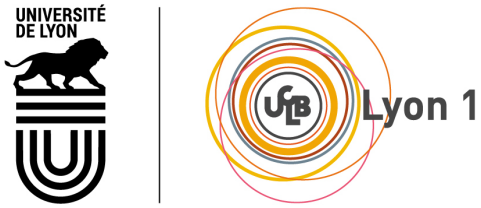
Evaluating the visibility threshold for a local geometric distortion on a 3D mesh and its applications

Georges Nader



N° d'ordre NNT : 2016LYSE1239

THÈSE DE DOCTORAT DE L'UNIVERSITÉ DE LYON

opérée au sein de

l'Université Claude Bernard Lyon 1

École Doctorale ED512
Infomaths

Spécialité de doctorat : Informatique
Discipline : Informatique Graphique

Soutenue publiquement le 22/11/2016, par :

Georges Nader

Calcul du seuil de visibilité d'une distorsion géometrique locale sur un maillage et ses applications

---

Devant le jury composé de :

Mme Luce Morin, Professeur, INSA Rennes	Présidente
M. Tamy Boubekeur, Professeur, Telecom ParisTech	Rapporteur
M. Patrick Le Callet, Professeur, Université de Nantes	Rapporteur
Mme Géraldine Morin, Maître de Conférences, INP Toulouse	Examinatrice
Mme Florence Denis, Maître de Conférences, Université Lyon 1	Examinatrice
M. Florent Dupont, Professeur, Université Lyon 1	Directeur de thèse
M. Franck Hétroy-Wheeler, Maître de Conférences, Grenoble INP	Co-directeur de thèse
M. Kai Wang, Chargé de Recherche, CNRS	Co-encadrant de thèse

Remerciements

Par les quelques lignes suivantes, je tiens à remercier toutes les personnes qui ont contribué directement ou indirectement à cette thèse en espérant de n'oublier personne.

Je tiens à remercier dans un premier temps mes directeurs de thèse, Pr. Florent Dupont, Dr. Franck Hétroy-Wheeler et Dr. Kai Wang pour m'avoir fait confiance et pour m'avoir créé un environnement de travail parfait pendant les trois années de thèse.

Je souhaite aussi remercier Pr. Patrick Le Callet et Pr. Tamy Boubekeur pour avoir accepté de relire mon manuscrit ainsi que les membres du Jury d'avoir accepté ce rôle. J'ai vraiment apprécié leurs remarques, leurs questions, et les discussions que nous avons pu avoir le jour de la soutenance. Cela m'a permis de prendre du recul envers mes travaux et d'ouvrir d'avantage mes horizons professionnels.

Je tiens à remercier aussi toutes les personnes qui ont participé aux nombreuses mesures et expériences utilisateurs menées durant cette thèse. Leurs patiences dans la salle d'expérience et le temps qu'ils m'ont accordés étaient essentiels pour pouvoir obtenir un résultat fiable et précis.

Vue que ma thèse était entre Lyon et Grenoble, je remercie les membres de l'équipe AGPIG du Gipsa-lab à Grenoble ainsi que ceux de l'équipe M2Disco du LIRIS à Lyon pour l'ambiance fraternelle qui est importante afin d'allèger le stress de la thèse. Un spécial merci à mes co-bureaux Anton Andreev et Bahram Ehsandoust de m'avoir supporté pendant ces trois ans. En plus, je remercie le personnel administratif des deux laboratoires (LIRIS et Gipsa-Lab) pour la gestion des nombreuses missions de déplacement entre Grenoble et Lyon.

Certainement, il ne faut pas oublier de remercier tous mes amis à Grenoble et ailleur pour le temps convivial qu'on a passé ensemble que ça soit autour d'un repas, ou bien pendant les nombreuses randonnées dans les montagnes, etc.

Un merci très très spécial à Talia pour toutes les belles mémoires.

Finalement, rien ne pourra exprimer ma gratitude envers mes parents. Je leurs remercie pour la confiance qu'ils m'ont accordé et leur soutien infini. Je suis sûr que ce travail n'existerait pas sans eux.

Résumé

Calcul du seuil de visibilité d'une distorsion géométrique locale sur un maillage et ses applications

---

Les opérations géométriques appliquées aux maillages 3D introduisent des distorsions géometirques qui peuvent être visibles pour un observateur humain. Dans cette thèse, nous étudions l'impact perceptuel de ces distorsions. Plus précisément, notre objectif est de calculer le seuil à partir duquel les distorsions géométriques locales deviennent visibles. Afin d'atteindre notre but, nous définissons tout d'abord des caractéristiques perceptuelles pour les maillages 3D. Nous avons ensuite effectué une étude expérimentale des propriétés du système visuel humain (sensibilité au contraste et effet du masquage visuel) en observant un maillage 3D. Les résultats de ces expériences sont finalement utilisés pour proposer un algorithme qui calcule le seuil de visibilité relatif à une distorsion locale. L'algorithme proposé s'adapte aux différentes conditions d'affichage (résolution et taille de l'écran), d'illumination et au type de rendu. Enfin, nous montrons l'utilité d'un tel algorithme en intégrant le seuil de visibilité dans le pipeline de plusieurs opérations géométriques (*ex*: simplification, subdivision adaptative).

Abstract

Evaluating the visibility threshold for a local geometric distortion on a 3D mesh and its applications

Geometric operations applied to a 3D mesh introduce geometric distortion in the form of vertex displacement that can be visible to a human observer. In this thesis, we have studied the perceptual impact of these geometric distortions. More precisely, our goal is to compute the threshold beyond which a local geometric distortion becomes visible. In order to reach this goal, we start by evaluating perceptually relevant properties on 3D meshes. We have then performed a series of psychophysical experiments in which we measured the visibility threshold relative to various properties of the Human Visual System (contrast sensitivity and visual masking). The results of these experiments allowed us to propose an algorithm that computes the visibility threshold relative to a local geometric distortion. This algorithm is capable of adapting to the different display condition of 3D meshes (resolution, display size, illumination condition and rendering). Finally, we showcase the utility of our work by integrating the developed perceptual method in several geometric operations such as mesh simplification and adaptive subdivision.

Contents

List of Figures

List of Tables

List of Abbreviations

2AFC	**Two A**lternative **F**orced **C**hoice
BPC	**B**its **P**er **C**oordinates
CSF	**C**ontrast **S**ensitivity **F**unction
DAME	**D**ehedral **A**ngle **M**etric **E**rror
FMPD	**F**ast **M**esh **P**erceptual **D**istance
HVS	**H**uman **V**isual **S**ystem
JND	**J**ust **N**oticeable **D**istortion
LGN	**L**ateral **G**eniculate **N**ucleus
LOD	**L**evel **o**f **D**etails
MSDM	**M**esh **S**tructural **D**istortion **M**easure
PDF	**P**robability **D**ensity **F**unction
PEST	**P**arameter **E**stimation by **S**equential **T**esting
RMS	**R**oot **M**ean **S**quare error
SSIM	**S**tructural **SIM**milarity
TPDM	**T**ensor-based **P**erceptual **D**istance **M**etric
VEP	**V**isual **E**quivalence **P**redictor
VDM	**V**isual **D**iscrimination **M**odel
VDP	**V**isual **D**ifference **P**redictor

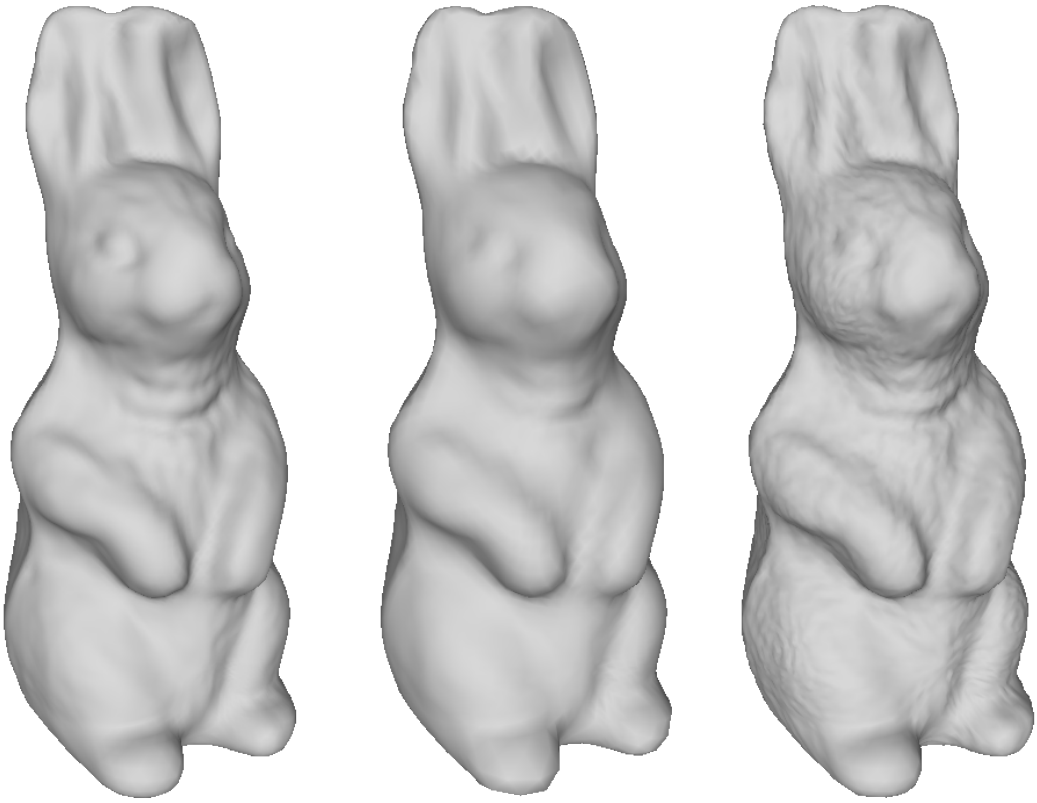
Chapter 1

Introduction

1.1 Context and Motivation

Three-Dimensional (3D) objects, most commonly represented by triangular meshes, are nowadays more and more used in a large number of applications spanning over different fields such as digital entertainment, cultural heritage, scientific visualization, and medical imaging. Moreover, the popularity of 3D objects is bound to drastically increase in the near future with the release of affordable and commercial virtual reality headsets and the current evolution of the web which revolves around the development of web3D technologies. 3D objects are usually created by artists via 3D modeling and sculpting tools or more recently obtained by scanning a real world object. In both cases, the raw 3D data cannot be directly used in a practical application. It is therefore common for 3D data to undergo various lossy operations in order to accommodate to the needs of these applications. For example, in a video game, the detailed 3D mesh created by the artist is simplified so that a real time and interactive visualization of the 3D world becomes possible. On the other hand, in a 3D web application, the 3D data usually need to be compressed in order to limit the bandwidth usage and reduce the data transfer time. Finally, a watermarking operation might be applied to the 3D mesh in order to limit any illegal or unauthorized duplication of the 3D object in question. These operations introduce geometric distortions in form of perturbation of vertex coordinates which might be visible to a human observer (Fig. 1.1). 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.

Although the importance of predicting the visibility of geometric distortions has been recognized within the computer graphics community [OHM+04, Fer08, FS09, MMBH10, LC10, TFCRS11, CLL+13], most existing geometry processing



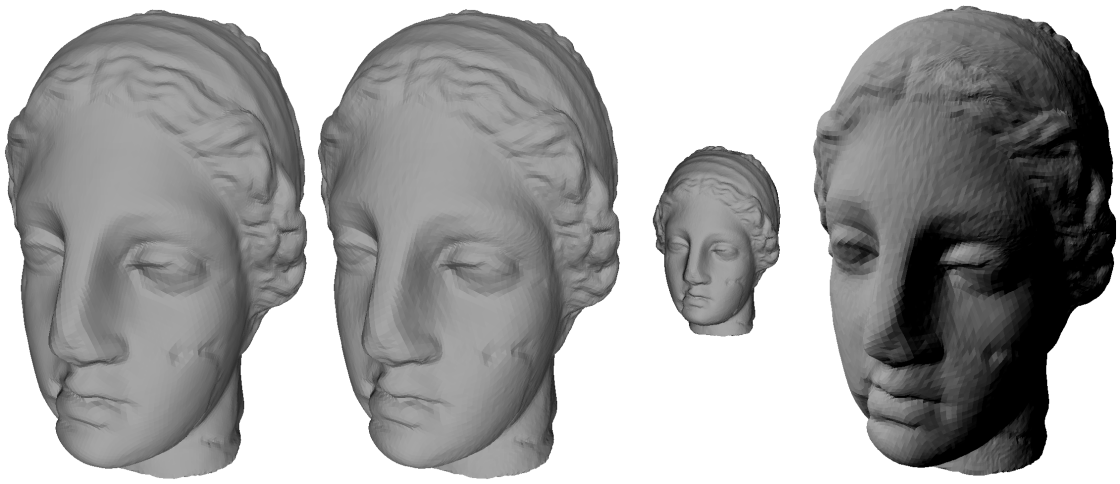
original

(a) Simplification
RMS = 4.3×10^{-4}

(b) Uniform Noise
RMS = 4.3×10^{-4}

FIGURE 1.1: The simplified and noisy versions of the Rabbit model have the same RMS value, but are not perceptually equivalent.

algorithms are driven and/or evaluated by geometric metrics like Hausdorff distance [ASCE02] or root mean square error (RMS) [CRS98] which do not correlate with human perception [CLL+13]. Figure 1.1 shows two distorted versions of the Rabbit model with the same RMS value that are perceptually different. Recently, a number of perceptually driven algorithms have been proposed [FSPG97, RPG99, WLC+03, CB05, CDGEB07, QM08, Lav09, MKRH11, TWC14, DFLS14, DFL+15]. The goal of these methods is to evaluate the perceptual impact of geometric distortion or to guide geometric operations. However, existing methods are usually based on assumptions about the general behavior of the Human Visual System (HVS) instead of taking advantage of the characteristics of its internal mechanism. Moreover, in most cases, the perceptual analysis of existing methods is carried out using geometric features such as surface curvature and surface roughness which are not necessarily perceptually relevant attributes. Consequently, these methods are in general neither easily applicable to models of different properties (size, details and density) nor capable of adapting to varying circumstances



original (a) (b) (c)

FIGURE 1.2: (a) The noise injected onto the original 3D mesh is slightly visible. (b) Increasing the viewing distance makes that same noise invisible. (c) Changing the light direction from front to top-left increases the perceived intensity of the noise.

of mesh usage (display characteristics, scene illumination and viewing distance). For instance, Fig. 1.2 showcases the effects of viewing distance and lighting conditions on the visibility of geometric distortions. While the noise injected onto the Venus model is slightly visible in Fig. 1.2.(a), it becomes invisible when the camera is moved farther away from the model in Fig. 1.2.(b). On the other hand, the injected noise appears to be clearly visible when the light direction is changed from front to top-left in Fig. 1.2.(c).

Early studies of the physiology of the Human Visual System (HVS) have shown that human vision is primarily sensitive to variation in light energy, *i.e.*, *contrast*, rather than its absolute magnitude [Wan95]. In other words, a visual pattern is only visible if its contrast value is above a certain threshold. Therefore, in the case of geometric distortions, a vertex displacement is visible to a human observer if it causes a change in contrast that is large enough for the HVS to detect its presence. In the field of image processing, the analysis of contrast information has been basis in many studies that are related to the visibility of pixel-based distortions [WB06, LK11, BLBI13]. In this thesis, we focus on using perceptual properties, such as contrast, to evaluate whether a geometric distortion is visible or not through an experimental study of the characteristics of the HVS, in particular *Contrast Sensitivity* and *Visual Masking*.

1.2 Objectives and Methodology

The goal of this thesis is to compute threshold beyond which a local geometric distortion becomes visible. The visibility threshold should take into account the various parameters that affect the visibility of geometric distortions such as the display specification (size and resolution), the rendering algorithm, the scene illumination, etc. The computed threshold can then be used to either predict the visibility of local geometric distortion or guide geometric operations. In order to achieve this goal we will:

1. Define perceptually relevant features for 3D meshes that are sensitive to the different parameters affecting the perception of 3D objects.
2. Perform a series of psychophysical experiments in order to study the properties of the HVS while observing a 3D mesh.
3. Use the results of these experiments to derive an algorithm that is able to compute the visibility threshold relative to local geometric distortions.

1.3 Thesis Organization

The work presented in this thesis focuses on computing the visibility threshold of local distortions on the surface of 3D meshes. In Chapter 2 we present the background on the properties of the HVS that are relevant to the research presented in this thesis. In Chapter 3 we discuss existing work on perceptually driven graphics techniques. Chapter 4 explains how perceptual properties are evaluated on a 3D mesh and presents a series of psychophysical experiments that were carried out in order to measure the visibility threshold and their results. Chapter 5 describes our method to evaluate the threshold beyond which a vertex displacement becomes visible. Finally, in Chapter 6 we showcase how our perceptual method can be integrated in geometric applications such as mesh simplification and adaptive subdivision.

Chapter 2

Background on the Human Visual System

Over the last two decades, perceptually driven methods have drawn more and more attention in both the computer graphics [OHM+04, Fer08, CLL+13] and image/video processing [WB06, LK11, BLBI13] communities. Much of the existing work is based on the visual mechanisms of the human vision. It is therefore important to understand the basic properties of the Human Visual System (HVS) relevant to these methods. This chapter is divided into two parts. The first part gives a brief overview about the process of vision which consists of analyzing information derived from incident light (Section 2.1.1). In particular, we focus on the physiology of the retinal receptive fields which gives us an understanding about the type of visual information that reaches the brain (Section 2.1.2). The second part focuses on the major characteristics of the HVS that are of relevance to our perceptual study (Section 2.2). More precisely, we discuss the contrast sensitivity and the visual masking characteristics of the HVS. For a more detailed discussion about human vision, we refer the reader to [Wan95, Pal99].

2.1 The Process of Vision

2.1.1 Overview

Vision is the process of extracting and analyzing the information that is contained in the incident light. This process (Fig. 2.1) starts in the eyes as the outside light enters through the pupil and gets focused on the retina, a light-sensitive tissue, with the help of the lens. The retina is composed of two types of photoreceptors cells: the *rods* and the *cones*. The rods are extremely sensitive to light as they can be excited by as little as one photon and provide the visual system with the necessary information for an achromatic vision in low illumination levels. The cones, on the other hand, are less sensitive to light but allow the HVS to have color vision. This is due to the existence of three types of cones which are sensitive to

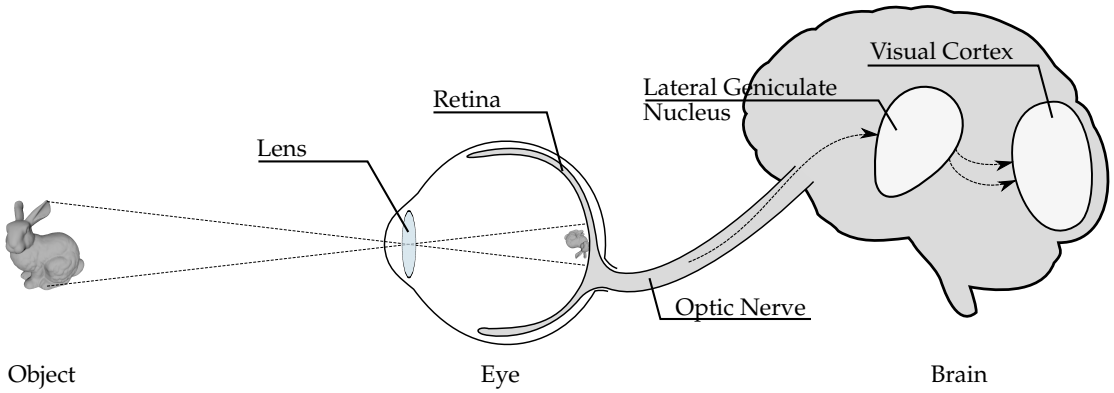
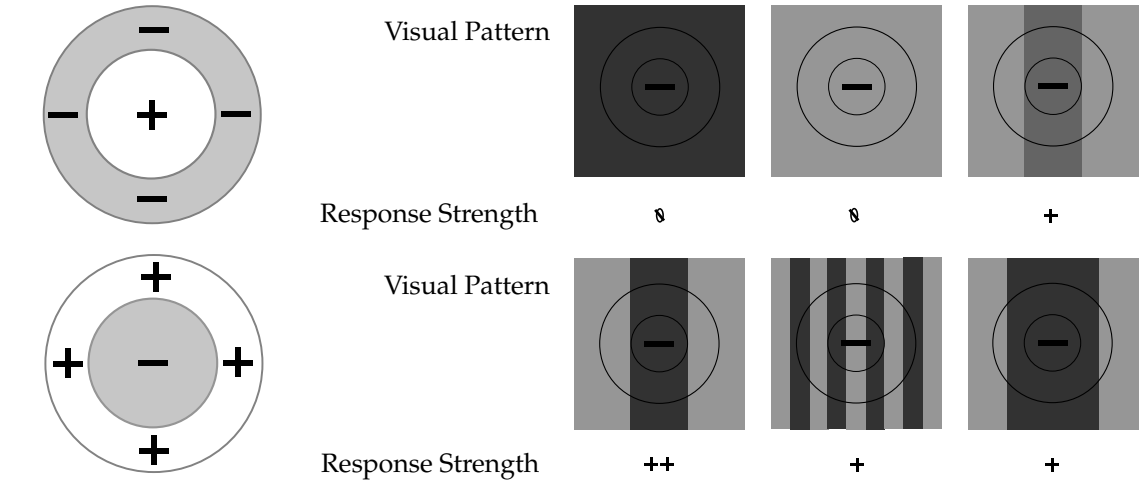


FIGURE 2.1: The light reflected by the observed objects enters the eyes and is focused onto the retina. The visual information captured by the retina is then tranfered by the optic nerve to the brain where it is finally processed.

a distinct (yet overlapped) interval of light frequency. The stimulation of these photoreceptor cells by the incident light creates an electrical signal which then reaches the *receptive fields* of the ganglion cells. These cells play a crucial role in visual perception as their goal is to encode the visual signal for a more efficient treatment of it. The visual information then travels through the visual pathways which lead to the *lateral geniculate nucleus* (LGN) and finally on to the *visual cortex* which is responsible for all the higher-level aspects of vision. While the details of the properties of the LGN and visual cortex are not within the scope of this thesis, it is interesting to know that one of the roles of the LGN is to control the amount of information that is allowed to pass to the visual cortex. In this work, we are mainly interested in the role of the receptive fields of the ganglion cells which provides us with the basic understanding of the characteristics of the HVS that are relevant to our perceptual study.

2.1.2 The Receptive Fields

The range of light intensity that we experience is huge. For example, the intensity of light coming from the sun is approximately 10 million times bigger than the intensity of moonlight. The first challenge the HVS faces is to be able to cope with this wide range of light intensity. More precisely, the challenge is to represent the visual information in an effective and meaningful way. This problem is solved in the early stages of the visual system at the receptive fields of the ganglion cells. Their role is to pre-process the visual information before passing it on to the brain. Early studies of the physiology of the HVS [CK66, CR68, BC69b] have shown that the receptive fields have a *center/surround* organization (Fig. 2.2). In other words, the light reaching the center of a ganglion cell's receptive field can



(a) center/surround organization of the ganglion receptive fields

(b) example of visual patterns and their corresponding neural response strength

FIGURE 2.2: The center-surround organization makes the human visual system sensitive to patterns of light rather than its absolute value. The neural response is at its peak when the visual signal lines up with the size of the receptive field.

either excite or inhibit the cell while the light in the surrounding region will have an opposite effect. This means that for a uniform visual stimulus, the inhibitory and excitatory signals will neutralize each other resulting in a weak neural response. On the other hand, if the visual stimulus consists of a pattern of dark and bright light, then the receptive fields will produce a strong neural response. The center/surround organization of the retinal receptive fields implies that information about the absolute value of light intensity is less important to the visual system than contrast information, *i.e.*, variation in light energy information. This sensitivity to patterns of light rather than to its absolute value is at the heart of the properties of the HVS that most perceptual methods rely on. In the following section we will detail these characteristics of the HVS.

2.2 Characteristics of the Human Visual System

As we mentioned earlier, the center/surround organization of the receptive fields of the ganglion cells makes the HVS primarily sensitive to variation of light energy rather than to its absolute value. This difference of light energy in a visual pattern in generally represented by its contrast value. Ultimately, a high contrast visual pattern should generate a strong neural response and a low contrast visual pattern should generate a weak neural response. So studying the characteristics

of the early stages of the HVS boils down to studying the perception of the visual pattern's contrast.

2.2.1 Contrast Sensitivity

The contrast threshold is the contrast value beyond which a visual stimulus becomes visible to a human. It represents the minimal amount of contrast required to generate an exitatory neural response. The value of that threshold is mainly dependent on three factors: spatial frequency, global illumination and retinal velocity.

Spatial Frequency. The *spatial frequency* is related to the size of the visual stimulus with respect to the size of one degree of the visual angle. It is expressed in terms of *cycles per degree* (cpd) which represents the number of times a visual stimulus can be repeated within one degree of the visual angle. The effect of the spatial frequency on the visibility threshold can be derived from the size of the receptor fields. While a pattern of dark and bright light will excite the ganglion cells, the neural response will be at its strongest if the size of the visual pattern lines up with the size of the center and surround region of the receptive field (Fig. 2.2). This means that the human visual system will be more sensitive to the spatial frequencies that correspond to the size of the receptive fields, generally between 2 and 5 cpd for humans, and less sensitive to the other spatial frequencies.

Global Luminance. The average energy of the light, *i.e.*, luminance, illuminating the observed scene also affects the contrast visibility threshold. In dark environments, the low energy light reaching the retina will trigger the *rods* photoreceptiors as the energy is not sufficient to provoke the *cones*. The difference in the source of the visual signal between low and bright lights, *i.e., rods* in low light and *cones* in bright light, causes the change in contrast visibility threshold when the light energy changes. In summary, physiological experiments show that at low luminance levels, the contrast threshold increases when the average luminance decreases, while it becomes relatively stable for luminance levels above $100\ \mathrm{cd/m^2}$ [Bar89].

Retinal Velocity. When the visual stimulus is in motion, its image on the retina will also move. The retinal velocity is therefore defined as the velocity of the retinal image of an object. This velocity is affected by the movement of the object and the eyes whose job is to track the moving stimulus in an attempt to stabilize the retinal image. The experiments of Kelly [Kel79a, Kel79b] have shown that the

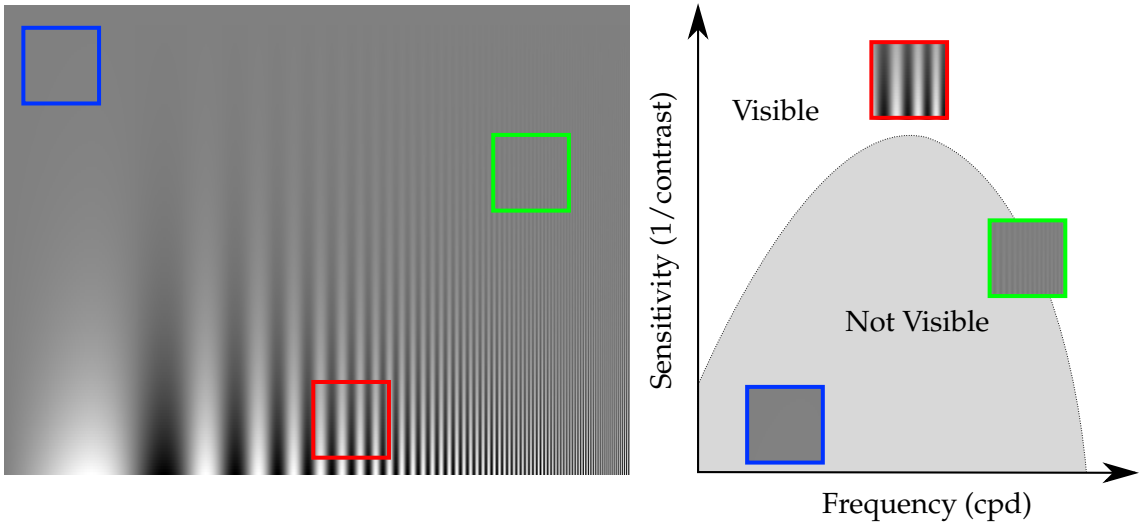
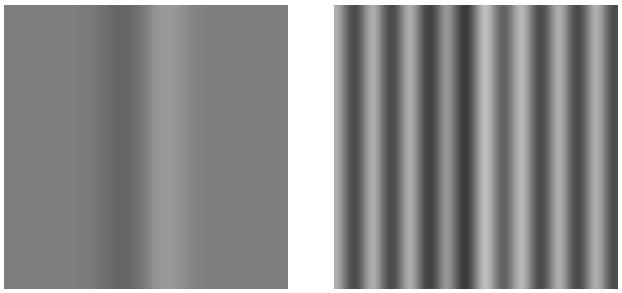


FIGURE 2.3: Left: The Campbell and Robson chart [CR68] that shows a sinusoidal visual stimulus whose contrast decreases vertically and frequency increases horizontally. Right: An example of the Contrast Sensitivity Function.

contrast visibility threshold is altered when the retinal velocity increases. For example, the range of sensitive spatial frequencies varies in general from between 2 and 5 cpd for a stationary stimulus to around 0.2 and 0.8 cpd for a stimulus whose image is moving at a 11 deg/s speed on the retina.

The reciprocal of the contrast visibility threshold is the contrast sensitivity. The *Contrast Sensitivity Function (CSF)* is a mathematical model that describes the evolution of the visibility threshold with respects to the aforementioned three parameters (spatial frequency, global luminance and retinal velocity). It was first introduced by Campbell and Robson [CK66] whose CSF model takes only into consideration the effects of spatial frequency and was later extended to consider global luminance levels [Bar89, Bar99]. The effects of object motion on the contrast sensitivity are taken into account in the model proposed by Kelly [Kel79b]. The CSF represents the visual system's band-pass filter characteristics when it comes to contrast sensitivity (Fig. 2.3). In general, it exhibits a peak between 2 and 5 cpd when the visual stimulus is stationary. The shape of the CSF (peak location and drop off slope) depends on the nature of the visual stimulus. For example, at high spatial frequencies the HVS is more sensitive to aperiodic visual patterns than to periodic ones [BC69a].



(a) (b)

FIGURE 2.4: (a) A visible sinusoidal stimulus on a grey background. (b) Having another visible sinusoidal stimulus as a background makes the original one less visible.

2.2.2 Visual Masking

Visual masking is a very important characteristic of human vision as it describes how the HVS handles interactions between different visual stimuli. Masking occurs when a stimulus (*target*) that is visible on its own cannot be detected due to the presence of another visible stimulus (*mask*). Figure 2.4 illustrates this effect. A visual signal, the *target*, that is visible on its own might become hard to notice when it is added to a visual pattern containing a visible visual stimulus, the *mask*. The effects of visual masking are caused by several factors. In particular, the visibility of the target visual stimulus is dependent on the contrast and the visual complexity of the mask.

Ledge and Foley [LF80] studied the contrast threshold necessary to detect the target when varying the contrast and frequency of the mask. One important observation that can be taken out from their experimental study about this aspect of the HVS is that the visibility threshold increases almost linearly with the contrast of the mask. The effects of masking can be mathematically described by a curve (Fig. 2.5) which possesses two asymptotic regions: the first one with a slope of zero and the second one with a positive slope of about 0.6 to 1 (depending on the stimulus) [Dal93]. The zero slope occurs for mask contrast values below the mask's visibility threshold as given by the CSF, indicating that there is no masking effect. When the mask is visible (its contrast above the value given by the CSF), the threshold for detecting the target lies of the second asymptotic region.

The experiments of Ledge and Foley focused on studying the visual masking effect of simple sinusoidal patterns. This means that their results do not account for the impact of the mask's *visual complexity* on the visibility threshold. The *visual complexity* is an equally important factor as a complex mask, which is *visually irregular*, would introduce some uncertainty to the observer's judgement

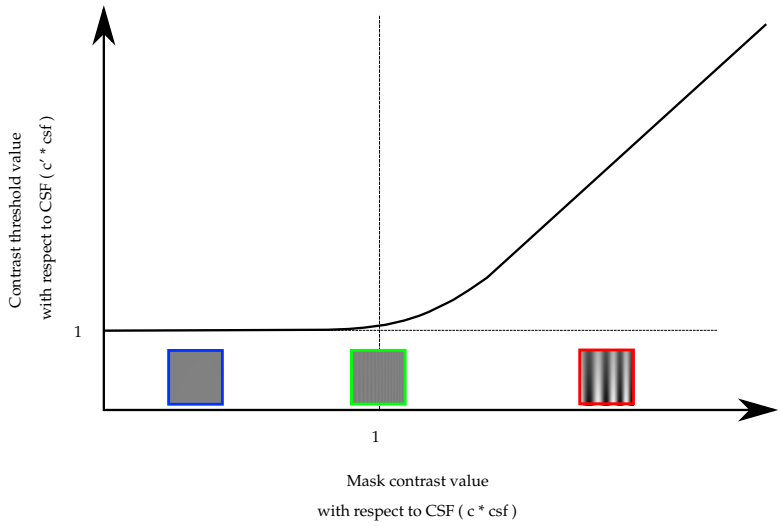
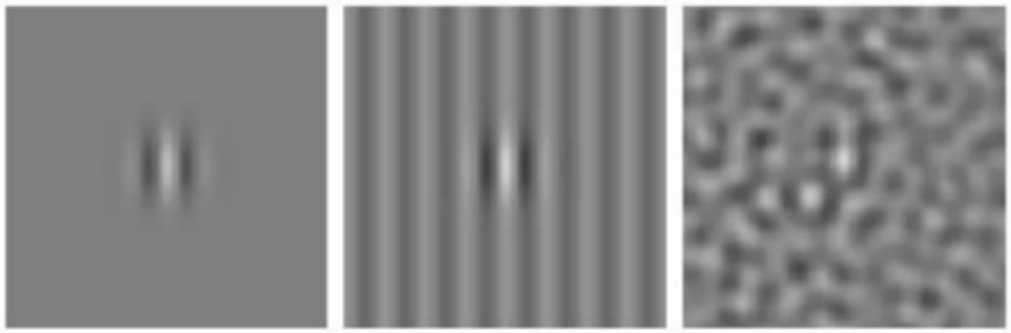


FIGURE 2.5: An example of a curve that describes the visual masking characteristic. When the contrast of the mask signal is visible (greater than 1) then the contrast threshold increases almost in a linear fashion.



(a) (b) (c)

FIGURE 2.6: The visual stimulus is visible when the background is either (a) uniform or (b) simple. However, the same stimulus becomes almost invisible when added to a (c) visually complex background.

and thus increasing the masking effect [WBT97]. For instance, Fig. 2.6 shows a visual stimulus that is visible on its own or when added to a relatively simple mask. However, when added to an irregular mask, the visual stimulus becomes harder to notice which indicates a change in the visibility threshold. The influence of the mask's visual complexity on the visibility threshold can be inferred by the *free-energy principle* theory [FKH06, Fri10]. By analyzing the incoming visual information, the HVS helps us understand the outside world. However, due to the sheer amount of input, the HVS cannot fully process all of the visual information [KP04]. Instead, the HVS tries to predict the visual information with an internal generative mechanism [KP04, FDK09]. The underlying idea behind the free-energy principle theory is that all adaptive biological agents have a tendency

to resist disorder [Fri10]. In other words, the HVS will try to extract as much information as possible from the coming visual information in order to minimize any uncertainty and avoid surprises (*i.e.*, information with uncertainty, which is often found in visually complex stimuli). This means that visual patterns with obvious statistical regularities are easier to be predicted and understood than those without any regularity (*i.e.*, complex patterns). As a result, the change in information (*i.e.*, contrast) in a regular visual pattern can be easily detected while it would be difficult to detect in an irregular, complex one [WBT97].

2.3 Summary

To summarize, the human visual system is sensitive to variation of light intensity (*i.e.*, contrast) rather than its absolute value. This is primarily due to the center/surround organization of the receptive fields of the ganglion cells as discussed in Section 2.1.2 (Fig. 2.2). As a consequence, the contrast value of a visual stimulus is at the heart of the early properties of the HVS, in particular *contrast sensitivity* and *visual masking*. Contrast sensitivity refers to the threshold beyond which a contrast becomes visible for a human observer. This threshold is affected by the spatial frequency, global luminance and retinal velocity of the visual stimulus and can be mathematically modelled by the contrast sensitivity function. Visual masking, on the other hand, explains how the HVS handles interactions between different visual stimuli. In other words, the visual masking describes the change in visibility threshold caused by the presence of another visible visual stimulus (mask). The amount by which the visibility threshold changes is proportional to the contrast value of mask and is affected by its visual regularity. Both the contrast sensitivity and especially the visual masking are at the center of many perceptual methods in computer graphics. The next chapter presents a detailed description of these methods.

Chapter 3

Perceptual Methods for Computer Graphics Applications

In the previous chapter, we have presented the theoretical and fundamental background about the early properties of the human visual system. This chapter focuses on the existing perceptual methods that have been proposed in the past two decades. We start by discussing the benefits of taking into account the perceptual characteristics of human vision in the design of computer graphics systems (Section 3.1). We then review in detail the most important methods in the field of perceptually adaptive graphics. Based on the general approach taken to perform the perceptual analysis, we group these methods into two categories: *Image-Based* (Section 3.2) and *Model-Based* (Section 3.3). Finally we compare these two approaches and discuss the limitations of current methods (Section 3.4).

3.1 The Role of Perception in Computer Graphics

One of the goals of computer graphics is to generate a 2D image for a human observer from the description of a 3D scene. In most cases, the scene is constituted of the following elements: (1) a geometric representation of the surface of an object, most commonly through a triangular mesh, (2) the material properties attributed to that surface and (3) the illumination information. In general, any computer graphics pipeline starts with an acquisition step where the scene's data is created. Usually, the 3D geometric data are then subject to various processing algorithms (compression, level of detail generation, ...) in order to accommodate for the need of the target application. Finally the scene's elements are passed on to a rendering algorithm which computes a 2D image. Ideally, we would like to generate a "perfect" image, *i.e.*, without any distortion, from this pipeline which is in practice highly unlikely as visual artifacts are bound to appear on the rendered image. The source of these artifacts may be the geometric operations applied to the 3D data or the rendering algorithm which is trying to simplify the computation in order to cope with hardware limitations. Consequently, in practice, the

best that we can hope for is that the computer graphics system is capable of generating a *perceptually effective* image, that is, an image that effectively provides the indended visual information [Fer03, TFCRS11]. The success of a computer graphics method, whether it is a rendering algorithm or a geometric operation, is therefore dependent on the perceptual effectiveness of the resulting image. In order to improve perceptual effectiveness of computer graphics, one approach consists in taking advantage of the characteristics of the human vision in order to guide computer graphics algorithms. The perceptual properties of the human visual system can thus be used as an optimization criterion in the design of computer graphics operations, in particular geometric ones.

Over the last two decades, the computer graphics community has recognized the importance of exploiting the perceptual properties of the HVS [OHM+04, Fer08, CLL+13] as perceptually motivated techniques have proven to be useful for several practical applications. The goal of this research is to propose a method that allows us to predict the visibility of a geometric distortion on a 3D triangular mesh. This is important since before rendering a 3D model, almost all raw geometry data are subject to several operations (*e.g.*, compression, watermarking, ...) that introduce geometric distortions in the form of vertex displacement which might be visible in the final 2D image. The perceptual analysis techniques that are related to the perception of surface material [Fle14, HFM16], non-photorealistic rendering [SD04, RBD06, CSD+09], physical simulation [HK15, HK16] and character animation [MNO07, DM08, LO11] are therefore not within the scope of this research. Our work is more focused on the perceptual methods that aim to guide or evaluate the output of the geometric operations applied to a 3D triangular mesh. These methods can be grouped into two categories *Image-Based* and *Model-Based* methods [LM15]. The first category concerns the algorithms where the perceptual analysis is carried out by analyzing rendered 2D image. On the contrary, the perceptual methods in the second category rely on a perceptual analysis that is performed on the geometric surface of the 3D object. In the rest of this chapter we will detail and discuss the most notable methods belonging to these two approaches.

3.2 Image-Based Methods

Since the 3D models are visualized on 2D displays, it seems logical to use the 2D rendered image in order to carry out the perceptual analysis. Using an image-based method for studying the perceptual impact of 3D distortions has its advantages. It allows researches to adapt the already established perceptual methods in

the field of image processing to computer graphics. More importantly, applying the perceptual analysis on the rendered 2D image will implicitly take into consideration any rendering or illumination method [LM15]. In this section we will first present the most important perceptual methods in the field of image processing that were adapted to computer graphics. We will then describe how image-based perceptual methods have been used for computer graphics applications.

3.2.1 Perceptual Methods in the Field of Image Processing

Two main approaches exist for designing a perceptual analysis on 2D images: a *top-down* approach which relies on hypotheses about the global behavior of the HVS and a *bottom-up* approach which tries to model the visual process of the HVS. Ultimately, a bottom-up method tries to build a computational system that mimics the way the HVS works [WB06]. On the contrary, in a top-down approach, the relationship between the input and the output is the thing that matters. This means that a top-down approach takes the visual signal as an input and outputs a result that is in agreement with the general behavior of the HVS without having to explicitly deal with the inner-working of the HVS.

3.2.1.1 Top-Down Perceptual Methods

Top-down approaches do not look to simulate the HVS. Instead, they are only concerned by outputting a result that is in accordance with the general behavior of the HVS. The advantage of these approaches is that they allow researchers to take into consideration complex aspects of the visual system that would rather be difficult to simulate in a bottom-up approach. Having the freedom to focus on the general behavior of the HVS rather than its internal mechanisms has led to the development of many top-down perceptual methods that either rely on studying whether the structure between a reference of a distorted image has changed [WB02, **?**, WBSS04] or focus more on analyzing if the distortion has caused a disruption in the information contained in the image [Fie87, Sim05, SB06]. In this section we focus on the methods that study the structure of an image as they have been popular in many computer graphics methods.

Structural Similarity. The visual signal contains information about the outside world that is analyzed by the visual system. This makes the HVS highly adapted and effective in extracting the key features in a natural image, *i.e.*, an image that represents the natural world. The idea behind *structural similarity* is that a change in the structure of an image, caused by a distortion, will be easily detected by the HVS. Wang *et al.* [WBSS04] proposed a metric (SSIM Index) whose goal is to

measure the perceptual quality of a distorted image by comparing its similarity to the reference image. Having two images x and y, respectively the reference and the distorted, the SSIM algorithm defines the task of comparing the similarity as the combination of a comparison of luminance, contrast and structure between x and y. The SSIM index is therefore defined as:

$$SSIM(x,y) = l(x,y)^{\alpha} \cdot c(x,y)^{\beta} \cdot s(x,y)^{\gamma} \,, \tag{3.1}$$

where $l(x,y)$, $c(x,y)$ and $s(x,y)$ are respectively the luminance, contrast and structure components and $\alpha > 0$, $\beta > 0$ and $\gamma > 0$ are parameters that control their relative importance. Wang *et al.* [WBSS04] defined the luminance, contrast and structure components as follows:

$$l(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad c(x,y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad s(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \tag{3.2}$$

where μ_x, σ_x and $\sigma_x\sigma_y$ represent respectively the mean, the standard deviation and the covariance of the pixel intensity. C_1, C_2 and C_3 are three constants added to avoid numerical instability when the compared image is dark ($\mu_x^2 + \mu_y^2$ close to 0) or when the image contains a uniform visual stimulus (σ_x and σ_y close to 0). In practice, it is preferable to apply the SSIM method locally on image patches (*e.g.* on a 11×11 window) rather than on the entire image. This will result in a quality map of the compared image which can then be aggregated into a single score. The SSIM approach has proven to be quite usefull in a number of image processing applications [WS04, WS05, WL11]. More importantly, the idea behind the SSIM Index has inspired the development of perceptual metrics for 3D meshes [LDGD+06, Lav11] which we will detail later in this chapter (Section 3.3).

Although top-down methods present a simple approach of taking into account complex properties of the HVS, they also have some drawbacks. First, the success of top-down methods is heavily dependent on the validity of the hypotheses they are based on, which are most of the time difficult to justify. For instance, while the SSIM index metric has proven to correlate well with the HVS [SSB06], there is no physiological evidence that would justify combining the contrast, luminance and structure elements using a multiplication [WB06]. Second, designing a top-down algorithm usually requires the inclusion of several abstract parameters to the model that are difficult to calibrate. For example, while the inclusion of the thee parameters α, β and γ to the SSIM model (Eq. (3.1)) has its benefits, it is however difficult to manually find a value that works the best on any type of image. These issues are much less likely to occur in bottom-up methods since

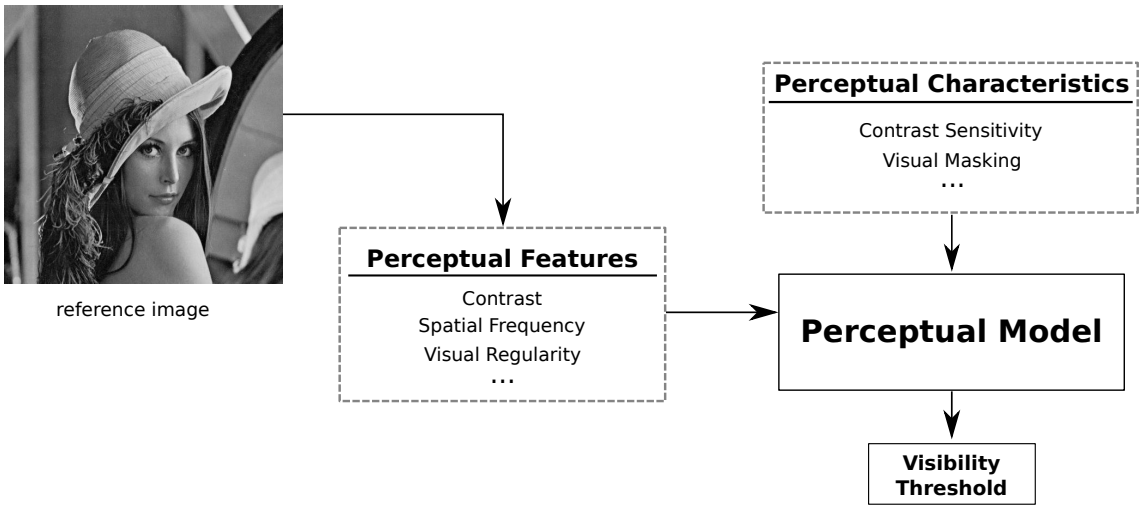


FIGURE 3.1: The general framework of bottom-up perceptual methods.

these approaches rely on simulating the different components of the HVS.

3.2.1.2 Bottom-Up Perceptual Methods

Bottom-up perceptual methods focus on studying each relevant component or feature of the HVS, such as contrast sensitivity (Section 2.3) and contrast masking (Section 2.4), and then combining them together in a computational model that mimics how the HVS works. This computational model can then be integrated in various image processing algorithms [CL95, LKW06, Lin06, LLP+10]. Bottom-up methods rely heavily on experimental results and physiological studies about the aspects of the human visual system [CK66, Kel79b, LF80, Wan95, WB06].

In general, most existing bottom-up methods try to compute a threshold map using mathematical models describing the characteristics of the HVS. This threshold refers to the maximum change in contrast a distortion is allowed to alter before it becomes visible. Most of bottom-up algorithms follow a framework similar to the one presented in Fig. 3.1. The first step consists of computing the perceptual properties, *i.e.,* luminance, contrast and spatial frequency, from a reference image. The luminance is generally obtained by converting the pixel values using a non-linear function. While there are many methods for estimating the contrast for natural images [Pel90] , the Michelson contrast [Mic27] is still adopted in most of the bottom-up methods. Finally, the spatial frequency is usually evaluated using a channel decomposition method such as Fourier decomposition [MS74] , local block-DCT transform [**?**], cortical transform [Wat87, Dal93]. In the second step, the computed perceptual properties are passed to a computational model representing the different properties of the HVS which then outputs a threshold that

can be used to guide various image operations [CL95, LKW06, Lin06, LLP+10]. In most cases, bottom-up algorithms take into consideration the contrast sensitivity, that is modeled by the CSF, and the visual masking aspects of the HVS. One of the most used masking models is the one proposed by Daly [Dal93]. This model, however, does not account for the effects of the signal's complexity on the visual masking (see Section 2.4). Recently several methods [WSL+13, DFL+15] have started to include the free-energy principle theory [Fri10] into the perceptual analysis for a more accurate simulation of visual masking.

A large number of bottom-up algorithms have been introduced in the past few decades. More notably, we mention Daly's Visual Difference Predictor (VDP) [Dal93] and the Visual Discrimination Model (VDM) [Lub95] since they are the basis of many image-based perceptual methods in computer graphics which we will detail in the next section. Both methods, although different in their technical details, aim to predict whether a difference between two images is visible or not. In summary, this is done by comparing the difference in contrast between a reference image and a distorted one with the threshold computed by the perceptual model.

3.2.2 Applying Image Processing Tools to Computer Graphics

One of the purposes of computer graphics is to render an image from a description of a 3D scene. However, since the 3D data are subject to various geometric operations that introduce geometric distortion to the model and due to the computational limitation of the hardware, it is practically impossible to obtain a "perfect" image. In consequence, computer graphics systems generally aim at generating a perceptually acceptable image by taking advantage of the properties of the HVS.

Early attempts to use perceptual elements consisted in adapting perceptual methods that were designed for image processing applications to computer graphics. Ferwerda *et al.* [FSPG97] first showed how Daly's VDP can be used to hide geometric visual artifacts with a texture. Furthermore, Bolin and Meyer [**?**] used a simplified version of the VDM metric [Lub95] to optimize the sampling operation in a ray-tracing algorithm. In the same context of perceptually guided rendering, Ramasubramanian *et al.* [RPG99] presented an iterative perceptual framework (Fig. 3.2) in order to reduce the computational cost of global illumination. In this case, the proposed perceptual method is able to define an automatic stopping criterion for the computationally demanding global illumination operation. The idea is to stop the rendering when the current iteration cannot produce a

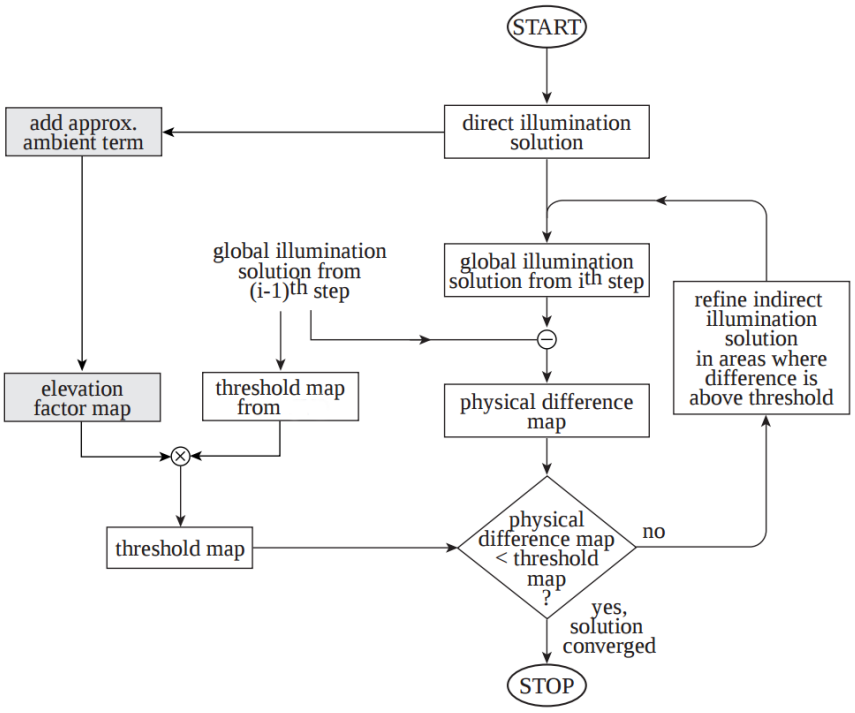


FIGURE 3.2: Overview of the perceptual rendering framework proposed by Ramasubramanian *et al.* [RPG99] (illustration extracted from [RPG99]).

visible change in the image. In other words, the algorithm stops when the physical difference between the image at the current iteration and the previous one is below the threshold map which is evaluated using Daly's VDP. The results of this method can further be improved by using a more complex perceptual model [MKRH11]. However, this approach can be over-conservative as it tends to overestimate the perceptual impact of none-disturbing visible distortions [RFWB07]. For that purpose, Ramanarayanan *et al.* [RFWB07] later introduced the concept of *visual equivalence* which considers two images equivalent if the viewer cannot tell them apart. In this work, the authors presented an experimental study which aimed at defining the elements that contributes to the definition of equivalency between two images. Using the results of this experiment, they then proposed a top-down perceptual metric, the visual equivalence predictor (VEP), which is based on machine-learning techniques to evaluate the equivalency of two images.

For the task of selecting the best level of detail (LOD) version of a 3D mesh, Reddy [Red97] analyzed pre-rendered images using the contrast sensitivity function. Later Dumont *et al.* [DPF03] proposed a system based on a decision theory approach that is capable in real-time of selecting the best LOD and texture resolution with the help of Daly's VDP. Another interesting approach is the one of Zhu

et al. [ZZDZ10] which consists of studying the visibility of fine geometric details using perceptual image metrics such as the VDP and SSIM to design a discrete LOD for visualizing 3D buildings.

Image-based perceptual methods have also been used for guiding mesh simplification. Lindstrom and Turk [LT00] first presented an image based simplification method. This algorithm works by rendering the model being simplified for various viewpoints and uses image-based metrics such as the VDM [Lub95] to guide the simplification. Luebke and Hallen [LH01] proposed a perceptual mesh simplification algorithm that uses a CSF model to estimate whether a local simplification operation will cause a visible change in contrast and frequency. Williams *et al.* [WLC+03] extended later this method to textured model. In both of these methods the simplification result depends on the chosen viewpoint. Qu and Meyer [QM08] used the masking function in [ZDL02] to compute a masking map taking into account the bump map and texture attached to the 3D mesh. This masking map is then used to drive the simplification process. Finally, Menzel and Guthe [MG10] combined a perceptual metric that takes into account the contrast and frequency of the 3D mesh on the rendered image with a geometric distance to decide whether to perform the edge collapse operation or not. The interesting point is that their method is able to handle different materials since the visual masking analysis is performed on an image-based bidirectional texture function (BTF).

3.3 Model-Based Methods

Apart from image-based methods, many algorithms have been developed that use the 3D geometry information for their perceptual analysis. Existing model-based perceptual methods for 3D meshes are based on observations about the general behavior of the human visual system while observing 3D models. These approaches rely on the 3D information of surface geometry in order to perform the perceptual analysis. More precisely, they mostly rely on roughness and curvature information of 3D geometry.

3.3.1 Roughness-Based Methods

3.3.1.1 Relation to Visual Masking

The earliest methods for evaluating the magnitude of geometric distortions were simple geometric distances like the Hausdorff distance [ASCE02] or the root mean square error (RMS) [CRS98]. These measures ignore the working principles of the

HVS and reflect the physical variation of the mesh geometry. They do not correlate with the human vision [CLL+13] (Fig. 1.1) and thus cannot be used to predict whether a geometric distortion is visible or not.

Motivated by the need of evaluating their compression algorithm, Karni and Gostman [KG00] combined the RMS measure with the average distance of the *geometric Laplacian* to obtain a visual metric capable of comparing two 3D objects A and B.

$$GL_1(A,B) = \alpha RMS(A,B) + (1-\alpha)\left(\sum_{i-1}^{n} ||GL(v_i^A) - GL(v_i^B)||^2\right)^{1/2}, \quad (3.3)$$

where $\alpha = 0.5$ and GL is the geometric Laplacian computed as follows:

$$GL(v) = v - \frac{\sum_{i\in n(v)} l_i^{-1} v_i}{\sum_{i\in n(v)} l_i^{-1}}, \quad (3.4)$$

where $n(v)$ is the set of indices of the neighbors of v, and l_i the Euclidean distance from v to v_i.

The idea behind mixing the geometric Laplacian with the RMS distance is that the former represents a local measure of smoothness. This means that the visual distance between two versions of a 3D model (original and distorted) given by the GL_1 is higher when the distortion causes a change in smoothness. Sorkine *et al.* [SCOT03] later proposed a small modification to the GL_1 distance: Setting the value of α to 0.15 instead of 0.5 gives geometric distortions on smooth regions a higher impact on the visual distance. Despite being better than simple geometric measures for computing the visual distance between two 3D models, the geometric Laplacian metric, GL_1, does not correlate well with the human perception [CLL+13]. However, the GL_1 distance and several other observations of visual artifacts produced by 3D watermarking techniques [DGECB05] suggest that the visibility of geometric distortions is related to the roughness of the surface. In other words, it was noted that the geometric distortions are more visible on a smooth surface than on a rough one [Lav09] (Fig. 3.3). This observation can indeed be explained by the visual masking effect of the human visual system (Section 2.2.2). For instance, the rough regions of a 3D model are more likely to generate a visually complex visual pattern in a computer graphics image and thus causing a visual masking effect. These observations have led to the development of many perceptual methods that rely on an estimation of surface roughness as their main perceptual tool.

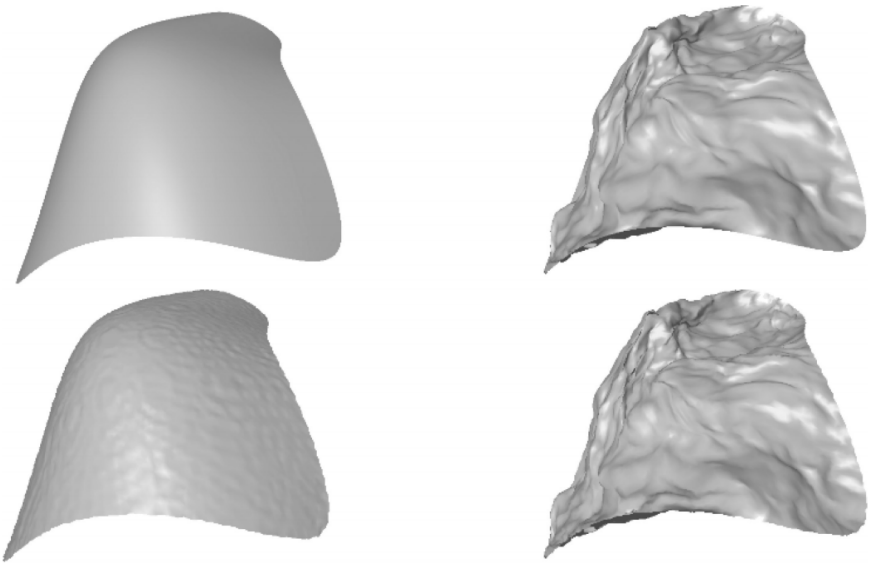


FIGURE 3.3: The roughness of a surface can affect the visibility of a geometric distortion. The geometric distortion injected into the smooth surface is visible while the one injected into the rough one is not visible (image extracted from [Lav09]).

3.3.1.2 Surface Roughness Measures

Following the idea that the effects of visual masking can be taken into consideration using the roughness value of the surface of a 3D object, many roughness estimations techniques have been proposed.

In order to evaluate the quality of the output of 3D watermarking algorithms, Corsini *et al.* [CDGEB07] presented two perceptual metrics, each of which is based on a different method for estimating surface roughness. The first method builds on the work of Wu *et al.* [WHTS01] which consists of using angles between two adjacent faces to evaluate surface roughness. This method proceeds as follows. At first, a per-face roughness measure is computed using angles between two adjacent faces, *i.e.*, dihedral angles. The idea here is that the face normal varies slowly when the surface is smooth while the opposite is true for a rough surface. Consequently, a smooth surface can be detected by the value of the dihedral angles which should be close to 0. Finally, a per-vertex value is obtained by combining the per-face roughness of the N-ring adjacent faces. The number of rings taken into account for computing the per-vertex roughness value controls the scale at which the surface roughness is being evaluated. The second method, first introduced in [DGECB05], is based on the idea that the difference between a detailed and a smoothed version of a 3D model is higher in rough regions than in smooth regions. Therefore, in this approach, computing the local per-vertex roughness boils down two steps (Fig 3.4). First, a smoothed version of the 3D

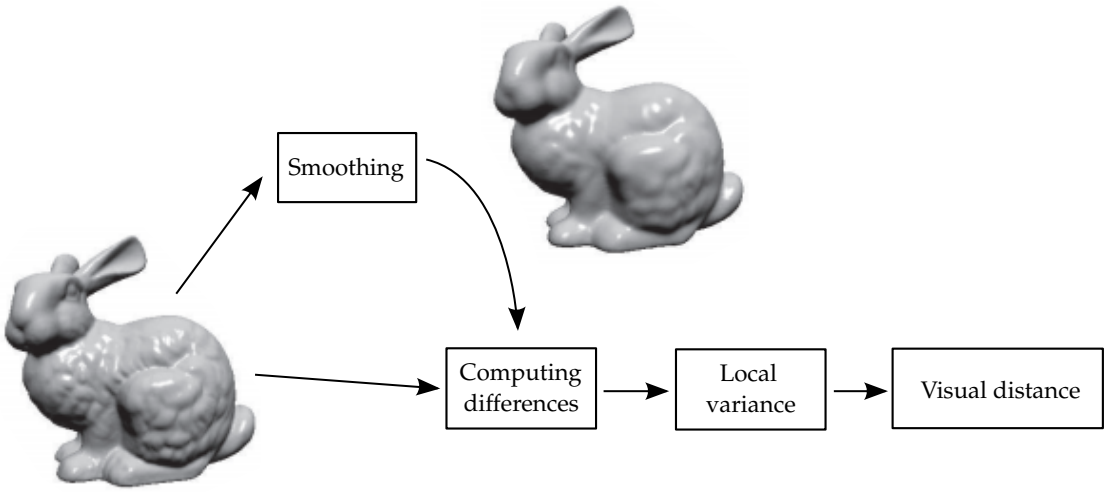


FIGURE 3.4: Overview of the roughness second estimation method presented in [CDGEB07].

model is computed, for example by using Taubin's smoothing operator [Tau95]). Second, the per-vertex difference is computed as:

$$d(v, v^s) = \text{proj}_{n^s}(v - v^s) \,, \tag{3.5}$$

where v^s is the smoothed vertex and $proj()$ indicates the projection of the vector $(v - v^s)$ on the vertex normal of the smoothed surface n^s. The local surface roughness is finally evaluated as the variance of that difference over an N-ring scale.

Lavoué [Lav09] presented a roughness evaluation algorithm based on computing the difference between the original and smoothed version of a 3D model. The method of Lavoué can be summarized by the following steps:

1. A smoothed version of the 3D mesh is generated.
2. The maximum curvature for each vertex of the original and smoothed mesh is computed
3. The average curvature over a local window is evaluated for each vertex.
4. The local roughness value is computed as the difference between the average curvature values of the original and smoothed models.

The local roughness measure of Lavoué improved upon the classification of the surface type. It was able to efficiently differentiate between three types of regions (smooth regions, rough regions and edge regions) on a 3D object (Fig. 3.5), each of which can be attributed a masking level. Furthermore, in his paper Lavoué

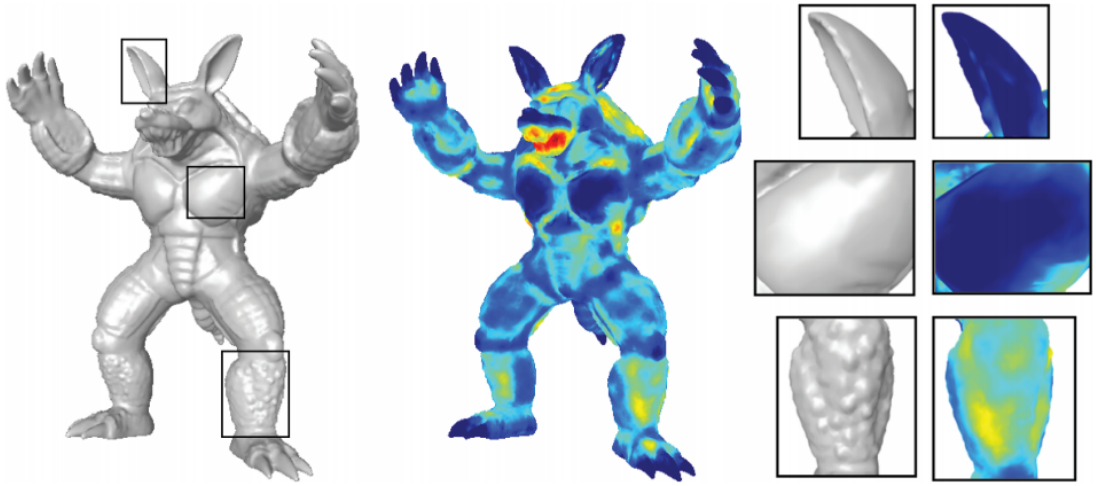


FIGURE 3.5: Surface roughness on the Armadillo model obtained with the roughness estimation method of Lavoué (image extracted from [Lav09]).

demonstrated the utility of the proposed roughness measure for two geometric operations: mesh compression and 3D watermarking. In summary, the roughness value is used to locally adapt the quantization operation of a compression algorithm. Instead of applying the same quantization level for the entire model, a higher level can be applied to the rough part as it can tolerate more geometric distortions. A similar idea was adapted to drive the 3D watermarking process as the strength of the watermark becomes dependent on the local roughness value.

In the interest of objectively evaluating the perceptual quality of a distorted mesh, Wang *et al.* [WTM12] proposed a perceptual metric (FMPD) that considers the visual masking effect using the roughness value of a surface. First the local roughness value is computed at each vertex as the Laplacian of the discrete Gaussian curvature which indicates whether the curvature is locally varying or not. The local roughness value is then modulated using a series of simple mathematical operations. The idea behind this modulation is to cause a small perceptual distance when a geometric distortion is located in a rough region and a big perceptual distance when the geometric distortion causes the smooth region to become rough, which mimics the effects of visual masking. The local roughness value is then integrated over the 3D mesh's surface to obtain a global roughness measure reflective of the overall roughness of the surface. Finally, the perceptual score relative to a distorted mesh is defined simply as the difference between the original and the distorted global roughness scores.

Finally, similarly to the first roughness measure proposed by Corsini *et al.* [CDGEB07], Váša and Rus [VR12] also rely on the values of dihedral angles to

measure the perceptual distance between two meshes. Their metric (DAME), is simply based on a weighted difference of the corresponding dihedral angles. Since the value of the angle between two adjacent faces is an indicator of surface roughness, then giving more weight to the difference of dihedral angle when its original value is small, *i.e.*, smooth surface, simulates the effects of visual masking. Recently, DAME was integrated into a 3D mesh compression algorithm [MVBH15] in order to perceptually drive the compression process by trying to minimize as much as possible the visible error caused by the absolute vertex displacement.

3.3.2 Curvature-Based Methods

In addition to surface roughness, the curvature of the surface has been another important geometric feature in the design of perceptual-driven methods as several observations [DFRS03, RBD06] have led to the conclusion that curvature information affects the intensity of the rendered image and thus affects the visual characteristics of a 3D model. In fact, it was noted that the human visual system is sensitive to strong variation in surface curvature [KKK02]. The surface curvature has, in particular, been used for assessing the visual quality of 3D models [Lav11, WTM12, TWC14, DFLS14] and mesh simplification through an estimation of visual saliency [HHO04, LVJ05, SLMR14].

3.3.2.1 Objective Quality Assessment for 3D Models

Following the concept of the structural similarity (SSIM) index [WBSS04] which consists of measuring the degradation of structural information in a 2D image, Lavoué *et al.* [LDGD+06] introduced a method for measuring the perceptual quality of 3D distorted meshes, MSDM. Instead of extracting the structural information using the luminance value in 2D images, the MSDM metric uses a combination of statistical analysis of surface curvature for that task. In order to compute the perceptual distance between two meshes X and X', a local perceptual distance between two local patches x and x' on the two meshes is first evaluated as follows (Fig. 3.6):

$$LMSDM(x, x') = \left(0.4 \cdot L(x, x')^3 + 0.4 \cdot C(x, x')^3 + 0.2 \cdot S(x, x')^3\right)^{1/3} , \qquad (3.6)$$

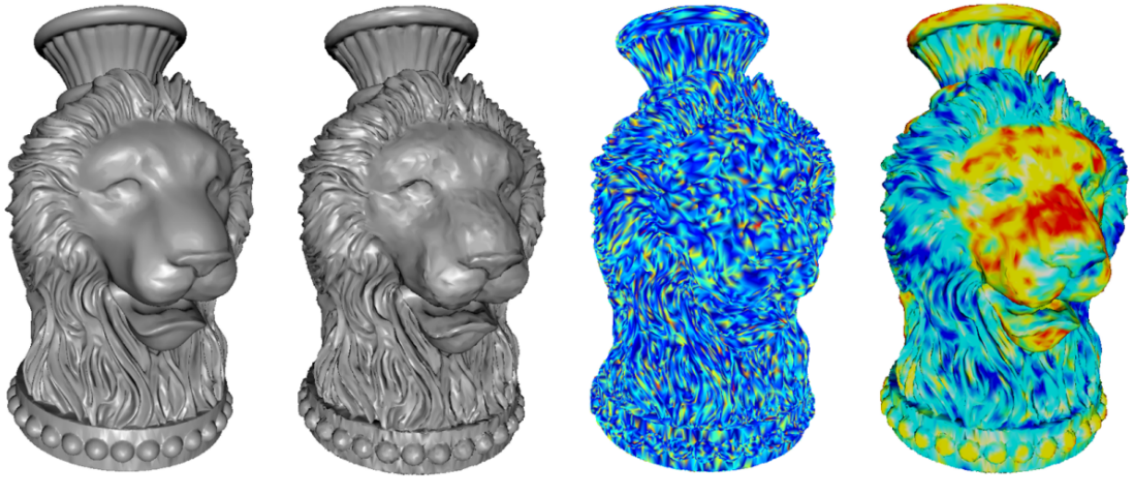


FIGURE 3.6: Local distortion maps of the Lion model obtained with the Hausdorff distance (left) and the MSDM2 metric (right) (image extracted from [Lav11]).

where L, C and S respectively correspond to the luminance, contrast and structure term of the SSIM index and are computed as:

$$L(x,x') = \frac{||\mu_x - \mu_{x'}||}{\max(\mu_x, \mu_{x'})}, C(x,x') = \frac{||\sigma_x - \sigma_{x'}||}{\max(\sigma_x, \sigma_{x'})}, S(x,x') = \frac{||\sigma_x\sigma_{x'} - \sigma_{xx'}||}{\sigma_x\sigma_{x'}}, \tag{3.7}$$

where μ_x, σ_x and $\sigma_{xx'}$ are respectively the mean, variance and covariance of the surface curvature over a local window of size ϵ around the vertex. It was recommended by the authors that ϵ would be equivalent to 0.5% the size of the model's bounding box. Finally the perceptual score between two models X and X' is obtained via a Minkowski pooling over the vertices of the 3D mesh as:

$$MSDM(X,X') = \left(\frac{1}{N}\sum_{i=1}^{N} LMSDM(x_i, x_i')^3\right)^{1/3}, \tag{3.8}$$

where N is the number of vertices. The MSDM metric was later improved by integrating a multiscale analysis of the perceptual distance and allowing the comparison of two meshes that do not share the same connectivity information [Lav11].

Different from MSDM and MSDM2 that only consider the amplitude of surface curvature, the TPDM perceptual metric [TWC14], proposed by Torkhani *et al.*, makes use of the principal curvature directions in its perceptual analysis. The surface normals vary the fastest in the direction of the maximum curvature and the slowest in the direction of the minimum curvature. Consequently, taking into

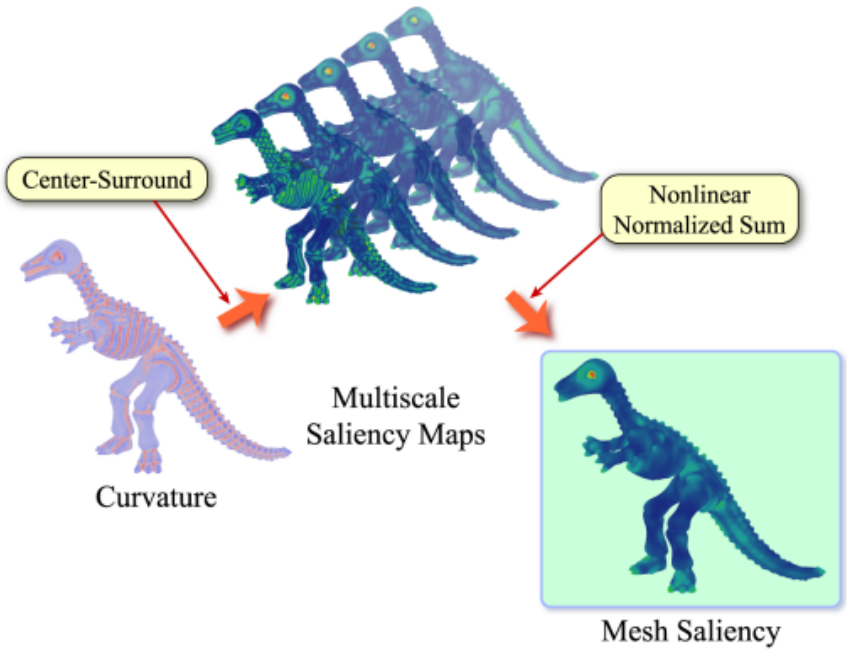


FIGURE 3.7: Overview of the method for computing the saliency as proposed by Lee *et al.* [LVJ05] (illutration extracted from [LVJ05]).

consideration the direction and amplitude of principal curvature provides a more detailed information about the structure of the surface. This is proven to be beneficial for the assessment of the perceptual quality of 3D triangular meshes by the TPDM metric.

3.3.2.2 Mesh Saliency and its Application to Mesh Simplification

There has been a large interest in the past few years to estimate the visual saliency on a 3D mesh as it has proven to be useful for several applications especially mesh simplification [HHO04, LVJ05, SLMR14]. Mesh saliency is a measure that tries to capture the visual importance of a region of a 3D mesh. In other words, a region with a high saliency value is more likely to attract the attention of the human viewer than a region with a low saliency value. Motivated by the experimental results of Howlett *et al.* [HHO04], in which the authors demonstrated the potential advantage of a saliency measure for geometric operations, Lee *et al.* [LVJ05] proposed an algorithm for computing the saliency on a 3D mesh. Their method is based on the idea that a region with a high saliency value stands out relative to its neighbors. The saliency estimation algorithm proceeds by computing a series of saliency maps at different scales which are then combined in a normalized non-linear sum (Fig. 3.7). In order to compute the saliency map at a scale σ, they proceed by the following steps. First the curvature is evaluated for each vertex

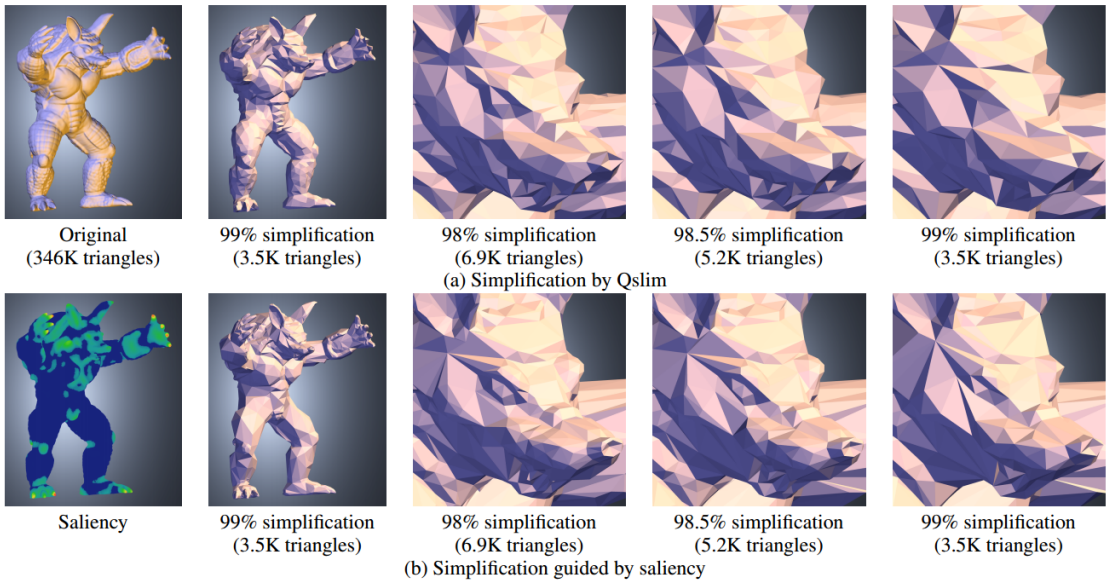


FIGURE 3.8: Results of the saliency guided simplification [LVJ05] (image extracted from [LVJ05]).

of the 3D mesh. The average curvature over a local window of sizes σ and 2σ is computed for each vertex. The saliency map value is finally obtained by computing the difference between the previously computed curvature averages. Lee *et al.* then demonstrated the utility of a saliency measure for the simplification of 3D models. They have modified the quadrics-based simplification method (Qslim) of Garland and Heckbert [GH97] by weighting the quadrics with mesh saliency. This resulted in a simplification procedure in which the visual important features of a 3D mesh are preserved (Fig. 3.8).

3.4 Image-Based vs. Model-Based Methods

In the last few years, perceptual methods have seen a rise in popularity [OHM+04, Fer08, CLL+13, LM15] as they have proven to be quite effective for a large number of applications. For instance, as detailed in this chapter, these approaches have been useful in the context of mesh rendering by either providing a criterion for selecting the best LOD for a certain scene [Red97, LH01, CB05, CSYB06] or by allowing for a more efficient management of resources in a physically-based rendering system [RPG99, RFWB07]. Perceptual methods have also provided the computer graphics community with several objective quality metrics [Lav11, WTM12, VR12, DFLS14, TWC14] that can be used to evaluate and debug existing geometric operations. Moreover, perceptually motivated approaches have proven to be especially helpful for the task of mesh simplification [WLC+03,

LVJ05, QM08, SLMR14] as they are able to preserve the visually important features of a 3D model. Despite all the aforementioned benefits a major issue remains when designing a perceptually oriented algorithm: *Is it better to perform the perceptual analysis on the 3D model or on the 2D rendered image?*

There are two sources of visual artifacts in a computer graphics system. First the geometric operation applied to the 3D model might introduce a visible geometric distortion and second the rendering algorithm may also cause some pixel-based artifacts. In theory, model-based approaches allow a more accurate analysis of geometric artifacts since the perceptual study is independent from any artifacts caused by the rendering algorithm. In this case, the perceptual analysis is carried out before generating the image, therefore, the visual impact of geometric and rendering distortions will not be mixed. Image-based methods have some advantages over current model-based ones. First and most importantly, since the perceptual analysis is carried out on the 2D image after the rendering step, it can easily cope with different rendering pipelines. This means that, since the analysis is done on the 2D rendered image, its results will adapt to the lighting condition, material properties, textures and rendering algorithm without having to alter the perceptual analysis procedure. In addition, these approaches also offer the choice between a view-dependent perceptual analysis, by taking one snapshot of the model, and a view-independent one, by taking multiple snapshots around the model, each of which can be useful for a range of applications.

However, in a comparative study to test the efficiency of perceptual image-based techniques in the case of computer generated images [CHM+12], Čadík *et al.* have concluded that image metrics are too sensitive for evaluating the visibility of distortions generated by a computer graphics pipeline. This is probably due to the difference in the type of both visual artifacts and images (artificial vs. natural) between the fields of computer graphics and image processing, for which these methods are designed. Moreover, many subjective studies have tried to compare between these two classes of methods in order to find out which is the more suitable for the task of estimating the perceptual impact of geometric distortions. This started with the experiments of Rogowitz and Rushmeier [RR01] which concluded that image-based metrics might not be suited for evaluating the quality of 3D model. This conclusion came as the result of a subjective experiment where the authors noticed that users have rated differently the artifacts cause by a simplification procedure when they observed the 3D model opposed to 2D still images. This difference is theorized to be due to the interactions when manipulating a 3D model. On the contrary, Cleju and Saupe have conducted a similar

experiment [CS06], but obtained conflicting results. The authors noticed that 2D metrics such as SSIM performed better than model-based ones when the simplification artifacts are beyond the visibility supra-threshold. Nevertheless, both of these experiments have some limitations as they only consider simplification artifacts and compare perceptual image metrics with non perceptual geometric ones (Hausdorff [ASCE02], RMS [CRS98]) since at that time more sophisticated geometric metrics (MSDM2 [Lav11], FMPD [WTM12], and DAME [VR12]) had not yet been developed. Recently and in the interest of providing a conclusive answer to the issue, Lavoué *et al.* [LLV16] realized a large study that compared the performance of state-of-the-art image-based methods with the state-of-the-art model-based methods for the task of evaluating the perceptual effects of a geometric distortion. In this study the authors took into consideration a large number of variables that affects the appearance of the 3D object in order to determine the parameters for which image-based methods performed the best. For instance, they have considered 2 rendering algorithms and 4 lighting conditions. In addition they used a large number of 3D models from the 3D Mean Opinion databases [LDGD+06, Lav09, VR12]. These databases contain 3D models with different distortion types (simplification, compression, filtering, ...) along with their corresponding mean subjective score. They have been used throughout the literature to evaluate the effectiveness of computer graphics perceptual metrics by computing the correlation between the metric results and the subjective scores. This study showed that, under the best parameters, the SSIM inspired metric of [WL11] along with the HDR-VDP2 [MKRH11] metric performed the best among image-based methods with a correlation greater than 60% with every database and scenario. Moreover, in simple scenarii, where each type of distortion or class of model is considered independently, image-based approaches are at their best with correlations close to 80% and even just surpassing 90% for simplification artifacts. However, this study also shows that perceptual image-based methods fall short in front of state-of-the-art model-based metrics. Indeed, the latter have a correlation that hovers around 85% regardless of the scenario.

Despite shown to be superior than image-based methods when it comes to evaluating the effects of geometric distortions on 3D objects, existing model-based methods have some issues. For instance, these metrics try to account for the perceptual characteristics such as visual masking and saliency using geometric measures defined on the 3D surface (surface roughness, curvature). The problem here is that these geometric features are not necessarily perceptually relevant attributes and their relation to human perception is based on general observations. Moreover, relying only on the surface geometry for the perceptual analysis makes

these methods unable of effectively adapting to a large number of important parameters that play an important role in the perception of rendered 3D objects such as illumination condition, the rendering procedure or display size and resolution.

3.5 Our Approach

In this thesis, different from all the methods mentioned in this chapter that either conduct the visibility analysis in a 2D space or completely rely on 3D geometric features, we present original algorithms whose goal is to compute the threshold beyond which a vertex displacement becomes visible, *i.e.,* the so-called Just Noticeable Distortion (JND) profile. Our approach is inspired by the image-based bottom-up framework (Section 3.2) which consists of trying to predict whether a change in contrast is visible or not. Hence, we start by computing appropriate perceptual properties (contrast, spatial frequency and visual regularity) on the meshe surface of a 3D object. These perceptual attributes should take into consideration the various parameters of mesh display (rendering, illumination, scale and display resolution) that generally affect their appearance. We then perform a series of psychophysical experiments to study the effects of contrast sensitivity and visual masking of the human visual system while observing a 3D model. The results of these experiments will allow us to propose a perceptual model that is able to predict whether a change in local contrast on a 3D mesh, induced by a local geometric distortion, is visible or not. This visibility model can afterwards be used to compute the threshold beyond which a vertex displacement becomes visible.

Chapter 4

Experimental Studies and Threshold Model

Existing model-based methods use measures such as surface roughness [CDGEB07, Lav09, WTM12, VR12] and surface curvature [LDGD+06, Lav11, TWC14] in order to carry out the perceptual analysis on a 3D mesh. However, these measures are not necessarily perceptually relevant. Our approach is inspired by the image-based bottom-up framework (Section 3.2) which consists of trying to predict whether a change in contrast is visible or not with the help of a mathematical model that simulates the perceptual properties of the HVS (Section 2.2). In this chapter, we define local perceptual properties for 3D meshes (*i.e.*, local contrast, spatial frequency and visual regularity) that are appropriate for a bottom-up evaluation of vertex displacement visibility. These perceptual properties allow us to study the effects of the contrast sensitivity and the visual masking in the 3D setting. In the following, we start by discussing the main experimental methods that were used throughout the literature in order to measure a certain perceptual threshold (Section 4.1). We then present both the proposed perceptual attributes and our experimental study regarding the visibility threshold in the case of flat-shaded (Section 4.2) and smooth-shaded (Section 4.3) 3D models.

4.1 Measuring the Visibility Threshold

The task of measuring the visual threshold is essential in vision science [CKT+99, WA05, Fer08, PB13] as it provides us with important information about the visual system. The results of such experiments have helped us understand and model the basic characteristics of the HVS which then can be used in several fields spanning from clinical studies [HHL+10] to bottom-up perceptual methods [Dal93, Red97, RPG99, LH01, LKW06, LK11].

In general, the experimental protocol for measuring a certain threshold follows the diagram presented in Fig. 4.1. The human subject is presented with task

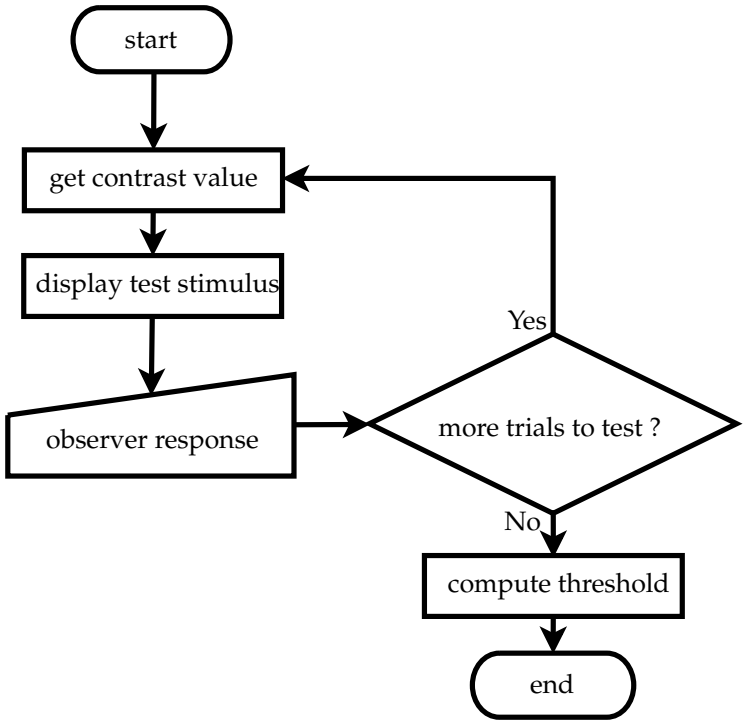


FIGURE 4.1: Diagram of a typical visibility threshold measurement experiment.

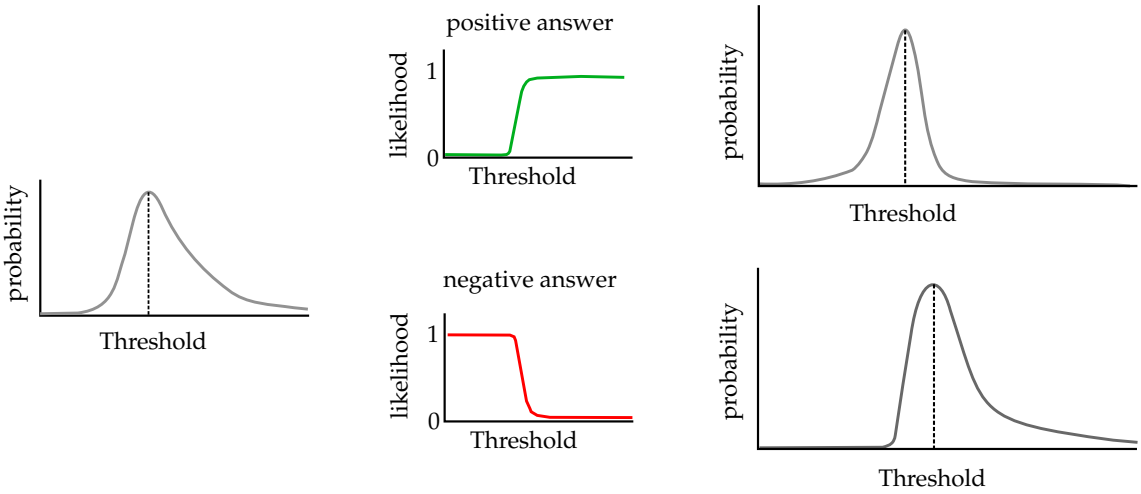
which usually consists in judging the visibility of some stimuli. For example, in a typical *Yes/No* task the subject is shown one image containing a visual stimulus and is asked to tell whether he can see it or not. On the other hand, in a *two alternative forced choice (2AFC)* type task the observer is presented with two images, one of which contains the stimulus. The observer is then asked to identify the image containing the stimulus. In order to measure the visibility threshold, the observer is tested over many trials. In each trial the stimulus is presented at a different contrast and the observer's answer is labeled either positive or negative. When the answer given by the observer indicates that he has seen the stimulus, for example a *Yes* answer in a yes/no type task or choosing the distorted stimulus in a 2AFC task, then it is labeled as *positive*. Otherwise, it is labeled as *negative*. The proportion of positive answers with respect to negative ones is then used to determine the contrast threshold.

There are many ways to select the contrast of the stimulus at which an observer is tested. They are grouped in two categories: (1) *non-adaptive* threshold methods where a set of contrasts is predetermined by the experimenter and (2) *adaptive* methods where the contrast of the stimulus on a certain trial is computed using the observer's response to the previous trials. While *non-adaptive* methods are easy to implement the main problem is their inefficiency. Before the experiment begins, the designer is required to perform a series of tests in order to

choose the set of contrasts at which the stimulus will be presented. In addition, measuring an accurate threshold with this method usually requires a large set of contrasts to be tested which will make the experiments long and therefore tiring for the subject. *Adaptive* methods, on the other hand, are more efficient and can converge to the threshold in as low as 20 trials. Many adaptive threshold methods have been developed throughout the literature. Starting with the *staircase* method [Cor62] where the contrast of the stimulus is either reduced or increased by a certain step after respectively a positive and negative answer. One of the biggest challenges of using this method is to determine the value at which the contrast is increased or decreased after each trial. This value plays an important role in the accuracy of the resulting threshold and efficiency of the experiment. A large value reduces the experiment time while a small value makes the resulting threshold more precise. Many strategies has been proposed for dealing with this issue such as using a fixed step [Cor62] or adapting its value according to the trial number [WL65]. Another type of adaptive methods, such as the PEST (Parameter Estimation by Sequential Testing) [TC67], groups a certain number of trials in several blocks. In these methods, instead of changing the contrast value after each trial, it is changed after each block with respect to the ratio of positive answers in that block. The motivation for doing that is to reduce the effects of *false positive* answers on the final threshold. For instance this is an important issue for 2AFC tasks where the subject has a 50% chance of choosing the distorted mesh even if he has not seen it. However, due to the large number of trials these methods are time consuming. Finally, there are ones that are based on the *maximum likelihood* method [Gre93] such as QUEST [WP83]. In this class of methods the contrast value and the response label are passed to a statistical model after each trial, which then outputs the next contrast value to test. This statistical model analyzes the user's response and outputs the most likely contrast threshold value so that the experiment is effective. it also takes into account the probability of false positive answers so that the measured threshold is as accurate as possible. In our experimental study we have used the QUEST method for measuring the visibility threshold, which we will detail in the next section.

4.1.1 The QUEST method

The idea of the QUEST method is to test at each trial the contrast value that is most likely to be the threshold. To do so, a probability density function (PDF) is first initialized on the contrast axis. In general, the PDF is assumed to be a Gaussian function whose parameters (mean and variance) are determined by the experimenter's prior knowledge about the threshold [WP83]. King-Smith *et al.* [KSGV+94] have also proposed to use a modified hyperbolic secant function as



(1) PDF at trial n, $P_n(T)$ (2) Likelihood function, $Q(r,T)$ (3) PDF at trial $n+1$, $P_{n+1}(T)$

FIGURE 4.2: Example of an iteration according to the QUEST procedure. The PDF of the contrast threshold at trial n is multiplied with the corresponding likelihood function in order to get the PDF for the trial $n+1$.

the PDF of the contrast threshold. The first stimulus is then presented to the observer with a contrast corresponding to the mode of the initial PDF, *i.e.*, the most likely threshold value. If the observer responds positively (contrast value is claimed to be visible by the observer), then the PDF is shifted towards lower contrast values otherwise it is shifted towards higher intensities. Watson and Pelly [WP83] have demonstrated that shift of the PDF can be done by applying Bayes' theorem. In consequence, the new PDF is computed as:

$$P_{n+1}(T) = Q(r,T) \cdot P_n(T)\ , \tag{4.1}$$

where P_n is the PDF of the contrast threshold T at the n^{th} trial and $Q(r,T)$ represents the likelihood of a getting positive answer ($r=1$) or a negative one ($r=0$) for the tested contrast value (T). This likelihood function is defined as:

$$Q(r,T) = \begin{cases} 1-\psi(T) & \text{if } r=0 \\ \psi(T) & \text{if } r=1 \end{cases}\ , \tag{4.2}$$

where $\psi(T)$ is the standard Weibull psychometric function [Wei51] (*i.e.*, $1-\delta-(1-\gamma-\delta)\cdot e^{-10^{3.5T}}$) whose parameters γ and δ reflect the probability of respectively false positive and false negative answers for a certain task. For example, in a *2AFC*-type task the probability of false positive answers is 50% and therefore γ is set to 0.5 while in a *Yes/No*-type experiment γ is usually initialized at 0.03 [Gre93, Tre95]. Multiplying the PDF with the likelihood function will shift position of the

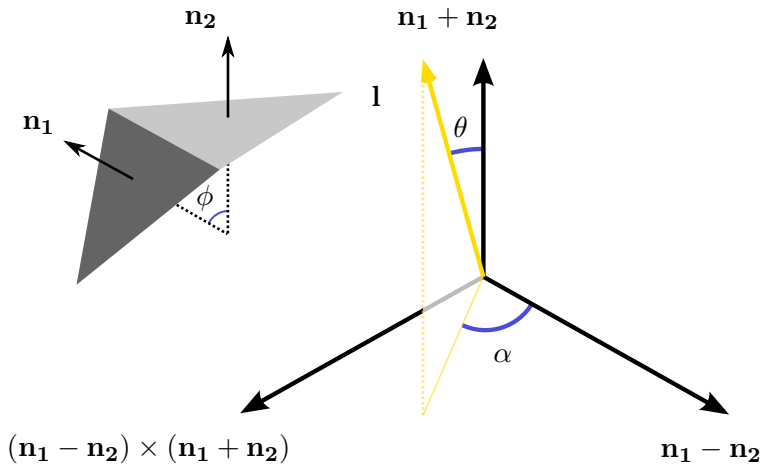


FIGURE 4.3: The contrast between adjacent faces is computed using the angle between their normals and the spherical coordinates of the light direction in the local coordinate system defined by the face normals.

PDF on the contrast axis and reduce its variance. The QUEST procedure usually stops when the PDF shift becomes negligible and the variance becomes too small. This is usually around 20 trials in a *Yes/No*-type experiment.

4.2 Measuring Contrast Threshold for Flat-Shaded Models

In this section we consider 3D meshes that are rendered with flat-shading algorithm. We start by describing a method for estimating the Michelson contrast (Section 4.2.1) and spatial frequency (Section 4.2.2) locally on a 3D mesh. We then present the psychophysical experiments (Section 4.2.3) that were aimed at measuring the visibility threshold in a flat-shaded setting. In addition, in this section, we limit our study to perfectly diffuse untextured surfaces that are illuminated with a directional white light.

4.2.1 Contrast Evaluation

The human visual system is primarily sensitive to variation in light energy, *i.e.*, contrast due to the center/surround organization of the receptive fields [Wan95]. In general, the most used contrast definition is that of Michelson where the contrast c is computed as:

$$c = \frac{L_{max} - L_{min}}{L_{max} + L_{min}}, \tag{4.3}$$

where L_{max} and L_{min} correspond to the luminance of the pixel with respectively the highest and lowest luminance value in a certain neighborhood.

In the case of a flat-shaded rendering, each face of the 3D mesh is attributed a single luminance value proportional to the cosine of the angle between its normal and the light direction. This means that the luminance value of each pixel belonging to a certain face is given by:

$$L = \max\left(\mathbf{l} \cdot \mathbf{n}, 0\right), \tag{4.4}$$

where $\mathbf{n}$ is the unit face normal and $\mathbf{l}$ is the light direction. In that setting, the local contrast is characterized between two faces as the contrast inside a face is 0 (all pixels have the same luminance level). The Michelson contrast between two adjacent faces is therefore defined by:

$$c = \frac{\parallel L_1 - L_2 \parallel}{L_1 + L_2} = \frac{\parallel \max\left(\mathbf{l} \cdot \mathbf{n_1}, 0\right) - \max\left(\mathbf{l} \cdot \mathbf{n_2}, 0\right) \parallel}{\max\left(\mathbf{l} \cdot \mathbf{n_1}, 0\right) + \max\left(\mathbf{l} \cdot \mathbf{n_2}, 0\right)}, \tag{4.5}$$

where $\mathbf{n_1}$ and $\mathbf{n_2}$ are the normals of the two adjacent faces. Under the circumstances where the inner products between the light direction and the two face normals are both positive, the above equation yields to the following equation:

$$c = \parallel \cos\alpha \cdot \tan\theta \cdot \tan\frac{\phi}{2} \parallel, \tag{4.6}$$

where α and θ are the spherical coordinates of the light direction in the local coordinate system defined by $\mathbf{n_1} - \mathbf{n_2}$, $\mathbf{n_1} + \mathbf{n_2}$ and their outer product (see Fig. 4.3). ϕ is the angle between the normals of the two faces. A detailed explanation of the transition between Eqs. (4.5) and (4.6) can be found in Appendix A.

Equation (4.6) shows how the contrast is affected by surface geometry and the scene illumination. The term $\tan\frac{\phi}{2}$ indicates the impact of surface geometry on the local contrast. On the one hand, if the surface is locally smooth ($\phi \approx 0°$), then the local contrast is minimal. On the other hand, if the surface is locally rough ($\phi \gg 0°$), then the local contrast tends to be high. In addition, the term $\cos\alpha \times \tan\theta$ describes how the light direction affects the local contrast. A grazing light direction will maximize the value of the contrast where θ is close to $90°$ and α is close to $0°$ or $180°$, while a light direction close to the normal direction ($\theta \approx 0°$) makes the contrast minimal.

4.2.2 Spatial Frequency

The spatial frequency is related to the size of the visual pattern, with respect to the size of one degree of the visual angle (Fig. 4.4) and is expressed in terms of cycles per degree (cpd). It is affected by the physical size of the object and the

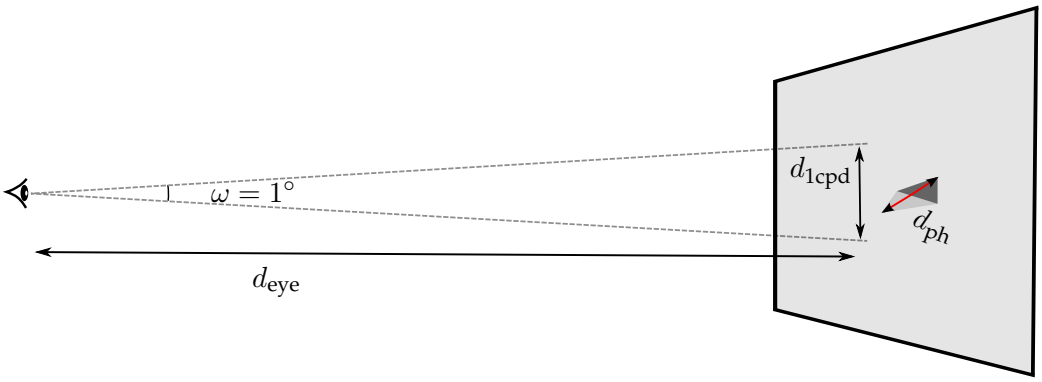


FIGURE 4.4: The spatial frequency is related to the size of the observed visual pattern (d_{ph}), with respect to the size of one degree of the visual angle ($d_{1\text{cpd}}$) on the display.

observer's distance to the object. In a flat-shaded environment, the visual stimulus is displayed on a screen and consists of the difference in luminance between a pair of adjacent faces. The perceived size of this stimulus depends then on the display's properties (resolution and size), the observer's distance to the display, the position of the model in the virtual 3D world and the size of the faces. So estimating the spatial frequency in the 3D setting requires first converting the size of the visual stimulus from its virtual value in the 3D world to its physical size on the display. As a consequence, we first evaluate the number of pixels that are occupied by the visual pattern. To do so, we start by computing the size of the visual stimulus in the virtual 3D world. It corresponds to the distance between the opposites vertices of two adjacent faces in a flat shaded mode. We then compute the number of pixels that the visual pattern occupies on the screen by applying a perspective projection.

Having evaluated the number of pixels, the physical size of the visual pattern is then computed using the display's properties (resolution and size) as:

$$d_{\text{ph}} = \frac{n_{\text{px}}}{\sqrt{{r_h}^2 + {r_v}^2}/s} \,, \tag{4.7}$$

where n_{px} is the number of pixels of the displayed visual pattern, r_h and r_v are the display's horizontal and vertical resolution and s its diagonal size. Finally the spatial frequency is estimated by:

$$f = \frac{d_{1\text{cpd}}}{d_{\text{ph}}} \,, \quad d_{1\text{cpd}} \approx d_{\text{eye}} \cdot \pi/180 \,, \tag{4.8}$$

where $d_{1\text{cpd}}$ is the size of one degree of the visual angle on the display and d_{eye} is the distance between the eye and the display. It is interesting to note the effects

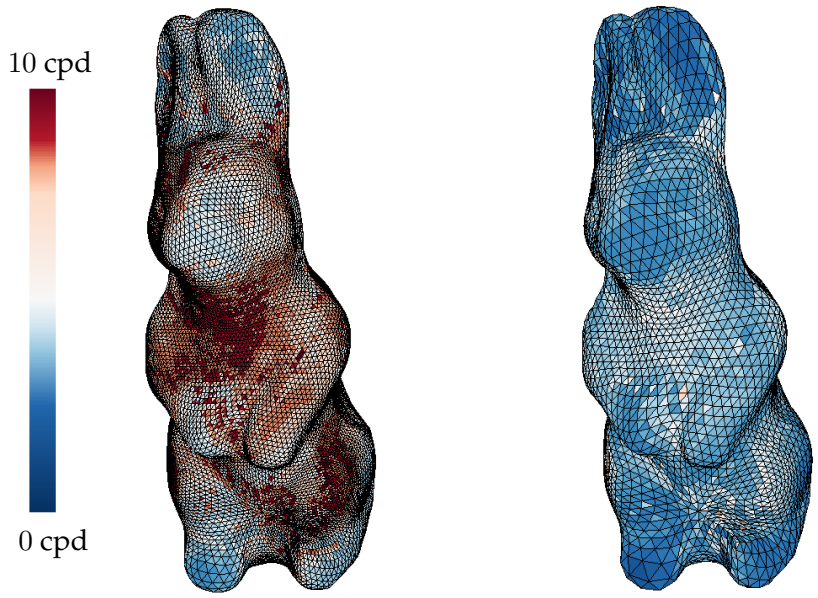


FIGURE 4.5: The vertex density has an impact on the spatial frequency.

of the vertex density of a 3D mesh on the perceived spatial frequency. While a dense model will most likely exhibit high frequency stimuli, a coarse model will show low frequency ones (Fig. 4.5).

4.2.3 Experimental Study

As the coordinates of a vertex on the surface of a 3D mesh change, the local contrast of the surface around this vertex changes also. The purpose of our experimental study is to measure the contrast threshold relative to the displacement of a vertex on the surface of a 3D model by studying the effects of contrast sensitivity and visual masking while observing a 3D model. We will start by describing the experimental protocol used for the task of measuring the visibility threshold relative to the contrast sensitivity and visual making of the visual system.

4.2.3.1 Experimental Protocol

Our first group of psychophysical experiments concerns measuring the visibility threshold relative to the effects of contrast sensitivity and visual masking in a flat shading rendering. To do so, we have proposed an experimental protocol that is similar to the usual visibility threshold measurement procedure (Fig. 4.1).

We have designed the task given to the participants to be as precise and efficient as possible at the same time. The efficiency of the task is an important criterion since long experiments will be both time consuming and more importantly tiring for the participants which might affect their performance. Hence, we have

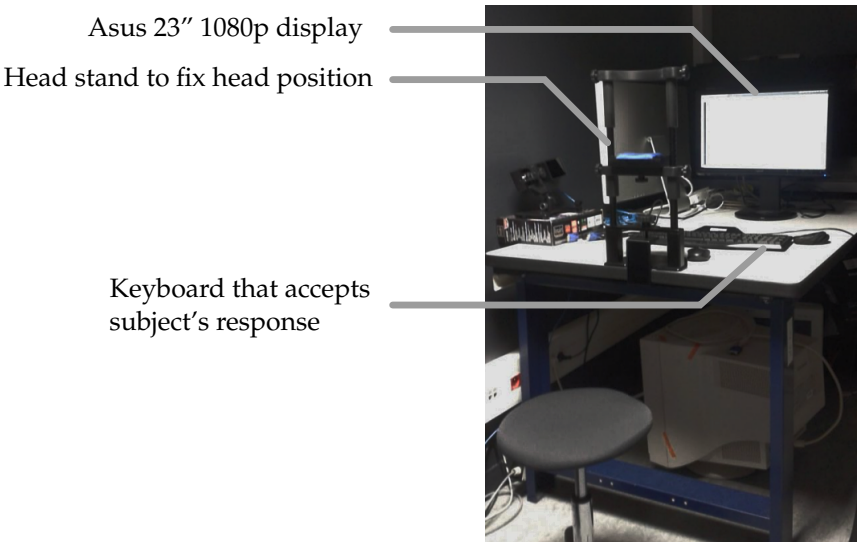


FIGURE 4.6: Experimental setup.

used the QUEST method to adjust the intensity of the stimulus after each trial and designed the task as follows. In a typical *Yes/No*-type task the subjects are shown one stimulus and are asked to indicate whether they are able to see it or not. The presence of only one stimulus means that the observer will judge it as visible (*Yes* answer) if its intensity is above a certain internal criterion [GS66]. The issue here is that this internal criterion might shift over time and is different for each subject and therefore could lead to a less accurate threshold. The solution for reducing the effects of the subjective criterion is to show two objects, one acting as a reference, so that the subject is able to base his answer on a comparison with this reference. For example, in a *2AFC*-type task, the subject is required to identify which of the two displayed objects is the distorted one. However, due to the high probability of false-positive answers in a *2AFC* the QUEST will slowly converge to the threshold which would make the experiments long and tiring [JW06]. Therefore, in our experiments, the task given to the observers was based on a slightly altered *Yes/No* procedure. Instead of showing one object, we have displayed two objects side by side on the screen, one of which exhibits a displaced vertex in its central area. The subjects where then asked to respond by *Yes* or *No* to the following question: *Are the two objects different?* If the vertex displacement is visible, then the objects will appear to be different and thus a *Yes* answer is expected. If the vertex displacement is not visible, then both objects will appear identical and a *No* answer is expected. Having a *Yes/No*-type task compared to a *2AFC*-type one makes measuring the threshold faster as the probability of a false-positive response is low [KSGV+94, JW06]. In addition, having a response based on a comparison of two objects increases the accuracy of the measured threshold compared to a typical Yes/No procedure as it reduces the effects of the internal

criterion based on which the subjects give their answers.

The experiments took place in a low illuminated laboratory environment (Fig. 4.6). The stimuli were displayed on an Asus 23-inch display in a low illuminated room. Screen resolution was 1920×1080. The stimuli were observed from a distance of 1 m, in order to allow us to measure the threshold for frequencies between 1 and 16 cpd (a closer screen would make high-frequency stimuli smaller than 1 pixel). 5 subjects participated in our experiments. All had normal or corrected-to-normal vision and were 22 to 26 years old. One of the participants was experienced in perceptual subjective evaluations and the other 4 were inexperienced. The participants repeated the experiment 4 times each on a different day and on a different time of day (morning, afternoon). No user interaction was allowed.

4.2.3.2 Contrast Sensitivity

FIGURE 4.7: Visual stimulus presented to the subjects for measuring the Contrast Sensitivity Function in the case of flat shading. Left: the reference plane. Right: a vertex is displaced in the central area of the plane.

Visual Stimulus In order to measure the CSF in the 3D setting, the natural visual stimulus consists of a vertex displaced from the surface of a regular plane whose local contrast is 0 (Fig. 4.7). The displacement of the vertex alters the normal of the adjacent faces and thus changes the contrast. In order to measure the threshold of different frequencies we change the vertex density of the plane, which alters the size of its faces. The threshold is measured for 8 spatial frequencies (1.12, 2, 2.83, 4, 5.66, 8, 11.30 and 16 cpd). An additional "dummy" frequency, whose data were not taken into account, was included at the beginning of each session to stabilize the subject's answers. In order to avoid any bias, frequency order was randomized for each observer in each session. The plane is tilted by 20° to give the observer a 3D feel.

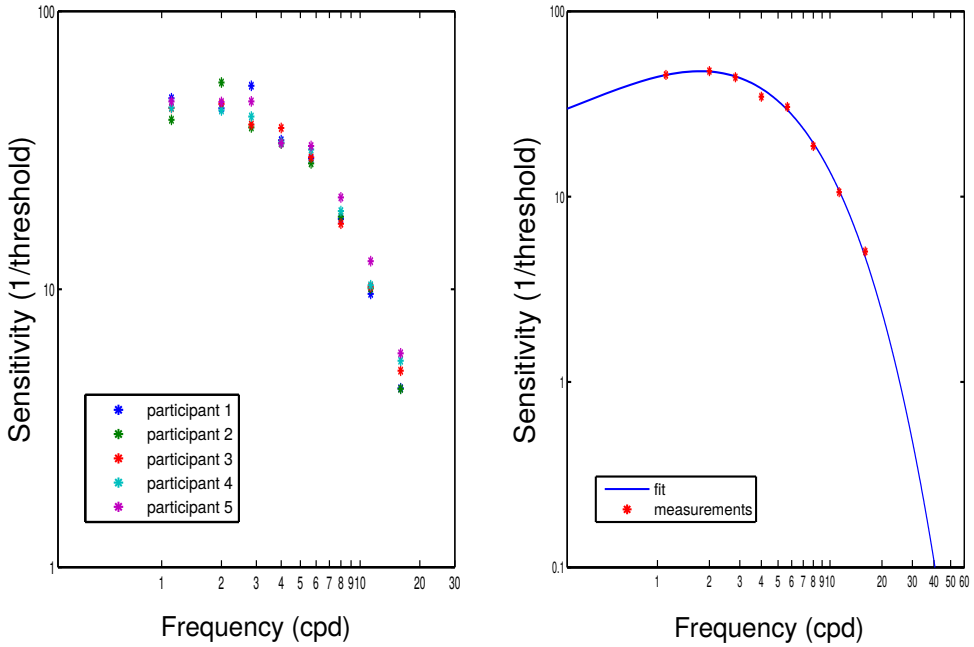


FIGURE 4.8: Left: plot of the mean sensitivity for each observer over each frequency. Right: plot of the subjects' mean sensitivity over each frequency fitted using Manos and Sakrison's mathematical model.

Results The results of this experiment are shown in Fig. 4.8. The displacement of a vertex causes a variation in contrast for multiple face pairs. We save the maximum contrast between the affected face pairs. The left panel of Fig. 4.8 plots the mean sensitivity for each observer over each frequency. The plot shows a high consistency between the participants: All of them exhibit a peak in sensitivity at 2 cpd and the drop off in sensitivity on either side of the peak is similar for all participants. The right panel of Fig. 4.8 shows the subjects' mean sensitivity over each frequency, fitted using Mannos and Sakrison's mathematical model [MS74] that is defined by:

$$\mathrm{csf}(f) = \left(1 - a + \frac{f}{f_0}\right) e^{-f^p}, \tag{4.9}$$

with $a = -15.13$, $f_0 = 0.0096$ and $p = 0.64$. The fit predicts a peak in sensitivity at around 2 cpd that drops rapidly at high frequencies. At low frequencies the drop in sensitivity is much slower than the one measured with a 2D contrast grating [BC69b, WA05]. This is probably due to the aperiodic nature of the visual stimulus [BC69b].

4.2.3.3 Visual Masking

Visual masking occurs when the visibility of stimulus (the *target*) is reduced due to the presence of another visible stimulus (the *mask*). Since the visibility of a

visual pattern depends on its spatial frequency, the masking threshold is different at each frequency. However, if we normalize the contrast values by the mask's CSF value, then the resulting threshold will be independent of the stimulus's spatial frequency [Dal93]. This normalization is achieved by the following

$$\tilde{c} = c \cdot \mathrm{csf}(f) \ , \tag{4.10}$$

where c and f are respectively the contrast and spatial frequency of a visual pattern. Ultimately, when $\tilde{c} \geqslant 1$ this means that the contrast is above the visibility threshold given by the CSF, otherwise ($\tilde{c} < 1$) the contrast is considered not visible. Therefore, measuring the masking effect can be done by changing the contrast value of a mask signal without the need to pay much attention to its spatial frequency. We have verified this hypothesis through a preliminary experiment where the contrast masking threshold for three different frequencies was the almost same after normalization by the corresponding CSF value. The results of this preliminary experiment can be found in Appendix B.

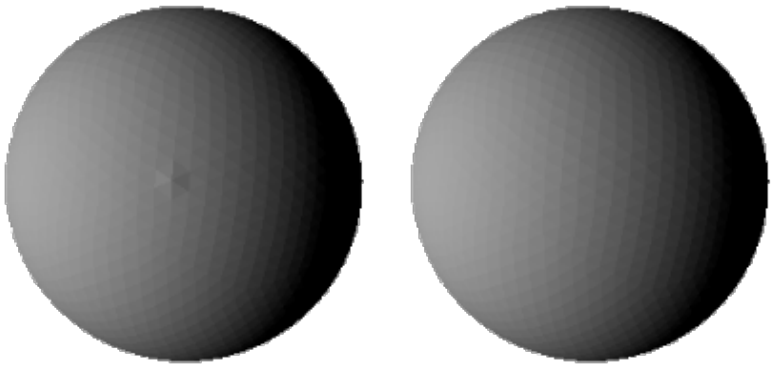


FIGURE 4.9: Visual stimulus for measuring threshold relative to the aspects of contrast masking in the case of flat shading. Left: a sphere approximated by an icosahedron subdivided 3 times from which a vertex is displaced. Right: the reference sphere.

Visual Stimulus In order to measure the threshold relative to the masking effect, the initial visual stimulus needs to exhibit a visible contrast (*i.e.*, $\tilde{c} \geqslant 1$). We then increase the initial contrast and measure the value needed to notice that change. In other words, if c is the initial contrast (mask signal) and c' is the increased value, we measure $\Delta c = c' - c$ (target signal) needed to discriminate between c and c'. The stimulus consists of a vertex displaced from a sphere approximated by a subdivided icosahedron (Fig. 4.9). The icosahedron is subdivided 3 times, which makes the contrast between two adjacent faces (stimulus of about 2 cpd) visible for an observer. This initial contrast represents the mask signal. Varying

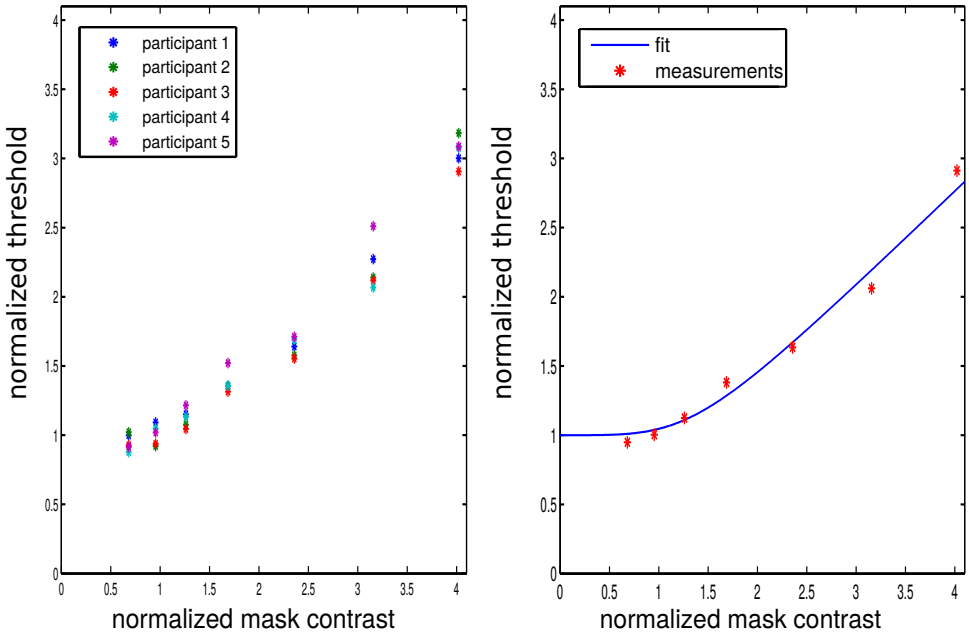


FIGURE 4.10: Left: plot of the normalized mean threshold for each observer over normalized mask contrast. Right: plot of the subjects' mean normalized threshold over each normalized mask contrast, fitted using Daly's mathematical contrast masking model.

the light direction modifies the value of the initial contrast between two adjacent faces. We measured the threshold relative to 7 normalized contrasts that were log-linearly spaced from 0.6 to 4.

Results The results of this experiment are shown in Fig. 4.10. The left panel plots for every participant the mean normalized threshold over the normalized contrast mask. For mask contrasts below the visibility threshold (normalized contrast of the mask lower than 1), the measured normalized threshold is close to 1. This indicates that the measured threshold refers to the one given by the CSF and that no masking has occurred. For mask contrasts above the visibility threshold, the measured normalized threshold is above the one given by CSF and lies close to the asymptotic region with a slope near 0.7. The right panel of Fig. 4.10 shows the subjects' mean threshold over each mask contrast fitted using Daly's mathematical masking model [Dal93] that is defined by:

$$\text{masking}(\tilde{c}) = \left(1 + (k_1 \times (k_2 \times \tilde{c})^s)^b\right)^{1/b}, \tag{4.11}$$

with $\tilde{c}$ the normalized threshold, and the fitted values $k_1 = 0.0078$, $k_2 = 88.29$, $s = 1.00$ and $b = 4.207$.

4.2.4 Contrast Threshold

Having performed a series of psychophysical experiments in order to study the effects of contrast sensitivity and visual masking in a 3D setting, we can now derive a computational model to evaluate the threshold T beyond which a change in contrast becomes visible for the human observer as follows:

$$T(c, f) = \frac{\text{masking}(c \cdot \text{csf}(f))}{\text{csf}(f)} , \tag{4.12}$$

where c is the original contrast and f the spatial frequency. The proposed threshold T can adapt to various parameters. When computing the local spatial frequency, it takes into consideration the size and resolution of the display as well as the vertex density of the mesh. The threshold T can also adjusts to the scene's illumination since it influences the initial contrast.

Furthermore, for estimating the probability of detecting a change in contrast, it is common in the field of visual perception to use a psychometric function (Eq. (4.13)) with a slope set to 3.5 [MT86].

$$p(\Delta c, T) = 1 - e^{(\Delta c/T)^{3.5}} , \tag{4.13}$$

where T is the contrast threshold and Δc is the change in contrast which corresponds to contrast difference before and after the displacement of a vertex. Δc is evaluated as:

$$\Delta c = \begin{cases} \parallel c' - c \parallel & \text{if } \text{sgn}(\mathbf{n_1} \cdot (v_4 - v_3)) \text{ does not change,} \\ c' + c & \text{if } \text{sgn}(\mathbf{n_1} \cdot (v_4 - v_3)) \text{ changes,} \end{cases} \tag{4.14}$$

where c and c' are respectively the contrast of the adjacent faces before and after the vertex displacement. We test whether the vertex displacement causes a change in convexity, which is reflected by a change in the sign of $\mathbf{n_1} \cdot (v_4 - v_3)$ in order to detect, for instance, the ambiguous case as shown in Fig. 4.11, where the displacement does not induce a change in the "conventional" contrast between the adjacent faces.

The method proposed in this section works only for models illuminated by a directional light and rendered with a flat-shaded algorithm as a result of using the limited contrast estimation method. In addition, the perceptual model used for computing the visibility of the geometric distortions does not take into account either the regularity of the visual pattern or the effects of global luminance on the CSF value. This is due to the simplified CSF and masking functions (Eqs. (4.9)

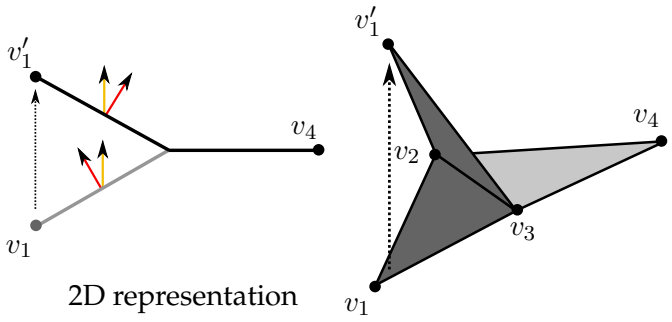


FIGURE 4.11: When the displacement of a vertex alters the convexity of two adjacent faces, the contrast might remain the same as long as the angle between the light direction (yellow arrow) and the face normal (red arrow) does not change.

and (4.11)) used in the threshold model and will most likely result in overestimating the perceptual impact of distortions in complex or dark regions of a 3D mesh. These limitations are taken into account in our second stage of our experimental study where we present a more complete threshold model for smooth shaded meshes.

4.3 Measuring Contrast Threshold for Smooth-Shaded Models

Building on the aforementioned work on computing the contrast visibility threshold in a flat-shaded environment, we now extend it to smooth-shaded models. To do so, we generalize the method of estimating the local contrast on a 3D model to smooth-shaded algorithms and different illumination types (directional and point light). We also extend our study of the contrast sensitivity to include the effects of the scene's global luminance. Moreover, based on the free-energy principle, we propose a method to compute the visual regularity of a rendered mesh which allows us to take into account its influence over the visibility threshold.

4.3.1 Contrast for a Smooth-Rendering Setting

In a smooth-shaded rendering algorithm, each point on a triangular face surface is attributed a luminance value. In consequence, each face of the triangular mesh exhibits a local contrast. Hence, in order to compute the contrast of a face, we need to find the points corresponding to the highest and lowest luminance values, L_{max} and L_{min} respectively.

In this section, we propose an analytical method to compute these points. This will allow us to estimate the Michelson contrast (Eq. (4.3)) for a given face. Let $F = \{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be a face of a 3D mesh and let $\mathbf{x}_i$ be a point belonging to F. The surface normal at $\mathbf{x}_i$ is obtained using a barycentric interpolation of vertex

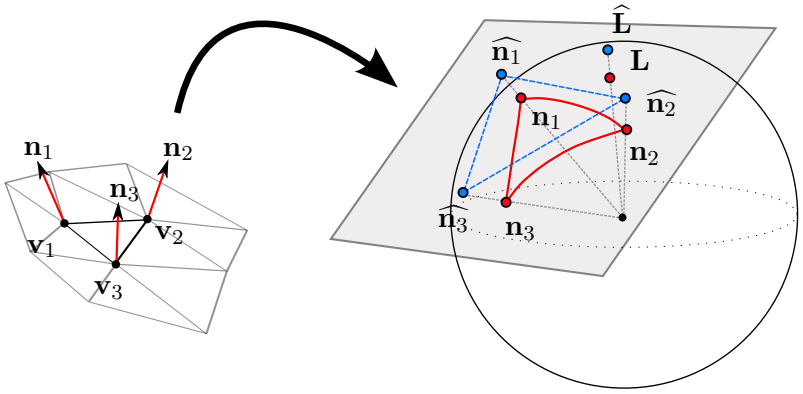


FIGURE 4.12: The projection of the normals, $[\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3]$, on the unit sphere and to the tangent plane allows us to compute the barycentric coordinates of the closest and farthest points to the light direction $\mathbf{L}$.

normals:

$$\mathbf{n}_{\mathbf{x}_i} = \frac{\mathbf{h}_{\mathbf{x}_i}}{||\mathbf{h}_{\mathbf{x}_i}||} \quad ; \quad \mathbf{h}_{\mathbf{x}_i} = N \times \mathbf{b}_{\mathbf{x}_i}, \tag{4.15}$$

where $N = [\mathbf{n}_1,\ \mathbf{n}_2,\ \mathbf{n}_3]$ is the matrix of vertex normals and $\mathbf{b}_{\mathbf{x}_i} = [\alpha_i,\ \beta_i,\ 1-\alpha_i-\beta_i]^T$ is the vector of barycentric coordinates of $\mathbf{x}_i$. In the case of a diffuse surface, the luminance attributed to $\mathbf{x}_i$ is proportional to the cosine of the angle between the surface normal $\mathbf{n}_{\mathbf{x}_i}$ at $\mathbf{x}_i$ and the light direction $\mathbf{L}$. So finding the brightest and darkest points of a face boils down to finding the points with respectively the smallest and biggest angle between the corresponding normal and light direction. This task can be achieved by computing their barycentric coordinates as explained below.

We first map the normals of all the points $\mathbf{x}_i \in F$ and the light direction $\mathbf{L}$ onto the unit sphere (Fig. 4.12). It is easy to prove that the set of normals of F forms a spherical triangle on the unit sphere as the normals of each edge of F correspond to a geodesic on the unit sphere. Let $\widehat{\mathbf{n}_{\mathbf{x}_i}}$ be the gnomonic projection of $\mathbf{n}_{\mathbf{x}_i}$ onto the tangent plane of the unit sphere at the centroid of the spherical triangle (Fig. 4.12) and let $\widehat{\mathbf{L}}$ be the projection of $\mathbf{L}$. The gnomonic projection is especially useful to our purposes since it projects geodesics to straight lines. In consequence, the points $\widehat{\mathbf{n}_{\mathbf{x}_i}}$ determine a euclidean triangle $\widehat{F}$ in the tangent plane. This means that finding the barycentric coordinates of the points with the smallest and biggest angles between the normal and light direction can be achieved by computing the barycentric coordinates of closest and farthest points between $\widehat{F}$ and $\widehat{\mathbf{L}}$. For $\mathbf{x}_i \in F$, the distance between corresponding $\widehat{\mathbf{n}_{\mathbf{x}_i}}$ and $\widehat{\mathbf{L}}$ can be expressed as:

$$d_{\mathbf{x}_i}(\alpha,\beta)^2 = ||\alpha\widehat{\mathbf{n}_3\mathbf{n}_1} + \beta\widehat{\mathbf{n}_3\mathbf{n}_2} + \widehat{\mathbf{L}\mathbf{n}_3}||^2\,, \tag{4.16}$$

where α, β are the barycentric coordinates of $\mathbf{x}_i$. The barycentric coordinates

relative to the point with the highest and lowest luminance value can finally be obtained by solving the following systems:

$$\begin{aligned} &\text{argmin } \{d_{\mathbf{x}_i}(\alpha,\beta)\}, \quad \alpha+\beta \leq 1 \text{ and } \alpha,\beta \in [0,1]\ ; \\ &\text{argmax } \{d_{\mathbf{x}_i}(\alpha,\beta)\}, \quad \alpha+\beta \leq 1 \text{ and } \alpha,\beta \in [0,1]\ . \end{aligned} \tag{4.17}$$

A detailed description of the solution of Eq. (4.17) can be found in Appendix A. Having computed the brightest and darkest points of a face, it is now possible to evaluate its Michelson contrast. The contrast computed according to the method described above is compatible with directional light sources. It is also possible to extend this method to point light sources by assigning to each point $\mathbf{x}_i \in F$ a light direction according to:

$$\mathbf{l}_{\mathbf{x}_i} = \frac{\mathbf{g}_{\mathbf{x}_i}}{||\mathbf{g}_{\mathbf{x}_i}||} \quad ; \quad \mathbf{g}_{\mathbf{x}_i} = \mathbf{x}_i - \mathbf{p} = M \times \mathbf{b}_{\mathbf{x}_i} - \mathbf{p} \tag{4.18}$$

where $\mathbf{l}_{\mathbf{x}_i}$ is the light direction at $\mathbf{x}_i$, $\mathbf{p}$ is the light position, $M = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3]$ is the matrix of vertex position and $\mathbf{b}_{\mathbf{x}_i}$ is vector of barycentric coordinates of $\mathbf{x}_i$. For the same reason, the mapping of the light directions on the unit sphere will form a spherical triangle as the light directions assigned to edges of the face correspond to a geodesic and thus creating a euclidean triangle when projected to the tangent plane. Finally, The distance between $\widehat{\mathbf{n}_{\mathbf{x}_i}}$ and $\widehat{\mathbf{l}_{\mathbf{x}_i}}$ on the tangent plane can be evaluated as:

$$d_{\mathbf{x}_i}(\alpha,\beta)^2 = ||\alpha(\widehat{\mathbf{n}_3\mathbf{n}_1} - \widehat{\mathbf{l}_3\mathbf{l}_1}) + \beta(\widehat{\mathbf{n}_3\mathbf{n}_2} - \widehat{\mathbf{l}_3\mathbf{l}_2}) + (\widehat{\mathbf{l}_3\mathbf{n}_3})||^2\ . \tag{4.19}$$

By solving Eq. (4.17) for the distance in Eq. (4.19) we can evaluate the Michelson contrast for 3D models illuminated by a point light.

Mapping the vertex normals on the unit sphere, makes it easy to understand how the shape of the surface affects the local contrast. As the curvature of the surface increases, the area of the spherical triangle increases. This makes the contrast value attributed to that face be potentially high as for a certain light direction the distance between the closest and farthest points on the triangle is most likely to be large. It is important to also note that the presented method is capable of adapting to simple rendering algorithms where the luminance is computed in a per-pixel basis so that the contrast inside a face remains the dominant local contrast. For example, Fig. 4.13 shows the contrast computed on a 3D mesh rendered with two different shadings: a regular smooth shading algorithm and a cell shading one. Notice how the contrast of the faces relative to the cell shaded rendering of the 3D model is 0 except for the ones where a transition in luminance occurs.

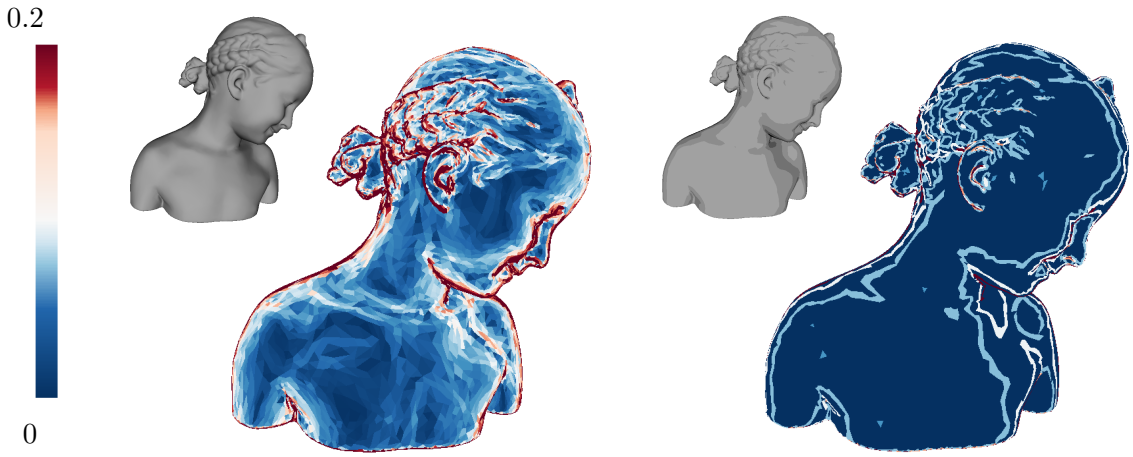


FIGURE 4.13: The Michelson contrast computed for each face of the Bimba model for a regular smooth shading algorithm (left) and cell shading rendering (right).

This method is also independent from the normal evaluation algorithm used to compute the normals of the vertices of the 3D model.

4.3.2 Regularity of the Visual Signal

The regularity of a visual pattern plays an important role in the ability of the visual system to distinguish between two visible contrasts. The effects of the visual regularity of a pattern on the contrast threshold can be explained by the *free-energy principle* theory [FKH06, Fri10]. The brain can easily and successfully predict the visual patterns of a regular stimulus while irregular visual stimuli are difficult to be predicted [KP04, FDK09]. Based on this fact, we can relate the visual regularity to the *prediction error* of a visual pattern.

We propose a computational model that aims to predict the local contrast value from the contrast information of its surrounding. The visual regularity can then be estimated from the residue between the actual contrast value and the predicted one. We suppose that the local contrast of a triangular face F, denoted by c, can be estimated using a linear combination of the local contrast of the three surrounding faces sharing an edge with F:

$$c' = x_1 c_1 + x_2 c_2 + x_3 c_3 \ , \tag{4.20}$$

where c' is the estimated contrast and c_1, c_2 and c_3 are the contrast values of the adjacent faces organized in a descending order. So in order to evaluate c' we must estimate the linear coefficients $[x_1, x_2, x_3]$. This can be achieved by solving

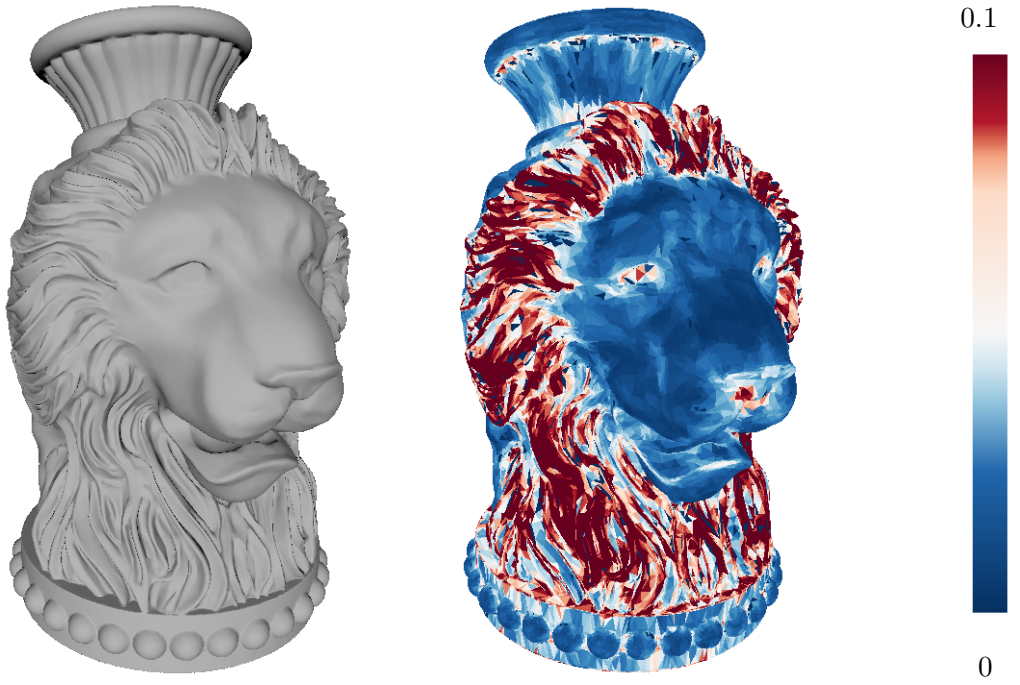


FIGURE 4.14: Visual regularity on the Lion-vase model.

the following linear system using the least square regression method:

$$\begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} \\ & \vdots & \\ c_{i,1} & c_{i,2} & c_{i,3} \\ & \vdots & \\ c_{n,1} & c_{n,2} & c_{n,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} c_1 \\ \vdots \\ c_i \\ \vdots \\ c_n \end{bmatrix} , \tag{4.21}$$

where c_i is the contrast value of the ith face within a predefined neighborhood centered at the current face F, $c_{i,1}, c_{i,2}, c_{i,3}$ are the contrast values of the corresponding adjacent faces and n is the total number of faces in the neighborhood. In practice we have used a neighborhood of a size equivalent to 3.5 cpd which corresponds to the most sensitive spatial frequency according to the contrast sensitivity function in order to estimate the value of $[x_1, x_2, x_3]$ for each face. Finally the visual regularity (closer to 0 means more regular) assigned to a face F is obtained by computing the absolute difference between the actual contrast and the estimated one:

$$r = |c - c'| \, . \tag{4.22}$$

Figure 4.14 shows the visual regularity for the Lion-vase model. Notice how the region containing the lion's mane is considered as visually irregular while the smooth face is visually regular.

4.3.3 Spatial Frequency

In the case of a smooth-shaded model, the local visual stimulus consists of a luminance patterns inside a triangular face. The frequency of this pattern is related to the distance between its brightest and darkest points. These points can be obtained using the method described in Section 4.3.1. Having computed the virtual distance between the brightest and darkest points or a face, we now proceed similarly to the case of flat-shaded models. First we apply the perceptive projection to get the number of pixels corresponding to this distance and then apply Eqs. (4.7) and (4.8) to obtain the frequency in cycles per degree.

Having defined local perceptual attributes for 3D models (contrast, frequency and visual regularity), we now present our experimental study that aims to measure the contrast threshold required to detect a change in the geometry of the mesh.

4.3.4 Experimental Study

In our second batch of psychophysical experiments, we measured the contrast threshold relative to the displacement of a vertex in case of a smooth-shaded environment. Compared with our previous experiments, we consider this time a more complex CSF model that takes into account not only the spatial frequency of the local stimulus but also the global luminance value of the rendered scene. In addition, we extend the masking model to include the effects of the regularity of a visual pattern on the contrast threshold.

4.3.4.1 Experimental Procedure

For these threshold measurements we have kept almost the same experimental procedure used in our previous experiments as described in Section 4.2.3. Two objects were displayed side by side on a screen, one of which exhibits a displaced vertex or a series of displaced vertices. The subjects were instructed to answer by *Yes* or *No* to whether the two objects on the screen appear identical. The magnitude of the vertex displacement was then adjusted after each response according to the QUEST procedure [WP83]. The experiments took place in the same experimental environment as the previous ones. 12 subjects took part in this series of experiments. 4 of them have participated in the previous measurements. Also no user interaction was allowed.

FIGURE 4.15: Increasing the vertex density of the plane would increase the spatial frequency of the visual stimulus.

4.3.4.2 Contrast Sensitivity

Visual Stimulus Similarly to the flat-shaded experiments, we measured the visibility threshold relative to the contrast sensitivity using a regular plane (Fig. 4.15). The difference this time is that we consider the effects of both spatial frequency and global luminance on the contrast threshold. To do so, we alter the mesh density to change the spatial frequency of the stimulus and we vary the lighting conditions (light energy) to change the global luminance level of the scene. The threshold was measured for 7 spatial frequencies (0.5, 2, 4, 5.66, 8, 11.3 and 16 cpd) and for 3 luminance levels ($180,\ 110$ and 33 cd/m^2).

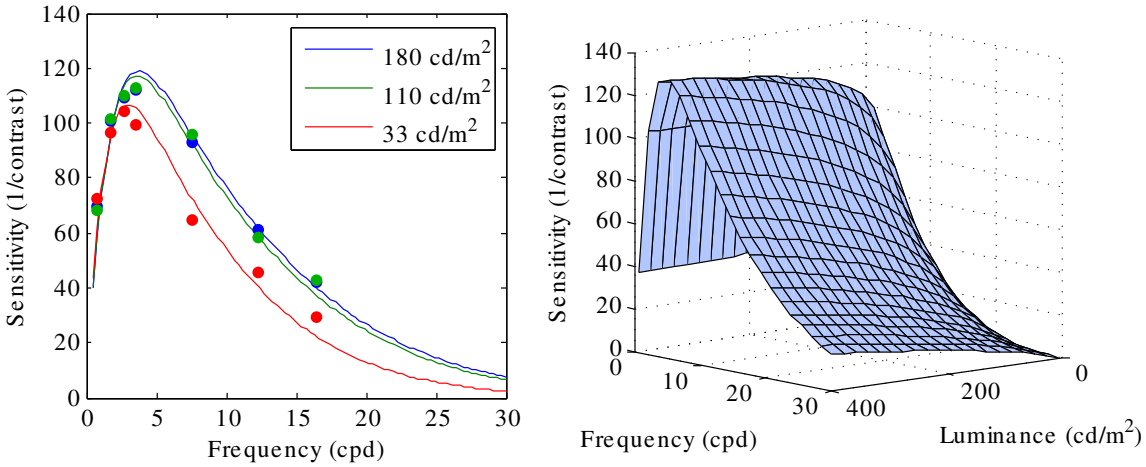


FIGURE 4.16: Left: plot of the subjects' mean sensitivity over each frequency and luminance level. Right: plot of the 3D fit of Barten's CSF model [Bar89].

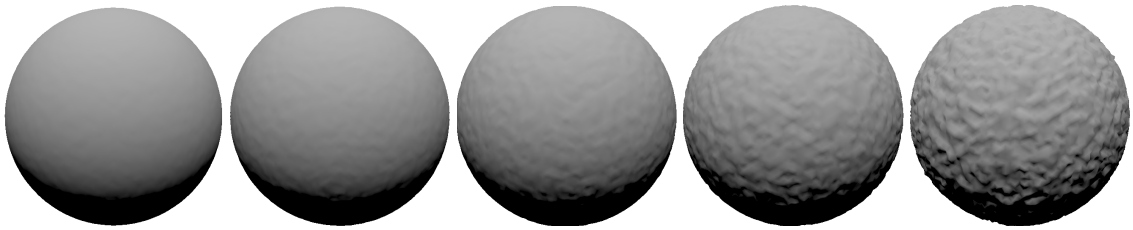
Results The results of these experiments are shown in Fig. 4.16. The plot shows a peak in sensitivity at around 3.5 cpd and a drop in sensitivity on either side of the peak. Additionally we can see that there is a decrease in sensitivity for low luminance while the sensitivity is relatively stable for a luminance level that is above 100 cd/m^2. The mean sensitivity over each frequency and luminance was then fitted to Barten's model [Bar89] which takes into consideration both the

frequency and the luminance and is defined by:

$$\begin{aligned} \text{csf}(f,l) = &\quad A(l) f e^{-B(l)f}\sqrt{1+ce^{B(l)f}}\,, \\ A(l) = &\quad a_0(1+0.7/l)^{a_1} \\ B(l) = &\quad b_0(1+100/l)^{b_1} \end{aligned} \tag{4.23}$$

with $a_0 = 125.42$, $a_1 = 0.09$, $b_0 = 0.343$, $b_1 = 0.17$ and $c = 0.19$.

4.3.4.3 Visual Masking



$r = 10^{-4}$ $r = 10^{-3}$ $r = 0.02$ $r = 0.1$ $r = 0.2$

FIGURE 4.17: Visual stimuli for measuring the visual masking threshold at different visual regularity levels.

Visual Stimulus Measuring the threshold relative to the visual masking aspect of the HVS requires a visual stimulus that exhibits a visible initial contrast (*i.e.*, above the CSF value) and certain visual regularity. We then gradually increase this initial contrast and measure the value needed to notice a change. Like previous experiments, we displace a series of vertices from a sphere approximated by a subdivided icosahedron. The icosahedron is subdivided 3 times which makes the contrast in each face visible for a human observer. By changing the light direction we can control the initial contrast value and by adding uniform noise to the sphere we can change its visual regularity (Fig. 4.17). We measure the masking threshold relative to 5 levels of visual regularity and 4 initial contrast values for each regularity level.

Results The results of these experiments are shown in Fig. 4.18. The plot shows the subjects' mean threshold for each of the visual regularity levels and initial contrast values. For the visible initial contrast whose normalized value is greater than 1 (normalization means multiplying by the corresponding CSF value, see Eq. (4.23)), the measured threshold lies close to an asymptote with a slope increasing with the value of r. This means the less the human visual system is capable of predicting the observed surface, the higher the slope of the asymptote. This result is consistent with the analysis of Daly [Dal93] which relates the value of the slope

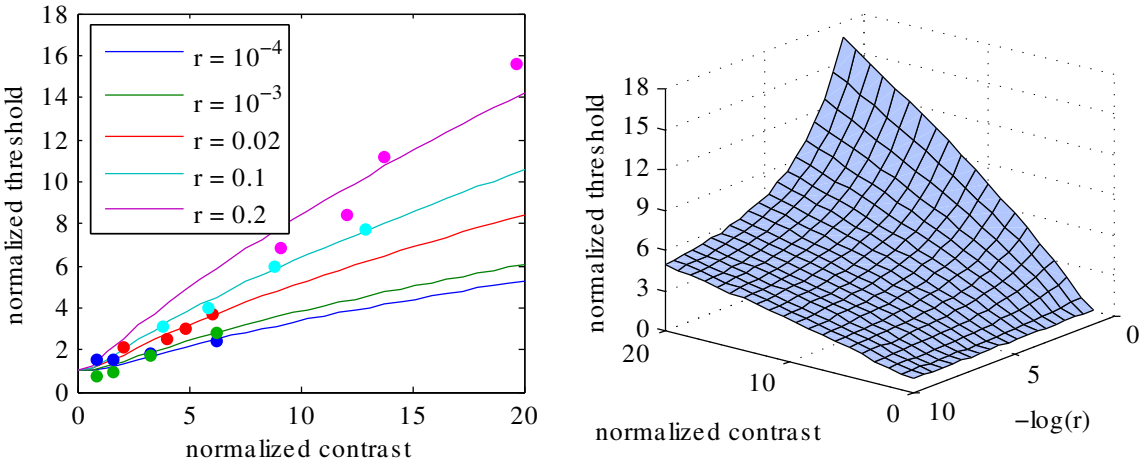


FIGURE 4.18: Left: plot of the subjects' mean threshold over each initial contrast value and visual regularity value. Right: plot of the 3D fit of the contrast masking model (Eq. (4.24)).

to the observer's familiarity with the observed stimulus. It also agrees with the observation that geometric distortions are more visible in smooth regions of the mesh than in rough ones [Lav09]. In order to take into consideration the visual regularity of a 3D mesh, we altered Daly's visual masking model by mapping the value of visual regularity to the value of the slope using an S-shaped psychometric function [Wei51], $s(r)$:

$$\begin{aligned} \text{masking}(\tilde{c}, r) &= \left(1 + \left(k_1 \cdot (k_2 \cdot \tilde{c})^{s(r)}\right)^b\right)^{1/b}, \\ s(r) &= (1-\delta) - (1-\gamma-\delta) \cdot e^{-10^{\beta(-\log(r)-\epsilon)}} \end{aligned} \tag{4.24}$$

with $\tilde{c}$ the normalized contrast, r the visual regularity and the fitted values $k_1 = 0.015$, $k_2 = 392.5$, $b = 4$, $\gamma = 0.63$, $\delta = -0.23$, $\beta = -0.12$ and $\epsilon = -3.5$.

4.3.5 Contrast Threshold

With results of these psychophysical experiments, we can compute the contrast threshold similarly to the case of flat-shaded rendering. However, since we have carried out a more precise threshold measurement in the case of smooth-shaded rendering mode by taking into account more parameters (luminance and visual regularity), the computed threshold T is now a function of four variables:

$$T(c, f, l, r) = \frac{\text{masking}(c \cdot \text{csf}(f, l), r)}{\text{csf}(f, l)}, \tag{4.25}$$

where c is the original value of the local contrast, f the local frequency and l and r correspond respectively to the global luminance of the scene and the regularity

of the visual pattern. Estimating the visibility of a change in local contrast is performed using Eq. (4.13). At first the change in contrast Δc (Eq. (4.14)) and the contrast threshold T (Eqs. (4.23) and (4.24)) are evaluated. They are then passed to a psychometric function which outputs the visibility probability.

4.4 Discussion

The results of this experimental study of the visibility threshold on 3D meshes and the relation between the defined local perceptual attributes (contrast, frequency and visual regularity) and the mesh properties (density, shape) can give us some interesting insights about the behavior of this threshold. First from the shape of the CSF and the relation between the mesh density and spatial frequency we can deduce that a 3D mesh will become more sensitive to geometric distortions when its density increases from a low value since it becomes easier for the human visual system to detect them. However, if the density of the model passes a certain value, then it becomes hard to detect the local geometric distortions because at high spatial frequencies the sensitivity of the visual system with respect to contrast, decreases. This result is in part inline with previous observations in computer graphics where it was noted that coarse mesh are better at hiding compression artifacts than dense ones [SCOT03]. This low sensitivity for coarse meshes will normally be located at curved coarse regions of a 3D mesh. This can be explained by two factors. The first is that the low frequency stimuli caused by the low density makes the human visual system less sensitive to contrast. The second is due to the curved shape of the surface which will potentially be reflected by high contrast values and thus create an important masking effect where a big change in contrast would be needed to notice the difference. In addition, our results justify the relation between mesh roughness and noise visibility on which most existing model-based methods are based. In fact, it is more likely to encounter complex visual pattern in rough regions which will increase the visibility threshold due to the increasing slope of the masking function.

Moreover, while we have used Barten's CSF mathematical model which is quite popular in the image/video processing communities, the sensitivity values and peak frequency positions that we have obtained are different from, for example, the ones computed with Daly's model [Dal93]. We think that is due to fact that models used in image-based methods are usually fitted using data from experiments where the visibility threshold is measured using a continuous sinusoidal signal. This is also in accordance with the observations of [CHM+12] in

which the authors' main objective is to test the efficiency of perceptual image-based techniques in the case of computer generated images. They concluded that image metrics are too sensitive for evaluating the visibility of computer graphics distortions. This is probably due to the difference in the types of visual artifacts caused by a geometric operation on the 3D mesh compared to the ones produced by an image processing operation.

The defined perceptual attributes and the experimental study have allowed us to propose a model whose goal is to compute the contrast visibility threshold. However, this threshold does not reflect the amount of displacement a vertex can tolerate but it rather indicates the maximum change in contrast a distortion can cause before it becomes visible. In the next chapter we will present an algorithm that will compute the vertex displacement threshold using the contrast visibility threshold model presented in this chapter.

Chapter 5

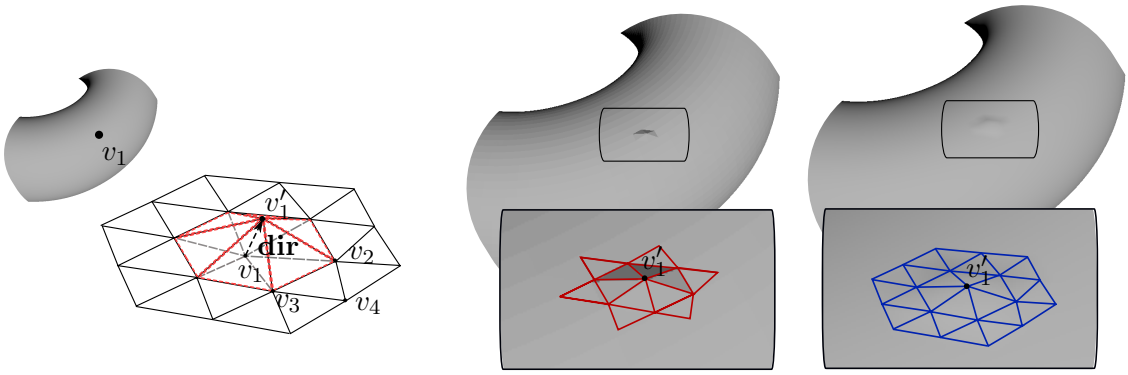
Just Noticeable Distortion Profile

The just noticeable distortion (JND) profile refers to the threshold beyond which a change in contrast becomes visible for the average observer [Lin06]. In the 3D setting, the JND is evaluated by computing the maximum displacement each vertex can tolerate. The displacement of a vertex in a given direction will probably cause a change in the normals of adjacent faces and a change in local density. This means that the displacement of a vertex probably alters the local perceptual properties (contrast, frequency) which will become visible at some point. In this chapter, we present a numerical method for computing the maximum displacement beyond which the local vertex distortion can be detected by the average human observer (Section 5.1). We then validate this computed threshold with a series of subjective experiments (Section 5.2)

5.1 Vertex Displacement Threshold

5.1.1 Visibility of a Vertex Displacement

The displacement of a vertex v_1 in a certain direction **dir** with a magnitude d alters the normals of the faces belonging to its one-ring neighborhood (Fig. 5.1). In addition, it will also cause a rotation in the normals of the vertices belonging to the one-ring neighborhood of v_1. As a result, if the 3D model is rendered in a flat-shaded mode, then this displacement is likely to alter the contrast and frequency of the surrounding pair of adjacent faces (Fig. 5.1.(b)). Similarly, in the case where the model is rendered with a smooth-shading algorithm, this vertex displacement also causes a change in contrast and frequency of faces that have at least one vertex in the 1-ring neighborhood of v_1 (Fig. 5.1.(c)). This alteration in the local perceptual attributes of a 3D mesh might cause the vertex displacement to be visible for a human observer.



(a) Vertex displacement (b) Flat shading (c) Smooth shading

FIGURE 5.1: (a) The displacement of a vertex v_1 in a direction **dir** causes the normals of the surrounding 1-ring faces (highlighted in red) to rotate. (b) In the case of a flat-shaded rendering, it will cause a change in contrast for surrounding pairs of faces *sharing a common edge* in 1-ring and 2-ring of the displaced vertex. (c) In the case of smooth-shaded rendering, it will cause a change in contrast in the faces having a vertex in the 1-ring neighborhood of v_1.

To estimate the visibility of a certain displacement, we start by evaluating the new normals of each of the effected faces. For instance, we express the new normal $\mathbf{n}'_1$ of the face $\{v_1, v_3, v_2\}$ (see Fig. 5.1.(a)) after displacing v_1 in a direction **dir** with a magnitude d by:

$$\begin{aligned} \tilde{\mathbf{n}}'_1 &= (v_1 - v_2) \times (v_3 - v_2) + d \cdot (\mathbf{dir} \times (v_3 - v_2)), \\ \mathbf{n}'_1 &= \frac{\tilde{\mathbf{n}}'_1}{\| \tilde{\mathbf{n}}'_1 \|}. \end{aligned} \tag{5.1}$$

Since none of the vertices of the second face $\{v_2, v_3, v_4\}$ in Fig 5.1.(a) is displaced, its normal direction does not change. We note that in the case of smooth shading, an additional step is required which consists of computing the new normals of the 1-ring vertices using the new normals of the 1-ring faces. Having computed the new normals we now evaluate the new contrast c' using the methods described in Sections 4.2.1 and 4.3.1 for respectively the set of affected face pairs in the case of flat and the set of affected faces in the case of smooth shaded rendering. The change in contrast Δc_i (Eq. (4.14)) along with the contrast threshold T_i is evaluated for each of the affected elements using respectively Eqs. (4.12) and (4.25) in the case of a flat or smooth shaded rendering. This change in contrast and contrast threshold are then passed to a probability function (Eq. (4.13)) which outputs the visibility likelihood for each affected element. The visibility of the displacement of v_1 in a direction **dir** with a magnitude d is then computed as:

$$\text{visibility}(v_1, \mathbf{dir}, d) = \max\{p(\Delta c_i, T_i)\}, \tag{5.2}$$

where $p(\Delta c_i, T_i)$ is the likelihood of detecting the change in contrast for the i^{th} affected pair of adjacent faces in the case of a flat shading or the i^{th} affected face in the case of smooth shading. Ultimately, the displacement of a vertex is considered visible when the change in contrast in at least one of the elements affected by this displacement becomes noticeable. In the following, we will describe an algorithm that allows us to efficiently find the threshold beyond which the displacement of a vertex becomes visible.

5.1.2 Evaluating the Vertex Displacement Threshold

In order to compute the threshold beyond which the displacement of a vertex v in a direction **dir** is visible, we proceed by the following steps. First, a list of the elements whose contrast is affected by the displacement is built (adjacent pairs of faces in the case of flat shading and faces having a vertex in the 1-ring neighborhood of v in the case of smooth shading). For each affected element, we start by computing its original perceptual properties (contrast, frequency and visual regularity) and the corresponding contrast threshold using Eq. (4.12) in the case of flat shading and Eq. (4.25) in a smooth shading setting. Then we gradually increase the displacement magnitude of v and compute its visibility as described in Section 5.1.1. Note that when the displacement causes a change in spatial frequencies (*e.g.*, in the case of a displacement in the local tangent plane of the vertex), we take into account the most sensitive frequency that results in a higher detection probability. Finally, the threshold is attributed to the magnitude where the vertex displacement visibility reaches a certain threshold. In practice we set the probability threshold to 0.95. To better understand this process, let us consider the two vertices v_1 and v_2 in Fig 5.2 in a flat shaded setting. Both vertices are displaced in their normal direction. The first vertex v_1 is situated on a rough region (initial contrast of all surrounding pairs of adjacent faces $>$ CSF threshold) and the second vertex v_2 on a smooth region (initial contrast $<$ CSF threshold). The displacement of v_1 and v_2 barely affects the spatial frequency of the surrounding face pairs as can be seen in the plots of the first row. The plots in the second row, show how displacing v_1 and v_2 in the normal direction affects the local contrast. The probability of detecting this change in contrast is shown in the plots in the third row. These plots show that v_2 is more sensitive and can tolerate less displacement than v_1. This is due to the different initial contrasts of the two vertices. The initial contrast around v_1 is above the CSF threshold. This implies that the visibility threshold is increased due to the masking effect, which explains the slow increase in detection probability. For v_2 all initial contrasts are below the CSF threshold. No masking should occur which means that once the contrast is above the CSF threshold the displacement should be visible. This is exactly

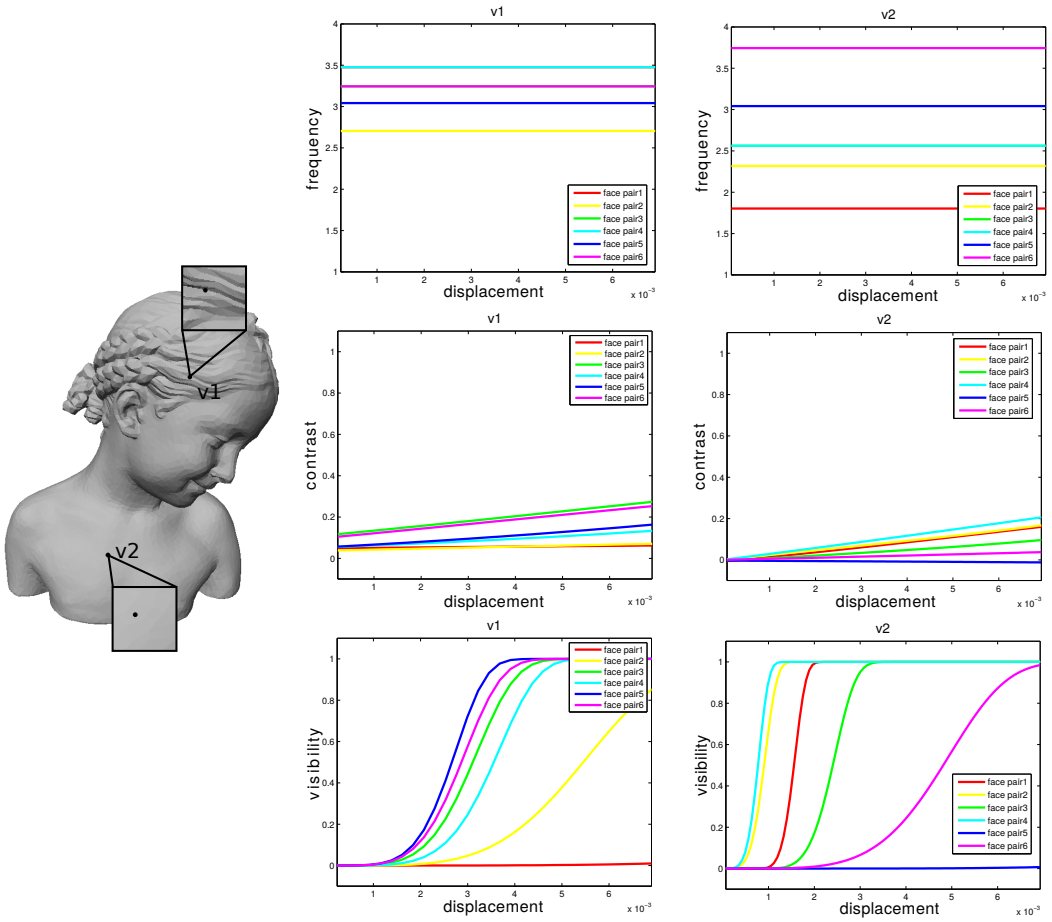


FIGURE 5.2: The evolution of the local perceptual properties and visibility, for two displaced vertices v_1 and v_2 on the Bimba model. Plots in the first row show the change in frequency, middle ones show the change in contrast and the bottom ones show the detection probability, of different pairs of affected adjacent faces of the two vertices. Note that some of the faces have the same spatial frequency, so the color curves overlap in the plots of the first row.

what we observe. When the contrast of *"face pair 4"* reaches the CSF level then the detection probability becomes close to 1. For the case of smooth shading, the exact same process is used to compute the displacement threshold. The only change in this case resides in the method to compute the local perceptual properties and the contrast threshold.

In the description above, we explain how to compute the displacement threshold by brute-force incremental step searching only for clarity purposes. However, it is important to note that as the vertex displacement increases, the contrast difference always increases as well. In addition, the psychometric function used to compute the probability to detect a change in contrast (Eq. (4.13)) is a monotone function. Therefore, in our implementation, we instead use a half-interval search algorithm to find the threshold (as described in Algorithm 1), which is simple yet

Algorithm 1: Half-interval search algorithm.

Data: **v**: vertex, **dir**: noise direction, **l**: light, th: visibility threshold, p: precision
Result: $dist$: displacement threshold

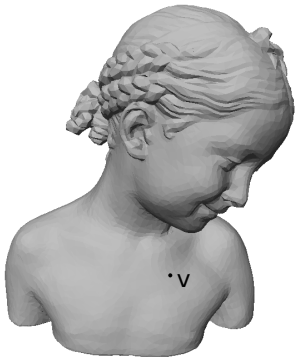
```
min = 0;
max = very_high_value;
dist = max;
visibility = compute_visibility(v, dir, l, dist);
while || visibility − th || > p do
    dist = (max - min) / 2 + min;
    visibility = compute_visibility(v, dir, l, dist);
    if visibility > th then
        max = dist;
    else
        min = dist;
    end
end
```

very fast and accurate. In our tests we have set the visibility threshold th to 0.95, the precision p to 0.005 and the parameter $very\_high\_value$ to $1/10th$ of the mesh bounding box.

Computing the displacement threshold requires an estimation of the spatial frequency, the local contrast and the visual regularity. This makes the computed displacement threshold capable of adapting to various display parameters. In particular, size and resolution of the display as well as the observer's distance to the screen and the model's distance to the viewpoint are inputs to the frequency estimation operation. In addition, the scene's illumination and the rendering mode affect the local contrast of the model. However, in an interactive setting where the light source is usually fixed relative to the viewpoint, the light direction varies with respect to the 3D mesh. It is therefore important to compute the displacement threshold independently of the light direction. We hereby propose a *light independent mode* for computing the displacement threshold.

5.1.2.1 Light Independent Mode

The algorithm presented in the previous section computes the vertex displacement threshold for a given light direction. However, it is useful to compute the displacement threshold independently from the light direction as in an interactive setting, the light direction usually varies with respect to the 3D mesh. To do so, we compute the threshold according to multiple light directions and then choose the smallest one. The light independent threshold can then be seen as the one corresponding to the worst possible illumination (*i.e.*, the light direction that makes the local vertex distortions the most visible).



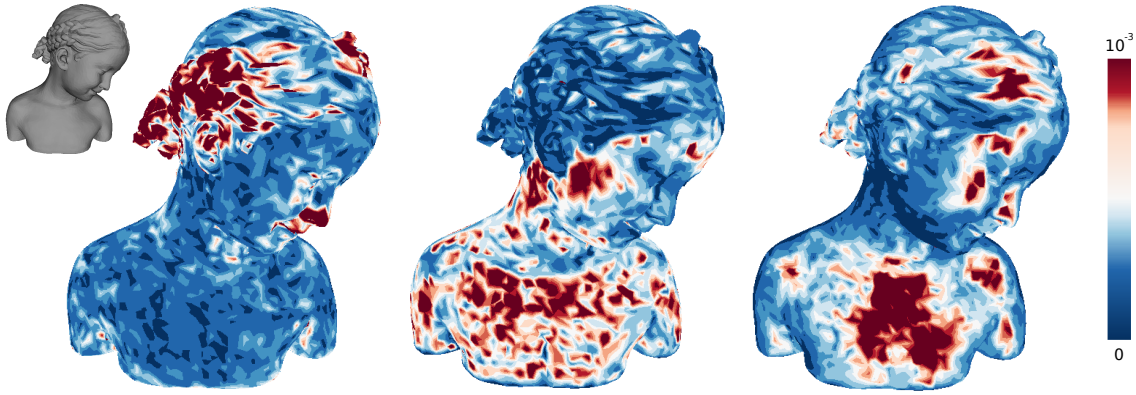
local light direction (α, θ)	vertex displacement threshold
$(10, 85)$	0.00195312
$(10, 55)$	0.003125
$(10, 25)$	0.0078125
$(100, 85)$	0.0015625
$(100, 55)$	0.00390625
$(100, 25)$	0.0069725
$(190, 85)$	0.00107422
$(190, 55)$	0.00234375
$(190, 25)$	0.0046875
$(280, 85)$	0.00107422
$(280, 55)$	0.0021875
$(280, 25)$	0.004375
$(0, 0)$	0.015625

FIGURE 5.3: The vertex displacement threshold in the normal direction of a vertex v computed for different light directions.

The set of all possible light directions belongs to the local sphere around a vertex. However, the local contrast is well defined when the dot product between the light direction and the normals is positive since otherwise its value is 0. This means that the set of all possible lights can be reduced to the local half sphere in the direction of the unit normal. Furthermore, we can intuitively say that the local contrast increases as the light direction gets close to the local tangent plane. This means that if the light direction is close to the base of the local half sphere, then a small displacement of a vertex will most likely cause a big change in contrast which might be visible. On the contrary, if the light direction is close to the normal direction of the displaced vertex, then even a big vertex displacement can only cause a small change in contrast which will make the displacement threshold value high. Figure 5.3 shows the vertex displacement threshold obtained from different light directions belonging to the half sphere of a vertex v. As previously explained, we notice that as the light direction approaches the base of the half sphere, the threshold gets smaller. This implies that the worst possible illumination is most of the time found near the base of the half sphere. Therefore, it is actually not necessary to densely sample the half sphere in order to obtain an accurate solution. It will be more efficient to concentrate the light samples near the base of the local half sphere.

5.1.3 Results

By computing the vertex displacement threshold relative to a certain direction for each vertex of a 3D mesh, we obtain the just noticeable distortion profile. Figure



(a) vertex displacement threshold for a displacement in the normal direction in a light-independent mode

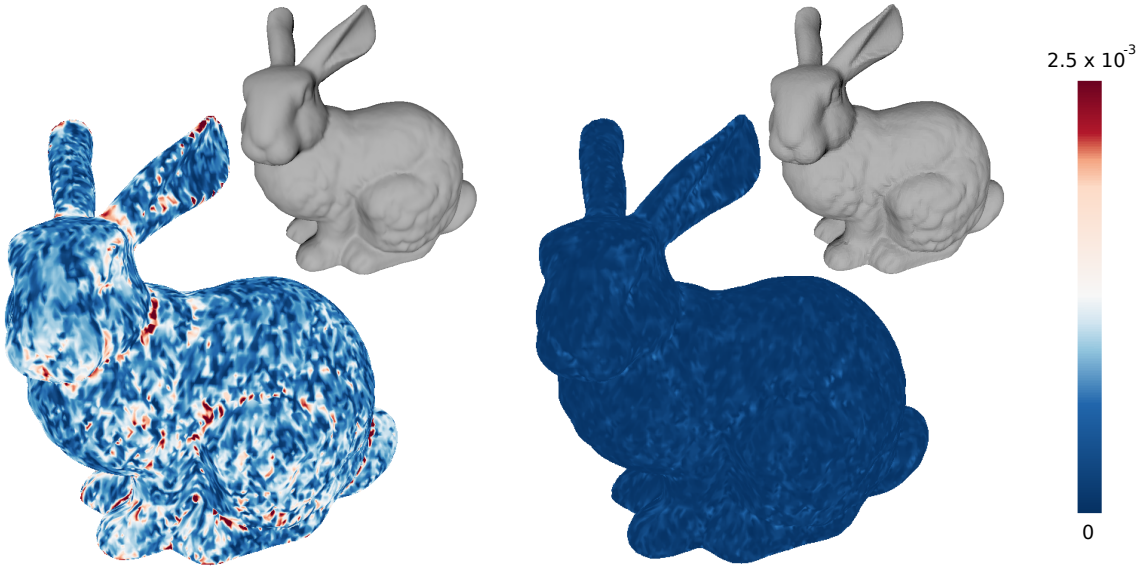
(b) vertex displacement threshold for a displacement in the tangent direction in a light-independent mode

(c) vertex displacement threshold for a displacement in the normal direction in a light-dependent mode

FIGURE 5.4: The vertex displacement profile for the Bimba model under different circumstances. (a) and (b) show color map representing the displacement threshold in a light-independent mode with respect to a displacement in the normal direction and tangent plane respectively. (c) shows the displacement threshold according to a displacement in the normal direction in a light dependant mode.

5.4 shows the JND profile for a mesh under various circumstances. Figure 5.4.(a) displays the JND profile relative to a displacement in the normal direction in a light independent mode. Due to the effects of contrast masking, the rough region of the model can tolerate more noise than the smooth part. This is not the case when the JND is computed relative to a displacement in the tangent direction (Fig. 5.4.(b)) where the smooth part can tolerate more displacement. This is because a displacement in the tangent plane for vertex in a smooth region will barely alter the normal of the surrounding faces and thus the local contrast will not be affected by the displacement, leading to a higher displacement threshold than in a rough region. We note here that in the case of a displacement in the tangent plane we make sure that the computed displacement threshold does not cause an auto-intersection. Figure 5.4.(c) shows the JND profile relative to a displacement in the normal direction when the light source is fixed at the viewpoint. As expected, we can see that the obtained threshold is maximal when the surface normals are in the same direction of the light as the contrast will increase slower compared to when the light direction is close to the local tangent plane.

Figure 5.5 presents side by side the JND profile for the Bunny model in a smooth shading mode and a flat shading mode. In general we have observed that the displacement threshold relative to a smooth shading rendering is 5 to 10 times bigger than the one relative to a flat shading mode. This difference is due to the way how the surface normals and contrast are computed in each mode. In a flat shading mode, the contrast is computed using the normals of a pair of faces,



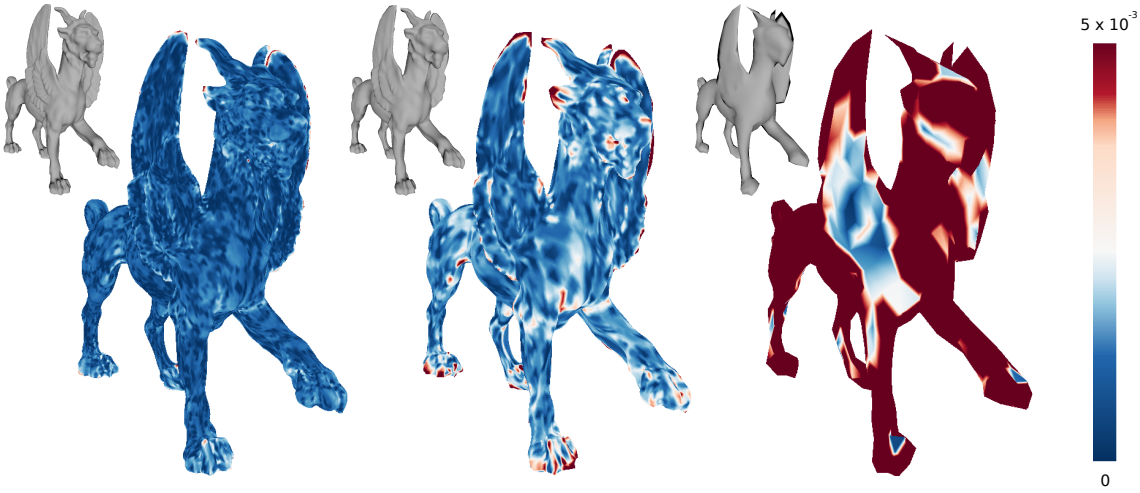
(a) vertex displacement threshold in the case of a smooth rendering mode

(b) vertex displacement threshold in the case of a flat rendering mode

FIGURE 5.5: The vertex displacement profile for the Bunny model in a smooth shading mode (a) and flat shading mode (b).

while in a smooth shading the contrast is evaluated using the vertex normals within a triangular face which reflects a smooth shading rendering. In general, the displacement of a vertex causes a bigger rotation in the normal direction of the faces adjacent to the displaced vertex in a flat shaded rendering, compared to the rotation of normal direction of the vertices in the 1-ring neighborhood computed in a smooth shaded rendering. In consequence, for the same displacement magnitude, the change in contrast is higher in a flat shading mode compared to a smooth shading mode which explains the lower visibility threshold for the former.

Figure 5.6 compares the vertex displacement threshold profile when the resolution of the 3D model changes. In general, as the density of a triangular mesh decreases the magnitude of the displacement a vertex that can tolerate, increases. In a dense mesh, the triangular faces are small compared to the ones belonging to a coarse mesh. As a result, rotation of the normals relative to a vertex displacement of the same magnitude is higher for a dense mesh. This means that, in general, as the displacement increases, the contrast varies slower in a low resolution mesh. In addition, the difference in density also affects the spatial frequency of the visual stimulus which further affects the visibility of a vertex displacement. As it can be deduced from the contrast sensitivity properties of the visual system that are modeled by the CSF, when the density of a 3D mesh increases from a low value it becomes easier for the visual system to detect the change in contrast

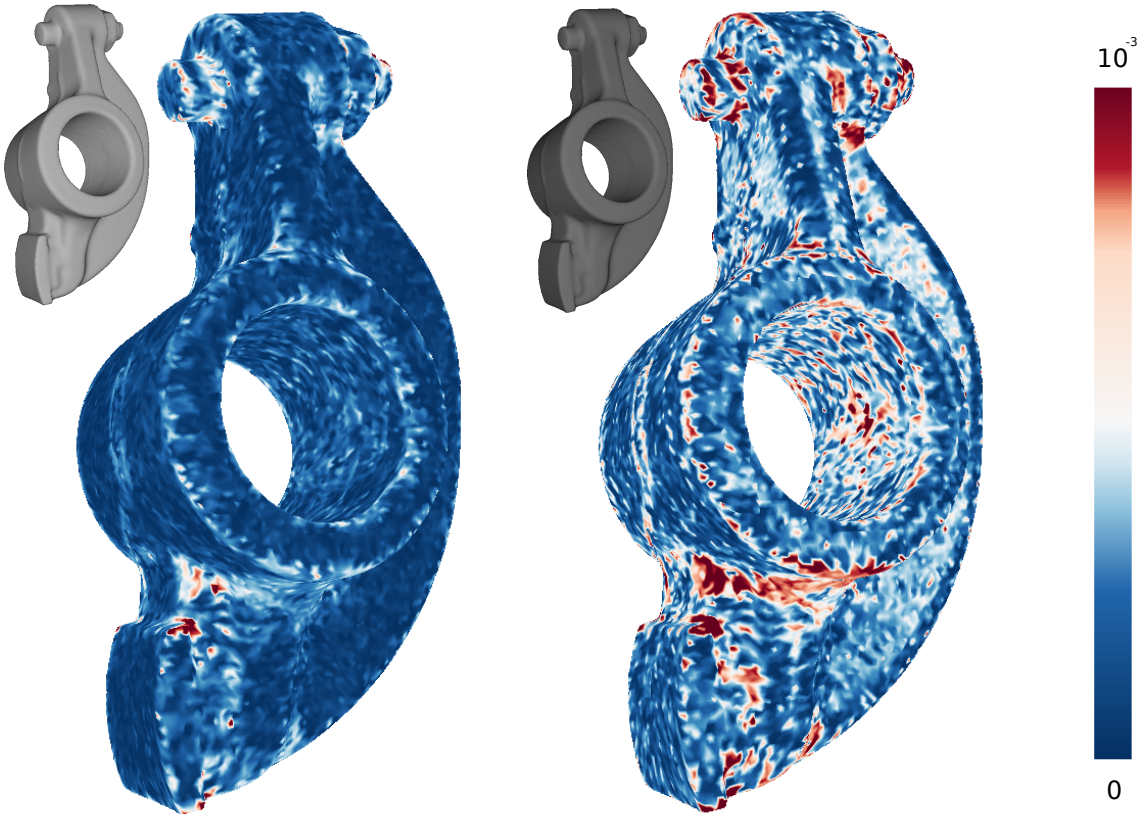


(a) Vertex displacement threshold for the 65k version of the Feline model

(b) Vertex displacement threshold for the 16k version of the Feline model

(c) Vertex displacement threshold for the 1k version of the Feline model

FIGURE 5.6: The vertex displacement threshold computed with the proposed algorithm is capable of adapting to the resolution of the 3D model.



(a) vertex displacement threshold with high energy light

(b) vertex displacement threshold with low energy light

FIGURE 5.7: In a smooth shading mode, the proposed threshold model takes into consideration the energy of the light illuminating the scene. For a point light source whose energy decreases proportionally to the square distance to the object, the scene becomes darker as the light source becomes farther. This increases the value of the displacement threshold.

caused by the displacement of a vertex. Both of these reasons make the vertex displacement threshold in high resolution models, in general, smaller than the threshold in low resolution models.

Finally in Fig. 5.7 we show how the intensity of the scene's illumination can affect the vertex displacement threshold; in the case where a 3D scene is illuminated with a point light whose energy decreases proportionally to the distance between the light source and the illuminated object. So if an object is far from the point light source (Fig. 5.7.(b)), the global luminance around that object is reduced, which causes an increase of the magnitude of the vertex displacement threshold. This boost in the value of the visibility threshold is mainly due to the CSF, which describes a reduction in contrast sensitivity when the global luminance is low.

5.1.4 Performance Analysis

Here we present some information about the theoretical and practical results concerning accuracy and the execution time of the vertex displacement threshold computation.

5.1.4.1 Threshold Accuracy

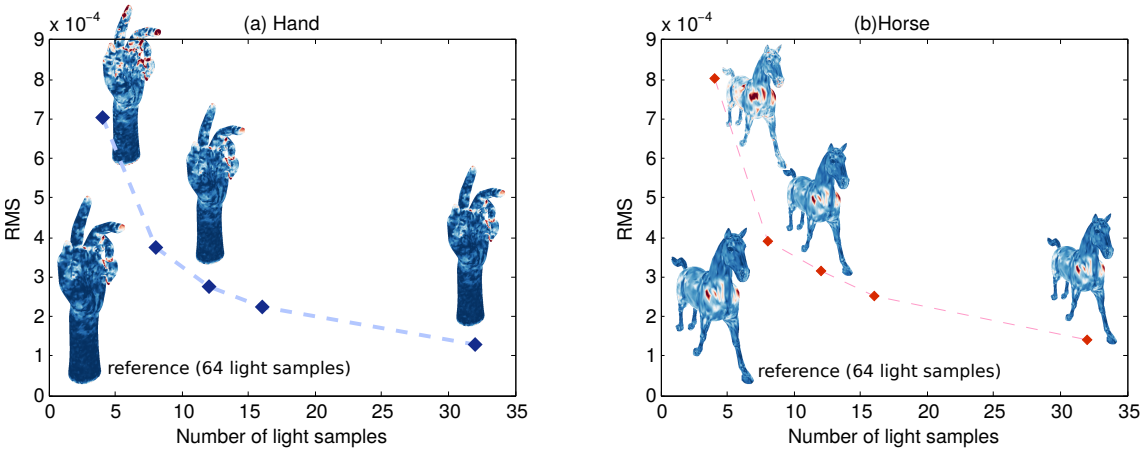


FIGURE 5.8: Effects of the number of light samples on the accuracy of the vertex displacement threshold in a light independent mode.

In a light-independent mode, we compute the vertex displacement threshold relative to several light directions sampled from a local half sphere around a vertex. We have observed that the algorithm begins to converge to an accurate displacement threshold value with 8 samples as it can be seen in Fig. 5.8, where the root mean square error (RMS), computed with regard to the displacement

threshold obtained with 64 light direction samples (shown in the leftmost corner of each graph, starts to stabilize beyond this point. In practice, we have used the 12-points sampling, similar to the one in Fig. 5.3 (excluding the point $(0, 0)$), which ensures a very good trade-off between threshold accuracy and algorithm speed according to our tests.

5.1.4.2 Theoretical Computational Complexity

A theoretical analysis of the proposed algorithm for computing the vertex displacement threshold shows that the complexity of computing the light independent-mode for one vertex is equivalent to:

$$O\left(L \times \log\left(\frac{x_{max}}{x_{precision}}\right)\right), \tag{5.3}$$

where L is the number of light samples and x_{max} and $x_{precision}$ are respectively the upper displacement bound and the precision used in the half-interval search algorithm (Algorithm 1 of this chapter). This means that the complexity for computing this threshold for an entire mesh is:

$$O\left(V \times L \times \log\left(\frac{x_{max}}{x_{precision}}\right)\right), \tag{5.4}$$

where V is the number of vertices. This shows that the execution time increases linearly with the number of vertices at a rate relative to the number of light samples and the precision of the search procedure.

5.1.4.3 Execution Time

Having adopted a half-interval search algorithm makes finding the JND threshold a very efficient operation. On average computing the vertex displacement threshold for a vertex in the light independent mode takes about 7×10^{-4} s. We have used an HP EliteBook 8570w with an i7-37400QM cpu (4 cores / 8 threads) and 16GB of RAM in our computation. As suggested by Eq. (5.4), when the number of vertices or light samples increases, we have observed that the execution time increases approximately in a linear way (Fig. 5.9). Figure 5.9 also shows that the computation of the vertex displacement threshold for a flat-shaded rendering is faster than a smooth-shaded one. For instance, for a model with approximately 200k vertices the computation of the vertex displacement threshold on the entire mesh took about 50s in a flat shading mode and about 95s in a smooth shading mode. This difference in computation time is due to the different normal update after each displacement iteration. In the case of a flat shading mode, the update

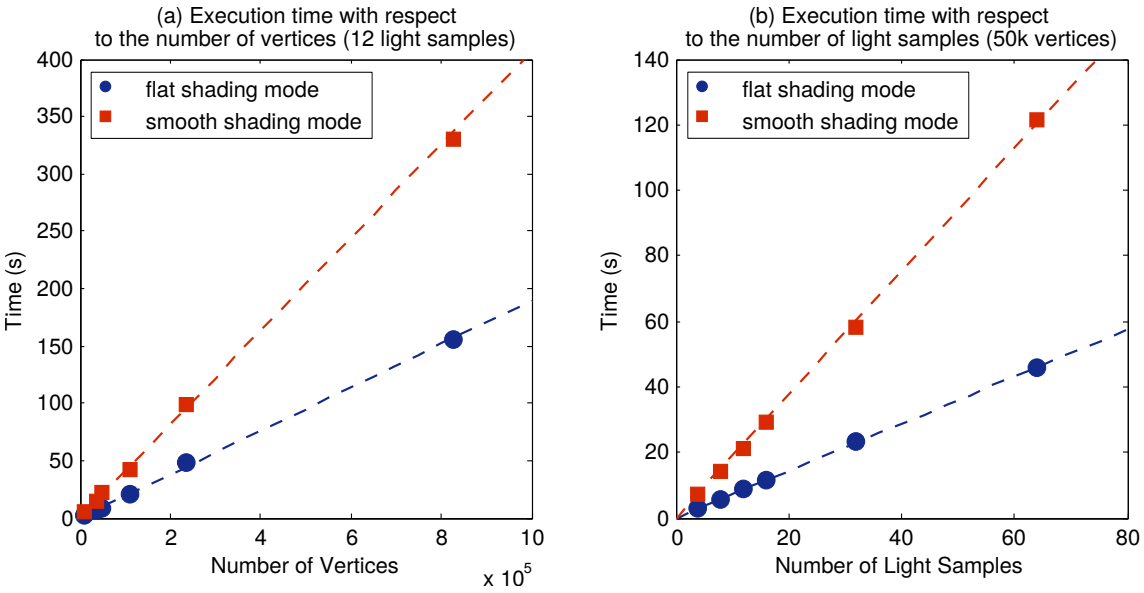


FIGURE 5.9: Vertex displacement threshold execution time.

of the normal direction of the 1-ring adjacent faces is straightforward, using Eq. (5.1). However, in a smooth shading mode, the normal update operation requires an additional step which consists of evaluating the normals of the vertices in the 1-ring neighborhood. In addition, since the computed displacement threshold of a vertex is independent from the threshold of other vertices then this operation can be computed on the entire mesh in a parallel way (with respect to each vertex). Using OpenMP and 8 threads, the algorithm performs about three to five times faster. For a model with 237k vertices, the vertex displacement threshold can be computed in about 18s instead of 52s.

5.2 Subjective Validation

In order to test the performance of a Just Noticeable Distortion profile, it is common in the image or video JND context to perform a subjective experiment where the visibility of a JND modulated random noise added to a series of images or videos is rated by several subjects [LLP+10, ZCZ+11, WSL+13]. A JND model should be able to maximize the amount of noise injected into the image or video while keeping it invisible; the best JND model being the one that is able to add the largest amount of invisible noise. We have conducted a series of subjective experiments where we have tested the performance of the proposed JND model in the flat and smooth shaded settings. We have compared the visibility and the quantity of vertex noise on altered 3D meshes, which were obtained by modulating the vertex noise in three different ways. The three types of noise modulation are:

- uniform random noise, *i.e.*, without any modulation;
- random noise modulated by the surface roughness;
- random noise modulated by the proposed JND model.

Surface roughness is an important candidate to test our JND model against since it is accepted in the computer graphics community that noise is less visible in rough regions [Lav09]. As discussed in the previous chapter, this observation can also be justified through the results of our experimental study.

5.2.1 Mesh Alteration

We injected noise into 3D meshes according to the following equation:

$$v_i^{'} = v_i + \mathrm{rnd} \times M\left(v_i\right) \times \mathbf{dir}_i, \tag{5.5}$$

where v_i is the i^{th} vertex of the initial mesh and $v_i^{'}$ is the corresponding noisy vertex. **dir** is the noise direction. rnd is a random value equal to either $+1$ or -1 and $M\left(v_i\right)$ represents the magnitude of the noise for v_i. It is defined as:

$$M\left(v_i\right) = \begin{cases} \beta_{unif} & \text{uniform noise,} \\ \beta_{rough} \times \mathrm{lr}\left(v_i\right) & \text{roughness modulated noise,} \\ \beta_{jnd} \times \mathrm{jnd}\left(v_i\right) & \text{JND modulated noise,} \end{cases} \tag{5.6}$$

where β_{unif}, β_{rough} and β_{jnd} regulate the global noise energy for each of the noise injection methods. $\mathrm{lr}\left(v_i\right)$ is the local surface roughness as defined in [WTM12] and $\mathrm{jnd}\left(v_i\right)$ is the JND value computed as explained in Section 5.1. In order to allow user interaction during the experiments, the JND value was computed independently from any light direction. It is important to note that in the case of a JND modulated noise, the value β_{jnd} of the global noise energy value should have some implication on the visibility of the injected noise. Therefore, if $\beta_{jnd} > 1$ the injected noise onto the 3D mesh should be visible, while if $\beta_{jnd} < 1$ the noise should remain undetected by the human observer. The following mesh models were used in our experiments:

- *Bimba* is a coarse model with a smooth and a rough part;
- *Horse* is a smooth model with a varying vertex density. Its head is densely sampled while its body is coarse.
- *Lion-vase* is a dense model with a mix of smooth and rough surfaces.
- *Venus* is a dense model with a smooth surface.
- *Dinosaur* is a dense model with a rough surface.

5.2.2 Validation of Flat-Shaded Vertex Displacement Threshold

Inspired by the literature on validation of image JND profiles, we have performed two subjective experiments whose goal is to test the accuracy of the computed vertex displacement threshold in the flat-shaded setting.

5.2.2.1 Experiment 1

The goal of the first experiment is to rate the visibility of noise on 3D models according to several noise modulation types and two global energy levels ($\beta_{jnd} = 1$ and $\beta_{jnd} = 2$). These levels correspond to a near-threshold noise and to a supra-threshold noise, respectively. For $\beta_{jnd} = 1$ the injected noise is supposed to be difficult to notice while for $\beta_{jnd} = 2$ the noise is expected to be visible. We then fix β_{unif} and β_{rough} such that for the meshes altered using our JND model, the maximum root mean square error (MRMS) [CRS98, ASCE02], a widely used purely geometric distance, is the biggest for each noise level. Indeed, the objective here is to show that our JND model is able to inject the highest amount of noise onto the mesh among the three methods, while producing the least visible one. In addition, we tested the performance of the JND model for noise in a random direction for each vertex and that in the normal direction for each vertex. To see the effects of light direction we ran the experiment twice: once with the light source in front of the model and another time with the light on top left of the model.

Experimental Protocol The first subjective experiment followed the *adjectival categorical judgement method* [Int12]. This procedure consists of displaying two 3D meshes side by side, the reference on the left and the noisy one on the right. The participants were asked to rate the visibility of the noise on a discrete scale from 0 to 5, 0 being the score attributed when the noise cannot be seen and 5 when the noise is clearly visible. 5 "dummy" models were included at the beginning of each session to stabilize subjective scores. The models were presented in a randomized order. To avoid any memory-based bias, two meshes derived from the same reference model were never displayed consecutively.

The experiment was conducted in a low illuminated environment. We used a 23-inch Asus screen with a 1920×1080 resolution to display the 3D models. The participants viewed the models from a distance of 50 cm. During the experiment, the two displayed meshes had a synchronized viewpoint and subjects could freely rotate around the displayed meshes. To encourage close examination of the displayed mesh, no score could be registered before 10 seconds of interaction occur. The initial viewpoint was manually set for all models. The light source

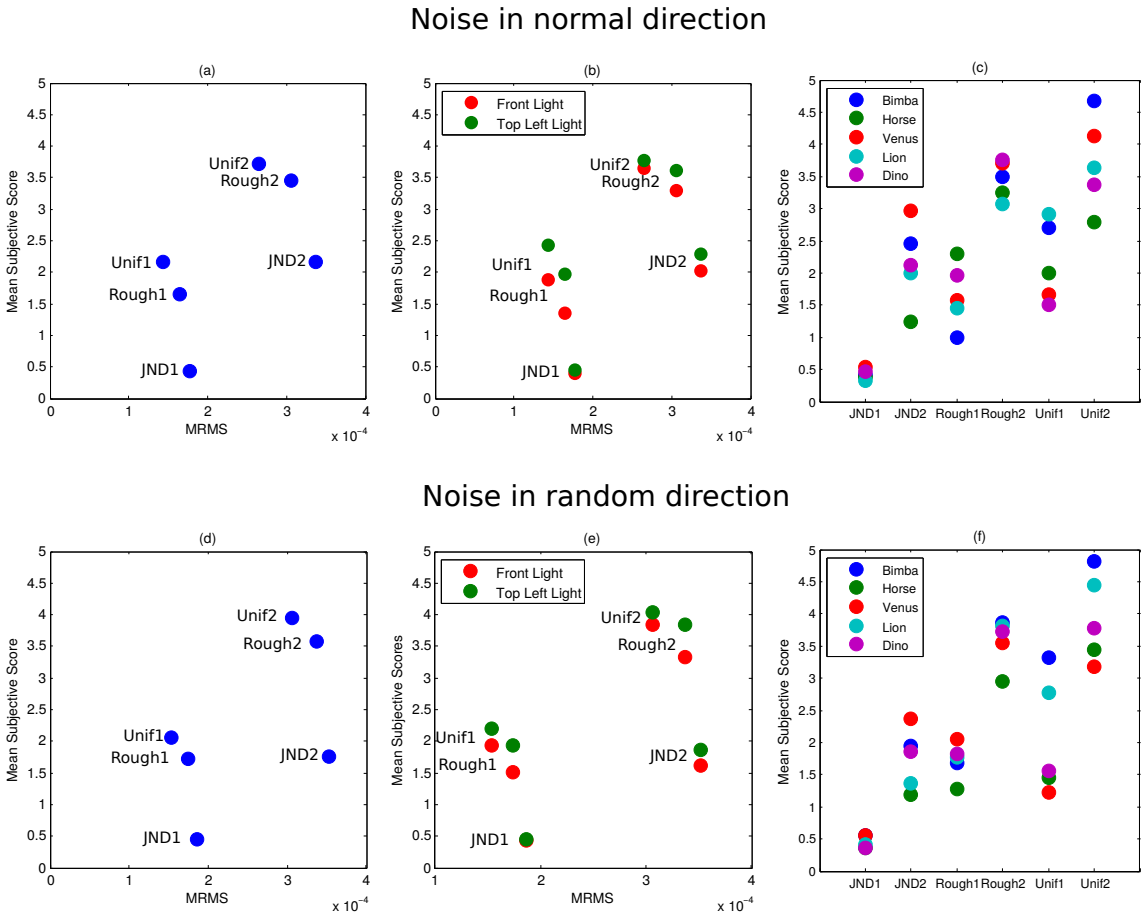


FIGURE 5.10: Mean subjective score values versus MRMS distance values. Plots (a) and (d) present, for different noise injections, the mean subjective scores over all test models and the two illumination settings. Plots (b) and (e) show the difference in mean subjective scores between the experiments in the two illumination settings. Plots (c) and (f) compare the mean subjective scores for the different models used in the experiments.

was fixed with reference to the camera position. A front and a top-left light directions were used. 12 subjects participated in these experiments. All of them had normal or corrected-to-normal vision and were between the age of 20 and 29.

Results After collecting the subjective scores, we have computed the mean score over each of the noise types. "JND 1" and "JND 2" refer to the models obtained by modulating the random noise with our JND model for near-threshold and supra-threshold levels, respectively. "Rough 1" and "Rough 2" refer to the ones obtained using the surface roughness measure and "Unif 1" and "Unif 2" to the ones with uniform random noise. Figure 5.10 displays the results of the subjective experiments. Plots (a) to (c) present the results for the noise in the normal direction and plots (d) to (e) the results for the noise in a random direction. Figures 5.10(a) and 5.10(d) show that the noise on the "JND 1" models was indeed difficult to detect

as the mean subjective score is about 0.45. Interestingly, the participants rated "Unif 1" and "Rough 1" models similarly to "JND 2" which refers to the supra-threshold noise level models that contain approximately twice the noise of "Unif 1" and "Rough 1". Plots (b) and (e) also show that "JND 1" models were perceived almost identically under both front and top-left illumination conditions. This is not the case for "Unif 1" and "Rough 1" models where the grazing light direction of the top-left illumination made the noise more apparent. It is also important to note that the visibility of the noise for "JND 1" models was almost identical for all models. This is not the case for "Rough 1" and "Unif 1" where the visibility of noise varied a lot for different models (see Figs. 5.10(c) and 5.10(f)). This is mainly due to the difference in mesh density between the models; high density models are in general more sensitive to noise than low density ones.

These results show that the proposed JND model is indeed able to add the largest amount of invisible noise onto the mesh surface among the three methods. Furthermore, the proposed JND model can accurately predict the visibility threshold for 3D meshes, taking into account the noise direction, the mesh characteristics and the scene illumination. However, the proposed model cannot accurately describe how the supra-threshold noise visibility (or annoyance) is perceived since it has not been designed for this purpose; the noise was perceived differently for each model in "JND 2" (Figs. 5.10(c) and (f)).

5.2.2.2 Experiment 2

The first experiment showed that the models with a JND modulated noise were rated the lowest on the visibility scale and could tolerate the biggest amount of distortions. In the following experiment we measure the global noise energy threshold beyond which the injected noise becomes visible for a 3D model. The idea behind this experiment is to find the minimum noise intensity (β_{unif}, β_{rough} and β_{jnd}) starting from which the participants notice the noise in the model and then compare their respective MRMS value. The JND modulated noise should have the highest amount of geometric distortion which is an indication that the proposed JND model is capable of effectively hiding a large amount of noise.

Experimental Protocol For this second experiment, we have adapted the same experimental procedure that we have used to measure the local contrast threshold in the studies of contrast sensitivity and visual masking (see Chapter 4). Two models were displayed on the screen, one of which has noise injected. The subjects had to answer by either *Yes* or *No* whether they saw the noise on one of the model. The intensity of the noise (β_{unif}, β_{rough} and β_{jnd}) is then adjusted using the

Models	Lion	Bimba	Horse	Dino	Venus
β_{jnd} (flat shading)	0.95	0.82	0.93	1.35	0.88

TABLE 5.1: Global noise energy value relative to JND modulated noise (β_{jnd}) in a flat shading setting.

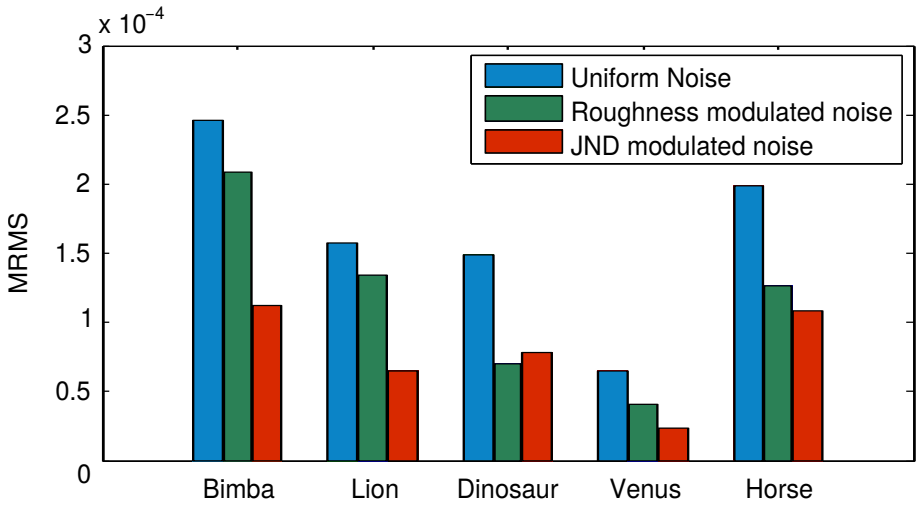


FIGURE 5.11: Plot of the MRMS induced by noise injection for three different types of noise at the same visibility level (under flat shading).

QUEST procedure [WP83]. The subjects were allowed to interact with the displayed models by rotating the camera around them. 5 new subjects participated in the experiment.

Results Table 5.1 and Fig. 5.11 display the results of this subjective experiment. Table 5.1 shows the mean measured intensity required to make JND modulated noise visible on a 3D mesh. We see that the measured β_{jnd} is close to 1 for all of the models, meaning that the proposed JND profile is able to accurately detect the threshold beyond which a noise is visible. Figure 5.11 shows that the MRMS value of the mesh model with JND modulated noise of just noticeable level is higher than those of the corresponding models with uniform noise or roughness modulated noise at the same visibility level. This means that the JND model is able to tolerate the highest amount of noise among the three candidates, which is what we expected.

5.2.3 Validation of Smooth-Shaded Vertex Displacement Threshold

In order to test the accuracy of the computed vertex displacement threshold in the smooth-shading setting, we have performed a subjective experiment similar to the second one for the flat-shading setting presented in Section 5.2.2.2. The goal

Models	Lion	Bimba	Horse	Dino	Venus
β_{jnd} (directional light)	0.87	0.91	1.06	1.15	1.09
β_{jnd} (point light)	0.91	0.89	0.97	1.11	1.05

TABLE 5.2: Global noise energy value relative to JND modulated noise (β_{jnd}) in a smooth shading setting.

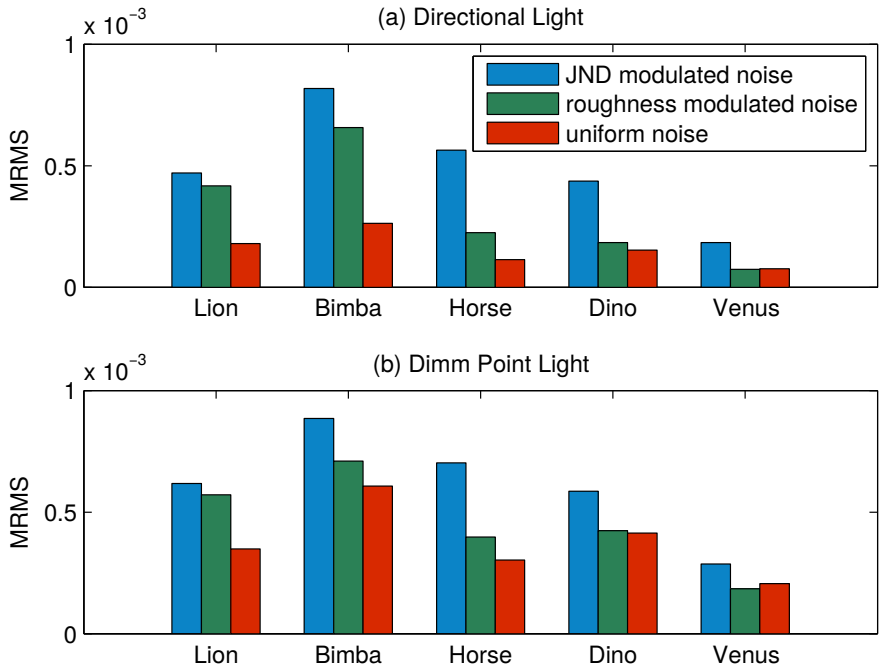


FIGURE 5.12: Plot of the MRMS induced by noise injection for three different types of noise at the same visibility level (under smooth shading).

here is to measure the global noise energy threshold beyond which the injected noise becomes visible on the 3D model.

Experimental Protocol We have followed the same protocol as described in Section 5.2.2.2. 12 new subjects participated in the experiment. The experiment was carried out with two lighting conditions: front directional light and front point light whose energy decreases proportionally to the square of the distance to the model.

Results Table 5.2 and Fig. 5.12 show the results of this subjective experiment. It is important to note that the value of the global noise intensity relative to the JND (β_{jnd}) is on average close to 1 for both types of illumination which indicates that the perceptual model was able to accurately predict the vertex displacement and adapt to the change in illumination conditions. Additionally, plots (a) and (b) in

Fig. 5.12 show the MRMS (maximum root mean square error [CRS98]) value of each model for the three noise intensity types. In all cases, the JND modulated models have the highest MRMS value indicating that our perceptual model is able to inject the highest amount of tolerable noise into the meshes. In addition, we point out that the models illuminated with the point light can tolerate more noise than the ones illuminated with the high energy directional one since the far distance of the point light will reduce the global luminance of the scene and thus reduces the sensitivity to contrast. We also note that the MRMS value of the models relative to the measured global energy threshold is higher when using a smooth-shaded rendering mode than under a flat-shaded rendering mode. The reason is that the change in contrast relative to the displacement magnitude increases slower for smooth-shaded surfaces. On average we have observed that smooth-shaded surfaces can tolerate 5 to 10 times (depending on the mesh's properties) more displacement noise than flat-shaded ones.

5.2.4 Further Comparisons and Examples

The series of subjective experiments that we have performed have shown that the proposed vertex displacement threshold algorithm is capable of accurately finding the maximum displacement magnitude a vertex can tolerate. This allows our perceptual JND model to inject the largest amount of tolerable noise onto a 3D mesh compared to modulating the vertex displacement magnitude with surface roughness or having a uniform displacement magnitude. The main advantage that our proposed JND model has over surface roughness measures is that it adapts to the mesh characteristics (density, size), the noise direction and the scene illumination. Figure 5.13 illustrates the visibility of vertex noise for three versions of a 3D model with the same RMS value and injected with respectively a JND modulated noise, a roughness modulated noise and a uniform noise in a smooth shaded setting and flat shaded setting. In the case of the Venus model (Fig. 5.13.(a)), a roughness modulated noise will concentrate the vertex distortions in the rough parts of the model and neglects its smooth parts, while a uniform noise will blindly displace all the vertices of the model with the same amount making it visible in the smooth areas at first. On the other hand the JND model will take advantage of all the vertices of the 3D mesh by adapting the magnitude of a vertex displacement to the local perceptual properties of this vertex and thus allowing a greater quantity of invisible noise to be added. The Horse (Fig. 5.13.(b)) is a model with mostly smooth regions, the rough regions are packed in the head's features. In addition, the head is densely sampled while the body is coarsely sampled. The JND model avoids adding noise in the dense head and takes advantage of the coarse body, while surface roughness measures

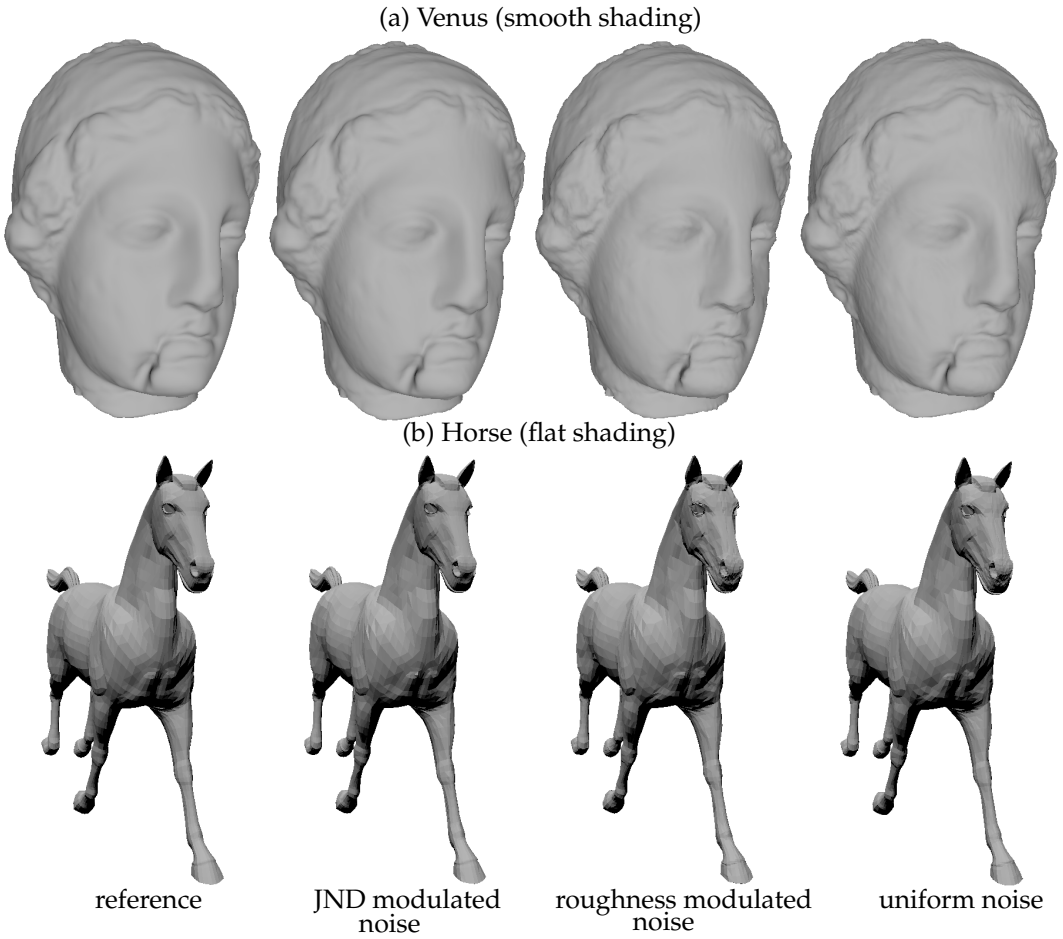
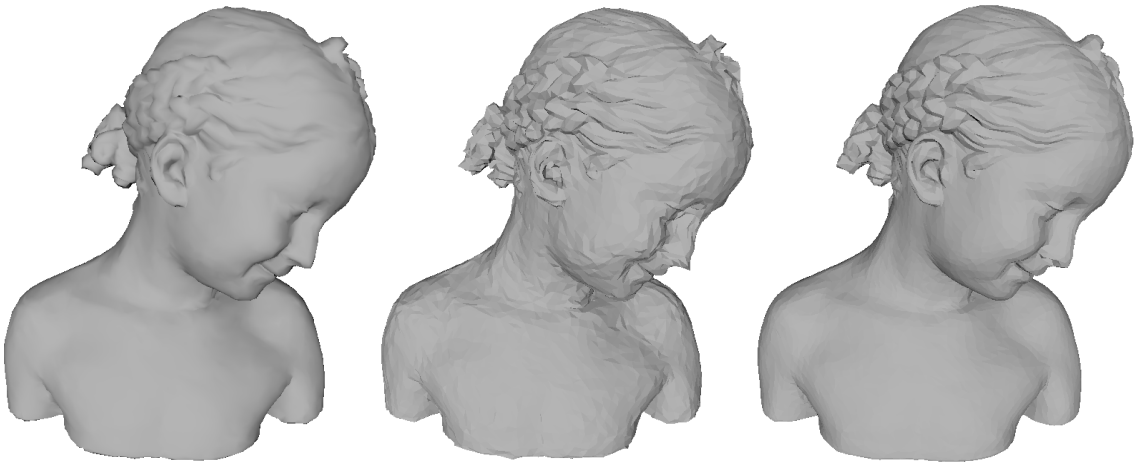


FIGURE 5.13: The visibility of vertex noise in three models having the same RMS value and injected with respectively a JND modulated noise, roughness modulated noise and uniform noise in smooth (for Venus) and flat shading (for Horse).

are not able to detect the difference in sampling. The noise is thus rather injected in the dense head features, which makes it visible. By changing the perceptual model used to compute the contrast threshold (Eqs. (4.12) and (4.25)) the computed vertex displacement threshold can adapt to either a flat shaded rendering or a smooth shaded rendering. For instance in Fig. 5.14, the vertex displacement threshold computed for smooth shaded mode will cause the JND modulated noise to be extremely visible if rendered with a flat shading algorithm. As we mentioned earlier, a 3D model rendered with a smooth shading method is capable of tolerating up to 10 times the amount of noise compared to a flat shading rendering.



(a) (b) (c)

FIGURE 5.14: (a) JND modulated noise computed for a smooth shaded mode added to the Bimba model and rendered in smooth shading. (b) JND modulated noise computed for a smooth shaded mode added to the Bimba model and rendered in flat shading. (c) JND modulated noise computed for a flat shaded mode added to the Bimba model and rendered in flat shading.

5.3 Discussion

5.3.1 Computing the Vertex Displacement Threshold for an Interactive Scene

In an interactive scene, the user can manipulate the model displayed on the screen. Therefore, the light direction and position of the model are susceptible to be changed over the course of the viewing session. Changing the light direction will change the local contrast. In addition, changing the position of the model can lead to a change in either local contrast or local spatial frequency. If the light is fixed relative to the view point and the model is rotated or translated according to the X and Y axis of the view frame, *i.e.*, distance from camera is not changed, then this will cause a change in the angles between the surface normal and the light direction. This is the same as fixing the model and changing the light direction and therefore causes a change in contrast. If the model is being scaled or translated according to the Z axis of the view frame, *i.e.*, distance from camera is changed, then this will cause a change in the perceived spatial frequency as size of the visual stimulus is altered. Regardless of the situation, a varying local contrast and/or local spatial frequency would change the threshold beyond which a displacement becomes visible.

First, to account for the change in contrast which is caused by changing the light direction with respect to the surface normal, either by fixing the light and

(a) reference (b) light dependent JND (c) light independent JND

FIGURE 5.15: A vertex noise equivalent to the JND level computed in a light dependant mode might become visible when the light direction is altered between the two rows. This is not the case for light independent JND.

changing the model's position or by fixing the model and changing the light direction, we compute the JND profile in the light independent mode. As explained in Section 5.1.2.1, this is done by computing the displacement threshold from a number of light directions sampled from a hemi-sphere around the local vertex and then choosing the lowest displacement value. The idea here is that the light independent threshold corresponds to the one relative to the "worst possible" light configuration. Figure 5.15 shows the Venus model injected with a vertex noise equivalent to the JND level computed with a light dependent and light independent mode. Notice how the noise relative to the light independent threshold remains invisible while the one related to the light dependent mode becomes visible when the light changed from a front direction (first row) to a top-right one

(a) reference

(b) vertex noise relative to a JND computed according to the most sensitive frequency

(c) vertex noise relative to a distance dependent JND

FIGURE 5.16: A vertex noise equivalent to the JND computed according to (b) the most sensitive frequency and (c) a certain fixed distance (that of the first row).

(second row).

Second, if in the interactive session the user is allowed to change the distance between the view point and the object by either changing the view distance or scaling the model, then the perceived spatial frequency is affected. Similarly to the first case, computing a threshold that works with any distance boils down to computing a threshold that works with the worst distance. Since a change in distance (or scale) affects only the spatial frequency then computing the JND relative to the worst possible frequency results in a threshold that is adaptive to various distances (or scales). This worst possible frequency is the one corresponding to the peak of the CSF curve since it is by definition the most sensitive frequency (*i.e.*, the frequency with the lowest visibility threshold). According to our experimental study we have found that the CSF peaks at around 3.5 cpd under smooth shading. Figure 5.16 shows the Venus model injected with a vertex noise at the JND threshold level computed according to the most sensitive frequency and that of a JND under a certain fixed distance. The Venus model is a dense mesh, therefore the frequencies relative to the first row of Fig. 5.16 are high ($\approx$10 cpd). As the

camera approaches the model, the size of the visual stimuli becomes bigger and therefore the frequency decreases. This means that as the model becomes closer to the camera, the frequency becomes more sensitive (according to the CSF) and thus should reduce the value of the visibility threshold. Indeed, when injecting in the model a vertex noise relative to a JND computed according to the original distance, it becomes visible when zooming in on the model (see Fig. 5.16.(c)). On the contrary, when computing the JND according to the most sensitive frequency, the noise remains invisible regardless of the distance since it has already been taken into account as the worst possible distance (see Fig. 5.16.(b)).

5.3.2 Comparison with an Image-Based Method: HDR-VDP2 [MKRH11]

In this section we compare the proposed vertex displacement threshold method to the HDR-VDP2 [MKRH11] image-based method. HDR-VDP2 is a popular perceptual metric that extends Daly's VDP to HDR images and its code is freely available at `http://hdrvdp.sourceforge.net/wiki/`. Figure 5.17 shows the visibility map given by the HDR-VPD2 method when comparing images of a reference Lion model with a distorted one from various viewpoints. The distorted model was obtained by injecting in the Lion model a vertex displacement noise that is below the threshold given by our method. To be more precise, the noise was injected according to the method described in Section 5.2.1 of this manuscript with $\beta_{jnd} = 0.85$. Therefore, according to our method the injected noise should be invisible which is the case if we visually compare the first two rows of Fig. 5.17. However, when looking at the visibility map given by the HDRVDP2 method we notice that the distortion visibility is particularly exaggerated in the rough regions of the Lion model. This exaggeration of distortion visibility on surface regions with high roughness and curvature can be attributed to a possible limitation of image-based approaches for estimating the visibility. A local geometric distortion will change the geometry of the object which will be reflected in a slight change of pixel positions in addition to the change in contrast. This change in pixel positions is considerable when the surface is curved because it can cause a local displacement of pixels without changing the global shape. Since in image-based methods, the perceptual analysis is in general done in a per-pixel basis, then these local pixel displacements caused by the geometric distortions are detected to be of high amplitude, leading to an exaggeration in visibility. Furthermore, image-based methods such as HDRVDP2 have an intrinsic limitation compared to our method, *i.e.*, for image-based methods it is necessary to generate the rendered 2D image before applying any perceptual analysis. This will ultimately make the integration of these methods in geometric operations a more complicated task and such methods also mix the visual effects of geometric

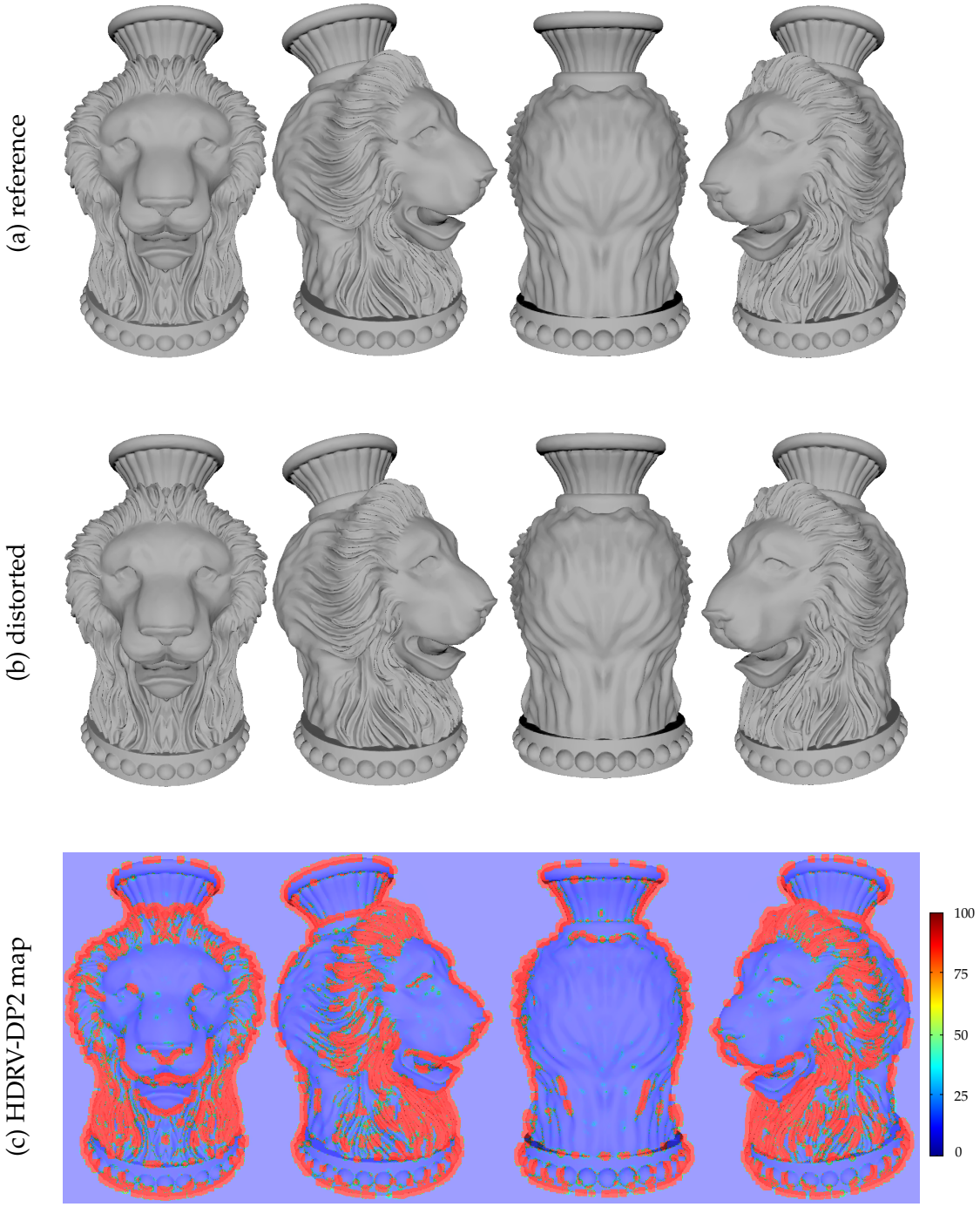


FIGURE 5.17: The HDR-VDP2 visibility map for a distorted Lion model whose noise is below the JND threshold.

distortions with rendering noise. Therefore we think that image-based methods are more suited for evaluating and/or guiding rendering algorithms rather than geometric operations.

5.4 Summary

In this chapter we have presented an algorithm that uses the perceptual model presented in Chapter 4 in order to compute the maximum displacement a vertex can tolerate before becoming visible. This algorithm takes as parameters the display setting (*i.e.*, screen resolution, size and brightness), the scene's illumination and the rendering method (flat or smooth shading). This computed vertex displacement threshold, therefore, adapts to these various parameters. As demonstrated in our subjective validations, the proposed perceptual model can effectively guide the injection of vertex noise into the 3D mesh while keeping it invisible at the same time. This can have a direct application in the case of 3D watermarking and geometric compression algorithms (*e.g.*, vertex coordinates quantization) as their corresponding performance usually relies on the degree of tolerable change in vertex coordinates. In the following chapter we showcase how the vertex displacement threshold and the proposed perceptual models can be integrated into various mesh processing operations. More precisely, we use the computed vertex displacement threshold to detect the optimal vertex quantization level for a certain mesh, guide simplification process and we finally use the perceptual models, *i.e.*, the CSF, to perform an adaptive subdivision on a coarse 3D model.

Chapter 6

Applications

The JND models of 2D images and videos have been used extensively throughout the literature to guide and perceptually optimize several image and video processing algorithms [CL95, LKW06, WN09]. In this chapter, we show how the proposed JND profile as well as the perceptual models can be integrated to several mesh processing algorithms. First, we use the JND profile to compare the global visibility of the geometric distortions caused by a vertex coordinate quantization operation and then use this comparison to automatically select the optimal vertex coordinates quantization level for a given mesh. Second, we use the computed vertex displacement threshold to guide the simplification of 3D meshes. Finally, we integrate the CSF model into an adaptive mesh subdivision pipeline.

6.1 Automatic Selection of Optimal Vertex Quantization Level

Vertex coordinates quantization is an important step in many mesh processing algorithms, especially in the case of 3D mesh compression. This operation introduces geometric distortions onto the original mesh that might be visible to a human observer. It would thus be useful to evaluate whether a vertex quantization noise is visible or not. This would allow us to find the optimal quantization level (in bits per coordinate, bpc) for any mesh. We consider the optimal quantization level to be the one with the lowest bpc, *i.e.*, having the highest noise engergy for which the quantization noise remains visible. The quantization noise should remain invisible and therefore the distorted mesh should remain visually indistinguishable from the original one. It is important to note that the optimal quantization level is different for each mesh due to the differences in geometric complexities, details and density.

The proposed JND model can provide a simple and automatic way to determine the optimal quantization level. The idea is to compute a global score which compares the model's JND profile to the magnitude of introduced noise. To do so, we start by computing the local displacement vectors caused by the quantization

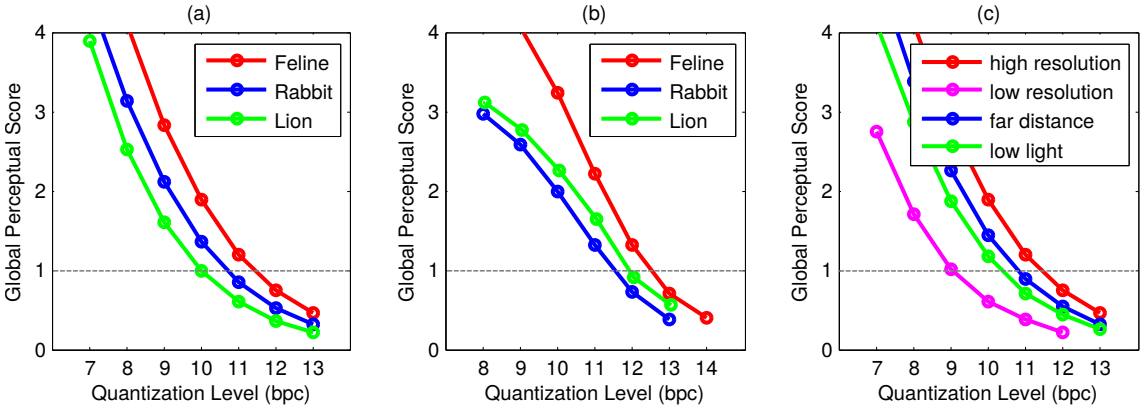


FIGURE 6.1: The perceptual score versus the quantization levels (in bpc) of three models rendered in (a) smooth shading and (b) flat shading. (c) shows the effects of the model resolution, object distance and light energy on the perceptual score of the Feline model when rendered in smooth shading.

operation as:

$$\mathbf{d}_i = \mathbf{v}_i^{'} - \mathbf{v}_i, \tag{6.1}$$

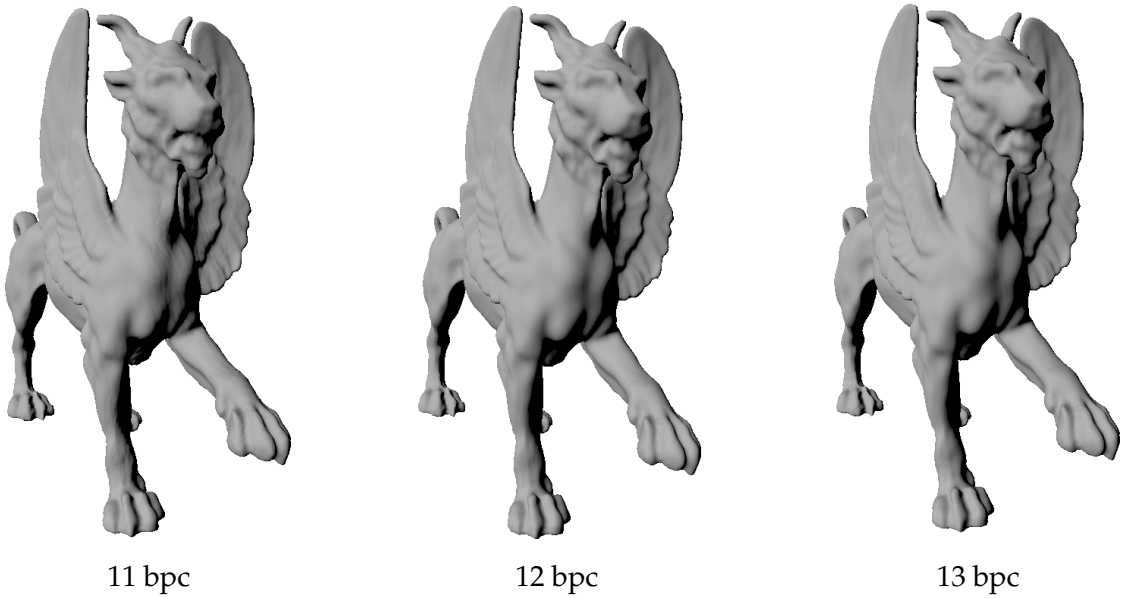
where $v_i^{'}$ and v_i are the i^{th} vertices of respectively the distorted mesh and the original one. The direction of $\mathbf{d}_i$ represents the noise direction relative to the i^{th} vertex. We then compute the JND profile of the original mesh with respect to the computed displacement direction at each vertex. This allows us to evaluate the visibility of the vertex displacement by comparing its magnitude to the computed displacement threshold as follow:

$$r_i = \frac{||\mathbf{d}_i||}{\text{jnd}\left(\mathbf{v}_i, \frac{\mathbf{d}_i}{||\mathbf{d}_i||}\right)}, \tag{6.2}$$

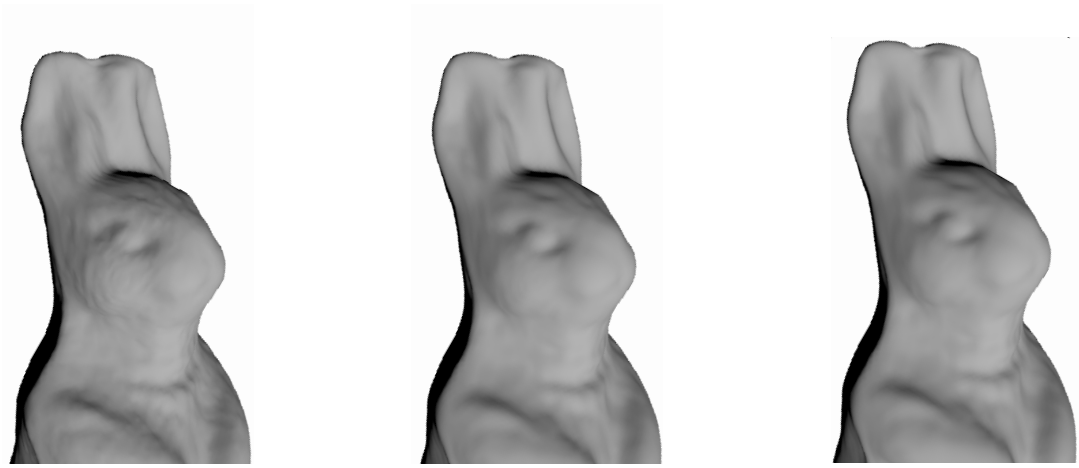
where $\text{jnd}\left(\mathbf{v}_i, \frac{\mathbf{d}_i}{||\mathbf{d}_i||}\right)$ represents the vertex displacement threshold of $\mathbf{v}_i$ in the direction of the local noise displacement, $\mathbf{d}_i$. Finally, we aggregate the local ratio values into a global visibility score using a Minkowski pooling technique as:

$$S = \left(\frac{1}{n}\sum_{i=1}^{n} r_i^p\right)^{(1/p)}, \tag{6.3}$$

where n is the number of vertices in the mesh and $p = 2$ is the Minkowski power. This score allows us to test whether the distortion introduced by the vertex quantization operation is globally visible. If $S \leq 1$, the noise magnitude is globally below the visibility threshold, which means that the distortion is not visible. On the contrary if $S > 1$, the distortion becomes visible as the noise magnitude is in general above the visibility threshold. The optimal quantization level would correspond to the one with the lowest bpc where the global visibility score is below 1.



11 bpc 12 bpc 13 bpc

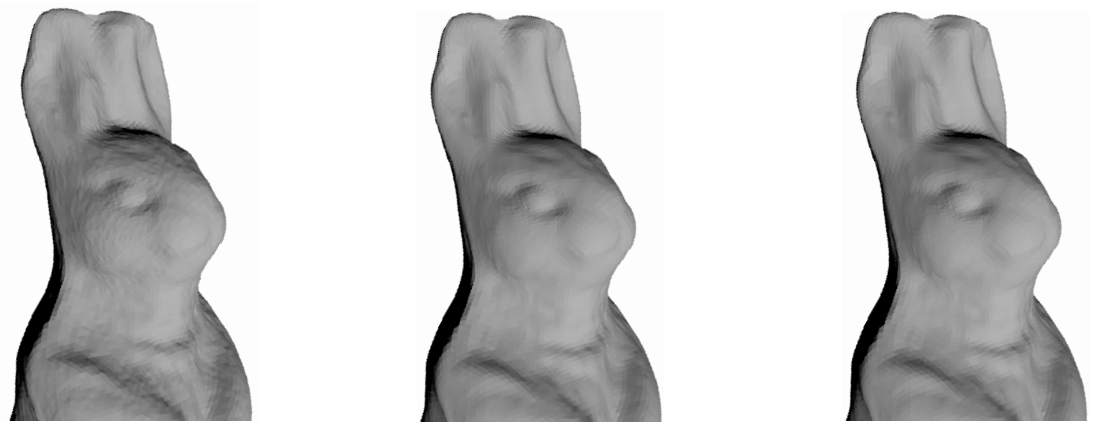


10 bpc 11 bpc 12 bpc

9 bpc 10 bpc 11 bpc

FIGURE 6.2: Quantized meshes with different quantization levels in a smooth shading setting. The middle column corresponds to the optimal quantization level (12, 11 and 10 bpc for respectively Feline, Rabbit and Lion). For better comparison between the models please refer to the electronic version of this manuscript.

12 bpc 13 bpc 14 bpc



11 bpc 12 bpc 13bpc

11 bpc 12 bpc 13 bpc

FIGURE 6.3: Quantized meshes with different quantization levels in a flat shading setting. The middle column corresponds to the optimal quantization level (13, 12 and 12 bpc for respectively Feline, Rabbit and Lion). For better comparison between the models please refer to the electronic version of this manuscript.

(a) 11 bpc (b) 11 bpc - low light (c) 11 bpc - distant view point

FIGURE 6.4: The high-resolution Feline model with an 11 bpc quantization level under different circumstances.

Figure 6.1 shows the global perceptual score versus the level of coordinates quantization for three meshes. According to the defined score the optimal quantization level is 10 bpc for the Lion model, 11 for the Rabbit model and 12 for the Feline model when rendered with a smooth shading algorithm. When the models are displayed with a flat shading algorithm the optimal quantization level, computed with the JND in flat shading mode, is 13 bpc for the Feline model and 12 bpc for the Rabbit and Lion models. This shows how our proposed JND profile can adapt to the shading algorithm as the optimal quantization level will adjust accordingly. In general, the higher optimal quantization level for the smooth shading algorithm indicates that it is less sensitive to vertex noise than flat shading rendering. This is consistent with the results of our subjective experiments where we noted that the vertex displacement threshold was higher for smooth shading than for flat shading (see Section 5.2). These results are compatible with the human observations as shown in Figs. 6.2, 6.3. In addition, the proposed global perceptual score can adapt to different circumstances of mesh usage such as view distance, light energy and mesh resolution (Fig. 6.4 (c)). By reducing the resolution of the Feline model, the optimal quantization level goes down from 12 bpc to 9 bpc while a distant view or low energy light makes the optimal quantization level become 11 bpc. By contrast, we cannot obtain all these results by thresholding the output of state-of-the-art mesh perceptual metrics used in computer graphics such as MSDM2 [Lav11] and HDR-VDP2 [MKRH11] (Fig. 6.5). In particular, the output of MSDM2 remains the same under different light energies, viewing distances and renderings (flat or smooth shading) as it relies entirely on geometric attributes in its perceptual analysis. As for HDR-VDP2, its score is exaggerated for low-resolution meshes indicating that it was not able to accurately

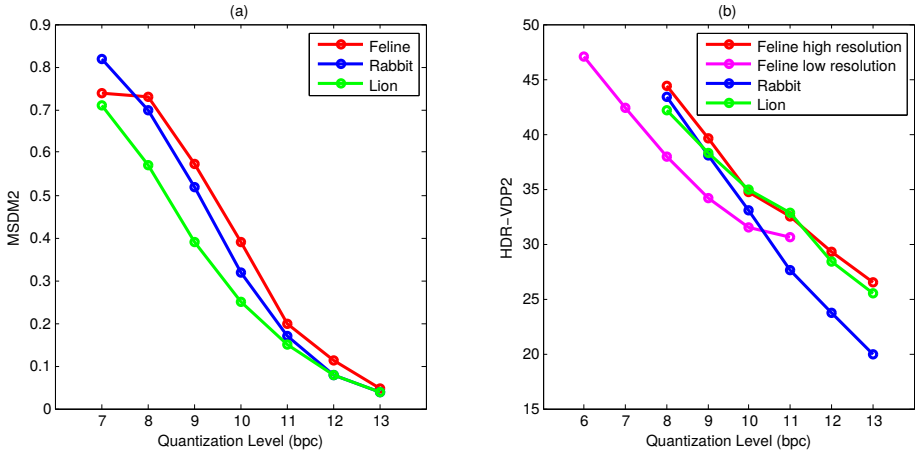


FIGURE 6.5: (a) The MSDM2 and (b) the HDR-VDP2 score versus the quantization levels (in bpc) in a smooth shading setting.

adapt to the change in mesh density. This is probably due to the fixed pixel window width used to perform the perceptual analysis, suggesting that the visibility of all distortions are evaluated relative to nearly the same spatial frequency. This is usually not necessarily the case in a 3D mesh where the vertex density might be variable and thus leading to the presence of visual distortions of different spatial frequencies.

6.2 JND Driven Mesh Simplification

The goal of mesh simplification algorithms is to reduce the number of vertices in a mesh to a certain degree by iteratively applying a simplification step (often via an edge collapse or vertex removal operation). Mesh simplification is usually used to efficiently display highly detailed models or to create multiple levels of details (LOD) of a mesh. Therefore it is required that the simplified mesh preserves the geometric features of the model as much as possible. To do so, most simplification methods tries to simplify the less important regions, *i.e.*, the ones with no geometric features, more often than the ones containing the geometric details. More specifically, a popular class of mesh simplification algorithms achieves this by the steps described below:

1. compute a *simplification cost* for each of the mesh edges (or vertices). This cost indicates whether a certain edge (or vertex) should be removed or not.

2. apply the simplification step, *i.e.*, edge collapse, to the edge with the lowest cost.

3. update the simplification cost.

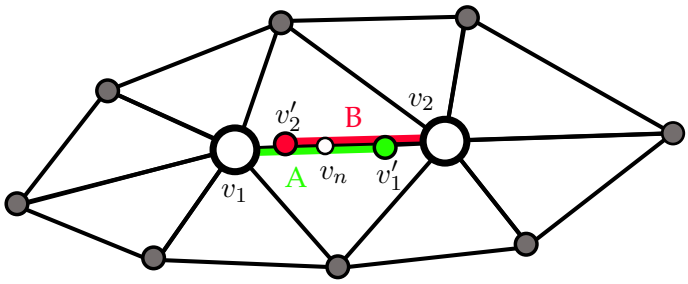


FIGURE 6.6: If $\mathbf{v_1v_2}$ and $\mathbf{v_1'v_2'}$ are in opposite directions, then the edge (v_1, v_2) can be collapsed to v_n without causing any visible distortion.

4. go back to step 2. until a certain number of vertices, edges or faces is obtained.

Ideally, the best visual quality can be obtained if the edge collapse operation is carried out starting from the one with the least visual impact. Throughout the literature, several perceptual and non perceptual methods have been proposed to compute a simplification cost that would best preserve the shape of mesh after simplification. However, existing perceptual methods either carry out the perceptual analysis on the rendered image [WLC+03, QM08, MG10] or rely on a top-down estimation of saliency [LVJ05, WSZL13, SLMR14]. Moreover, none of the existing algorithms propose a method to automatically control the quality of the resulting output; the simplification is usually carried out until a manually prescribed number of edges, vertices or faces is reached. In this section, we show how we can use our proposed JND model to define both the simplification cost for each edge and more importantly a stopping criterion that is able to automatically control the quality of the simplified mesh.

Edge Cost In an edge collapse operation, an edge (v_1, v_2) is removed and is replaced by a vertex v_n (Fig. 6.6). This can be seen as if the vertices v_1 and v_2 moved towards the new vertex v_n. Using our JND model we analyze the visibility of displacing v_1 and v_2 along the edge (v_1, v_2). Let A (resp. B) be a part of (v_1, v_2) bounded by v_1 and v_1' (resp. v_2 and v_2') (see Fig. 6.6) where v_1' (resp. v_2') is the vertex obtained by displacing v_1 (resp. v_2) by exactly the JND value in the direction of $\mathbf{v_1v_2}$ (resp. $\mathbf{v_2v_1}$). This means that replacing v_1 (resp. v_2) by a vertex belonging to A (resp. B) will not cause any visible distortion. In order to apply an edge collapse that is invisible to a human observer, we need to find a new vertex v_n such that $v_n \in A \cap B$. This requires that the vectors $\mathbf{v_1v_2}$ and $\mathbf{v_1'v_2'}$ should be in opposite directions so that $A \cap B \neq \emptyset$. Otherwise, if $\mathbf{v_1v_2}$ and $\mathbf{v_1'v_2'}$ are in the same direction, then we have $A \cap B = \emptyset$, making the distortion caused by the edge collapse visible. This analysis leads us to define the simplification cost which reflects

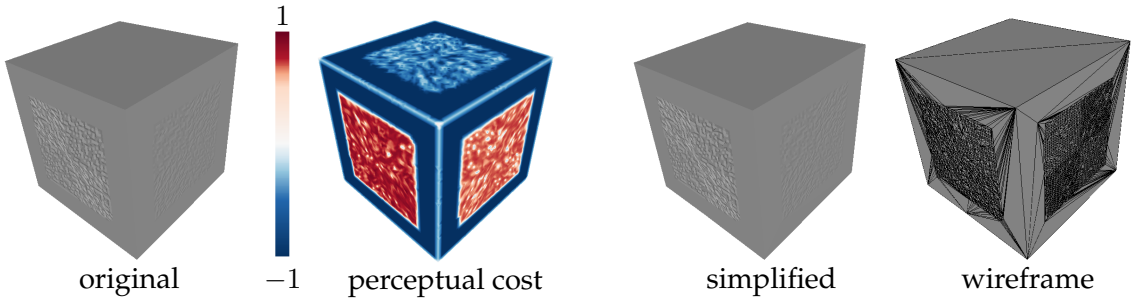


FIGURE 6.7: The simplification algorithm guided by our simplification cost, computed for a flat shading redering, keeps the visible features on the right and left sides of the cube and simplifies the invisible features on the top side.

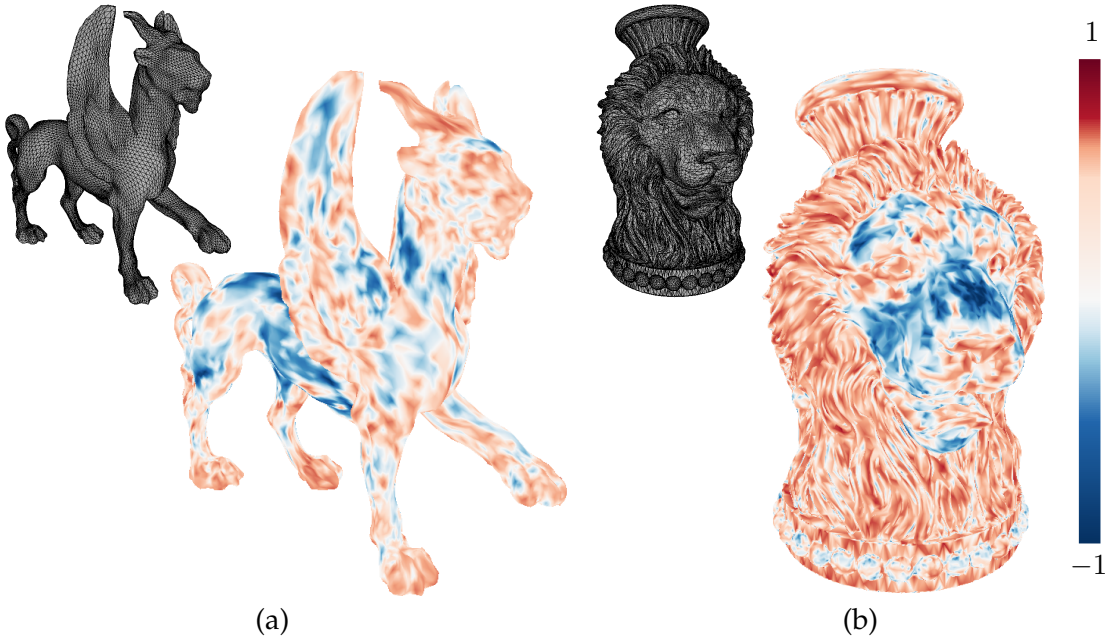


FIGURE 6.8: The perceptual simplification cost for (a) the Feline and (b) the Lion models under a smooth shading mode.

the perceptual impact of an edge collapse by:

$$c = \frac{\mathbf{v_1v_2}.\mathbf{v_1'v_2'}}{||\mathbf{v_1v_2}||^2}. \tag{6.4}$$

The value of our simplification cost c varies between $[-1, 1]$. If $c < 0$ then the collapse operation does not affect the visual fidelity of the model. If $c > 0$ then the edge collapse will be visible. Figure 6.7 shows the simplification cost on a cube under flat shading where we have injected a random noise of different intensity on each of its sides. The noise on the top side is below the JND threshold. On the right side, the noise is barely visible as it is just above the JND threshold and on the left side is injected a visible noise. The simplification cost of the edges belonging to the top side is below 0 as the injected noise is under the JND

threshold while it is above 0 on the right and left sides of the cube. The simplification algorithm guided by our simplification cost will keep the visible features on the right and left sides of the cube and simplify the non visible features on the top side. Figure 6.8 shows some additional results concerning the proposed perceptual cost. Notice how, for both the Feline and Lion models, the proposed perceptual simplification cost is high for the parts where there are geometric details on the surface mesh. This suggests that the simplification algorithm guided by this cost will tend to preserve these details by prioritizing the simplification of regions that do not contain fine geometric features.

Vertex Placement Having defined the simplification cost of an edge, we now should decide how the position of the new vertex v_n is computed. In order to get the "optimal" position we have found that minimizing the following energy produces very good results:

$$\arg\min \left\{ \left(\frac{||\mathbf{v_1 v_n}||}{\mathrm{jnd}_{\mathrm{v_1}}} \right)^p + \left(\frac{||\mathbf{v_2 v_n}||}{\mathrm{jnd}_{\mathrm{v_2}}} \right)^p \right\}, \tag{6.5}$$

where $\mathrm{jnd}_{\mathrm{v_1}}$ (resp. $\mathrm{jnd}_{\mathrm{v_2}}$) is the JND threshold of v_1 (resp. v_2) in the direction of $\mathbf{v_1 v_2}$ (resp. $\mathbf{v_2 v_1}$) and p is the energy's order. For $p = 2$, solving Eq. (6.5) yields to:

$$||\mathbf{v_1 v_n}|| = ||\mathbf{v_1 v_2}|| \times \frac{\mathrm{jnd}^2_{\mathrm{v_1}}}{\mathrm{jnd}^2_{\mathrm{v_1}} + \mathrm{jnd}^2_{\mathrm{v_2}}}, \tag{6.6}$$

where $||\mathbf{v_1 v_n}||$ and $||\mathbf{v_2 v_n}||$ represent respectively the distances by which v_1 and v_2 are being displaced. The idea behind minimizing this energy is to make the displacement of v_1 and v_2 adaptive to their corresponding JND values.

Stopping Criterion The value of the defined simplification cost varies between $[-1, 1]$. A negative value of the perceptual cost indicates that the collapse operation will remain unnoticed by a human observer while a positive value means that the collapse will be visible. So if we collapse the edges according to an increasing order of perceptual impact, then we can control the visual quality of the output mesh. Intuitively, by stopping the simplification process when all the edges have a positive perceptual impact we obtain a most simplified mesh that is visually similar to the detailed input. For instance, in Fig. 6.9 we show a simplified tetrahedron obtained using the proposed perceptual simplification method in the case of flat shading and smooth shading, and compare it with the traditional simplification method of [LT98]. In the case of flat shading, our method preserves the hard edges on the model so that the resulting mesh looks identical to the original one. The same result can also be achieved with [LT98], however

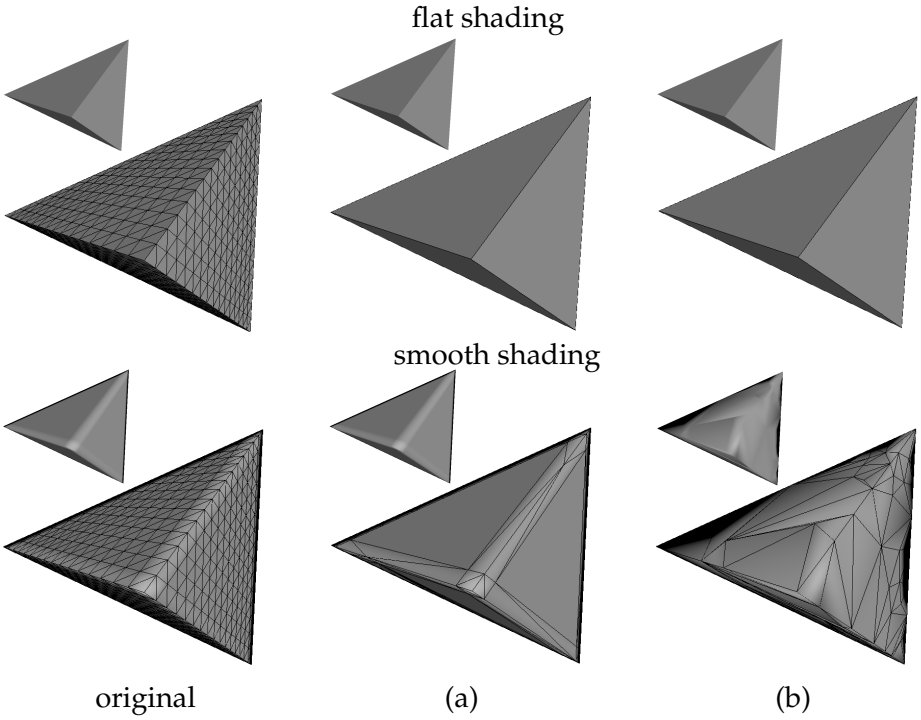
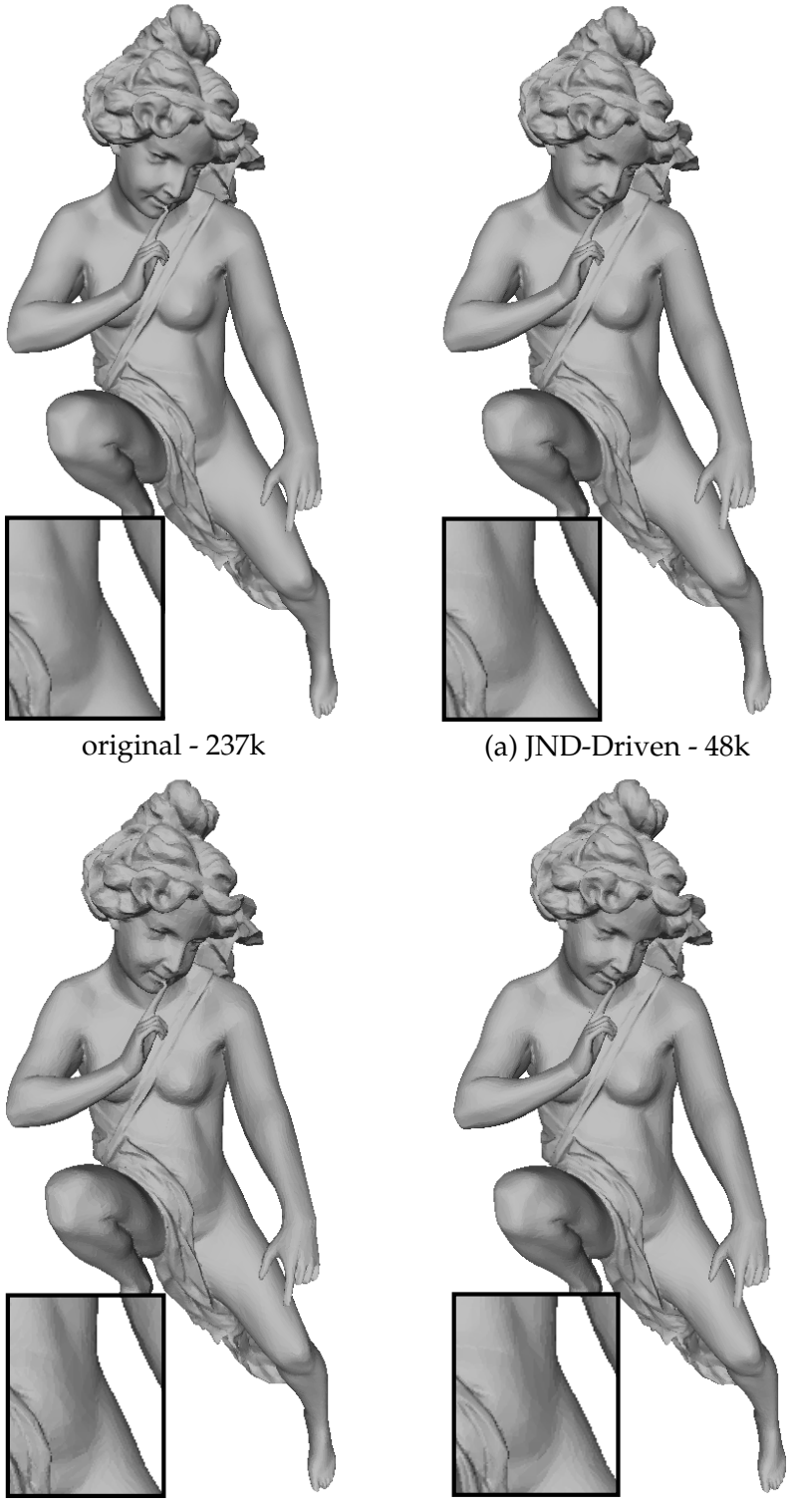


FIGURE 6.9: (a) The perceptually driven simplification is able to prevent perceptually relevant edges from being collapsed as it outputs a model that is visually similar to the original one in the case of flat and smooth shading. (b) The output model using the method of Lindstrom and Turk [LT98] with the same number of vertices is visually different from the original mesh under smooth shading.

the major difference is that this algorithm requires that the user explicitly input the number of edges beyond which the simplification would stop. This is not the case with our method as the simplification operation is stopped automatically according to the defined perceptual stopping criterion. In addition, our method can adapt to the rendering algorithm. When rendered in smooth shading, the simplification process does not collapse the edges that are adjacent to the hard edges of the tetrahedron as they have a big influence on the interpolation of luminance and thus have a high perceptual impact. This is however not the case for [LT98] which focuses on preserving the geometry of the object without taking into consideration the shading. Figure 6.10 shows a very dense 3D mesh rendered in flat shading which is then simplified with the JND-driven simplification method. The resulting simplified mesh (Fig. 6.10.(a)) has 80% less vertices and is visually very similar to the original version. Removing 5% more vertices beyond the JND level introduces slightly visible distortions to the model (Fig. 6.10.(b)). In addition, simplifying the model using Lindstrom and Turk's method [LT98] (edge collapse with a different cost) to the same number of vertices as the JND-driven simplification also results in slightly visible distortions (Fig. 6.10.(c)). Additionally, since the proposed perceptual model is capable of adapting to the various parameters



original - 237k (a) JND-Driven - 48k

(b) 45.5k (c) Lindstrom and Turk [LT98] - 48k

FIGURE 6.10: (a) The JND-driven mesh simplification process outputs a model that is visually very similar to the flat shaded original model. (b) Removing 5% more vertices will introduce slightly visible distortions to the simplified model. (c) The simplified model by using the method of Lindstrom and Turk [LT98] to the same number of vertices as the JND-driven simplification. (b) and (c) contain slightly visible distortions, especially on the belly and thighs.

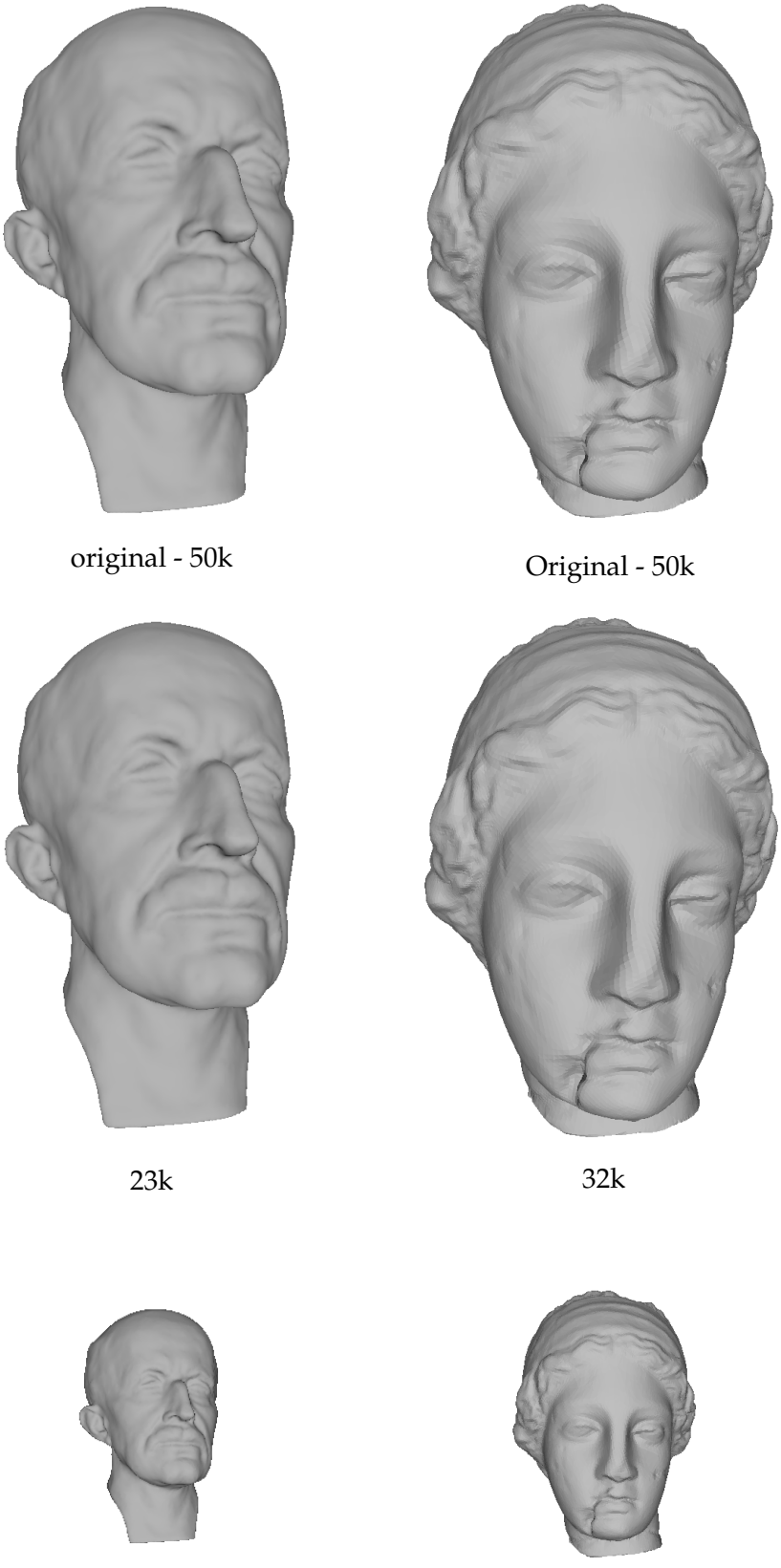


FIGURE 6.11: LODs generated using the perceptually driven simplification method at different viewing distances for the Max Plank (smooth shading) and Venus (flat shading) models. The degree of simplification automatically adapts to the viewing distance.

of the display (size, resolution and brightness) and to the distance between the model and the viewpoint, then the perceptual impact of an edge will adapt as well. In consequence the degree of simplification applied to a detailed mesh will depend on those parameters. This is important as it will be possible to generate multiple LODs by simply specifying the distances between the 3D model and the viewpoint (see Fig. 6.11).

6.3 Perceptual Adaptive Mesh Subdivision

In an adaptive mesh refinement setting, the subdivision operation is applied to the faces that fulfil a certain condition. In most cases, the subdivision is applied to the faces that are relatively close to the viewpoint or are part of the mesh silhouette. Moreover, the subdivision process is usually stopped when a certain polygon budget is reached. In general, the goal of adaptive mesh refinement methods is to display a coarse model in a way that appears *visually smooth*. Intuitively, this can be achieved if the normal vectors used for the shading computation produce a smooth visual pattern. In other words, we may consider that *visual smoothness* is achieved if the interpolation between the brightest and darkest luminance level inside a face is unnoticeable to a human observer under a smooth shading rendering. As a result, we can use the proposed perceptual model in order to test whether this interpolation is visible or not. This test can therefore be done by simply normalizing the contrast value by the CSF value:

$$\tilde{c} = c \cdot \mathrm{csf}(f, l) \ , \tag{6.7}$$

where $\tilde{c}$ is the normalized contrast, c is the face's contrast value, f is the corresponding frequency and l is the global luminance. Hence, the criterion to subdivide a certain face could be defined when the corresponding $\tilde{c}$ value is above a certain threshold:

$$\tilde{c} > \alpha. \tag{6.8}$$

One major advantage of this perceptual criterion, is that the value of $\tilde{c}$ will change if the spatial frequency relative to a face changes. This means that the proposed subdivision criterion will automatically adapt to the original density of the 3D model and its distance to the viewpoint. In a conservative way, a mesh appears to be visually smooth if the local contrast of its faces is not visible as this means that the interpolation between the darkest point and the brightest point of the face will not be noticed by the observer. This implies that the subdivision operation is applied as long as $\tilde{c}$ is greater than 1.

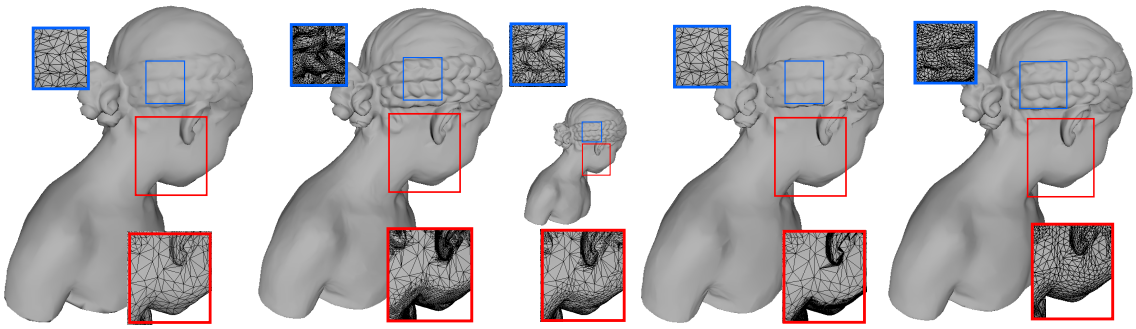


FIGURE 6.12: The perceptual subdivision process converges at around 30K vertices when the model is (a) close from the viewpoint and at 11K when it is (b) far away. (c) A silhouette subdivision criterion would concentrate the subdivisions in the contour region leaving the rough part of the model intact. (d) A proximity subdivision criterion would subdivide the face in already visually smooth areas where the subdivision will not have any visual impact.

We have tested this perceptual subdivision criterion using Boubekeur and Alexa's Phong tessellation method [BA08] with a loop subdivision scheme. Figure 6.12 shows a coarse version of the Bimba model that was subdivided using the proposed perceptually driven method. The subdivided version exhibits an increase in density in rough regions which usually contain faces with visible contrast, *i.e.*, $\tilde{c} > 1$. The subdivision process automatically stops around 30k vertices, when the local contrast in all the faces is invisible, and the resulting mesh appears to be visually smooth (Fig. 6.12 (a)). By increasing the view distance, the spatial frequency of the observed visual stimuli is increased. Consequently, this causes a changes in the visibility threshold and thus affects the degree of subdivision needed to obtain a visually smooth model. Therefore, in this case, the adaptive subdivision automatically stops after fewer subdivisions (Fig. 6.12 (b)). In addition, we have compared our perceptual subdivision criterion to a silhouette-based and proximity-based subdivision criterion. We have stopped the subdivision process when they reached the same number of vertices as in Fig. 6.12 (a). The silhouette subdivision criterion (Fig. 6.12 (c)) consists of subdividing the faces that are part of the mesh's contour so that it appears smooth. In that case, the subdivisions are concentrated in the silhouette and could potentially leave rough parts of the model untouched. In a proximity subdivision criterion (Fig. 6.12 (d)), the faces that are closer to the viewpoint are subdivided. This will most likely result in a waste of computing resources as the algorithm will subdivide faces that are already visually smooth and therefore the operation will not have much visual impact.

6.4 Summary

In this chapter we have integrated the proposed perceptual model into several mesh processing algorithms. For instance, we have used the JND profile to compare the geometric distortions caused by vertex coordinate quantization operation which allowed us to detect the optimal quatization level for a given mesh. In addition, we have used our perceptual method to guide mesh simplification. This is done by computing an edge collapse cost using the JND threshold that indicates whether the collapse operation is visible or not. Finally we have used the CSF model to define a perceptual criterion for the adaptive mesh subdivision task. This series of applications showcases the usefulness of our perceptual models in the case of guiding geometric operation. In particular, we think that the main advantage, that our method can offer, is that it is possible to adapt the output of the geometric operation to the various parameters (screen resolution and size, viewpoint distance, rendering type, illumination, mesh density). This advantage is due to the fact that our proposed perceptual analysis takes these parameters into consideration while computing the displacement threshold.

Chapter 7

Conclusion

It is common that a 3D mesh undergoes some lossy operations such as 3D compression, simplification and watermarking which introduce geometric noise to the underlying 3D object in the form of vertex displacement. In some cases, this noise can be visible to the human end user. Controlling the visibility of these geometric distortions is an important issue as, for human-centered applications, it can have a direct impact on the quality of experience of the user. It is therefore interesting to be able to evaluate whether the vertex displacement caused by the geometric operations is visible or not.

7.1 Summary of Contributions

In this thesis, we have presented our work which focused on evaluating the threshold beyond which a local geometric distortion becomes visible on a 3D mesh. This was achieved with the help of an experimental study of the properties of the human visual system. To reach our goal, we have started by evaluating perceptually oriented attributes (such as local contrast and spatial frequency) on 3D models. We have then performed a series of psychophysical experiments where we have measured the threshold needed for a human observer to detect a change in contrast induced by a change in vertex position under different viewing circumstances. The results of these experiments allowed us to derive a computational model that evaluates the threshold beyond which a change in contrast on a 3D mesh becomes visible. This model can adapt to the various display parameters (resolution, size and brightness), to the size and density of the triangular mesh as well as to directional and point light illumination. Using this perceptual model we have then presented an algorithm that computes the vertex displacement threshold relative to a certain direction, *i.e.*, the Just Noticeable Distortion (JND) profile. Finally, we have illustrated the utility of the proposed JND profile and the perceptual model in several applications such as vertex coordinates quantization, mesh simplification and adaptive mesh subdivision.

The contributions of this thesis are summarized as follows.

Evaluating Local Perceptual Attributes on 3D Triangular Meshes

The properties related to the contrast are essential when studying the visibility of a certain visual pattern. Therefore we have started our perceptual study by evaluating the perceptual attributes related to this aspect of the human perception. More specifically, we have proposed a method to compute the local contrast in the case of flat and smooth shaded rendering. In both cases, our methods for computing the contrast highlight the effects of surface geometry and lighting condition on the value of the local contrast. Moreover, we have evaluated the spatial frequency and shown that it is affected by the size and vertex density of the observed mesh. Finally, inspired by the free energy principle, we have computed the regularity of a visual pattern on the 3D mesh. These perceptually related attributes allowed us to take into consideration the contrast sensitivity and visual masking effects of the visual system when it comes to its capacity of discriminating between two visual patterns.

Measuring the Visibility Threshold

The displacement of a vertex causes a change in contrast in the region surrounding the vertex being displaced. Therefore we have performed a series of psychophysical experiments that aims at measuring the thresholds related to the contrast sensitivity and visual masking aspects of the visual system. For measuring the *Contrast Sensitivity Function (CSF)* we have measured the contrast threshold required for observing a displaced vertex on the surface of a plane. By changing the density of the displayed plane, we have been able to measure the threshold at different frequencies. Measuring the threshold related to the *Visual Masking* requires detecting the threshold beyond which a human observer can notice a difference between two visible contrasts. Therefore for these measurements we have displaced a vertex on the surface of a sphere as its curved surface generates a visible contrast. These experiments were carried out both in the case where the displayed mesh is rendered by a flat shading algorithm and smooth shading one. In the latter case we have taken into account in our measurements the effects of luminance on the CSF threshold and the effects of visual regularity on the visual masking. Finally, the results of these experiments have allowed us to propose a perceptual model to obtain the contrast threshold which is then used to compute the probability of noticing a difference given a certain change in contrast.

Computing the Just Noticeable Distortion Profile for 3D Meshes

Having computed the probability of detecting a change in contrast caused by a vertex displacement, we have proposed an algorithm for the evaluation of the vertex displacement threshold, *i.e.*, JND profile, for 3D meshes. The computed threshold is capable of adapting to the varying circumstances of mesh usage such as the display parameters, light direction and viewing distance since they affect the local perceptual properties. In the case of an interactive scene where the light and the viewing distance are likely to be changed through the course of the viewing session, we have proposed a way to compute this threshold independently from the light direction and with respect to the most sensitive spatial frequency. We have tested the performance of the computed JND profile via a series of subjective experiments where the participants had to rate the visibility of JND modulated random noise added to a number of 3D models with different geometric features. The results of these experiments show that our perceptual model can accurately predict the visibility threshold of local vertex distortions.

Integrating the Perceptual Model into Various Geometry Processing Operations

In Chapter 6 of this thesis we have demonstrated the utility of the proposed perceptual model in a number of geometric applications. First, we have used the JND profile to compare two meshes: a reference and a distorted version. The idea was to compute the distortion on the distorted model in terms of JND units then to deduce a global score that is indicative of the visibility of the vertex noise introduced by the distortion. This mesh comparison method was used to select the optimal quantization level (in bits per coordinates), for any mesh. Second, we used the vertex displacement threshold to guide edge-collapse-based mesh simplification. More precisely, the vertex JND allowed us to compute a simplification cost that indicates whether the edge collapse operation would be visible or not. The most important aspect of this application is that, since the JND value can adapt to the display parameters and view distance, then the cost will adapt as well. In addition, using our perceptual cost, we can define a stopping criterion that would automatically terminate the simplification operation. This aspect is particularly useful for generating a LOD. We have also shown that not only the JND threshold can be useful for geometric applications but also there is a potential in using the local contrast value in combination with the proposed perceptual model such as the CSF. For instance, we have presented a perceptual criterion for adaptively subdividing a 3D triangular mesh, which uses the normalized local contrast value. Similarly to the simplification operation, the main advantage of

our perceptual criterion is that it makes possible to automatically stop the subdivision operation once the subdivision operation will not add any visual smoothness to the displayed mesh.

7.2 Perspectives

In this thesis, we have presented a perceptual framework for studying the visual impact of geometric distortions on the surface of 3D triangular mesh. In the future, we think that this work can be evolved according to three axes.

Perceptual Attributes

Computing the visibility of vertex noise relies heavily on perceptual attributes defined on the 3D mesh. In its current state, our method works for untextured diffuse surfaces that are illuminated by one light source and rendered with either a simple flat-shading or smooth-shading algorithm. This is due to restrictions that we put to the algorithms used to compute the contrast. In addition, since contrast is evaluated locally, *i.e.*, between two faces in a flat shading setting and on a single face for the case of smooth shading, then our visibility prediction is also limited to local distortions. Therefore, we think that in order for our method to support more advanced lighting such as environment maps, complex surface materials and large-scale geometric distortions, it would require a more general definition of contrast that could be based on a thorough and non-trivial analysis of the rendering algorithm and that can be evaluated at multiple scales. Textures on the other hand can be taken into account by trying to combine the contrast due to the shading with the contrast of the texture. Additionally it would be interesting to rethink the way we define the spatial frequency on a 3D mesh. In its current state we have used the traditional unit of *cpd*, which represents the number of cycles of dark and bright patterns of light that can fit in one degree of the visual angle, to express the spatial frequency. While this frequency unit is more natural in the case of images with sinusoidal patterns, we explained in Section 4.2.2 how it can relate to the density of a 3D mesh. For instance, a noise injected onto a dense mesh will exhibit visual pattern with higher frequencies than when injected on a coarse mesh. This leads to the idea of expressing the spatial frequency in terms of *vertices per degree (vpd)* which represents the number of vertices in one degree of the visual angle and thus can be more natural in the case of 3D meshes.

Our proposed perceptual framework is now capable of evaluating whether a local distortion is visible or not. It was not designed to test whether or not a

visible distortion is perceptually more annoying than another visible one. For instance, the value of our proposed perceptual simplification cost is bounded between -1 and 1 where a positive value indicates that the edge collapse operation would be visible. However, since our perceptual analysis is focused on the visibility of a distortion then it is possible to obtain a simplification cost of value 1 for both an edge whose collapse is visible but not annoying and an edge whose collapse is visible and not tolerable. This is due to our focus on low-level perceptual attributes such as contrast in this work. This is why it would be important and interesting to exploit the utility of the proposed contrast algorithm in order to define high-level visual attributes such as mesh saliency or visual attention.

Finally, it was important to estimate the regularity of a visual pattern in order to have a more accurate modeling of the visual masking of the human vision. Our method for estimating that measure was inspired from the free energy principle theory which suggests that the visual system is actively trying to predict the visual patterns in order to minimize surprise. Therefore we have used a simple linear system to predict the local contrast using its surrounding and thus compute a value of visual regularity. However, even if this simple linear system works well in practice, it would be interesting to test whether higher-order methods would give better results.

Perceptual Model

A further development of the perceptual models would make the perceptual analysis more accurate. For instance, modifying Daly's masking model to incorporate the effects of regularity of the visual system on the visibility threshold (Section 4.3) allowed for a better treatment of visual masking and resulted in a more precise visibility threshold. An interesting extension of the existing perceptual model would be to include the effects of velocity on the computed visibility threshold. As explained in Section 2.2 the contrast sensitivity of the HVS is affected by the velocity of a moving visual stimulus. We have conducted a preliminary study, which can be found in Appendix C, to measure the contrast sensitivity for moving 3D objects. This would allow us to expand this current perceptual model to support dynamic moving 3D meshes in addition to static ones. Moreover, as it stands, the proposed framework does not take into account the color attributed to the mesh or the illumination as it focuses on white lumination levels. An idea for extending this method to color would require conducting psychophysical experiments to measure the contrast threshold for each color channel as the sensitivity is different for various light frequencies [Kel83]. We can then apply the same perceptual analysis as the one described in this manuscript for each

channel in order to compute the visibility threshold. On the long term, we also think that expanding the perceptual model to take into account the characteristics of the human visual system that are related to the orientation and scale selectivity would be interesting. These aspects would be important especially for developing methods that take into consideration the higher-level properties of human vision. However, this step first requires a multi-scale definition of contrast for 3D meshes.

Applications

In the end, the goal behind developing a perceptual method for 3D meshes is to integrate it into geometry processing so that it would be possible to control the quality of its output or to evaluate and compare existing geometric methods. The applications showcased in this thesis present an example about the potential of using perceptual methods in geometric application and can be further improved. For instance, the presented simplification cost reflects whether an edge collapse would be visible or not if the position of the new vertex is located on the collapsed edge. We think it would be more interesting to develop a perceptual simplification cost that overcomes this limitation. Also further improvements can be done to the perceptual criterion proposed for the adaptive subdivision operation which suggests that the mesh is visually smooth if the local contrast is not visible. While this assumption can be considered as correct, it is however too conservative as it is possible to have a visible local contrast that is visually smooth. One possible direction to improve this application is to adapt the threshold for applying the subdivision to the visual regularity of the surface. Intuitively, if a surface is visually regular (usually a smooth surface) then the threshold for subdividing can be high while for a visually complex surface (usually a rough surface) it is better to have a low threshold value.

Moreover, it would be interesting to use this perceptual method to develop an objective perceptual quality measure for 3D meshes. A geometric distortion would generally introduce a change in local perceptual attributes, most notably contrast, to the 3D mesh. Using our method, it is possible to evaluate this change in contrast and then convert it in terms of visibility threshold units. For example, for the vertices where the change in contrast is visible the value would be greater than 1 otherwise it will be lower than 1. The same can also be done for visual regularity as a distortion would also cause a change in that attribute (*e.g.*, a noise on a smooth surface makes it more complex). Having a measure for a change in contrast and another for a change in visual regularity can be indicative of a change in structure and would allow us to develop, in the future, a perceptual

quality metric for 3D meshes that is similar to the SSIM metric for images. A preliminary test of a quality metric using our perceptual method is presented in Appendix D and it shows some promising results.

Appendix A

Computational Details for Contrast Estimation

A.1 Contrast Estimation for Flat-Shaded Surfaces

We provide here the details of the transition between Eqs. (4.5) and (4.6) (see Section 4.2.1).

$$c = \frac{\| \max(\mathbf{l}\cdot\mathbf{n_1}, 0) - \max(\mathbf{l}\cdot\mathbf{n_2}, 0) \|}{\max(\mathbf{l}\cdot\mathbf{n_1}, 0) + \max(\mathbf{l}\cdot\mathbf{n_2}, 0)}, \quad c = \| \cos\alpha \cdot \tan\theta \cdot \tan\frac{\phi}{2} \|$$

In the case where the inner products between the light direction and the two face normals are both positive, the first equation above becomes:

$$\begin{aligned} c &= \frac{\| \mathbf{l}\cdot\mathbf{n_1} - \mathbf{l}\cdot\mathbf{n_2} \|}{\mathbf{l}\cdot\mathbf{n_1} + \mathbf{l}\cdot\mathbf{n_2}} \\ c &= \frac{\| \mathbf{l}\cdot(\mathbf{n_1} - \mathbf{n_2}) \|}{\mathbf{l}\cdot(\mathbf{n_1} + \mathbf{n_2})} \\ c &= \frac{\| \cos\angle(\mathbf{l}, \mathbf{n_1} - \mathbf{n_2}) \|}{\| \cos\angle(\mathbf{l}, \mathbf{n_1} + \mathbf{n_2}) \|} \cdot \frac{\| \mathbf{n_1} - \mathbf{n_2} \|}{\| \mathbf{n_1} + \mathbf{n_2} \|}. \end{aligned}$$

The ratio between the norms of the vectors $\mathbf{n_1} - \mathbf{n_2}$ and $\mathbf{n_1} + \mathbf{n_2}$ is evaluated by:

$$\frac{\| \mathbf{n_1} - \mathbf{n_2} \|}{\| \mathbf{n_1} + \mathbf{n_2} \|} = \sqrt{\frac{1-\cos\phi}{1+\cos\phi}} = \tan\frac{\phi}{2},$$

where ϕ is the angle between $\mathbf{n_1}$ and $\mathbf{n_2}$ (see Fig 4.3). The cosines of the angles between l and $\mathbf{n_1} - \mathbf{n_2}$ and between l and $\mathbf{n_1} + \mathbf{n_2}$ are respectively equivalent to:

$$\begin{aligned} \cos\angle(\mathbf{l}, \mathbf{n_1} - \mathbf{n_2}) &= \cos\alpha \cdot \sin\theta \\ \cos\angle(\mathbf{l}, \mathbf{n_1} + \mathbf{n_2}) &= \cos\theta, \end{aligned}$$

where α and θ are the spherical coordinates of the light direction in the local coordinate system defined by $\mathbf{n_1} - \mathbf{n_2}$, $\mathbf{n_1} + \mathbf{n_2}$ and their outer product (see Fig.

4.3). This finally leads to:

$$c = \| \frac{\cos\alpha \cdot \sin\theta}{\cos\theta} \cdot \tan\frac{\phi}{2} \|$$
$$c = \| \cos\alpha \cdot \tan\theta \cdot \tan\frac{\phi}{2} \| .$$

A.2 Contrast Estimation for Smooth-Shaded Surfaces

We provide the details on how the optimization problem in Eq. (4.17) of the manuscript can be solved. First we rewrite Eq. (4.16) as:

$$d_i(\alpha,\beta)^2 = ||\alpha\mathbf{A} + \beta\mathbf{B} + \mathbf{C}||^2 , \tag{A.1}$$

where $\mathbf{A} = \widehat{\mathbf{n}_3\mathbf{n}_1}$, $\mathbf{B} = \widehat{\mathbf{n}_3\mathbf{n}_2}$, and $\mathbf{C} = \widehat{\mathbf{L}\mathbf{n}_3}$. The reason behind changing the notation of Eq. (4.16) is to be able to solve Eq. (4.17) for Eqs. (4.16) and (4.19) in a similar way. For Eq. (4.19) we will have $\mathbf{A} = \widehat{\mathbf{n}_3\mathbf{n}_1} - \widehat{\mathbf{l}_3\mathbf{l}_1}$, $\mathbf{B} = \widehat{\mathbf{n}_3\mathbf{n}_2} - \widehat{\mathbf{l}_3\mathbf{l}_2}$, and $\mathbf{C} = \widehat{\mathbf{l}_3\mathbf{n}_3}$.
By developing Eq. (A.1) we obtain:

$$d_i(\alpha,\beta)^2 = \alpha^2\mathbf{A}.\mathbf{A} + \beta^2\mathbf{B}.\mathbf{B} + 2\alpha\beta\mathbf{A}.\mathbf{B} + 2\alpha\mathbf{C}.\mathbf{A} + 2\beta\mathbf{C}.\mathbf{B} + \mathbf{C}.\mathbf{C}. \tag{A.2}$$

Equation (A.2) represents a paraboloid, so solving Eq. (4.17) boils down to finding the minimum and maximum points on that paraboloid such that $\alpha + \beta \leq 1$ and $\alpha, \beta \in [0, 1]$. The minimum point is computed by :

$$\alpha = \frac{\mathbf{C}.\mathbf{B}\cdot\mathbf{A}.\mathbf{B} - \mathbf{C}.\mathbf{A}\cdot\mathbf{B}.\mathbf{B}}{\mathbf{B}.\mathbf{B}\cdot\mathbf{A}.\mathbf{A} - \mathbf{A}.\mathbf{B}\cdot\mathbf{A}.\mathbf{B}} , \beta = -\frac{\mathbf{C}.\mathbf{B} + \alpha\cdot\mathbf{A}.\mathbf{B}}{\mathbf{B}.\mathbf{B}}. \tag{A.3}$$

If the computed α and β do not respect the minimization constraints, then their values are adjusted accordingly:

$$(\alpha,\ \beta) = \begin{cases} \left(0,\ \min\left(-\frac{\mathbf{C}.\mathbf{B}}{\mathbf{B}.\mathbf{B}},\ 0\right)\right) & \text{if } \alpha < 0 \\ \left(\min\left(-\frac{\mathbf{C}.\mathbf{A}}{\mathbf{A}.\mathbf{A}},\ 0\right),\ 0\right) & \text{if } \beta < 0 \\ \left(\min\left(\frac{\mathbf{C}.\mathbf{B}\cdot\mathbf{A}.\mathbf{B} - \mathbf{C}.\mathbf{A}\cdot\mathbf{B}.\mathbf{B}}{\mathbf{B}.\mathbf{B}\cdot\mathbf{A}.\mathbf{A} - \mathbf{A}.\mathbf{B}\cdot\mathbf{A}.\mathbf{B}}, 1\right),\ 1-\alpha\right) & \text{if } \alpha+\beta > 1 \end{cases}$$

It is easy to show that the maximum distance will always correspond to $(\alpha,\beta) = (0,1),\ (1,0)$ or $(0,0)$. Having computed α and β it is simple to compute the position or normal of the corresponding point using the barycentric coordinates $[\alpha, \beta, 1-\alpha-\beta]$.

Appendix B

Measuring the Masking Threshold Independently from the Spatial Frequency

As we mentioned in Chapter 4, the effects of the spatial frequency on the visibility of a visual pattern makes the masking threshold value different at each frequency. However, it is not practical and is time consuming to measure the masking threshold at different frequencies. In addition, it would be more difficult to model this effect based on per-frequency data. Therefore if possible, it would be more interesting to measure the masking threshold independently from the spatial frequency of the visual signal. In his paper [Dal93], Daly has suggested that it would be possible to model the masking effect regardless of the frequency if the contrast values were normalized by the CSF value. The idea here is that instead of dealing with absolute contrast values which will depend on the frequency, the masking model will handle normalized contrast values that are independent from the frequency. To test whether it is possible to follow this idea by normalizing the contrast value by the CSF (Eq. (4.10)), we have performed a preliminary experiment in which we measured the masking threshold at three different frequencies (1.1cpd, 2.63cpd and 5.59cpd) and in a flat shading setting.

Experimental Protocol For this preliminary experiment we followed the same experimental procedure as described in Section 4.2.3 of the manuscript. To summarize, two icospheres were displayed on the screen side by side, one of which exhibits a vertex displacement. The subjects responded by *Yes* or *No* to whether they can see a difference between the two displayed spheres which will then affect the displacement magnitude using the QUEST method. In order to measure the masking threshold at three different frequencies we have used three icospheres each with a different subdivision level (Fig B.1).

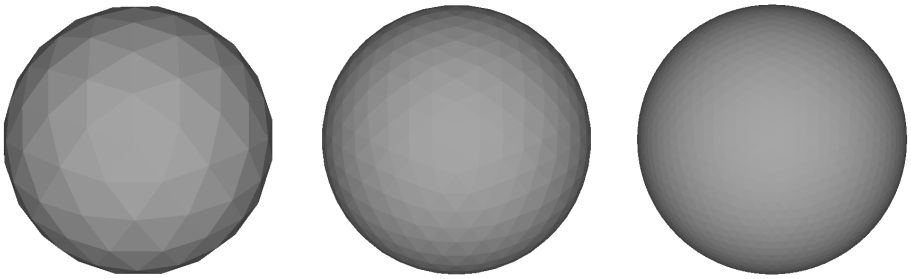


FIGURE B.1: By changing the subdivision level of an icosphere we were able to measure the masking threshold at different spatial frequencies.

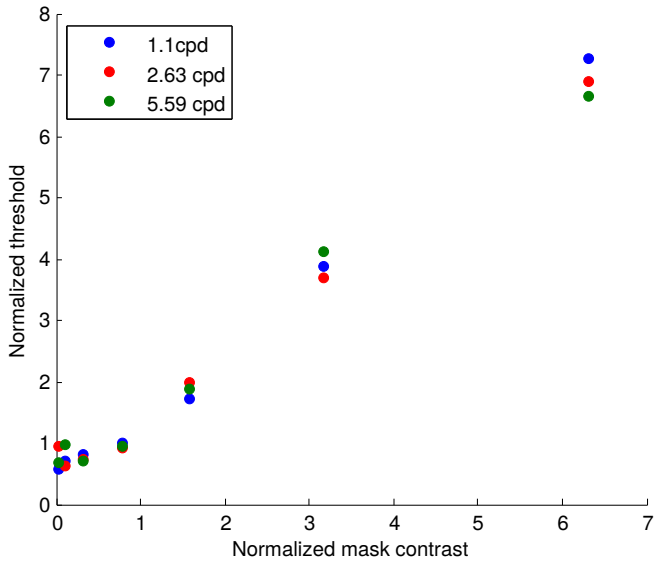


FIGURE B.2: Normalizing the contrast threshold by the corresponding CSF value allows us to measure the masking effects independently from the spatial frequency of the mask.

Results Figure B.2 shows the results of this preliminary experiment with both the mask contrast and measured threshold normalized by the corresponding CSF value. When the mask contrast is not visible (normalized value < 1) the measured threshold lies on a horizontal line while when the mask contrast increases beyond the visibility threshold (normalized value > 1) the measured threshold increases almost linearly. Normalizing the measured threshold by multiplying it with the corresponding CSF value makes the data from different frequencies very close to each other. Therefore, it would be possible to measure the threshold relative to the visual masking effect independently from the spatial frequency of the displayed stimulus, just as suggested by Daly in this seminal paper [Dal93].

Appendix C

Measuring the Dynamic Aspects of the Contrast Sensitivity Function

As detailed in Chapter 2, the threshold related to the contrast sensitivity aspect of the HVS is affected by the velocity at which an object in moving. However, in our experimental study (see Chapter 4) we did not take this effect into account. The reason is that we were more focused on computing the visibility threshold for static meshes rather than dynamic ones. Nevertheless, during the last few months of the thesis we have carried out a preliminary study that aims at measuring the contrast sensitivity for a moving 3D object.

C.1 Retinal Velocity and Eye Movement

Early physiological studies in human vision stated that the visibility threshold is affected by the retinal velocity of the object, *velocity of the object's image on the retina* [Kel79a, Kel79b]. These studies showed that as the retinal velocity increases the CSF curve is translated to the left. In other words, this means that the HVS becomes less sensitive to mid and high frequencies as a visual stimulus moves faster across the retina. Before measuring the CSF for a moving object, it is important to differentiate between the retinal velocity and the physical velocity of a moving object. In general, the HVS will try to track a moving object by following it with the eyes so that its image on the retina is as stable as possible. Therefore, it is important to take into account the movement of the eyes in the measurements of dynamic CSF. Having the value of the eye's velocity, the retinal velocity can be evaluated by the following:

$$v_r = ||v - v_{\text{eye}}||, \tag{C.1}$$

where v is the physical velocity of the object and v_{eye} is the eye's velocity.

There has been little attention given towards modeling the movement of the eyes throughout the literature of CSF modeling. Nevertheless, it is suggested that the eye's movements can be divided into three categories [Dit73]:

- The natural eye drift. This corresponds to the minimum velocity of the eyes. It has been shown that even if the eyes are intentionally fixated on a certain position they will be drifting at a very slow speed of about 0.15cpd/s [Kel79a].
- The smooth eye pursuit. This corresponds to movement of the eye when tracking a moving object.
- The saccadic movement. This corresponds to the almost instantaneous movements of the eye when the moving object is fast enough so that it cannot be tracked. It is considered that the contrast sensitivity during the saccadic movement is 0.

In [Dal01], Daly has proposed the following model to evaluate the velocity of the eyes:

$$v_{eye} = \min\left(\alpha \cdot v + v_{min}, v_{max}\right), \tag{C.2}$$

where α is the tracking efficiency and $v_{min} = 0.15$ deg/s, $v_{max} = 80$ deg/s are respectively the minimum eye velocity due to the drift and the maximum eye velocity before transitioning to saccadic movements. To the best of our knowledge, there have not been any consensus about the value of α in perceptual methods for video processing. Despite having a measured value of around 0.82 [Dal01], many algorithms for computing JND of a video use a value of 0.90 or 0.98 [JLK06, WN09, AvMS10]. In addition, Yee *et al.* [YPG01] adapted the value of α to the saliency value arguing that a salient object would attract the attention of an observer and therefore makes its tracking more efficient.

C.2 Experimental Study

It is clear that measuring the dynamic aspect of the Contrast Sensitivity Function (CSF), requires as well measuring the velocity at which the eye is moving. This would allow us to have a relatively accurate idea about the retinal velocity of the observed moving 3D object.

Experimental Protocol We have adapted the experimental protocol presented in Section 4.1 of the thesis to the dynamic setting. Instead of displaying on the screen two models side by side, we now display them one on top of the other. The reason for that is to allow a large horizontal movement of the 3D object. Similarly

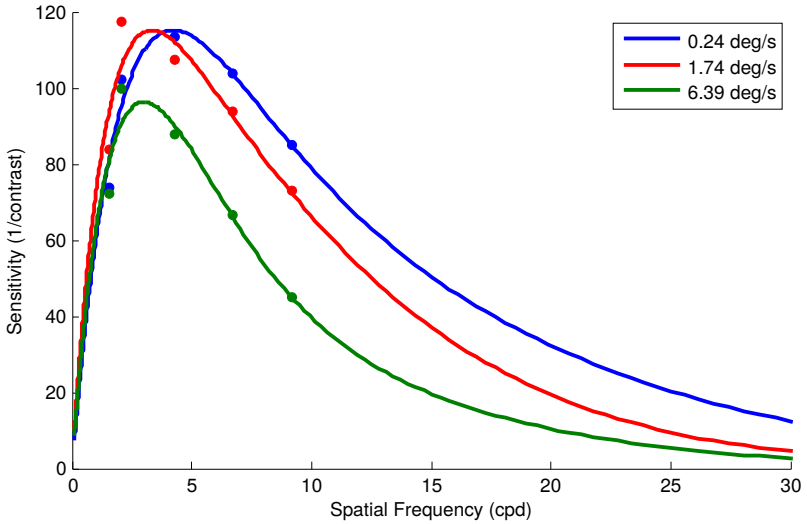


FIGURE C.1: As the retinal velocity of the 3D object increases, the contrast sensitivity is reduced and the CSF curve shits to the left.

to the static CSF measurements (see Section 4.3.4), one of the displayed planes exhibited a displaced vertex in its central area. The subjects answered by *Yes* or *No* to whether they noticed a difference between the displayed planes. The magnitude of the displaced vertex is then adjusted using the QUEST method. The CSF was measured for three different physical velocities (1 deg/s, 10 deg/s and 25 deg/s) corresponding to a slow, medium and relatively fast object movement. 5 subjects participated in these experiments which took place in a low illuminated laboratory environment (see Fig. 4.6). The head of the subjects was fixed on a head-stand to restrict its movement. An eye tracker was placed in front of the display to register the position where the observer is looking at. This will allow us to evaluate the eye movement for each subject and therefore compute the retinal velocity using Eq. (C.1).

Results We have used an eye tracker in order to measure the corresponding retinal velocity of the moving planes. Our data indicated that the observer's eyes were tracking the planes with an average velocity of 0.76 deg/s, 8.26 deg/s and 18.61 deg/s for respectively the 1 deg/s, 10 deg/s and 25 deg/s plane velocities. This leads to the retinal velocities of 0.24 deg/s, 1.74 deg/s and 6.39 deg/s. It is interesting to also note, that the average tracking efficiency deduced from these measurements is $\alpha = 0.76$ which is close to the value reported by Daly [Dal01]. We think that the difference lies within the nature of the task presented to the subjects. In Daly's experiments, the subjects were looking at one moving stimulus while in our case the subjects were looking at two planes in order to compare their appearances. The more complex task given to subjects in our preliminary

experiments might explain the slight reduction in the value of the tracking efficiency. The CSF results of these experiments are shown in Fig C.1. The plot shows the subjects' mean contrast sensitivity for each frequency for the three different retinal velocities. As expected, as the velocity increases the CSF shifts to the left; the CSF peak is shifted from about 4 cpd at 0.24 deg/s to about 2.9 at 6.39 deg/s. Moreover, for the relatively high velocity of 6.39 deg/s there is a considerable loss in contrast sensitivity observed.

Appendix D

Towards a Contrast-Based Perceptual Metric for 3D Meshes

As discussed in Chapter 3, the goal of many existing perceptual methods is to evaluate the perceptual quality of a distorted mesh. This could be a useful tool to evaluate and compare the performance of geometric algorithms. However, all existing methods rely on a purely geometric analysis for their perceptual evaluation which makes them not able to adapt to various display parameters. Therefore, we have also tried to use our perceptual analysis to evaluate the quality of a distorted mesh. In this method we have used the perceptual model that is suited for smooth shading as state-of-the-art subjective quality evaluation experiments were conducted by rendering the model with a smooth shading. The idea for evaluating the perceptual quality is relatively simple: Having a reference and a distorted model, we first evaluate the change in contrast (Eq. (4.14)) caused by the geometric distortion in a per-face basis. We then compute the per face threshold using Eq. (4.25). Computing the ratio between the change in contrast and threshold gives us a local distortion map. We then compute this distortion map for different light directions. In practice we have sampled the light direction from a sphere around the model. We finally aggregate the local distortion maps into a single score using a Minkowski pooling.

Table D.1 shows the Pearson correlation of our proposed contrast-based perceptual metric along with existing model-based methods (MSDM2 [Lav11], FMPD [WTM12] and DAME [VR12]) and image-based ones (IW-SSIM [WL11] and HDR-VDP2 [MKRH11]). We notice that compared to image-based methods our method performs better for all databases. However, it is not the case compared to model-based methods. It is clear that MSDM2 and FMPD are better when it comes to the *Masking* and *General Purpose* databases. A closer look at the results reveals some interesting insights that can explain the behavior of our method and would help improve its performance. In the *General Purpose* database there are two types of distortions: a *random noise* type and a *surface smoothing* type. By isolating the

	Masking [Lav09]	General Purpose [LDGD+06]	Compression [VR12]
ours	**0.76**	**0.75**	**0.91**
MSDM2 [Lav11]	0.87	0.81	0.89
FMPD [WTM12]	0.80	0.89	0.88
DAME [VR12]	0.58	0.75	0.93
IW-SSIM [WL11]	0.67	0.73	0.74
HDR-VDP2 [MKRH11]	0.52	0.67	0.84

TABLE D.1: Performance comparisons of our proposed perceptual quality metric prototype, existing model-based methods and image-based ones.

noisy models from the smoothed ones and computing the Pearson correlation scores of our method we get a correlation of 0.88 for the vertex noise type distortion and 0.80 for the smoothing one. Looking more closely at the scores given by our method we notice that there seems to be a difference in the range of the objective score between a vertex noise type distortion and a vertex smoothing type distortion. More precisely, it appears that the score for the surface smoothing distortion is being exaggerated. There are several hypotheses that could explain this behavior. First, our method is designed to handle local distortion types which is not the case for the smoothing distortion. A multi-scale contrast method might therefore help with this issue. Second and more importantly, in its current state our method is solely based on a measured difference of contrast. However as argued in [HP13] the quality attributed to an image is not just a function of the change in contrast but is also related to the clarity of information which is generally conveyed by its structure. Intuitively, a change in the visual regularity can be indicative of a change in structure. Indeed, upon further investigation, we notice that a weak vertex noise and a weak surface smoothing will have an almost similar effect on the local contrast, but only the vertex noise will cause a change in the structure of the surface. Therefore, we think that it would be interesting to test whether integrating the visual regularity to our method in a similar fashion a a SSIM-based method [WBSS04, WL11] would improve the performance of this quality metric. As for the *Compression* database our method performs better, but remains slightly behind DAME. The reason for this good performance in this case comes from the magnitude of the compression distortion included in this database. In fact, we have noticed that a large number of meshes in that database has a compression distortion that is barely visible, *i.e.*, close to the visibility threshold. This explains the high performance of our perceptual model as it is designed to compute the visibility threshold and therefore should be accurate in comparing distortions whose magnitude is close to it.

Appendix E

Résumé en Français

Les objets 3D deviennent de plus en plus utilisés dans des applications diverses allant des jeux vidéo à la visualisation scientifique. Avant que ces objets 3D soient affichés sur un écran, ils sont souvent soumis à plusieurs opérations géométriques. Par exemple, dans les applications interactives, les modèles géométriques détaillés sont simplifiés afin de garantir un affichage rapide des données 3D. Cependant, ces traitements géométriques introduisent des distorsions sous forme de perturbation des positions des sommets. Vu que ces distorsions peuvent affecter la qualité visuelle de l'objet 3D affiché, il est donc important de quantifier leur visibilité. La difficulté de ce problème est que la visibilité des distorsions géométriques sur un objet ne peut pas être déterminée en fonction de leur magnitude vue qu'il y a plusieurs paramètres externes qui entrent en jeu comme : illumination de la scène, résolution et taille de l'écran, point de vue, etc. Cela rend les métriques géométriques traditionnelles, comme la RMS ou la distance de Haussdorf, inefficace pour accomplir cette tâche. Le but du travail effectué dans cette thèse est d'évaluer le seuil de visibilité d'une distorsion géométrique. En d'autres termes, on cherche à évaluer le déplacement maximal qu'un sommet peut tolérer avant de devenir visible tout en prenant en compte les différents paramètres qui contribuent à la visibilité de ce déplacement.

Au cours des deux dernières décennies, les méthodes perceptuelles sont devenues de plus en plus populaires dans la communauté graphique [CLL+13]. Ces méthodes ont prouvé leurs utilités dans l'évaluation de la qualité visuelle des modèles 3D [CDGEB07, Lav11, VR12, TWC14] et pour optimiser des processus de plusieurs traitements géométriques comme la compression [MVBH15] et la simplification [LVJ05, QM08, MG10]. Les techniques perceptuelles existantes se basent sur des hypothèses sur le comportement global du système visuel en observant des objets 3D. Par exemple, il est accepté dans la communauté graphique que les artefacts visuels sont moins visibles dans les régions rugueuses que dans les régions lisses d'un maillage 3D [Lav09]. Plusieurs méthodes ont donc été développées pour estimer la rugosité d'une surface en utilisant des propriétés

géométriques comme la courbure [Lav09, WTM12]. Puisque les méthodes perceptuelles actuelles basent leur analyse sur des propriétés géométriques, sans signification perceptuelle formellement justifiée, il est donc difficile d'adapter le résultat de ces méthodes aux différents facteurs affectant la perception des maillages 3D (taille, densité, illumination, affichage, ...).

Différent de toutes les méthodes actuelles, on a décidé d'adopter une approche qui se base sur une analyse perceptuelle où on cherche à modéliser le fonctionnement interne du système visuel humain relatif à la visibilité des stimuli visuels afin d'atteindre notre but. Ce type de méthodes a été prouvé être efficace en traitement d'image [Dal93, LK11]. D'une manière simplifiée, les rayons lumineux provenant d'un objet arrivent dans l'œil puis passent au cerveau à travers la rétine et le nerf optique [Wan95]. Ce qui est intéressant et important dans ce processus concernant notre but est ce qui se passe au niveau de la rétine. Plus précisément, la structure physiologique des cellules ganglionnaires de la rétine fait en sorte que le système visuel humain soit plutôt sensible aux stimuli visuels ayant une large variation de luminance. Cette variation de luminance est généralement exprimée en matière de contraste. Donc plus le contraste est fort, plus les détails d'un stimulus visuel seront visibles. Cela fait que l'étude des propriétés reliées à la perception du contraste est essentielle au calcul du seuil de visibilité. En particulier, on se base sur les deux aspects suivants du système visuel humain :

- **La sensibilité au contraste (CSF)** qui indique le seuil de contraste à partir duquel un stimulus visuel devient visible [CK66]. Notamment, le seuil de visibilité dépend principalement de la fréquence spatiale et la luminance globale du stimulus observé.
- **Le masquage visuel** qui reflète la capacité du système visuel humain de distinguer entre deux stimuli visibles. Cela est principalement en fonction de la valeur du contraste de ces stimuli [LF80] et leur régularité visuelle [WBT97].

Une distorsion géométrique locale consiste à faire déplacer un sommet du maillage dans une certaine direction. Ce déplacement cause un changement de la direction des normales des faces adjacentes ce qui provoque un changement des propriétés perceptuelles locales, à savoir le contraste et la fréquence spatiale sur un maillage. En se basant sur l'analyse des aspects du système visuel humain présentés ci-dessus, on a proposé dans cette thèse une analyse perceptuelle qui a pour but de calculer le seuil de visibilité d'une distorsion géométrique. En

d'autre terme, on a présenté un algorithme qui a pour but de calculer le déplacement maximal qu'un sommet peut tolérer avant que ce déplacement devienne visible. Notre méthode se résume de la manière suivante :

1. On évalue les attributs perceptuels (contraste, fréquence, luminance globale et régularité visuelle) relatifs à l'étude de la visibilité d'un stimulus visuel sur un maillage (Sections 4.2.1 et 4.3.1).

2. Les attributs perceptuels ainsi calculés, sont utisés comme entrés à un modèle perceptuel qui calcule le seuil de visibilité et qui est obtenu suite à une étude expérimentale (Sections 4.2.3 et 4.3.4). Le modèle perceptuel proposé prend en considération la sensibilité au contraste et le masquage visuel.

3. Ayant calculé le seuil de visibilité, on a proposé un algorithme efficace afin d'évaluer le déplacement maximal qu'un sommet peut tolérer (Chapitre 5). Le seuil de déplacement d'un sommet ainsi calculé s'adapte aux différents facteurs pouvant affecter la visibilité des distorsions dans un environnement 3D (distance au point de vue, illumination,...).

Ayant réussi à calculer le seuil à partir duquel le déplacement d'un sommet devient visible, on a ensuite intégré l'analyse perceptuelle proposée dans différents algorithmes de traitement géométrique afin de montrer son potentiel (Chapitre 6). Premièrement, on a utilisé le seuil de visibilité calculé pour étudier la visibilité d'un "edge collapse" lors de la simplification d'un maillage 3D. Ceci nous a permis de proposer un coût perceptuel de simplification qu'on a utilisé pour privilégier la simplification des arêtes n'ayant pas un effet perceptuel sur la qualité visuelle du maillage. En plus, ceci a encore permis de proposer un critère d'arrêt automatique à l'opération de simplification ce qui a un large potentiel dans la gestion des différents niveaux de détails d'un objet 3D. Deuxièmement, on a utilisé le modèle perceptuel proposé pour définir un critère de subdivision adaptative de maillage 3D. L'idée ici est de subdiviser une face seulement si le contraste local de celle-ci est visible. Vue la visibilité d'un contraste est en fonction de la fréquence spatiale, notre méthode est capable d'adapter le degré de subdivision en fonction de la distance au point de vue et de l'illumination d'une manière automatique.

Pour résumer, dans cette thèse, nous avons proposé une méthode perceptuelle qui permet de calculer le déplacement maximal qu'un sommet puisse tolérer avant qu'il ne devienne visible pour un observateur humain. Différent de toutes les méthodes existantes en informatique graphique, notre approche se base sur un modèle perceptuel provenant d'une étude expérimentale sur le système visuel humain lors de l'observation d'un maillage 3D. Nous avons ensuite intégré

ce modèle dans plusieurs algorithmes de traitements géométriques pour montrer l'utilité d'une telle analyse perceptuelle en informatique graphique. Cependant, dans son état actuel, notre méthode est limitée à l'étude de la visibilité des distorsions géométriques locales se trouvant sur des maillages éclairés par une lumière simple et affichés à l'aide d'un rendu "flat" ou "smooth" basique. Cela est dû aux algorithmes simplifiés utilisés pour estimer les propriétés perceptuelles locales sur un maillage comme le contraste ou la fréquence. L'extension de notre analyse perceptuelle pour les cas complexes nécessite donc une généralisation des méthodes d'évaluation des propriétés perceptuelles. Un autre point important est que notre modèle perceptuel proposé est plutôt adapté à l'étude de la visibilité des distorsions locales et n'est pas capable d'évaluer son impact perceptuel une fois cette distorsion est au-dessus du seuil de visibilité. Cela vient du fait que notre analyse perceptuelle se base sur des propriétés de bas niveau du système visuel humain comme la sensibilité au contraste et le masquage visuel. Il serait donc intéressant d'étendre notre modèle aux aspects perceptuels de haut niveau pour permettre une analyse plus rigoureuse des distorsions visibles.

Bibliography

[ASCE02] N. Aspert, D. Santa-Cruz, and T. Ebrahimi. MESH: Measuring error between surfaces using the Hausdorff distance. In *Proc. of IEEE International Conference on Multimedia & Expo*, pages 705–708, 2002.

[AvMS10] Tunç Ozan Aydin, Martin Čadík, Karol Myszkowski, and Hans-Peter Seidel. Video quality assessment for computer graphics applications. *ACM Transactions of Graphics.*, 29(6):161:1–161:12, 2010.

[BA08] Tamy Boubekeur and Marc Alexa. Phong tessellation. *ACM Transactions of Graphics*, 27(5):141, 2008.

[Bar89] Peter G J Barten. The Square Root Integral (SQRI): A new metric to describe the effect of various display parameters on perceived image quality. In *Proc. of SPIE*, volume 1077, pages 73–82, 1989.

[Bar99] Peter GJ Barten. *Contrast Sensivity of the Human Eye and its Effects on Image Quality*. SPIE Publications, 1999.

[BC69a] C T Blakemore and Fergus W Campbell. Adaptation to spatial stimuli. *The Journal of Physiology*, 200(1):11P, 1969.

[BC69b] C T Blakemore and Fergus W Campbell. On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *The Journal of Physiology*, 203(1):237–260, 1969.

[BLBI13] Azeddine Beghdadi, M-C Larabi, Abdesselam Bouzerdoum, and Khan M Iftekharuddin. A survey of perceptual image processing methods. *Signal Processing: Image Communication*, 28(8):811–831, 2013.

[CB05] Irene Cheng and Pierre Boulanger. A 3D perceptual metric using just-noticeable-difference. In *Proc. of Eurographics Short Papers*, pages 97–100, 2005.

[CDGEB07] Massimiliano Corsini, Elisa Drelie Gelasca, Touradj Ebrahimi, and Mauro Barni. Watermarked 3D mesh quality assessment. *IEEE Transactions on Multimedia*, 9(2):247–256, 2007.

[CHM+12] Martin Cadík, Robert Herzog, RafałMantiuk, Karol Myszkowski, and Hans-Peter Seidel. New measurements reveal weaknesses of image quality metrics in evaluating graphics artifacts. *ACM Transactions on Graphics*, 31(6):1–10, 2012.

[CK66] Fergus W Campbell and JJ Kulikowski. Orientational selectivity of the human visual system. *The Journal of Physiology*, 187(2):437, 1966.

[CKT+99] Thom Carney, Stanley A. Klein, Christopher W. Tyler, Amnon D. Silverstein, Brent Beutter, Dennis Levi, Andrew B. Watson, Adam J. Reeves, Anthony M. Norcia, Chien-Chung Chen, Walter Makous, and Miguel P. Eckstein. Development of an image/threshold database for designing and testing human vision models. volume 3644, pages 542–551, 1999.

[CL95] Chun-Hsien Chou and Yun-Chin Li. A perceptually tuned subband image coder based on the measure of just-noticeable-distortion profile. *IEEE Transactions on Circuits and Systems for Video Technology*, 5(6):467–476, 1995.

[CLL+13] Massimiliano Corsini, Mohamed-Chaker Larabi, Guillaume Lavoué, Oldřich Petřík, Libor Váša, and Kai Wang. Perceptual metrics for static and dynamic triangle meshes. *Computer Graphics Forum*, 32(1):101–125, 2013.

[Cor62] Tom N Cornsweet. The staircase-method in psychophysics. *The American journal of psychology*, 75(3):485–491, 1962.

[CR68] Fergus W Campbell and JG Robson. Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology*, 197(3):551, 1968.

[CRS98] P. Cignoni, C. Rocchini, and R. Scopigno. Metro: Measuring error on simplified surfaces. *Computer Graphics Forum*, 17(2):167–174, 1998.

[CS06] Ioan Cleju and Dietmar Saupe. Evaluation of supra-threshold perceptual metrics for 3d models. In *Proceedings of the 3rd symposium on Applied perception in graphics and visualization*, pages 41–44, 2006.

[CSD+09] Forrester Cole, Kevin Sanik, Doug DeCarlo, Adam Finkelstein, Thomas Funkhouser, Szymon Rusinkiewicz, and Manish Singh. How well do line drawings depict shape? *ACM Transactions of Graphics*, 28(3):28:1–28:9, 2009.

[CSYB06] Irene Cheng, Rui Shen, Xing Dong Yang, and Pierre Boulanger. Perceptual analysis of level-of-detail: The JND approach. In *Proc. of IEEE International Symposium on Multimedia*, pages 533–540, 2006.

[Dal93] Scott Daly. The visible differences predictor: An algorithm for the assessment of image fidelity. In Andrew B. Watson, editor, *Digital Images and Human Vision*, pages 179–206. MIT Press, Cambridge, MA, USA, 1993.

[Dal01] Scott Daly. *Engineering Observations from Spatiovelocity and Spatiotemporal Visual Models*, pages 179–200. Springer US, 2001.

[DFL+15] Lu Dong, Yuming Fang, Weisi Lin, Chenwei Deng, Ce Zhu, and Hock Soon Seah. Exploiting entropy masking in perceptual graphic rendering. *Signal Processing: Image Communication*, 33:1–13, 2015.

[DFLS14] Lu Dong, Yuming Fang, Weisi Lin, and Hock Soon Seah. Objective visual quality assessment for 3D meshes. In *Proc. of International Workshop on Quality of Multimedia Experience*, pages 1–6, 2014.

[DFRS03] Doug DeCarlo, Adam Finkelstein, Szymon Rusinkiewicz, and Anthony Santella. Suggestive contours for conveying shape. *ACM Transactions on Graphics*, 22(3):848–855, 2003.

[DGECB05] E. Drelie Gelasca, T. Ebrahimi, M. Corsini, and M. Barni. Objective evaluation of the perceptual quality of 3D watermarking. In *Proc. of IEEE International Conference on Image Processing*, volume 1, pages 241–244, 2005.

[Dit73] Robert William Ditchburn. *Eye-movements and visual perception.* Clarendon Press, Cambridge, UK, 1973.

[DM08] Zhigang Deng and Xiaohan Ma. Perceptually guided expressive facial animation. In *Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 67–76, 2008.

[DPF03] Reynald Dumont, Fabio Pellacini, and James A Ferwerda. Perceptually-driven decision theory for interactive realistic rendering. *ACM Transactions on Graphics*, 22(2):152–181, 2003.

[FDK09] Karl J Friston, Jean Daunizeau, and Stefan J Kiebel. Reinforcement learning or active inference? *Public Library of Sicence One*, 4(7):e6421, 2009.

[Fer03] James A. Ferwerda. Three varieties of realism in computer graphics. In *Proc. SPIE*, volume 5007, pages 290–297, 2003.

[Fer08] James A. Ferwerda. Psychophysics 101: How to run perception experiments in computer graphics. In *ACM SIGGRAPH 2008 Classes*, pages 87:1–87:60, 2008.

[Fie87] David J. Field. Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A*, 4(12):2379–2394, 1987.

[FKH06] Karl Friston, James Kilner, and Lee Harrison. A free energy principle for the brain. *Journal of Physiology*, 100(1):70–87, 2006.

[Fle14] Roland W. Fleming. Visual perception of materials and their properties. *Vision Research*, 94:62 – 75, 2014.

[Fri10] Karl Friston. The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2):127–138, 2010.

[FS09] Roland W. Fleming and Manish Singh. Visual perception of 3D shape. In *Proc. of ACM SIGGRAPH Courses*, pages 24:1–24:94, 2009.

[FSPG97] James A. Ferwerda, Peter Shirley, Sumanta N. Pattanaik, and Donald P. Greenberg. A model of visual masking for computer graphics. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, pages 143–152, 1997.

[GH97] Michael Garland and Paul S Heckbert. Surface simplification using quadric error metrics. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pages 209–216, 1997.

[Gre93] David M Green. A maximum-likelihood method for estimating thresholds in a yes–no task. *The Journal of the Acoustical Society of America*, 93(4):2096–2105, 1993.

[GS66] David M. Green and John A. Swets. *Signal Detection Theory and Psychophysics*. Wiley, 1966.

[HFM16] V. Havran, J. Filip, and K. Myszkowski. Perceptually motivated brdf comparison using single image. *Computer Graphics Forum*, 35(4):1–12, 2016.

[HHL+10] Fang Hou, Chang-Bing Huang, Luis Lesmes, Li-Xia Feng, Liming Tao, Yi-Feng Zhou, and Zhong-Lin Lu. qcsf in clinical application: efficient characterization and classification of contrast sensitivity functions in amblyopia. *Investigative ophthalmology & visual science*, 51(10):5365–5377, 2010.

[HHO04] Sarah Howlett, John Hamill, and Carol O'Sullivan. An experimental approach to predicting saliency for simplified polygonal models. In *Proceedings of the 1st Symposium on Applied Perception in Graphics and Visualization*, pages 57–64, 2004.

[HK15] Donghui Han and John Keyser. Effect of appearance on perception of deformation. In *Proceedings of the 14th ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, pages 37–44, 2015.

[HK16] Donghui Han and John Keyser. Effect of low-level visual details in perception of deformation. *Computer Graphics Forum*, 35:375–383, 2016.

[HP13] Andrew M. Haun and Eli Peli. Is image quality a function of contrast perception? In *IS&T/SPIE Electronic Imaging*, page 86510C–86510C, 2013.

[Int12] International Telecommunication Union. Rec. BT.500: Methodology for the Subjective Assessment of the Quality of Television Pictures, 2012.

[JLK06] Yuting Jia, Weisi. Lin, and Ashraf A. Kassim. Estimating just-noticeable distortion for video. *IEEE Transactions on Circuits and Systems for Video Technology*, 16(7):820–829, 2006.

[JW06] Frank Jäkel and Felix A Wichmann. Spatial four-alternative forced-choice method is the preferred psychophysical method for naïve observers. *Journal of Vision*, 6(11):13–13, 2006.

[Kel79a] D. H. Kelly. Motion and vision. i. stabilized images of stationary gratings. *J. Opt. Soc. Am.*, 69(9):1266–1274, Sep 1979.

[Kel79b] DH Kelly. Motion and vision. II. stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America*, 69(10):1340–1349, 1979.

[Kel83] D. H. Kelly. Spatiotemporal variation of chromatic and achromatic contrast thresholds. *Journal of Optical Society of America*, 73(6):742–750, 1983.

[KG00] Zachi Karni and Craig Gotsman. Spectral compression of mesh geometry. In *Proceedings of SIGGRAPH*, pages 279–286, 2000.

[KKK02] Sun-Jeong Kim, Soo-Kyun Kim, and Chang-Hun Kim. Discrete differential error metric for surface simplification. In *Proceedings of Pacific Graphics*, pages 276–283, 2002.

[KP04] David C Knill and Alexandre Pouget. The bayesian brain: the role of uncertainty in neural coding and computation. *Trends in Neurosciences*, 27(12):712–719, 2004.

[KSGV+94] P Ewen King-Smith, Scott S Grigsby, Algis J Vingrys, Susan C Benes, and Aaron Supowit. Efficient and unbiased modifications of the quest threshold method: theory, simulations, experimental evaluation and practical implementation. *Vision research*, 34(7):885–912, 1994.

[Lav09] Guillaume Lavoué. A local roughness measure for 3D meshes and its application to visual masking. *ACM Transactions on Applied Perception*, 5(4):1–23, 2009.

[Lav11] Guillaume Lavoué. A multiscale metric for 3D mesh visual quality assessment. *Computer Graphics Forum*, 30(5):1427–1437, 2011.

[LC10] 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.

[LDGD+06] Guillaume Lavoué, Elisa Drelie Gelasca, Florent Dupont, Atilla Baskurt, and Touradj Ebrahimi. Perceptually driven 3d distance metrics with application to watermarking. In *Proc. SPIE*, volume 6312, pages 63120L–63120L–12, 2006.

[LF80] Gordon E Legge and John M Foley. Contrast masking in human vision. *Journal of Optical Society of America*, 70(12):1458–1471, 1980.

[LH01] David Luebke and Benjamin Hallen. *Perceptually Driven Simplification for Interactive Rendering*, pages 223–234. London, UK, 2001.

[Lin06] Weisi Lin. Computational models for just-noticeable difference. In H R Wu and K R Rao, editors, *Digital Video Image Quality and Perceptual Coding*, pages 281–303. CRC Press, London, UK, 2006.

[LK11] Weisi Lin and C-C Jay Kuo. Perceptual visual quality metrics: A survey. *Journal of Visual Communication and Image Representation*, 22(4):297–312, 2011.

[LKW06] Zhen Liu, Lina J Karam, and Andrew B Watson. JPEG2000 encoding with perceptual distortion control. *IEEE Transactions on Image Processing*, 15(7):1763–1778, 2006.

[LLP+10] Anmin Liu, Weisi Lin, Manoranjan Paul, Chenwei Deng, and Fan Zhang. Just noticeable difference for images with decomposition model for separating edge and textured regions. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11):1648–1652, 2010.

[LLV16] Guillaume Lavoué, Mohamed-Chaker Larabi, and Libor Vasa. On the efficiency of image metrics for evaluating the visual quality of 3D models. *IEEE Transactions on Visualization and Computer Graphics*, 2016.

[LM15] Guillaume Lavoué and Rafał Mantiuk. *Quality Assessment in Computer Graphics*. Springer International Publishing, 2015.

[LO11] Michéal Larkin and Carol O'Sullivan. Perception of simplification artifacts for animated characters. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization*, pages 93–100, 2011.

[LT98] P. Lindstrom and G. Turk. Fast and memory efficient polygonal simplification. In *Proc. of IEEE Visualization Conference*, pages 279–286, 1998.

[LT00] Peter Lindstrom and Greg Turk. Image-driven simplification. *ACM Transactions on Graphics*, 19(3):204–241, 2000.

[Lub95] Jeffrey Lubin. A visual discrimination model for imaging system design and evaluation. *Vision models for target detection and recognition*, 2:245–357, 1995.

[LVJ05] Chang Ha Lee, Amitabh Varshney, and David W Jacobs. Mesh saliency. *ACM Transactions on Graphics*, 24(3):659–666, 2005.

[MG10] Nicolas Menzel and Michael Guthe. Towards perceptual simplification of models with arbitrary materials. *Computer Graphics Forum*, 29(7):2261–2270, 2010.

[Mic27] Albert Abraham Michelson. *Studies in optics*. University of Chicago Press, 1927.

[MKRH11] Rafat Mantiuk, Kil Joong Kim, Allan G Rempel, and Wolfgang Heidrich. HDR-VDP-2: a calibrated visual metric for visibility and quality predictions in all luminance conditions. *ACM Transactions on Graphics*, 30(4):40:1–40:13, 2011.

[MMBH10] Ann McNamara, Katerina Mania, Marty Banks, and Christopher Healey. Perceptually-motivated graphics, visualization and 3d displays. In *ACM SIGGRAPH 2010 Courses*, pages 7:1–7:159, 2010.

[MNO07] Rachel McDonnell, Fiona Newell, and Carol O'Sullivan. Smooth movers: Perceptually guided human motion simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pages 259–269, 2007.

[MS74] James Mannos and David J Sakrison. The effects of a visual fidelity criterion of the encoding of images. *IEEE Transactions on Information Theory*, 20(4):525–536, 1974.

[MT86] Melanie J Mayer and Christopher W Tyler. Invariance of the slope of the psychometric function with spatial summation. *Journal of Optical Society of America A*, 3(8):1166–1172, 1986.

[MVBH15] Stefano Marras, Libor Váša, Guido Brunnett, and Kai Hormann. Perception-driven adaptive compression of static triangle meshes. *Computer-Aided Design*, 58:24–33, 2015.

[OHM+04] Carol O'Sullivan, Sarah Howlett, Yann Morvan, Rachel McDonnell, and Keith O'Conor. Perceptually adaptive graphics. In *Eurographics state of the art reports*, volume 4, pages 1–24, 2004.

[Pal99] Stephen E Palmer. *Vision Science: Photons to Phenomenology*. MIT Press, Cambridge, MA, USA, 1999.

[PB13] Denis G Pelli and Peter Bex. Measuring contrast sensitivity. *Vision Research*, 90:10–14, 2013.

[Pel90] Eli Peli. Contrast in complex images. *Journal of the Optical Society of America A*, 7:2032–2040, 1990.

[QM08] Lijun Qu and Gary W. Meyer. Perceptually guided polygon reduction. *IEEE Transactions on Visualization and Computer Graphics*, 14(5):1015–1029, 2008.

[RBD06] Szymon Rusinkiewicz, Michael Burns, and Doug DeCarlo. Exaggerated shading for depicting shape and detail. *ACM Transactions of Graphics*, 25(3):1199–1205, 2006.

[Red97] Martin Reddy. *Perceptually modulated level of detail for virtual environments*. PhD thesis, University of Edinburgh, UK, 1997.

[RFWB07] Ganesh Ramanarayanan, James Ferwerda, Bruce Walter, and Kavita Bala. Visual equivalence: towards a new standard for image fidelity. *ACM Transactions on Graphics*, 26(3):76:1–76:11, 2007.

[RPG99] Mahesh Ramasubramanian, Sumanta Pattanaik, and Donald Greenberg. A perceptually based physical error metric for realistic image synthesis. In *Proc. of ACM SIGGRAPH*, pages 73–82, 1999.

[RR01] Bernice E. Rogowitz and Holly E. Rushmeier. Are image quality metrics adequate to evaluate the quality of geometric objects? In *Proc. of SPIE Human Vision and Electronic Imaging*, pages 340–348, 2001.

[SB06] Hamid R Sheikh and Alan C Bovik. Image information and visual quality. *IEEE Transactions on Image Processing*, 15:430–444, 2006.

[SCOT03] Olga Sorkine, Daniel Cohen-Or, and Sivan Toledo. High-pass quantization for mesh encoding. In *Proceeding of Symposium on Geometry Processing*, pages 42–51, 2003.

[SD04] Anthony Santella and Doug DeCarlo. Visual interest and npr: An evaluation and manifesto. In *Proceedings of the 3rd International Symposium on Non-photorealistic Animation and Rendering*, pages 71–150, 2004.

[Sim05] Eero P. Simoncelli. *Statistical Modeling of Photographic Images*, pages 431–441. Elsevier Inc., 2005.

[SLMR14] Ran Song, Yonghuai Liu, Ralph R. Martin, and Paul L. Rosin. Mesh saliency via spectral processing. *ACM Transactions on Graphics*, 33(1):1–17, 2014.

[SSB06] Hamid R Sheikh, Muhammad F Sabir, and Alan C Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on image processing*, 15(11):3440–3451, 2006.

[Tau95] Gabriel Taubin. A signal processing approach to fair surface design. In *Proceedings of SIGGRAPH*, pages 351–358, 1995.

[TC67] MiM Taylor and C Douglas Creelman. Pest: Efficient estimates on probability functions. *The Journal of the Acoustical Society of America*, 41(4A):782–787, 1967.

[TFCRS11] William Thompson, Roland Fleming, Sarah Creem-Regehr, and Jeanine Kelly Stefanucci. *Visual Perception from a Computer Graphics Perspective*. A. K. Peters, Ltd., June 2011.

[Tre95] Bernhard Treutwein. Adaptive psychophysical procedures. *Vision research*, 35(17):2503–2522, 1995.

[TWC14] Fakhri Torkhani, Kai Wang, and Jean-Marc Chassery. A curvature-tensor-based perceptual quality metric for 3D triangular meshes. *Machine Graphics & Vision*, 23(1–2):59–82, 2014.

[VR12] Libor Váša and Jan Rus. Dihedral angle mesh error: a fast perception correlated distortion measure for fixed connectivity triangle meshes. *Computer Graphics Forum*, 31(5):1715–1724, 2012.

[WA05] Andrew B. Watson and Albert J. Ahumada. A standard model for foveal detection of spatial contrast ModelFest experiment. *Journal of Vision*, 5(9):717–740, 2005.

[Wan95] Brian A Wandell. *Foundations of Vision*. Sinauer Associates, Sunderland, MA, USA, 1995.

[Wat87] Andrew B Watson. The cortex transform: rapid computation of simulated neural images. *Computer vision, graphics, and image processing*, 39(3):311–327, 1987.

[WB02] Zhou Wang and Alan C Bovik. A universal image quality index. *IEEE signal processing letters*, 9(3):81–84, 2002.

[WB06] Zhou Wang and Alan C Bovik. Modern image quality assessment. *Synthesis Lectures on Image, Video, and Multimedia Processing*, 2(1):1–156, 2006.

[WBSS04] Zhou 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.

[WBT97] Andrew B Watson, Robert Borthwick, and Mathias Taylor. Image quality and entropy masking. In *Electronic Imaging'97*, pages 2–12. International Society for Optics and Photonics, 1997.

[Wei51] Waloddi Weibull. Wide applicability. *Journal of Applied Mechanics*, 103:33, 1951.

[WHTS01] Jian-Hua Wu, Shi-Min Hu, Chiew-Lan Tai, and Jia-Guang Sun. An effective feature-preserving mesh simplification scheme based on face constriction. In *Proceedings od Pacific Graphics*, pages 12–21, 2001.

[WL65] GB Wetherill and H Levitt. Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 18(1):1–10, 1965.

[WL11] Zhou Wang and Qiang Li. Information content weighting for perceptual image quality assessment. *IEEE Transactions on Image Processing*, 20(5):1185–1198, 2011.

[WLC+03] Nathaniel Williams, David Luebke, Jonathan D. Cohen, Michael Kelley, and Brenden Schubert. Perceptually guided simplification of lit, textured meshes. In *Proc. of ACM Symposium on Interactive 3D Graphics*, pages 113–121, 2003.

[WN09] Zhenyu Wei and K.N. Ngan. Spatio-temporal just noticeable distortion profile for grey scale image/video in DCT domain. *IEEE Transactions on Circuits and Systems for Video Technology*, 19(3):337–346, 2009.

[WP83] Andrew B. Watson and Denis G. Pelli. QUEST: a Bayesian adaptive psychometric method. *Perception & Psychophysics*, 33(2):113–120, 1983.

[WS04] Zhou Wang and Eero P Simoncelli. Local phase coherence and the perception of blur. In *Advances in Neural Information Processing Systems*, pages 1435–1442, 2004.

[WS05] Zhou Wang and EP Simoncelli. Translation insensitive image similarity in complex wavelet domain. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 2, pages 573–576, 2005.

[WSL+13] Jinjian Wu, Guangming Shi, Weisi Lin, Anmin Liu, and Fei Qi. Just noticeable difference estimation for images with free-energy principle. *IEEE Transactions on Multimedia*, 15(7):1705–1710, 2013.

[WSZL13] Jinliang Wu, Xiaoyong Shen, Wei Zhu, and Ligang Liu. Mesh saliency with global rarity. *Graphical Models*, 75(5):255–264, 2013.

[WTM12] Kai Wang, Fakhri Torkhani, and Annick Montanvert. A fast roughness-based approach to the assessment of 3D mesh visual quality. *Computers & Graphics*, 36(7):808–818, 2012.

[YPG01] Hector Yee, Sumanita Pattanaik, and Donald P Greenberg. Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Transactions on Graphics*, 20(1):39–65, 2001.

[ZCZ+11] Yin Zhao, Zhenzhong Chen, Ce Zhu, Yap-Peng Tan, and Lu Yu. Binocular just-noticeable-difference model for stereoscopic images. *IEEE Signal Processing Letters*, 18(1):19–22, 2011.

[ZDL02] Wenjun Zeng, Scott Daly, and Shawmin Lei. An overview of the visual optimization tools in jpeg 2000. *Signal Processing: Image Communication*, 17(1):85–104, 2002.

[ZZDZ10] Qing Zhu, Junqiao Zhao, Zhiqiang Du, and Yeting Zhang. Quantitative analysis of discrete 3d geometrical detail levels based on perceptual metric. *Computers & Graphics*, 34(1):55–65, 2010.

Author's Papers

Papers Related to the Thesis

International Journals

[1] NADER G., WANG K., HÉTROY-WHEELER F., DUPONT F.: visual contrast sensitivity and discrimination for 3D meshes and their applications. *Computer Graphics Forum*, vol. 35, no. 7, pp. 497-506, 2016 (Pacific Graphics 2016 - Best Paper Award)

[2] NADER G., WANG K., HÉTROY-WHEELER F., DUPONT F.: just noticeable distortion profile for flat-shaded 3D mesh surfaces. *IEEE Transactions on Visualisation and Computer Graphics*, vol. 22, no. 22, pp. 2423-2436, 2016

National Conferences

[3] NADER G., WANG K., HÉTROY-WHEELER F., DUPONT F.: Modèle perceptuel pour la détection du seuil de visibilité des distorsions sur un maillage 3D. *28èmes journées de l'Association Française en Informatique Graphique (AFIG15)*, 25-27 Novembre 2015, Lyon (France) (3rd Best paper award).

Talks in Scientific Journeys

[4] NADER G., WANG K., HÉTROY-WHEELER F., DUPONT F.: Mesure de la différence perçue sur des maillages surfaciques. *De l'acquisition à la compression des objets 3D (AC3D'14)*, 11-13 Juin 2014, Ile de Porquerolles (France).

Other

International Journals

[5] NADER G., GUENNEBAUD G., MELLADO N: Adaptive multiscale analysis for point-based surface editing. *Computer Graphics Forum*, Vol. 33, no. 7, pp. 171-179 , 2014 (Pacific Graphics 2014)