
10	0	0	0	0	0	12	5	5	2	2
10.5	0	0	0	0	0	0	4	5	5	5
11	0	0	0	0	0	0	0	9	8	9
11.5	0	0	0	0	0	0	0	0	9	11
12	0	0	0	0	0	0	0	0	0	15
12.5	0	0	0	0	0	0	0	0	0	0



Quantisation level of a or c	Quantisation level of b or d									
	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5
8	0	39	30	19	17	9	3	1	0	0
8.5	0	0	26	17	14	8	4	2	0	0
9	0	0	0	21	18	11	1	1	1	2
9.5	0	0	0	0	14	8	9	2	1	1
10	0	0	0	0	0	14	8	4	7	4
10.5	0	0	0	0	0	0	10	7	3	7
11	0	0	0	0	0	0	0	9	11	7
11.5	0	0	0	0	0	0	0	0	17	16
12	0	0	0	0	0	0	0	0	0	17
12.5	0	0	0	0	0	0	0	0	0	0



Quantisation level of a or c	Quantisation level of b or d									
	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5
8	0	52	42	30	19	12	6	2	0	0
8.5	0	0	40	29	20	12	6	2	0	0
9	0	0	0	30	20	12	8	2	1	0
9.5	0	0	0	0	21	12	6	2	0	1
10	0	0	0	0	0	10	6	2	0	0
10.5	0	0	0	0	0	0	6	2	0	0
11	0	0	0	0	0	0	0	0	3	0
11.5	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	4
12.5	0	0	0	0	0	0	0	0	0	0



Quantisation level of a or c	Quantisation level of b or d									
	8	8.5	9	9.5	10	10.5	11	11.5	12	12.5
8	0	54	42	30	20	12	6	2	0	0
8.5	0	0	42	29	20	12	6	2	0	0
9	0	0	0	30	20	12	6	2	1	1
9.5	0	0	0	0	19	12	6	2	0	0
10	0	0	0	0	0	12	6	1	0	0
10.5	0	0	0	0	0	0	5	2	0	0
11	0	0	0	0	0	0	0	3	1	0
11.5	0	0	0	0	0	0	0	0	1	0
12	0	0	0	0	0	0	0	0	0	1
12.5	0	0	0	0	0	0	0	0	0	0

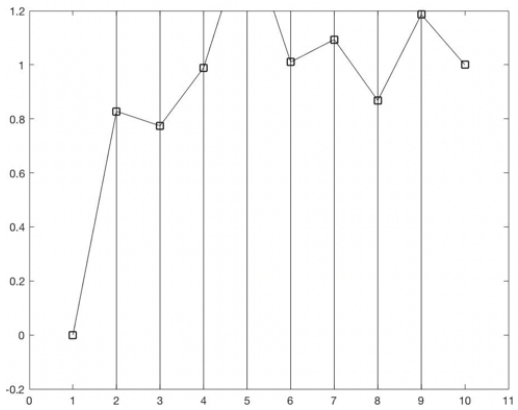


Figure 6.11: From top to bottom: participants 1, 3, 5, and 12. The tables on the left, show the number of times a pair of quantisations was chosen by that participant. The corresponding MLE curves are shown on the right.

Chapter 7

Conclusions

In this thesis, we have presented a research study in the area of perceptual 3D mesh quality. Our main objective was the experimental study of the visual impact of the quantisation of the vertex coordinates of a 3D triangular mesh. We accomplished this goal through conducting four experiments to evaluate the participants perception of quantised meshes, and estimate discrimination thresholds, that is, estimating the quantisation levels after which the quantised mesh becomes visually indistinguishable from the original. Results obtained by these experiments established links and correlations between discrimination thresholds and a number of factors related to quantisation processes, rendering methods, and most importantly, characteristics of the mesh geometry.

7.1 Summary of Contributions

The contributions of this manuscript can be summarized as follows:

- **Design, implementation and analysis of an objective experiment to compare between two quantisation methods, rounding and dithering.**

  The experimental study focuses on the comparison between two different quantisation methods. Rounding, sets all the bits above the quantisation level to zero. Dithering, on the other hand, all bits above the quantisation level are

considered as having a random value. The experimental results showed that dithering has a higher quantisation threshold than rounding while the difference between the two methods is small, around one bit per vertex coordinate, it is nevertheless statistically significant and that was not proven before.

- **Designing an experiment to study the psychology of observer performance over the quality of 3D models.**

  In the experiment we used MLDS for modeling the perceived visual differences when the vertices of a triangle mesh are quantized at various levels. The aim was to predict threshold scale where the original model become indistinguishable from the quantised one and by that we use this range to compress models which is better than a single value.

- **Build a relationship between the quantisation thresholds and the size of the mesh and rendering algorithms.**

  We design an a two-alternative forced-choice psychophysical experiment. Our results indicate that meshes with larger number of triangles require higher level of quantisation. As an intuitive clarification for this, we observe that higher resolution 3D models have more details, which require higher quantisation levels to be represented with high quality. Moreover, larger number of triangles resolutions means smaller and skinny triangles, the normals of which are more sensitive to the spatial perturbations caused by the quantisation and thus, the rendering process, which is based on normal information, is affected more severely. Moreover, as expected, renderings based on reflectance models with higher specular component require higher quantisation level, the reason again being that normal perturbations are easier to perceive in glossy rendering with a strong specular component. Briefly, by proving the interactions between the geometric characteristics and rendering method we could go further with this study.

- **Fit a psychometric function.**

  To conclude our research on different measurments that affect the quantisation

level and prove it one by one. We, instantly, followed the design of experiment 3, yet each mesh consist of about 100K vertices and used one rendering condition. The results had proven that there is a clear correlation between the discrimination threshold and the filesizes of the compressed meshes.

In a nutshell, this thesis work highlights the importance of exploiting the relationship between geometry, material and lighting information of 3D models and the psychology of integrating human in the quality evaluating task of visual perception. As a matter of fact, future research will have to continue investigating these relationship and the effect of integrating human vision processes on the visual quality of 3D models.

7.2 Perspective

Several research perspectives appear at the end of this research.

We found that the discrimination thresholds are affected by a multitude of factors.

Even it might be intuitive and small, we found that the blockiness has higher quantisation level than high frequency noise which was proven to be perceptually stronger in the first experiment from comparing two quantisation methods. Similarly to rendering methods, we found an observable impact on quantisation level from experiment 2 and 3. While we experimented only with common rendering methods, further investigation could be done on methods that decrease the threshold such as flat rendering, or sophisticated methods that increase it such as reflection lines.

We found that the relationship between quantisation thresholds and geometric properties of the mesh was evident.

Important factors were identified to be the size of the mesh (number of triangles) in experiment 4 while geometry was also important in experiment 3.

We found that quantisation can actually be beneficial.

In literature we found that 16 bits per vertex coordinate is enough for most cases as the standard quantisation level [71], [49], when they report results in compression, or watermarking. Yet, we found it best to use 8 to 10 bits.

The design of the experiment is very important.

Very complex designs for objective experiment are problematic and that is what we faced in MLDS experiment , while simpler ones seem to work better for untrained users as in other experiments.

7.2.1 Directions for future work

Taken altogether, the data presented in this thesis provide evidence of the existence of statistically significant relationships between quantisation thresholds, geometric characteristics of the mesh and rendering settings, we could not quantify them, that is, we did not have enough data to produce a formula relating, for example, the number of triangles or the light information in the mesh with the quantisation threshold, which we intend for future plan. Particularly, we still need more progression in understanding the interaction between geometry, material and rendering methods, and also the connections between these interactions and human visual system.

Additionally, we need to consider if these factors apply to animations and video as well as images ?. Will similar factors have the same impact on their visual quality and results as well as the fixed images ?. We need to consider the real-time interactions as an important factor when measuring the quality of video.

The fields of visual quality of 3D graphical data are still challenging. Indeed, all the above factors should be considered to deliver efficient quality in future. Overall, we admit that the research in this area has a promising future.

7.2.2 Relevant publications

Almutairi, Aeshah, Toni Saarela, and Ioannis Ivrissimtzis. "A user study on quantisation thresholds of triangle meshes." The Computer Graphics and Visual Computing (CGVC), 2017.

Almutairi, Aeshah, Toni Saarela, and Ioannis Ivrissimtzis. "A perceptual difference scaling study on quantized 3D models." The ACM Symposium on Applied Perception (ACM SAP), 2018.

Aeshah Almutairi, Ioannis Ivrissimtzis, and Toni Saarela. "Imperceptibility

thresholds in quantised 3D triangle meshes." The 4th International Conference on Image and Graphics Processing, 2021.

Aeshah Almutairi, Ioannis Ivrissimtzis, and Toni Saarela."Quality perception and discrimination thresholds in quantised triangle meshes." The International Workshop on Image Processing (IWIP), 2021 .

Appendix A

Appendix A

In chapter 2 We present the results of a user study on estimating a quantisation threshold above which the quantised triangle mesh is perceived as indistinguishable from its unquantised original. The experiment focuses on the comparison between two different quantisation methods: rounding, in which all bits above the threshold are put to zero; and dithering, in which all bits above the threshold are randomised. We used three different mesh models, mesh 1 Cube, mesh 2 is Eight and mesh 3 is Max-Planck.

The following figures show the results of the mesh models (Cube, Eight or Max-Planck) and algorithms (rounding or dithering).

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Threads	20	100.0%	0	0.0%	20	100.0%

			Statistic	Std. Error
Threads	Mean		9.4590	.67855
	95% Confidence Interval for Mean	Lower Bound	8.0388	
		Upper Bound	10.8792	
	5% Trimmed Mean		9.1878	
	Median		8.9000	
	Variance		9.209	
	Std. Deviation		3.03458	
	Minimum		6.40	
	Maximum		17.40	
	Range		11.00	
	Interquartile Range		3.23	
	Skewness		1.299	.512
	Kurtosis		1.189	.992

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Threads	.184	20	.074	.855	20	.006

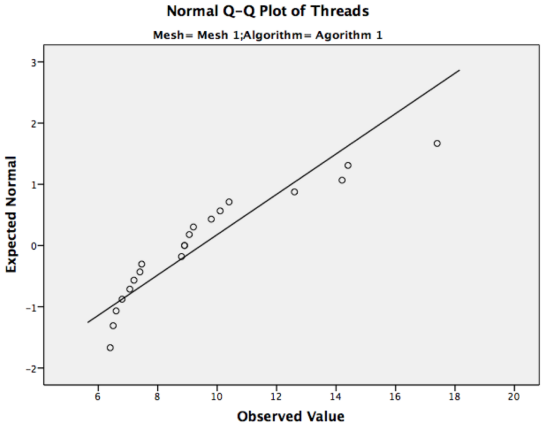


Figure A.1: The result of Cube with rounding.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Threads	20	100.0%	0	0.0%	20	100.0%

			Statistic	Std. Error
Threads	Mean		10.0745	.65506
	95% Confidence Interval for Mean	Lower Bound	8.7034	
		Upper Bound	11.4456	
	5% Trimmed Mean		9.7811	
	Median		9.2000	
	Variance		8.582	
	Std. Deviation		2.92952	
	Minimum		7.30	
	Maximum		18.13	
	Range		10.83	
	Interquartile Range		4.15	
	Skewness		1.389	.512
	Kurtosis		2.004	.992

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Threads	.178	20	.097	.838	20	.003

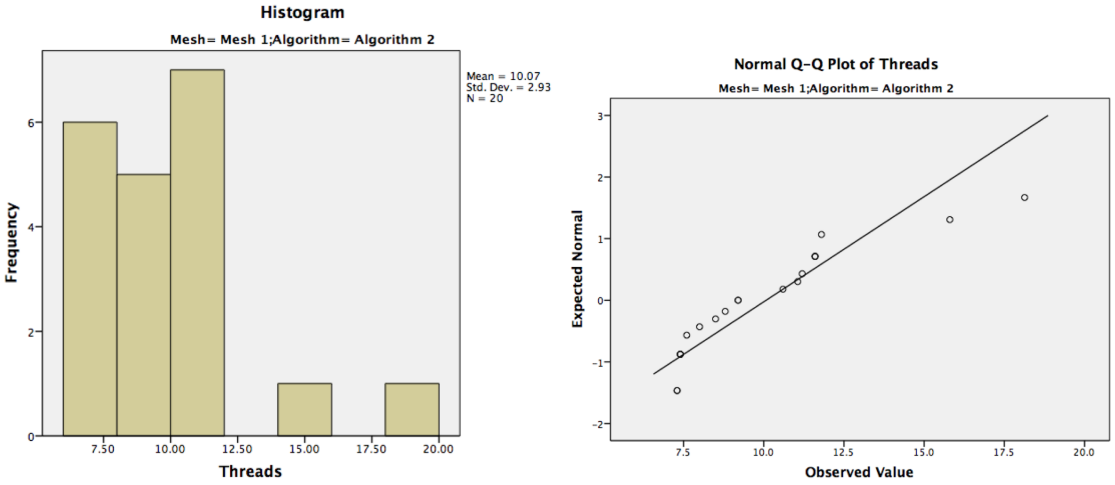


Figure A.2: The result of Cube with dithering.

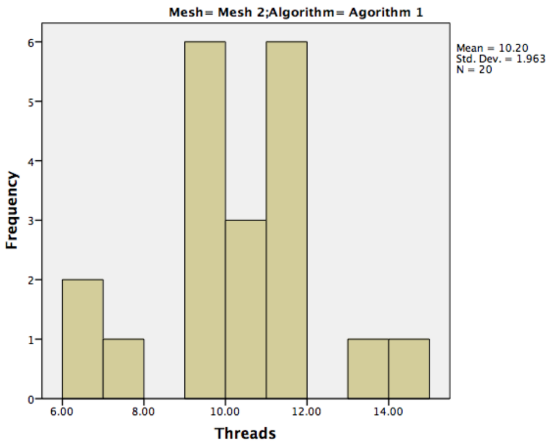
Case Processing Summary[a]

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Threads	20	100.0%	0	0.0%	20	100.0%

			Statistic	Std. Error
Threads	Mean		10.2015	.43888
	95% Confidence Interval for Mean	Lower Bound	9.2829	
		Upper Bound	11.1201	
	5% Trimmed Mean		10.1739	
	Median		10.1300	
	Variance		3.852	
	Std. Deviation		1.96272	
	Minimum		6.50	
	Maximum		14.40	
	Range		7.90	
	Interquartile Range		1.87	
	Skewness		.032	.512
	Kurtosis		.595	.992

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Threads	.130	20	.200*	.951	20	.376



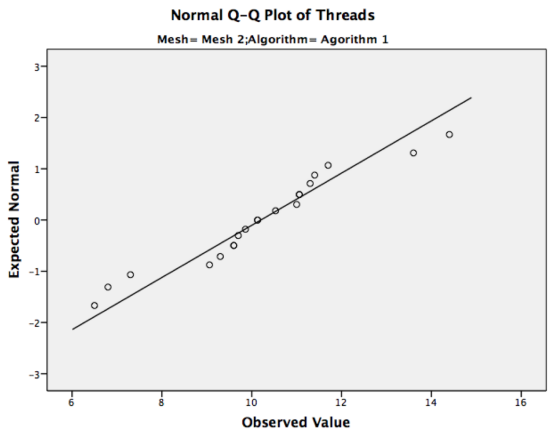


Figure A.3: The result of Eight with rounding.

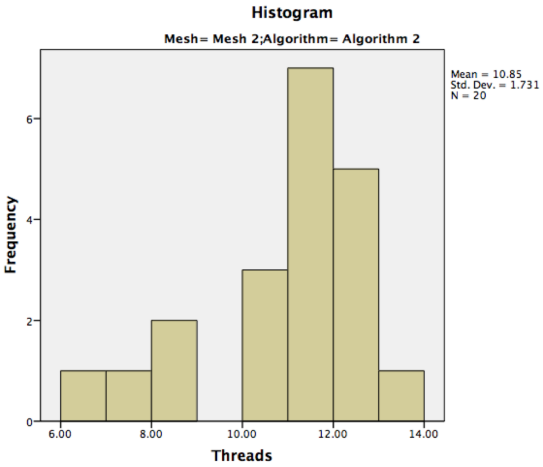
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Threads	20	100.0%	0	0.0%	20	100.0%

			Statistic	Std. Error
Threads	Mean		10.8475	.38706
	95% Confidence Interval for Mean	Lower Bound	10.0374	
		Upper Bound	11.6576	
	5% Trimmed Mean		10.9528	
	Median		11.3800	
	Variance		2.996	
	Std. Deviation		1.73097	
	Minimum		6.60	
	Maximum		13.20	
	Range		6.60	
	Interquartile Range		1.65	
	Skewness		-1.182	.512
	Kurtosis		.814	.992

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Threads	.198	20	.039	.881	20	.018



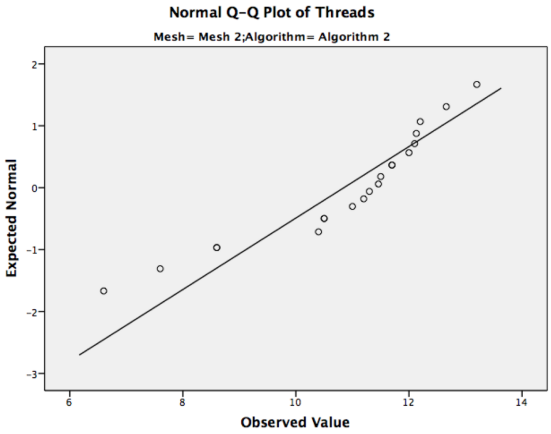


Figure A.4: The result of Eight with dithering.

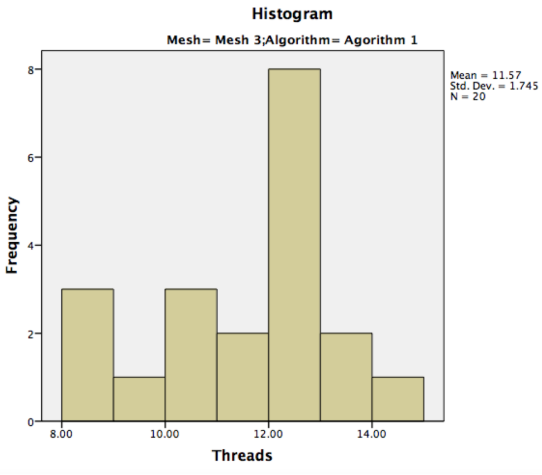
Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Threads	20	100.0%	0	0.0%	20	100.0%

			Statistic	Std. Error
Threads	Mean		11.5700	.39022
	95% Confidence Interval for Mean	Lower Bound	10.7533	
		Upper Bound	12.3867	
	5% Trimmed Mean		11.5556	
	Median		12.0000	
	Variance		3.045	
	Std. Deviation		1.74510	
	Minimum		8.50	
	Maximum		14.90	
	Range		6.40	
	Interquartile Range		2.50	
	Skewness		-.222	.512
	Kurtosis		-.613	.992

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Threads	.147	20	.200*	.960	20	.539



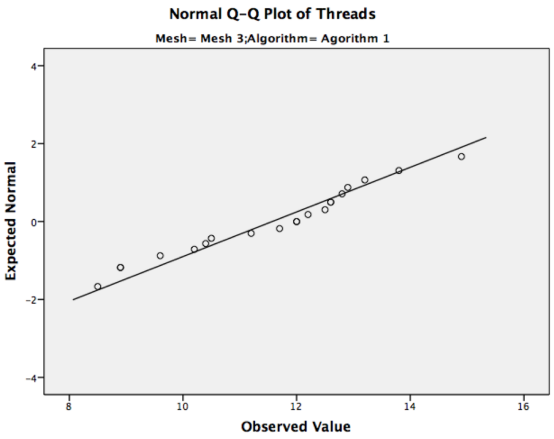


Figure A.5: The result of Max-Planck with rounding.

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Threads	20	100.0%	0	0.0%	20	100.0%

			Statistic	Std. Error
Threads	Mean		13.0200	.64173
	95% Confidence Interval for Mean	Lower Bound	11.6768	
		Upper Bound	14.3632	
	5% Trimmed Mean		12.7500	
	Median		12.5000	
	Variance		8.236	
	Std. Deviation		2.86992	
	Minimum		9.20	
	Maximum		21.70	
	Range		12.50	
	Interquartile Range		2.00	
	Skewness		1.914	.512
	Kurtosis		4.255	.992

Tests of Normality

	Kolmogorov-Smirnov			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Threads	.297	20	.000	.794	20	.001

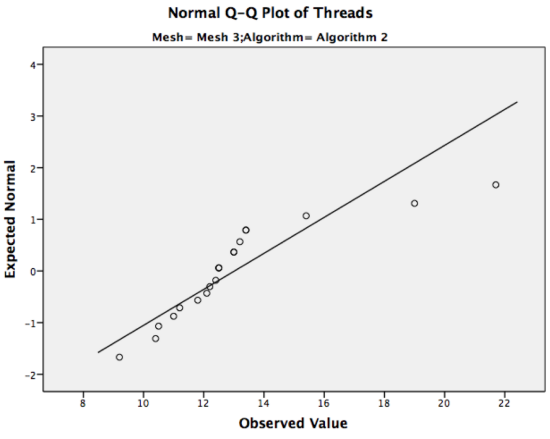


Figure A.6: The result of Max-Planck with dithering.

Bibliography

[1] Methodology for the subjective assessment of the quality of television pictures. *Technical Report. Recommendation ITU-R BT.500-11*, 1998.

[2] A. Almutairi, T. Saarela, and I. Ivrissimtzis. A user study on quantisation thresholds of triangle meshes. In *Computer Graphics and Visual Computing*. The Eurographics Association, 2017.

[3] A. Almutairi, T. Saarela, and I. Ivrissimtzis. Imperceptibility thresholds in quantised 3d triangle meshes. In *Proceedings of the The 4th International Conference on Image and Graphics Processing (ICIGP)*, 2021.

[4] RJ. Baker and S. Rosen. Evaluation of maximum-likelihood threshold estimation with tone-in-noise masking. *British Journal of Audiology*, 35(1):43–52, 2001.

[5] R. E. Bank and R. Smith. Mesh smoothing using a posteriori error estimates. *SIAM Journal on Numerical Analysis*, 34(3):979–997, 1997.

[6] M. Botsch, M. Pauly, L. Kobbelt, P. Alliez, and B. Lévy. Geometric Modeling Based on Polygonal Meshes. 2008. This document is the support of a course given at the Eurographics 2008 conference (Crete, Greece, April 14-18).

[7] C. Boucheny, G. Bonneau, J. Droulez, G. Thibault, and S. Ploix. A Perceptive Evaluation of Volume Rendering Techniques. *ACM Transactions on Applied Perception*, 5(4):1–24, January 2009.

[8] C. Charrier, K. Knoblauch, L. Maloney, A. Bovik, and A. Moorthy. Optimizing multiscale ssim for compression via mlds. *IEEE Transactions on Image Processing*, 21(12):4682–4694, 2012.

[9] C. Charrier, K. Knoblauch, A. Moorthy, A. Bovik, and L. Maloney. Comparison of image quality assessment algorithms on compressed images. In *Image Quality and System Performance VII*, volume 7529, page 75290B, 2010.

[10] C. Charrier, L. Maloney, H. Cherifi, and K. Knoblauch. Maximum likelihood difference scaling of image quality in compression-degraded images. *JOSA A*, 24(11):3418–3426, 2007.

[11] D. Chen, X. Tian, Y. Shen, and M. Ouhyoung. On visual similarity based 3d model retrieval. In *Computer graphics forum*, volume 22, pages 223–232. Wiley Online Library, 2003.

[12] M. Chen and H. Jaenicke. An information-theoretic framework for visualization. *IEEE Transactions on Visualization and Computer Graphics*, 16(6):1206–1215, Nov 2010.

[13] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia. MeshLab: an Open-Source Mesh Processing Tool. In Vittorio Scarano, Rosario De Chiara, and Ugo Erra, editors, *Eurographics Italian Chapter Conference*. The Eurographics Association, 2008.

[14] T. Cornsweet. The staircase-method in psychophysics. *The American journal of psychology*, 75(3):485–491, 1962.

[15] C. Correa, R. Hero, and K. Ma. A comparison of gradient estimation methods for volume rendering on unstructured meshes. 2011.

[16] M. Corsini, E. Gelasca, T. Ebrahimi, and M. Barni. Watermarked 3-d mesh quality assessment. *IEEE Transactions on Multimedia*, 9(2):247–256, 2007.

[17] M. Corsini, M. Larabi, G. Lavoué, O. Petřík, L. Váša, and K. Wang. Perceptual metrics for static and dynamic triangle meshes. In *Computer Graphics Forum*, volume 32, pages 101–125. Wiley Online Library, 2013.

[18] P. Cosman, K. Oehler, E. Riskin, and R. Gray. Using vector quantization for image processing. *Proceedings of the IEEE*, 81(9):1326–1341, 1993.

[19] M. J. Curtis, R. A. Bond, D. Spina, A. Ahluwalia, S. Alexander, M. A Giembycz, A. Gilchrist, D. Hoyer, P. A. Insel, A. A. Izzo, et al. Experimental design and analysis and their reporting: new guidance for publication in bjp. *British journal of pharmacology*, 172(14):3461–3471, 2015.

[20] S. Daly. Visible differences predictor: an algorithm for the assessment of image fidelity. In *Human Vision, Visual Processing, and Digital Display III*, volume 1666, pages 2–15. International Society for Optics and Photonics, 1992.

[21] F. Devinck and K. Knoblauch. A common signal detection model accounts for both perception and discrimination of the watercolor effect. *Journal of Vision*, 12(3):19–19, 03 2012.

[22] WJ. Dixon. The up-and-down method for small samples. *Journal of the American Statistical Association*, 60(312):967–978, 1965.

[23] J. Dompierre, M. Vallet, P. Labbé, and F. Guibault. An analysis of simplex shape measures for anisotropic meshes. *Computer Methods in Applied Mechanics and Engineering*, 194(48-49):4895–4914, 2005.

[24] W. H. Ehrenstein and Addie Ehrenstein. *Psychophysical Methods*, pages 1211–1241. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[25] K. Emrith, M. Chantler, P. Green, L. Maloney, and A. Clarke. Measuring perceived differences in surface texture due to changes in higher order statistics. *Journal of the optical society of america a-Optics image science and vision*, 27(5):1232–1244, May 2010.

[26] A. M. Eskicioglu and P. S. Fisher. Image quality measures and their performance. *IEEE Transactions on Communications*, 43(12):2959–2965, 1995.

[27] Z. Fan and R. De Queiroz. Maximum likelihood estimation of jpeg quantization table in the identification of bitmap compression history. In *Proceedings*

2000 International Conference on Image Processing (Cat. No. 00CH37101), volume 1, pages 948–951. IEEE, 2000.

[28] D. Farrington, D. Gottfredson, L. Sherman, and B. Welsh. The maryland scientific methods scale. *Evidence-based crime prevention*, pages 13–21, 2002.

[29] J. Fischer and D. Whitney. Serial dependence in visual perception. *Nature neuroscience*, 17(5):738–743, 2014.

[30] R. Fleming, R. Dror, and E. Adelson. Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, 3(5):3–3, 2003.

[31] R. Fleming, F. Jäkel, and L. Maloney. Visual perception of thick transparent materials. *Psychological science*, 22(6):812–820, 2011.

[32] A. Forouzan and G. Hill. Data communications and networking, by behrouz. *Forouzan.-*, 2006.

[33] L A. Freitag and P. Knupp. Tetrahedral mesh improvement via optimization of the element condition number. *International Journal for Numerical Methods in Engineering*, 53(6):1377–1391, 2002.

[34] L. A Freitag and C. Ollivier-Gooch. A comparison of tetrahedral mesh improvement techniques. Technical report, Argonne National Lab., IL (United States), 1996.

[35] M. Fritsche, P. Mostert, and F. de Lange. Opposite effects of recent history on perception and decision. *Current Biology*, 27(4):590–595, 2017.

[36] M. A. Garcıa-Pérez. Forced-choice staircases with fixed step sizes: asymptotic and small-sample properties. *Vision research*, 38(12):1861–1881, 1998.

[37] A. Gargallo-Peiró, X. Roca, J. Peraire, and J. Sarrate. Defining quality measures for mesh optimization on parameterized cad surfaces. In Xiangmin Jiao and Jean-Christophe Weill, editors, *Proceedings of the 21st International Meshing Roundtable*, pages 85–102, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg.

[38] A. Gersho. Principles of quantization. *IEEE Transactions on Circuits and Systems*, 25(7):427–436, 1978.

[39] A. Gersho and R. M. Gray. *Vector quantization and signal compression*, volume 159. Springer Science & Business Media, 2012.

[40] A. Gersho and R.M. Gray. *Vector Quantization and Signal Compression*. The Springer International Series in Engineering and Computer Science. Springer US, 2012.

[41] R. M. Gray and D. L. Neuhoff. Quantization. *IEEE transactions on information theory*, 44(6):2325–2383, 1998.

[42] R. M. Gray and D. L. Neuhoff. Quantization. *IEEE Transactions on Information Theory*, 44(6):2325–2383, 1998.

[43] P. Green. On use of the em algorithm for penalized likelihood estimation. *Journal of the Royal Statistical Society: Series B (Methodological)*, 52(3):443–452, 1990.

[44] J. Guo, V. Vidal, A. Baskurt, and G. Lavoué. Evaluating the local visibility of geometric artifacts. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception*, SAP '15, pages 91–98, New York, NY, USA, 2015. ACM.

[45] Y. Guo, S. Lin, Z. SuXiaonan, , and L. WangYang Kang. A 3d model perceptual feature metric based on global height field. 2016.

[46] S. Hahmann. Visualization techniques for surface analysis. *Data visualization techniques*, pages 49–74, 1999.

[47] S. Haque and V. Govindu. Global mesh denoising with fairness. In *2015 International Conference on 3D Vision (3DV)*, pages 46–54, Los Alamitos, CA, USA, oct 2015. IEEE Computer Society.

[48] L. D. Harmon and B. Julesz. Masking in visual recognition: effects of two-dimensional filtered noise. *Science*, 180(91):1194–1197, 1973.

[49] M. Isenburg, I. Ivrissimtzis, S. Gumhold, and H. Seidel. Geometry prediction for high degree polygons. In *Proceedings of the 21st spring conference on Computer graphics*, pages 147–152. ACM, 2005.

[50] I. Ivrissimtzis. Effects of noise on quantized triangle meshes. In *International Conference on Mathematical Methods for Curves and Surfaces*, pages 274–284. Springer, 2008.

[51] H. Jänicke and M. Chen. A salience-based quality metric for visualization. In *Proceedings of the 12th Eurographics / IEEE - VGTC Conference on Visualization*, EuroVis'10, pages 1183–1192, Chichester, UK, 2010. The Eurographs Association & John Wiley & Sons, Ltd.

[52] J. Kajiya. New techniques for ray tracing procedurally defined objects. *ACM SIGGRAPH Computer Graphics*, 17(3):91–102, 1983.

[53] A. Kalaiah and A. Varshney. Statistical geometry representation for efficient transmission and rendering. *ACM Trans. Graph.*, 24(2):348–373, April 2005.

[54] A. Katz and V. Sankaran. Mesh quality effects on the accuracy of cfd solutions on unstructured meshes. *Journal of Computational Physics*, 230(20):7670–7686, 2011.

[55] A. Kaufman and K. Mueller. *Overview of Volume Rendering*, volume 7, pages 127–XI. 12 2005.

[56] K. Knoblauch, C. Charrier, H. Cherifi, J. Yang, and LT. Maloney. Difference scaling of image quality in compression-degraded images. *Perception ECVP abstract*, 27:0–0, 1998.

[57] K. Knoblauch and L. Maloney. *Modeling psychophysical data in R*, volume 32. Springer Science & Business Media, 2012.

[58] P. Knupp. Achieving finite element mesh quality via optimization of the jacobian matrix norm and associated quantities. part ii—a framework for volume mesh optimization and the condition number of the jacobian matrix. *International Journal for Numerical Methods in Engineering*, 48(8):1165–1185, 2000.

[59] P. Knupp. Algebraic mesh quality metrics. *SIAM journal on scientific computing*, 23(1):193–218, 2001.

[60] P. M. Knupp. Achieving finite element mesh quality via optimization of the jacobian matrix norm and associated quantities. part i—a framework for surface mesh optimization. *International Journal for Numerical Methods in Engineering*, 48:401–420, 05 2000.

[61] W. Kraaij and W. Post. Task based evaluation of exploratory search systems. In *Proc. of SIGIR 2006 Workshop, Evaluation Exploratory Search Systems, Seattle, USA*, pages 24–27, 2006.

[62] D. KWeiskopf and G Erlebacher. *Overview of flow visualization*, pages 127–XI. 12 2005.

[63] G. Lavoué. A local roughness measure for 3d meshes and its application to visual masking. *ACM Transactions on Applied perception (TAP)*, 5(4):21, 2009.

[64] G. Lavoué. A multiscale metric for 3d mesh visual quality assessment. In *Computer Graphics Forum*, volume 30, pages 1427–1437. Wiley Online Library, 2011.

[65] G. Lavoué, I. Cheng, and A. Basu. Perceptual quality metrics for 3d meshes: towards an optimal multi-attribute computational model. In *2013 IEEE International Conference on Systems, Man, and Cybernetics*, pages 3271–3276. IEEE, 2013.

[66] G. Lavoué and M. Corsini. A comparison of perceptually-based metrics for objective evaluation of geometry processing. *IEEE Transactions on Multimedia*, 12(7):636–649, 2010.

[67] H. Levitt. Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical society of America*, 49(2B):467–477, 1971.

[68] M. Levoy and T. Whitted. The use of points as a display primitive. 2000.

[69] P. Lindstrom and G. Turk. Image-driven simplification. *ACM Transactions on Graphics (ToG)*, 19(3):204–241, 2000.

[70] T. Liu, M. Chen, Y. Song, H. Li, and B. Lu. Quality improvement of surface triangular mesh using a modified laplacian smoothing approach avoiding intersection. *PLoS One*, 12(9):e0184206, 2017.

[71] A. Maglo, F. Lavoué, G.and Dupont, and C. Hudelot. 3d mesh compression: Survey, comparisons, and emerging trends. *ACM Comput. Surv.*, 47(3):44:1–44:41, 2015.

[72] L. Maloney and J. Yang. Maximum likelihood difference scaling. *Journal of Vision*, 3(8):5–5, 2003.

[73] M. C. Morrone, D. C. Burr, and J. Ross. Added noise restores recognizability of coarse quantized images. *Nature*, 305(5931):226–228, 1983.

[74] T. Munson. Mesh shape-quality optimization using the inverse mean-ratio metric. *Mathematical Programming*, 110(3):561–590, 2007.

[75] G. Obein, K. Knoblauch, and F. Viéot. Difference scaling of gloss: Nonlinearity, binocularity, and constancy. *Journal of Vision*, 4(9):4–4, 08 2004.

[76] M. Olkkonen and D. Brainard. Joint effects of illumination geometry and object shape in the perception of surface reflectance. *i-Perception*, 2(9):1014–1034, 2011.

[77] Y. Pan, I. Cheng, and A. Basu. Quality metric for approximating subjective evaluation of 3-d objects. *IEEE Transactions on Multimedia*, 7(2):269–279, 2005.

[78] J. Park and S. Shontz. An alternating mesh quality metric scheme for efficient mesh quality improvement. *Procedia Computer Science*, 4:292–301, 2011.

[79] V. Paulun, T. Kawabe, S. Nishida, and R. Fleming. Seeing liquids from static snapshots. *Vision research*, 115:163–174, 2015.

[80] D. Pelli and B. Farell. Why use noise? *JOSA A*, 16(3):647–653, 1999.

[81] J. Peng, C. Kim, and C. Kuo. Technologies for 3d mesh compression: A survey. *Journal of Visual Communication and Image Representation*, 16(6):688–733, 2005.

[82] M. H. Pinson and S. Wolf. Comparing subjective video quality testing methodologies. In *Visual Communications and Image Processing 2003*, pages 573–582. International Society for Optics and Photonics, 2003.

[83] R. Pollak, A. Palazotto, and T. Nicholas. A simulation-based investigation of the staircase method for fatigue strength testing. *Mechanics of materials*, 38(12):1170–1181, 2006.

[84] N. Prins et al. *Psychophysics: a practical introduction*. Academic Press, 2016.

[85] N. Prins and F. Kingdom. Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the palamedes toolbox. *Frontiers in Psychology*, 9:1250, 2018.

[86] D. Quevedo and G. Goodwin. Audio quantization from a receding horizon control perspective. In *Proceedings of the 2003 American Control Conference, 2003.*, volume 5, pages 4131–4136. IEEE, 2003.

[87] D. Roberts and I. Ivrissimtzis. Quality measures of reconstruction filters for stereoscopic volume rendering. *Computational Visual Media*, 2(1):19–30, 2016.

[88] L. Roberts. Picture coding using pseudo-random noise. *IRE Transactions on Information Theory*, 8(2):145–154, 1962.

[89] J. Robertson and M. Kaptein. An introduction to modern statistical methods in hci. In *Modern Statistical Methods for HCI*, pages 1–14. Springer, 2016.

[90] B. Rogowitz and H. Rushmeier. Are image quality metrics adequate to evaluate the quality of geometric objects. *Proc SPIE*, 11 2001.

[91] B. Rogowitz and H. Rushmeier. Are image quality metrics adequate to evaluate the quality of geometric objects? In *Human Vision and Electronic Imaging*

VI, volume 4299, pages 340–348. International Society for Optics and Photonics, 2001.

[92] R. M. Rose, D. Teller, and P. Rendleman. Statistical properties of staircase estimates. *Attention, Perception, & Psychophysics*, 8(4):199–204, 1970.

[93] H. Rushmeier, B. Rogowitz, and C. Piatko. Perceptual issues in substituting texture for geometry. In *Human Vision and Electronic Imaging V*, volume 3959, pages 372–383. International Society for Optics and Photonics, 2000.

[94] L. Sharan, Y. Li, I. Motoyoshi, S. Nishida, and E Adelson. Image statistics for surface reflectance perception. *J. Opt. Soc. Am. A*, 25(4):846–865, Apr 2008.

[95] B. Shelton, M. Picardi, and D. Green. Comparison of three adaptive psychophysical procedures. *The Journal of the Acoustical Society of America*, 71(6):1527–1533, 1982.

[96] W. Sheppard. On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant division of a scale. *Proceedings of the London Mathematical Society*, 1(1):353–380, 1897.

[97] W. F. Sheppard. On the calculation of the most probable values of frequency-constants, for data arranged according to equidistant division of a scale. *Proceedings of the London Mathematical Society*, s1-29(1):353–380, 11 1897.

[98] S. Silva, B. Santos, C. Ferreira, and J. Madeira. A perceptual data repository for polygonal meshes. In *Visualisation, 2009. VIZ'09. Second International Conference in*, pages 207–212. IEEE, 2009.

[99] O. Sorkine, D. Cohen-Or, and S. Toledo. High-pass quantization for mesh encoding. In *Eurographics/ACM SIGGRAPH symposium on Geometry processing*, pages 42–51. The Eurographics Association, 2003.

[100] H. Strasburger. Converting between measures of slope of the psychometric function. *Perception & psychophysics*, 63(8):1348–1355, 2001.

[101] K. Thung and P. Raveendran. A survey of image quality measures. In *2009 International Conference for Technical Postgraduates (TECHPOS)*, pages 1–4, 2009.

[102] F. Torkhani, K. Wang, and J. Chassery. Perceptual quality assessment of 3d dynamic meshes. *Image Commun.*, 31(C):185–204, February 2015.

[103] E. Tosun, Y. I. Gingold, J. Reisman, and D. Zorin. Shape optimization using reflection lines. In *Proceedings of the Fifth Eurographics Symposium on Geometry Processing*, SGP '07, pages 193–202, Aire-la-Ville, Switzerland, Switzerland, 2007. Eurographics Association.

[104] C. Touma and C. Gotsman. Triangle mesh compression. *Proc. Graphics Interface*, pages 26–34, 1998.

[105] J. K. Udupa, H. Hung, and K. Chuang. Surface and volume rendering in three-dimensional imaging: A comparison. *Journal of Digital Imaging*, 4(3):159, 1991.

[106] K. Vanhoey, B. Sauvage, P. Kraemer, and G. Lavoué. Visual quality assessment of 3d models: On the influence of light-material interaction. *ACM Transactions on Applied Perception*, 15:1–18, 10 2017.

[107] K. Wang and M. Marek-Sadowska. Power/ground mesh area optimization using multigrid-based technique [ic design]. In *2003 Design, Automation and Test in Europe Conference and Exhibition*, pages 850–855. IEEE, 2003.

[108] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli. Image quality assessment: From error visibility to structural similarity. *Image Processing, IEEE Transactions on*, 13:600 – 612, 05 2004.

[109] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.

[110] B. Watson, A. Friedman, and A. McGaffey. Using naming time to evaluate quality predictors for model simplification. In *Proceedings of the SIGCHI*

conference on Human factors in computing systems, pages 113–120. ACM, 2000.

[111] B. Watson, A. Friedman, and A. McGaffey. Measuring and predicting visual fidelity. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 213–220. ACM, 2001.

[112] B. Watson, A. Friedman, and A. McGaffey. Measuring and predicting visual fidelity. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '01, page 213–220. ACM, 2001.

[113] A. Watt. *3D Computer Graphics*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 1993.

[114] J. Willis. *A framework for task-based learning*, volume 60. Longman Harlow, 1996.

[115] N. Woodhouse. *Geometric quantization*. Oxford university press, 1997.

[116] D. Yamanaka, Y. Ohtake, and H. Suzuki. The sinogram polygonizer for reconstructing 3d shapes. *IEEE Transactions on Visualization and Computer Graphics*, 19:1911–1922, 2013.

[117] Y. Yang, N. Peyerimhoff, and I. Ivrissimtzis. Linear correlations between spatial and normal noise in triangle meshes. *IEEE Transactions on Visualization and Computer Graphics*, 19(1):45–55, 2012.

[118] Y. Yang, N. Peyerimhoff, and I. Ivrissimtzis. Linear correlations between spatial and normal noise in triangle meshes. *IEEE transactions on visualization and computer graphics*, 19(1):45–55, 2013.

[119] Z. Yildiz, A. Bulbul, and T. Capin. A framework for applying the principles of depth perception to information visualization. *ACM Trans. Appl. Percept.*, 10(4):19:1–19:22, October 2013.