Perceptual metrics for static and dynamic triangle meshes

M. Corsini[1], M.C. Larabi[2], G. Lavoué[3], O. Petřík[4], L. Váša[4], K. Wang[5]

[1]Visual Computing Laboratory, ISTI-CNR, Italy
[2]Université de Poitiers, CNRS, XLIM-SIC UMR 7252, France
[3]Université de Lyon, CNRS, Insa-Lyon, LIRIS UMR 5205, France
[4]Centre of Computer Graphics and Visualization, University of West Bohemia, Czech Republic
[5]GIPSA-lab, CNRS UMR5216, France

Abstract

Almost all mesh processing procedures cause some more or less visible changes in the appearance of objects represented by polygonal meshes. In many cases, such as mesh watermarking, simplification or lossy compression, the objective is to make the change in appearance negligible, or as small as possible, given some other constraints. Measuring the amount of distortion requires taking into account the final purpose of the data. In many applications, the final consumer of the data is a human observer, and therefore the perceptibility of the introduced appearance change by a human observer should be the criterion that is taken into account when designing and configuring the processing algorithms.

In this review, we discuss the existing comparison metrics for static and dynamic (animated) triangle meshes. We describe the concepts used in perception-oriented metrics used for 2D image comparison, and we show how these concepts are employed in existing 3D mesh metrics. We describe the character of subjective data used for evaluation of mesh metrics and provide comparison results identifying the advantages and drawbacks of each method. Finally, we also discuss employing the perception-correlated metrics in perception-oriented mesh processing algorithms.

Categories and Subject Descriptors (according to ACM CCS):   Models And Principles [H.1.2]: User/Machine Systems—Human Factors

1. Introduction

With technological advances in telecommunication, hardware design and multimedia, the use of 3D data is now well established in several industrial domains, like digital entertainment, scientific visualization, computer-aided design, architecture and many others. The 3D content is mostly represented by polygonal meshes, or sequences of polygonal meshes (i.e. dynamic meshes), which may be associated with colour information or texture maps. For its transmission, protection, visualization or manipulation, this 3D content is subject to a wide variety of processing operations such as compression, filtering, simplification, watermarking, and so forth. These operations introduce distortions which may alter the visual quality of the 3D content; this is a critical issue, as these processing operations are often targeted at human-centred applications with viewing as the intended use.

A main problem is that most existing processing algorithms (e.g. simplification, watermarking, compression) are driven and/or evaluated by simple metrics like Hausdorff distance and root mean square error (RMS), which are not correlated with human vision. For instance, the three distorted models on the right in Figure 1 are all associated with the same RMS distance from the original model (on the left); however, the respective visual quality of each of them is very different. Hence, some *objective* quality metrics have been introduced; their goal is to produce a score that predicts the *subjective* visual quality (or the visual impact of the distortion) of a distorted 3D model with respect to a reference (distortion-free) model. These objective scores should be statistically consistent with those of human observers. Such perception-oriented metrics are of major importance for Computer Graphics; they provide a whole new paradigm for the evaluation, control and optimisation of many kinds of processing operations.

In the field of 2D image processing, the research on objective quality assessment metrics is highly developed, and
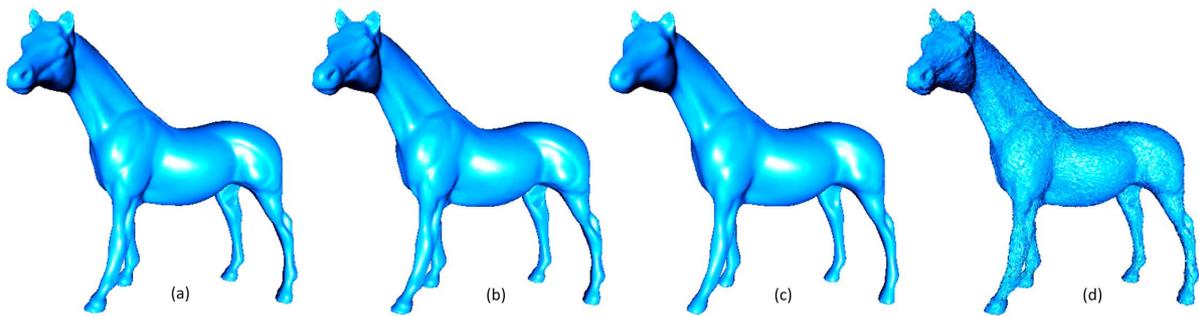
Figure 1: *Original and distorted versions of the Horse model, all associated with the same maximum root mean square error (MRMS = $1.05 \times 10{-3}$). (a) Original model. Results after (b) watermarking from Wang et al. [WLDB11], (c) Laplacian smoothing [Tau00], (d) Gaussian noise addition.*

some of the quality metrics for 3D meshes build on the concepts originally proposed in the context of image quality evaluation. Hence, the next section presents quality metrics and perceptually-related works for 2D images. Then, sections 3 and 4 respectively present these topics for 3D static and dynamic meshes. Section 5 attempts to evaluate and compare these metrics, while section 6 focuses on two applications for which perceptual metrics are highly relevant: compression and watermarking.

2. Human perception and metrics for 2D images

This section is constructed around two parts. In the first part, an overview is given of the main characteristics of human perception that have been widely exploited in recent years. It allows one to have a better understanding of the major phenomena of the Human Visual System (HVS), such as sensitivity to contrast, visual masking, and so on. The second part tackles the very active field of quality assessment of 2D images by highlighting the different families of metrics and the ways the HVS is integrated in the developed models.

2.1. Human perception

Understanding human perception and cognition, and modeling the Human Visual System (HVS) behavior is an essential step for developing image-based applications [Wan95]. This allows one to take advantage of the end-user perception in order to hide or highlight specific details and thus evaluate the perceived quality of an image or an image sequence [SPC04, Win02].

The HVS perceives a stimulus depending on its colour/intensity, orientation, and also on its spatial distribution. This important phenomenon caused by the visual cortex allows one to avoid capturing useless information (e.g. a white and black grating at a high spatial frequency will be seen as a grey stimulus). Figure 2 has been introduced by Campbell and Robson [CR68] to explain the phenomenon.

It represents a sine-wave stimulus varying in contrast on the y-axis and in spatial frequency on the x-axis. One can determine his own contrast sensitivity by identifying the different points beyond which the stimulus cannot be distinguished from the background.

Several works have been focused on the study of this characteristic, leading thus to the definition of the *contrast sensitivity function (CSF)* used in the construction of many algorithms (metrics) and systems in the imaging field. Generally, a band-pass filter characterises the luminance CSF with a peak frequency between 4 and 6 cpd (cycles by degree) and a cut-off around 30 cpd. One of the most popular analytical models was introduced by Mannos and Sakrison [MS74] in the 70s for the development of the first image quality metric for encoded monochrome images. Another simple model has been provided by Movshon and Kiorpes [MK88] as a three parameter exponential function. Daly proposed in his Visual Difference Predictor (VDP) [Dal93] a CSF model using several parameters, including radial spatial frequency (orientation), luminance levels, image size, image eccentricity, and viewing distance, allowing one to take into account a wide range of viewing conditions. This results in an anisotropic band-pass CSF giving more sensitivity to horizontal and vertical spatial frequencies in comparison to diagonal frequencies. Another complete and complex model has been pro-
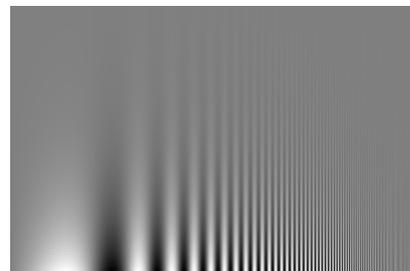

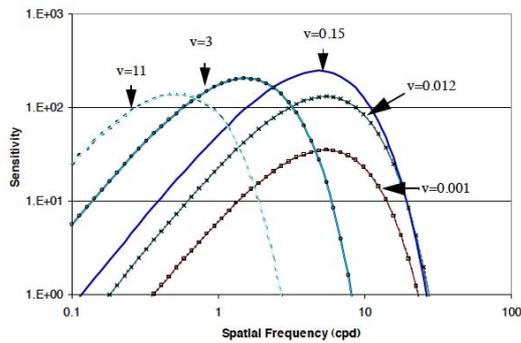**Figure 2:** *Campbell and Robson chart [CR68].*

Figure 3: *Velocity dependent Contrast Sensitivity Function as defined empirically by Kelly [Kel79].*

posed by Barten [Bar99] starting from a large amount of psychophysical data. It takes into account four parameters, including mean luminance, spatial frequency, stimulus size and pupil diameter. This flexibility of Barten's CSF comes with the price of model complexity and, in contrast to Daly's CSF, it is incapable of predicting orientation effects.

The previous CSF models are based on the detection threshold related to the detection of a stimulus by the HVS. However, when addressing quality assessment tasks, the HVS performs more than simple detection because it needs to discriminate between two stimuli (for a full reference evaluation) or between the provided stimulus and an implicit reference (for a no-reference evaluation). For those tasks, an estimation of the discrimination threshold for the construction of CSF is more appropriate. This was the focus of the work performed by Larabi et al. [LBF06], from which they proposed a CSF model constructed after extensive psychophysical experiments.

Few works have been dedicated to chromatic CSFs. It is admitted that chromatic mechanisms are of a low pass behavior with cutoff frequencies lower than those of luminance. This behavior is partially explained by the fact that edge detection/enhancement does not occur in the chromatic dimension [Fai05].

In addition to the sensitivity to spatial frequency, the human visual system is sensitive to motion. The CSF is thus dependent on the velocity of the stimulus, as demonstrated by Kelly [Kel79] through experiments measuring threshold contrast for viewing travelling sine waves. Figure 3 shows the variation of the CSF function of the velocity.

Visual masking defines the reduction in the visibility of one stimulus due to the simultaneous presence of another. This phenomenon is strongest when both stimuli have the same or similar frequency, orientation, and location [SPC04]. There are two types of visual masking. First, the luminance masking caused by the brightness sensitivity of the HVS, which is maximised on a distortion with a medium

background intensity and reduced when the distortion happens on a very low or very high intensity background. Second, the texture masking pointing out the maximised visibility of a stimulus on homogeneous regions rather than on textured ones [Wan95].

Visual masking has been widely used in image/video compression, watermarking, computer graphics, quality assessment and so on. For instance, Ferwerda et al. proposed a visual masking model allowing one to predict the influence of one visual pattern on another [FSPG97]. In a different field, Kutter et al. designed a vision-based masking model for spread-spectrum image watermarking [KW02]. Finally, Daly exploited both luminance and contrast masking for the definition of VDP [Dal93]. The developed models have been used by several authors in order to take into account this particularity of the HVS.

Another aspect of human perception that has been widely explored and used in the last few years is related to visual saliency. This property is fundamental in the exploration of the surrounding visual world. Analysis of visual attention is considered a very important element in human perception because of its suitability in various computer vision applications. Eye tracking is the main way to studyi and understand this property. One of the most famous and often-cited studies was performed by Yarbus in the 60s [Yar67]. The aim of this study was the illustration of differences in visual exploration paths during interpretation of a given scene. Therefore, Yarbus showed Ilya Repin's painting to several observers and assigned to them different viewing tasks. The visual paths of these observers is reported on Figure 4. Yarbus noted that the observation of stationary objects such as images, for example, translates into a sequence of saccades and fixations on key/interest points of the observed object. The eye moves between locations with the most information. The duration of a fixation is then proportional to the quantity of conveyed information. From Figure 4, one can notice that the visual path corresponding to a free exploration is different from the path obtained when subjects were asked to judge the material status of the family and which itself is different from the path when they were asked to guess the age of different individuals.

Visual saliency models try to mimic the human visual system in order to reproduce the saliency property on an image or a video sequence. Most of the proposed models in the literature are static and do not take into account motion which represent important information. There are two types of computational models for saliency depending on what the model is driven by: a bottom-up saliency using low-level features (e.g. contrast) [IKN98, BT06, HZ08, RvdLBC08, MPG*09] and a top-down saliency focusing on tasks/semantics [TOCH06, KTZC09]. Of course, top-down algorithms are more complex than bottom-up ones but they allow one to take into account high-level features such as faces and texts. Hence, it has been demonstrated recently
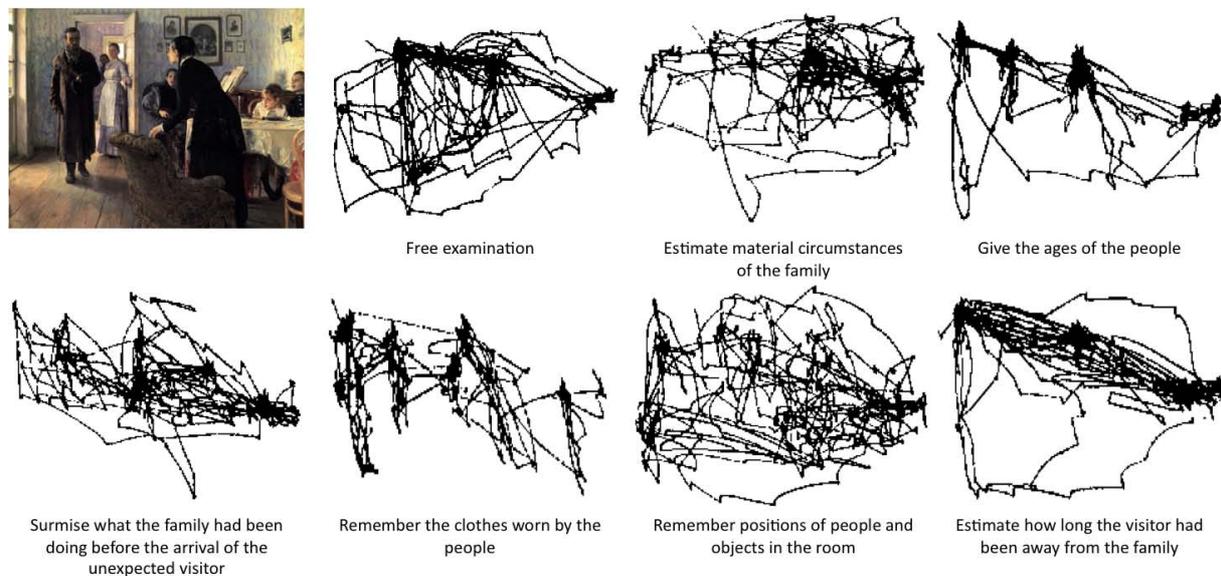
Figure 4: *Experiments performed by Yarbus [Yar67] on how the task given to a person influences the eye movement.*

that the latter attract the human gaze independently of the assigned task [CFK09].

Beyond what has been discussed already, the phenomena related to colour perception are confusing by their number and specificity. Hence, colour appearance has captured the attention of many researchers for decades. One of the major problems is related to the WYSIWYG (What You See Is What You Get) paradigm in many applications such as printing, textiles, multimedia, cinema, etc. Depending on the media, colour does not always seem similar, as, for example, the printed image could be different from the same displayed on a screen. The perception of a colour stimulus is partly dependent on the environment's properties, such as background colour and lighting conditions. In order to ensure the invariability of the perceived colour and its quality at the same time, the CIE (Commission International de l'Eclairage / International Lighting Commission) developed several models such as the CIE Lab [Sch07], CIECAM97 and CIECAM02 [Fai05], the most accomplished and stable one. The objective of a colour appearance model (CAM) is to correct the colour stimulus to ensure its invariability in different environments. This correction is illustrated in Figure 5 where the same colour stimuli are perceived as different (Figure 5-a) because of the use of different backgrounds. This phenomenon is known as simultaneous contrast and is corrected by CIECAM02 and the results allow one to have a closer colour stimulation (Figure 5-b).

In order to extend colour appearance models to images, Tulet et al. [TLF08] conducted psychophysical experiments to understand the effect of spatial frequencies on the appearance of a complex stimulus. This study resulted in a com-

putational model allowing the rendering of complex stimuli, i.e. images, based on their spatial structure.

There are other characteristics that can be taken into account in the framework of image quality assessment, such as luminance adaptation, simultaneous contrasts, temporal sensitivity, binocular rivalry/compensation [BL12] and so on.

2.2. Quality metrics for 2D images

Image quality assessment has attracted many researchers in the last decade. This has resulted in the development of hundreds of quality metrics for various applications and types of images. Generally, image quality metrics can be classified into three categories, including full-reference (FR), reduced reference (RR) and no-reference (NR), according to the availability of the original image [Kee02, WSB03]. FR and RR metrics require at the quality evaluation stage that full or partial information on both images is present, the reference and the distorted one. RR metrics are very challenging because they are used for applications where the original image is not available, as is the case of all transmission/broadcast systems. Several metrics have been proposed in recent years and are described in [NLF10, WS05]. NR metrics are distortion-based; the specialised metric looks for a specific artifact in the image and evaluates the level of annoyance introduced by that distortion without any cue of the original, as those described in [PLR*04, MK05, BS06].

In this section, the focus is put on the FR metrics because they are the most successful metrics and the most addressed ones. Several benchmarks have been made in order to study the performance of these metrics with regard to human judg-
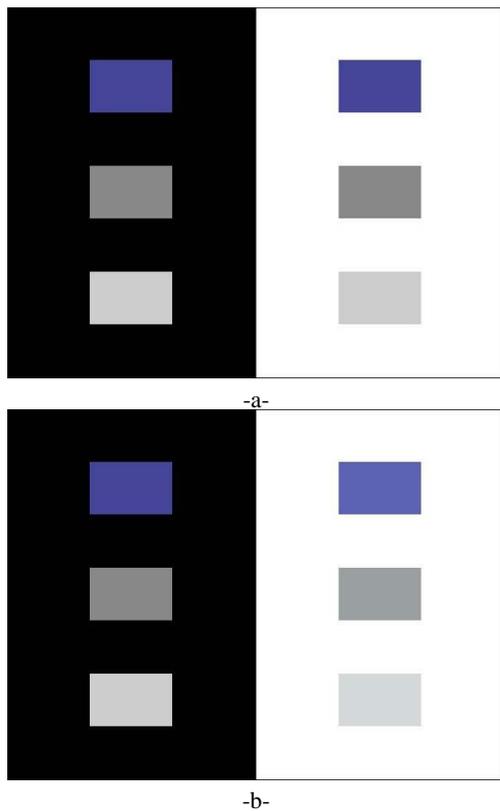
Figure 5: *Colour appearance problem: a- before CIECAM02 correction and b- after CIECAM02 correction.*

ment. An important effort has been made by the development of web-applications dedicated to FR metrics (http://www.qualimage.net), including a benchmark service, an online quality assessment using selected metrics and a documentation service [NLF11]. Before starting the description of the signal-oriented metrics, it is important to talk about the colour fidelity metrics. Several experiments, such as the one performed by Yendrikhovskij [Yen98], showed that in the cognitive process, the image quality is going through a kind of fidelity measurement between the colours of the scene and the memory colours like sky, grass, etc. Therefore, the CIE proposed two colour difference equations named ΔE94 and ΔE2000 respectively in 1994 and 2000 [BAR*91]. The second equation is more complex and takes into account more phenomena than the first one. Besides these colour difference equations, Zhang and Wandell proposed an extension to the CIELAB Δ*E* [ZW97] named s-CIELAB. After a colour transform into an antagonist colour space, each component of the image is filtered using a separable 2D spatial filter. Finally, an extension called x-CIELAB has been proposed [RLFM08], introducing a filter using an anisotropic contrast sensitivity function (CSF) built by estimating the perception threshold.

The signal-oriented metrics do not take into account any comprehensive HVS model with regard to quality evaluation. They are mainly based on a signal modeling process and are often application-dependent (compression, transmission, etc.). The metrics falling into this category are often suitable for real-time applications because of their low complexity. The most common simple metric is still the peak signal-noise ratio (PSNR) metric for the balanced compromise it provides between its complexity and performance. The often-cited SSIM (Structural SIMilarity) index, introduced by Wang and Bovik [WBSS04], exploits an important aspect of HVS perception linked to structural information. With a more theoretical definition, the VIF [SB06] has been developed as an extension of the information fidelity criterion (IFC) [SBd05] with the aim to quantify the loss of image information due to the distortion process and explore the relationship between image information and visual quality.

The second type of full-reference metrics uses a single-channel modeling of the human visual system (HVS). In this context, the HVS is seen as a spatial filter whose characteristics are given by the contrast sensitivity function (CSF), for example. The first metric developed under this framework is that of Mannos and Sakrison [MS74]. The principle of this metric is to weight the spectrum of the error image between the original image and the degraded one, using a CSF obtained from psychophysical experiments based on the detection of sinusoidal gratings. Chandler and Hemami proposed an efficient metric called VSNR for quantifying the visual fidelity of natural images based on near-threshold and supra-threshold properties of human vision [CH07]. It is based on visual masking and visual summation for detecting distortions and uses low-level features if it is beyond supra-thresholds.

The perceptual metrics represent an interesting approach in the evaluation of image quality. A summary of various studies carried out in this context shows that these metrics are modeled on the operation of the HVS and use the perceptual factors that are known to have a direct influence on the visibility of distortions [PS00]. A generic block diagram of these metrics is given in Figure 6.

The flowchart starts with a colour conversion allowing one to transpose both reference and impaired images into a perceptual colour space. At this point, an emphasis is usually placed on the luminance component because it is believed that the performance gain, generated by the consideration of colour, is far from balancing the complexity induced by the
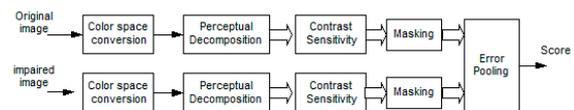

Figure 6: *Block diagram of perceptual metrics.*

processing of the chrominance channels. Then, a perceptual decomposition (multi-channel decomposition) is applied in order to take into account the spatial-frequency sensitivity of the HVS. The most used decompositions are those of Daly [Dal94], Lubin [Lub93] and Watson [Wat87] and the output of this block results in a set of luminance images. For each of these images, a local contrast is calculated at each point. The masking block aims to exploit the masking abilities of the HVS described in the previous section. Its role is to specify for each sub-band and for each point the variation of the visibility threshold when the masking effect is taken into account. These values allow one to keep only the errors located above the threshold and thus contributing to the estimation of the final quality. Finally, the pooling stage is designed to reduce this dimensionality of the computed data. Generally, the pooling is performed in two steps. The first where the error images spread across all frequency channels are combined into a single error image (the frequency pooling). The second step is dedicated to the spatial pooling and is to combine the spatial errors in a final measure that represents the score given by the algorithm to the impaired image.

The most representative examples of such a structure are the Visible Difference Predictor (VDP) introduced by Daly [Dal93] and the metric proposed by Karunasekera and Kingsbury [KK93]. Perceptual metrics, as described, are generic metrics. They can be used for any type of impairments and are known to have a high rate of correlation with subjective scores. However, the use of increasingly complex models of the HVS tends to increase their computational complexity.

3. Metrics for static 3D meshes

In this section we provide an overview of the perceptually-motivated metrics developed over the years by CG researchers for static 3D meshes. Mesh simplification, perceptually-driven rendering, and evaluation of specific geometry processing algorithms, such as compression and watermarking, are the main fields of applications of this type of perceptual metric. First we provide a discussion concerning some important issues about the properties of the perceptual metrics. Then, we review in detail the most important metrics developed during recent years.

3.1. View-dependent and view-independent metrics

It is convenient to categorise the perceptual metrics for static meshes in two well-separated categories: the *image-based* ones and the *geometry-based* ones. We found this categorisation very important since many times the domain where the perceptual metric works and the relative perceptual mechanisms involved are not sufficiently emphasised. The metrics which belong to the first category work in image space by applying the perceptual mechanisms of the HVS to a still image generated through rendering techniques from the 3D

data. This means that these metrics are *view-dependent*. Usually, where the view-dependency is a limit for the specific application, the image-based metrics are evaluated on a set of images created using different views of the 3D objects. We underline that this approach is not completely reliable, due to the fact that accurate perceptual studies conducted by Rogowitz and Rushmeier [RR01] demonstrated that, in general, the visual perception of a set of images of a certain 3D object is different from that perceived by a human observer of the 3D model in a graphics application.

In the second category, the perceptual metrics work by analyzing the geometry of the 3D models to predict perceptual impairments or evaluate other perceptual quality aspects, making the evaluation *view-independent*. Hence, the geometry of the model is the domain of this type of metric. In this case, more complex perceptual mechanisms are involved, such as the role of movement of the 3D object in a computer animation video or by user interaction. Also, the feedback between the movement of the object and the actions of the user should be considered, even if this aspect is in general not taken into account by the existing metrics.

3.2. Black-box and mechanistic approaches

As raised in section 2, there are two different approaches to developing perceptual metrics: *mechanistic* (i.e. HVS-based) and *black-box* (i.e. signal-oriented). The mechanistic approach takes into account the complex mathematical models of the psychophysical and physiological mechanisms of the HVS in order to develop the perceptual metric while the black-box approach does not rely on how the visual system works but attempts to define a function that, given the visual stimulus as input, is able to predict how much some specific visual artifacts will be perceived by a human observer; this approach is preferable when it is difficult to determine how to integrate the different visual stimuli involved. In Computer Graphics both mechanistic and black-box approaches have been used for the development of image-based perceptual metrics. The model-based perceptual metrics, instead, usually employ the black-box approach due to the complexity of the perceptual mechanisms involved.

3.3. Image-based perceptual metrics

Concerning *perceptually-based mesh simplification*, Lindstrom and Turk [LT00] proposed to render the model being simplified from several viewpoints and use a fast image quality metric to evaluate the impact of the simplification. Lindstrom [Lin00] proposed using a perceptual image metric based on a simplified version of the Sarnoff Model [Lub95]. The final result is a simplified 3D model which minimises the perceived visual differences in the image set. This algorithm is particularly effective for textured 3D models.

Luebke and Hallen [LH01] developed a perceptually-based simplification algorithm based on a simplified version

of the Contrast Sensitivity Function. They map the change resulting from a local simplification operation to a worst-case contrast and a worst-case frequency and then determine whether this operation will be imperceptible; their model also takes into account silhouette changes. Their method was then extended by Williams et al. [WLC*03] to integrate texture and lighting effects. These latter approaches are view-dependent; however, they consider the 3D geometry information.

More recently, Qu and Meyer [QM08] considered the visual masking properties of 2D texture maps to drive simplification and remeshing of textured meshes. Two perceptual metrics are proposed to evaluate the potential masking effect of the surface signals (textures, bump maps, etc); one based on the Sarnoff VDM [Lub95] and another based on the visual masking function employed to optimise the quantisation in the JPEG2000 [ZDL02]. The perceptually-based remeshing algorithm is driven by the masking map computed in the parametric space of the textures. Not only texture maps but also bump maps are taken into account for its calculation. The final remeshing can be view-independent or view-dependent depending on the visual effects considered. For example, specular reflection introduces a view-dependent effect. The simplification-driven algorithm takes into account an average masking importance map that emerges from the analysis of the 3D object from several viewpoints.

The objective of *perceptually-driven rendering* is to determine, according to the location of the observer, the amount of accuracy to use during the rendering, for example changing the Level Of Detail (LOD) of certain models or reducing/augmenting sampling density in ray-tracing rendering systems. One of the first studies of this kind was that of Reddy [Red97], which analysed the frequency content in several pre-rendered images to determine for each model the best LOD to use in a real-time rendering system.

Ferwerda et al. [FSPG97] proposed a perception model for Computer Graphics with particular attention to the masking model, extending the original Daly *VDP* operator, which demonstrates how surface texture can hide some visual artifacts of the geometry in given shading conditions, in particular polygonal tessellation.

Bolin and Meyer [BM98] used a perceptual model to optimise the sampling for ray-tracing algorithms. The visual differences operator developed by Bolin and Meyer is a simplified version of the Sarnoff VDM [Lub95], modified to take into account also the chromatic aberration effect to deal with colour images. This is achieved by considering the variations of chromatic/achromatic CSF. This operator is used to drive adaptively the sampling in a ray tracing framework.

Ramasubramanian et al. [RPG99] proposed a rendering framework to considerably reduce the overhead of incorporating a perceptual metric into a rendering system. First, they evaluated a perceptual *threshold map* taking into account the direct illumination of the scene and then this map is used to

add indirect illumination, which is usually the most computational expensive task in a global illumination rendering system.

Another interesting approach is that of Dumont et al. [DPF03], which proposed a real-time rendering system capable of optimising the performance in terms of image quality and frame rate, taking appropriate decisions. The proposed framework is based on a decision-theory approach. According to decision theory, it is possible to formalise the problem as maximising the utility of certain choices, i.e. rendering actions, given a set of constraints. Constraints take into account resource limitations. The rendering actions consider the approximation the system can make, such as choosing a LOD or deciding the resolution of a texture. The perceptually-based utility metrics used to select texture resolution and evaluate mesh elements for radiosity computation are based on the VDP version developed by Ramasubramanian et al. [RPG99] due to its accuracy and computational efficiency.

Recently, perceptual evaluation has been moved to a higher level of investigation concerning visual mechanisms. For example, Ramanarayanan et al. [RFWB07] proposed the new concept of *visual equivalence*; images are said to be visually equivalent if they convey the same impressions of scene appearance. In this work, the authors explore how the perception of geometry, material and illumination in a scene are affected by lighting environment changes.

Some other interesting methods have recently been proposed: Zhu et al. [ZZDZ10] studied the relationship between the viewing distance and the perceptibility of model details using 2D metrics (VDP and SSIM); Aydin et al. [AvMS10] and Herzog et al. [HCA*12] introduce quality metrics focusing respectively on video and images synthetically generated by computer graphics rendering techniques. Finally, Bosc et al. [BPL*11] introduce a quality metric for synthesised views generated from depth-image-based rendering (3DTV application).

3.4. Model-based perceptual metrics

The main limitation of the image-based metrics in the context of Computer Graphics applications is that, as just previously mentioned, in general, as demonstrated by the experiments conducted by Rogowitz and Rushmeier [RR01], the perceived degradation of still images may not be adequate to evaluate the perceived degradation of the equivalent 3D model. In their work, they demonstrated that the subjects evaluated differently the quality of a simplified 3D model if an animation or a set of static frames of the same animation were used. The main reason is that the object's movement introduces changes in the perception of differences that are difficult to integrate in the perceptual metric.

One of the first attempts to integrate image movement, visual attention and saliency was the work of Yee et

al. [YPG01], which combined the many aspects in a final map called the *aleph map*, used during the rendering of the computer animation. Myszkowski [Mys02] proposed an extension of the VDP for quality evaluation of computer-generated animations and applied such metrics to speed-up global illumination rendering. The application of these spatio-temporal perceptual metrics in the context of 3D model visual fidelity evaluation has, to our knowledge, never been investigated. This is an interesting direction for future research in object-based perceptual metrics.

Model-based metrics are used in different contexts. One of these is to control mesh simplification algorithms, in order to reduce the number of vertices while preserving the visual appearance. Kim et al. [KKK02] stated that human vision is sensitive to curvature changes and proposed a *Discrete Differential Error Metric* (DDEM). In a different way, Howlett et al. [HHO04] drove their simplification to emphasise visually salient features determined through an eye tracking system. Lee et al. [LVJ05] follow a similar approach, but automatically extract the saliency from the input mesh by computing multiresolution curvature maps.

Recently, several researchers have investigated the use of black-box perceptual metrics for the evaluation of specific artifacts. Karni and Gotsman [KG00], in order to evaluate properly their compression algorithm, consider the *Geometric Laplacian*, which represents a measure of the smoothness of each vertex. Starting from the Geometric Laplacian, they derived a visual metric to compare two 3D objects (abbreviated as GL_1 in Table 1). Subsequently, Sorkine et al. [SCOT03] proposed a different version of this metric (GL_2), which assumes slightly different values of the parameters involved.

Tian and AlRegib [TA04] and Pan et al. [PCA05] proposed simple quality metrics dedicated to optimising the transmission of textured meshes; their metrics respectively rely on geometry and texture deviations [TA04] and on texture and mesh resolutions [PCA05]. Their results underline the fact that the perceptual contribution of image texture is, in general, more important than the model's geometry.

Following the idea that a measure of the visual artifacts produced by watermarking should be based on the amount of *roughness* introduced on the surface, Corsini and Drelie Gelasca et al. [CGEB07] proposed two perceptual metrics for quality evaluation of watermarking algorithms (abbreviated as $3DWPM_1$ and $3DWPM_2$ in Table 1). The watermarking visual impairment is evaluated by considering the increment of total roughness between the original model and the watermarked model. Two ways to measure model roughness were proposed. The first one [CDGE05] is a roughness measure based on a variant of the method by Wu et al. [WHST01], based on statistical considerations about the dihedral angles, i.e. the angle between the normals of two adjacent faces. In order to take into account the *scale* of the roughness, the per-face roughness is turned into a per-vertex

roughness and rings of different size (1-ring, 2-ring, etc.) are considered during roughness evaluation. The total roughness of the 3D object is the sum of the roughnesses of all vertices. The second method by Drelie Gelasca et al. [DGECB05] is based on the consideration that visual artifacts should be better perceived on smooth surfaces. So, a smoothing algorithm is applied to the mesh and then the roughness is evaluated as the variance of the differences between the smoothed version of the model and its original version.

In the ambit of quality evaluation of 3D watermarking algorithms, Lavoué et al. [LDD*06] proposed a perceptually-inspired metric called the *Mesh Structural Distortion Measure* (*MSDM*). This metric follows the concept of structural similarity introduced for 2D image quality assessment by Wang et al. [WBSS04]: differences of curvature statistics (mean, variance, covariance) are computed over corresponding local windows from both meshes being compared. A global measure between the two meshes is then defined by a Minkowski sum of the distances over the local windows (one local window per vertex is considered). A multi-resolution improved version, named *MSDM*2, has recently been proposed in [Lav11]. It provides better performance and allows one to compare meshes with arbitrary connectivities. These metrics are available online within the MEPP platform http://liris.cnrs.fr/mepp/.

Recently, Bian et al. [BHM08, BHM09] developed a geometry-based perceptual metric (abbreviated as *SF* in Table 1) based on the *strain energy*, i.e. a measure of the energy which causes the deformation between the original and the processed mesh. The idea is that the more the mesh is deformed, the higher is the probability that the observer perceives the difference between the processed and the original mesh. The strain energy calculation on the mesh is simplified by considering that each mesh element (a triangular mesh is assumed) is perturbed along its plane. It is important to underline that this metric is suitable for small deformations only. The perceptual distance between the original model and the perturbed one is defined as the weighted average of the strain energy over all triangles of the mesh, normalised by the total area of the triangular faces. Bian et al. tested some variants of this metric by choosing different weights, but from their experimental results they concluded that the unweighted version gave results similar to the tested variants; hence it is preferable due to its simplicity.

4. Metrics for dynamic 3D meshes

The approaches to dynamic mesh distortion evaluation can be generally separated into two main classes: the ones based on some static mesh distortion metric, which is applied in a per-frame fashion, and the ones specifically tailored to the case of dynamic meshes.

The first class inherits all the problems of the original metrics, i.e. any metric that fails to correlate with human per-

ception in the static case will most likely also fail when applied to the dynamic case. Moreover, with the addition of animation, a new class of temporal artifacts may appear, and a proper metric probably should detect this kind of artifact as well. However, temporal artifacts of course cannot be detected by a static mesh metric applied in per-frame fashion.

A typical example of such an artifact might be a smooth distortion, such as adding one period of a sine wave to the X coordinates of a particular frame. If the amplitude of the sine is small, then such a distortion will be almost unnoticeable on the static mesh. A subsequent frame might be influenced by the same kind of distortion, only this time using a cosine. Both frames contain a distortion that is hard to notice on its own. However, in a playing animation, the effect of oscillating between sine and cosine distortion is probably quite visible.

4.1. Static mesh metrics applied on dynamic meshes

As mentioned before, any of the metrics for static meshes presented in previous sections can be applied on dynamic meshes in the per-frame fashion, using the per-frame result sum, average or maximum as a result. Some authors display the result of some particular static mesh metric for each frame in the form of a time dependency graph.

Early papers on dynamic mesh compression, such as [Len99] and [IR03], have used average SNR to evaluate the amount of distortion caused by the lossy encoding. Later, after publication of the Metro tool, metrics based on Hausdorff distance became more popular. Some papers – [MZP06], [AK09] – show temporal development of RMSE or its average, while others – [HKL09] – show the temporal development of Hausdorff distance.

The common problem of all these metrics is the lack of correlation with human perception, which has already been identified in one of the first works on dynamic mesh compression by Lengyel [Len99]. The work of Lee et al. [LKT*07], where the sum of Discrete Shape Operator differences is used (similar to the metric in [KG00]), is one of the few exceptions, where a perceptually motivated static mesh metric has been used for dynamic mesh comparison. But, even in this case, the metric cannot capture any temporal artifacts that may arise in dynamic mesh processing.

4.2. KG error

A metric used quite commonly in dynamic mesh compression is the KG error, proposed by Karni and Gotsman in [KG04]. The metric is designed specifically for animated triangle meshes. It works on matrices describing original and distorted meshes, where columns of the matrices describe trajectories of respective vertices of the animation. Having a matrix M describing the original animation sequence, and a matrix M' describing the distorted version,

the metric uses the Frobenius norm of the matrix difference $\|M - M'\|$ and produces a normalised version (for details see [KG04]) of this value as the result. Therefore, having function $AMSE(M, M')$ that computes the average mean squared error between animations represented by matrices M and M', the KG error can be rewritten in the form of function $KG(M, M') = f(M, AMSE(M, M'))$. Due to this fact, one might expect that the KG error metric will show the same insufficiencies as any other averaged static metric based on MSE.

4.3. D_a error

Another metric designed specifically for animated meshes has been proposed by Jang et al. in [JKJ*04]. This metric cannot be expressed in terms of per-frame static mesh metric results, because it works on ribbons formed by error vectors in subsequent frames. An error vector is a vector connecting the original and distorted position of a particular vertex in a particular frame. The D_a error metric works on a simplified version of error vectors, taking always only a single coordinate into account. The error vectors associated with a particular vertex in two subsequent frames form a ribbon-like structure in 2D space (coordinate + time), and the D_a error metric computes the area of this ribbon and uses it as a contribution of the particular vertex to the overall error. The metric obtains the contributions from all the vertices and all pairs of subsequent frames of the animation, finally normalising the result by dividing by the largest span in all directions.

Although the metric is defined in a form that is only applicable to dynamic meshes, its relation to perceptual difference is not clear, and the design is not based on any perceptual experiment that would support it. Moreover, there are at least two intuitive flaws in the metric that indicate that its relation to perception is rather vague. These are:

1. Preference to oscillating vertices. The metric uses a different formula for straight and twisted vertices in order to correctly compute the ribbon area. However, this leads to a smaller contribution from vertices oscillating around a central position (an obvious temporal artifact) than from vertices that are constantly dislocated in time.
2. Lack of rotation invariance. Due to the per-coordinate processing, the metric produces different results in coordinate systems that are rotated with respect to each other. However, a natural expectation is that a metric result should be translation and rotation invariant.

In spite of these limitations, the D_a metric has been used in some papers dealing with dynamic mesh compression ([MSK*05], [MSKW06], [MSK*06]), and it even became part of the MPEG-4 standard in the form of Animation Framework eXtension Core Experiments Description [ISO].

4.4. 4D Hausdorff distance

A modification of Hausdorff distance that goes beyond averaging the metric over all the frames has been proposed in [VS06] by Váša and Skala. The metric works in a 4D space, where the fourth dimension is the time of the animation. A triangle in two subsequent frames forms a 4D prism in the 4D space. These prisms are coherently subdivided into 4D tetrahedra for easier manipulation, and the metric works on these tetrahedra in a manner equivalent to Hausdorff distance evaluation in 3D.

The main advantage of the metric is that it is able to detect temporal proximity of surfaces, which was not possible with the metrics based on static mesh comparison. However, the metric requires an additional parameter, a constant relating the spatial and temporal distances. The proper value of such a constant is not easy to obtain. It could probably be determined in a subjective experiment; however, the authors suggest using only the usual (average) speed of vertices in the animation as a relating factor.

Apart from that, the metric also has other disadvantages, such as high computational cost, high memory requirements, and also insufficiencies related to the core idea of Hausdorff distance and its lack of correlation with human perception. Therefore, the metric has not been used in practice.

4.5. STED error

The first, and so far the only attempt at a perceptual metric for dynamic meshes is the STED error proposed by Váša and Skala [VS11]. It is based on the observation that perception of distortion is related to local and relative changes rather than global and absolute changes of vertex positions. The metric works on edges as basic primitives, and computes the relative change in length for each edge of the mesh in each frame of the animation. Subsequently, for each vertex, the standard deviation of relative edge lengths is computed within a topological neighbourhood of the vertex. This deviation is then used as a contribution of the vertex to the spatial part of the error metric, assuming that high local deviation relates to higher local distortion and thus higher perceived error.

The metric also attempts to capture temporal artifacts by working with virtual temporal edges: that is, edges that connect position of a vertex in two subsequent frames. The difference between original temporal edge length and distorted temporal edge length is then again used as a contribution to the temporal part of the error metric. The metric normalises the contributions of temporal edges using the speed of the vertex in a local temporal window, thus taking into account that "shaking" artifacts are more noticeable in areas that are static or moving slowly.

Finally, the result is taken as a hypotenuse of the spatial and temporal parts of the error. The metric has several parameters, such as the width of the topological neighbourhood

over which the contribution to spatial error is computed, or a relating constant used in combining the spatial and temporal parts. These parameters were set in order to obtain the highest possible correlation with the results of a subjective experiment that was carried out as a part of the work.

The STED error measure can be evaluated using a command-line utility that can be downloaded from the following URL: http://compression.kiv.zcu.cz.

4.6. Simplification evaluation

While STED is primarily designed for situations, such as compression, where the original and the distorted version of the mesh both have an equal number of vertices and the same connectivity, Larkin and O'Sullivan [LO11] focused on the perception of distortion introduced by simplification of animated meshes of human characters. In this case, the distorted version of the mesh has a lower number of vertices than the original, which may cause visible artifacts. The authors identified three types of artifacts caused by simplification:

- *texture* (errors due to the interpolation of texture coordinates)
- *lighting* (errors due to the interpolation of normals)
- *silhouette* (errors in the silhouette of the mesh)

They performed a user study to determine the influence of each of these artifacts in static and dynamic cases on the perception of the simplified mesh. The results of the study show that the silhouette artifacts are the most easily identified by human observers, while the other two types have a rather minor effect. The results also indicate that the animation of the mesh itself does not change the perception of artifacts, given that the mesh stays in the same location on the screen. Movement of the mesh throughout the screen, however, might have an impairing influence on the perception of error, as described by McDonnel et al. [MNO07]. Using these findings, Larkin and O'Sullivan devised a render-based metric to evaluate the distortion caused by simplification [LO11].

The metric is designed to quantify the perceptual change of the mesh silhouette. Since the subjective experiment did not prove that animation changes the perception of errors, the metric only works on a single frame of the animation (a static mesh). Similarly to the static mesh metric by Lindstrom and Turk [LT00], it compares a series of renders of the mesh from different viewing angles. To speed up the process, the space occupied by the mesh is voxelised into voxels small enough to cover one degree of visual angle on the screen and only voxels intersecting the mesh surface are used. For each such voxel, a render targeted at this voxel and its neighbourhood is performed with the camera looking along the silhouette. The renders of corresponding voxels of the original and the simplified mesh are then compared using a perception-correlated image metric – *PerceptualDiff*

by Yee and Newman [YN04]. This metric returns the number of pixels that may be perceived as different between the meshes. These pixel counts are averaged over all the surface voxels to create the output value of the metric. The output depends on the resolution and the field of view angle of the renders, the size of the screen and the viewer distance from the screen, which are the parameters of the metric.

The authors claim the metric to be a good indicator of the perceived distortion caused by simplification of animated meshes of human characters. They, however, do not provide any values of correlation with subjective experiment results to prove this claim. They also note that the subjective tests were performed from a single point of view, while the metric considers any point of view on the mesh.

Still regarding simplification, Bulbul et al. [BKCG10] proposed a saliency estimator for dynamic meshes, similar to the static mesh saliency estimator from Lee et al. [LVJ05]; this estimator integrates several features such as colour, geometry and motion and was used to drive a simplification algorithm (i.e. salient regions are more preserved).

4.7. Fidelity of physical simulations

Dynamic mesh data sets can be created by various methods, one of which is physical simulations. Such animations have unique characteristics, which make them an important part of dynamic mesh processing research. For example, while skeletal animations are relatively easy to compress using skinning- or clustering-based compression algorithms, using similar techniques on animations based on physical simulations might not be as effective and a different method may produce better results.

In the case of simulation-generated dynamic scenes, besides evaluating the perceptual difference of a distorted physical simulation animation from a reference simulation, we can also evaluate the perceptual plausibility of the physics in the simulation. Based on a series of subjective experiments, O'Sullivan et al. [ODGK03] proposed a design of a visual fidelity metric for physically-based simulations of colliding objects. The metric estimates a probability $P(A)$ of perceiving a simulation as implausible:

$$P(A) = f\left(P_{angular}(A), P_{momentum}(A), P_{spatiotemporal}(A)\right), \tag{1}$$

where $P_{angular}$ is the probability of spotting an error in the post-collision angles of the objects, $P_{momentum}$ is the probability of perceiving the post-collision speeds of the objects as implausible and $P_{spatiotemporal}$ is the probability of seeing a gap between the colliding objects at the time of the collision or a delay between the collision and the subsequent movement. Details on the definition of these probabilities can be found in [ODGK03]. The authors do not describe the combining function f, as they claim to have insufficient subjective experiment data to do so. Thus, they consider the three components separately.

5. Subjective databases and evaluation

Perceptual metrics presented above aim at *predicting* the visual quality of a 3D (or 3D+t) model as perceived by a human observer. This perceived quality can also be directly and quantitatively assessed by means of subjective tests; in such tests, human observers directly give their opinion or some ratings about the perceived quality of a corpus of distorted models; a mean opinion score (MOS) is then computed for each distorted object reflecting its average quality as appreciated by the observers. The correlation between these subjective Mean Opinion Scores and the objective scores computed by the metrics provides an excellent indicator of the performance of these metrics and a very good way to compare them quantitatively.

This section presents the protocols usually used in subjective tests, the existing MOS databases and some evaluation and comparison results regarding existing metrics.

5.1. Subjective test

Practically and whatever the type of media (image, video or 3D models), the design of a subjective test is composed of the following steps:

1. A database is constructed containing different objects (reference objects and distorted versions).
2. A subjective experiment is conducted where human observers directly give their opinion or some ratings about the perceived distortions of the database objects. A mean opinion score (MOS) is then computed for each distorted object of the corpus: $MOS_i = \frac{1}{n}\sum_{j=1}^{n} m_{ij}$, where MOS_i is the mean opinion score of the ith object, n is the number of test subjects, and m_{ij} is the score (in a given range) given by the jth subject to the ith object.
3. Since some observers may have used the rating scale differently, a normalisation of the MOS values is usually conducted, followed by a filtering of possible outlier subjects according to the recommendation of the I.T.U. (International Telecommunication Union) [Rec02]. The reliability of the MOS may also be checked by computing the 95% confidence intervals or the intraclass correlation coefficient.
4. The correlation is computed between the mean opinion scores of the objects and their associated metric's values; usually two correlation coefficients are considered: the Spearman Rank Order Correlation, which measures the monotonic association between the MOS and the metric values and the Pearson Linear Correlation Coefficient, which measures the prediction accuracy. The Pearson correlation is computed after performing a non-linear regression on the metric values, usually using a logistic or a cumulative Gaussian function. This serves to optimise the matching between the values given by the objective metric and the subjective opinion scores provided by the

subjects. This step allows the evaluation to take into account the saturation effects typical of human senses.

As raised recently by Ebrahimi [Ebr09], the design of subjective tests producing reliable and reproducible MOS is a delicate task which depends on several ingredients:

- The environment, i.e. type of monitors, viewing distances, lighting conditions.
- The material, i.e. the test objects. The choice of the corpus of models is critical to obtain reliable results; to be able to generalise the results, the corpus should contain different kinds of models and different types of distortions and not focus on a specific scenario. In case of distortions of different types, the range of their visual impacts have to be correctly balanced. It is also usually better to present worst case models (i.e. anchor conditions) to allow the observers to calibrate their ratings.
- The methodology, i.e. how to present the distorted models and how to rate them. A lot of protocols exist for presenting the stimuli and rating them, for instance, the distorted model can be displayed together with its original version (Simultaneous Double Stimulus) or alone (Single Stimulus). The rating can be categorical adjectival (bad, poor, fair, good, excellent), categorical numerical (1,2,3,4,5) or on a continuous scale (e.g. $\in [0,100]$) and it can concern either a quality or an impairment.
- The analysis of the data, i.e. how to make sure that MOS are significant.

For image and video quality assessment, the International Telecommunication Union has made recommendations for test conditions and methodology [Rec99] [Rec02] [Rec07]. However, they cannot be transposed directly for 3D object quality assessment. In particular, 3D (3D+t) model rendering involves a whole set of supplementary parameters [CGEB07]: the background, the light source, the material and texture and the level of interactions. Currently, no normalised recommendation exists for designing subjective tests involving 3D (3D+t) models. The next subsection describes existing subjective databases (Corpus + MOS) and details more particularly the three publicly-available ones.

5.2. Existing subjective databases

5.2.1. 3D static mesh

Several authors have made subjective tests involving 3D models [RRP00] [WFM01] [RR01] [PCA05] [LDD*06] [CGEB07] [SSF07] [SSFM09] [Lav09]. Their experiments have different purposes and use different methodologies. Bulbul et al. [BCLP11] recently provided a nice overview and comparison of their environments, methodologies and materials.

Subjective tests from Watson et al. [WFM01] and Rogowitz and Rushmeier [RR01] focused on a mesh simplification scenario; their test databases were created by applying different simplification algorithms at different ratios on
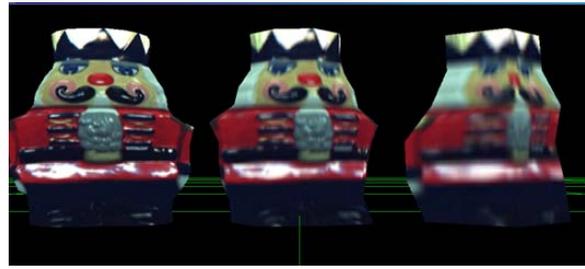

Figure 7: *Evaluation interface for the subjective test of Pan et al. [PCA05]. The observers were asked to compare the target stimulus (centre) with the two referential stimuli (left and right) and assign it one of the following ratings: very poor (1), poor (2), fair (3), good (4), very good (5). Reprinted from [PCA05].*

several 3D models. The purposes of their experiments were respectively to compare image-based metrics and geometric ones to predict the perceived degradation of simplified 3D models [WFM01] and to study if 2D images of a 3D model are really suited to evaluate its quality [RR01].

Rushmeier et al. [RRP00] and Pan et al. [PCA05] also considered a simplification scenario; however, their 3D models were textured. These experiments provided useful insights on how resolution of texture and resolution of mesh influence the visual appearance of the object. Pan et al. [PCA05] also provided a perceptual metric predicting this visual quality and evaluated it quantitatively by studying the correlation with subjective ratings from their experiment.

Corsini et al. [CGEB07] proposed two subjective experiments focusing on a watermarking scenario; the material was composed of 3D models processed by different watermarking algorithms introducing different kinds of artifacts. The authors then used the Mean Opinion Scores to evaluate the effectiveness of several geometric metrics and proposed a new perceptual one (see section 3) to assess the quality of watermarked 3D models.

Below, we describe in more detail the subjective databases from [LDD*06] [Lav09] [SSF07], since 3D models and MOS are publicly available:

- The LIRIS/EPFL General-Purpose Database [LDD*06] was created at the EPFL, Switzerland. It contains 88 models with between 40K and 50K vertices generated from 4 reference objects (Armadillo, Dyno, Venus and Rocker-Arm). Two types of distortion (noise addition and smoothing) are applied with different strengths and at four locations: uniformly (on the whole object), on smooth areas, on rough areas and on intermediate areas. These distortions aim at simulating the visual impairment of generic geometric processing operations (compression, watermarking, smoothing). 12 observers participated to the subjective evaluation; they were asked to provide a

score reflecting the degree of perceived distortion between 0 (identical to the original) and 10 (worst case). The resulting MOS were originally used to evaluate the performance of the MSDM perceptual metric (see section 3).

- The LIRIS Masking Database [Lav09] was created at the Université of Lyon, France. It contains 26 models with between 9K and 40K vertices generated from 4 reference objects (Armadillo, Bimba, Dyno and Lion) specifically chosen because they contain significantly smooth and rough areas. The only distortion is noise addition applied with three strengths. However, it is applied either on smooth or rough regions. The specific objective of this database was to evaluate the *visual masking* effect. It turns out that the noise is indeed far less visible on rough regions. Hence, the metrics should follow this perceptual mechanism. 11 observers participated to the subjective evaluation. The data resulting from this as well as the previous subjective experiment can be downloaded from `http://liris.cnrs.fr/guillaume.lavoue/data/datasets.html`.

- The IEETA Simplification Database [SSF07] was created at the University of Aveiro, Portugal. It contains 30 models generated from 5 reference objects (Bunny, Foot, Head, Lung and Strange) from 2K to 25K vertices. The reference models have been simplified using three different methods and two levels (20% and 50% of the original number of faces). 65 observers participated in the subjective evaluation; they were asked to provide a score from 1 (very bad) to 5 (very good). Along with this rating, in another phase of the test, the observers were also asked about their preference among several simplified models presented together; this can also constitute highly relevant information, which is, however, more difficult to exploit. The same authors have recently done another subjective experiment using a larger corpus of models [SSFM09]. However, only preferences were collected. The data resulting from this subjective experiment can be downloaded from `http://www.ieeta.pt/~sss/repository/`.

5.2.2. 3D dynamic mesh

To the best of our knowledge, the only experiment dealing with error perception in dynamic meshes was the one performed by Váša and Skala [VS11] in their work proposing the STED metric. Their setting used 5 dynamic meshes (chicken, dance, cloth, mocap and jump), each in 9 versions, using different kinds of both spatial and temporal distortion of varying types (random noise, smooth sinusoidal dislocation of vertices, temporal shaking and results of various compression algorithms). Overall, there were 170 evaluators; however, most of them only evaluated one or at most two datasets, i.e. for each of the five datasets there were 37-49 subjective evaluations. The users were asked to rate the amount of perceived distortion on scale of 0-10. The users had all the versions (including the original) available at the same time (running on 10 computers), and they were asked to use the whole scale of evaluation.

5.3. Evaluation results

Databases and Mean Opinion Scores produced by the subjective tests presented above constitute an excellent basis for comparing and evaluating existing perceptual metrics, by studying the correlation between the MOS and the metric's values.

5.3.1. 3D static mesh

For model-based metrics (i.e. relying on the geometry), a recent study [LC10] has provided an extensive quantitative comparison of existing metrics by computing Pearson and Spearman correlations with MOS from the LIRIS Masking Database and the LIRIS/EPFL General-Purpose Database. These results were updated by the recent study from [Lav11], which also provided correlation values on the IEETA Simplification Database. Table 1 summarises these correlation results. Most of the existing metrics cannot be applied to evaluating simplification distortions because they need the compared objects to share the same connectivity – [KG00] [SCOT03] [BHM09] [LDD*06] – or the same level of details – [CGEB07].

As the table shows, the most recent MSDM2 metric provides very good results on all the databases; 3DWPM metrics also have a correct behaviour. On the contrary, the classical geometric distances, like Hausdorff and RMS, provide a very poor correlation with human judgement. An important point to raise is that the General-purpose and Masking databases represent quite difficult scenarios (several different models, several types of distortion, non-uniform distortion); in simpler scenarios (one single uniform distortion, like uniform noise addition, for instance), even simple geometric distances are able to correlate with the human judgment; for instance, for the Simplification database (only one type of distortion), the Hausdorff and RMS metrics provide correct results. Moreover, in a purely watermarking scenario, 3DWPM metrics have been shown to provide very good results [CGEB07].

Unfortunately, image-based metrics have not been quantitatively tested on these public databases, whereas several authors [WFM01] [CS06] have shown that, in a simplification scenario, they provide very good results, better than simple geometric distances. As also raised by Bulbul et al. [BCLP11], it would be very interesting to compare quantitatively these image-based metrics to the most effective model-based ones.

5.3.2. 3D dynamic mesh

The user opinions gathered by Váša and Skala in [VS11] were evaluated using similar tools as described for the case

	General Purpose [LDD*06]		Masking [Lav09]		Simplification [SSF07]	
	Spearman	*Pearson*	*Spearman*	*Pearson*	*Spearman*	*Pearson*
Hausdorff	13.8	1.3	26.6	4.1	49.4	25.5
RMS	26.8	7.9	48.8	17.0	64.3	34.4
GL_1 [KG00]	33.1	12.6	42.0	15.7	N/A	N/A
GL_2 [SCOT03]	39.3	18.0	40.1	14.7	N/A	N/A
SF [BHM09]	15.7	0.5	38.6	2.4	N/A	N/A
$3DWPM_1$ [CGEB07]	69.3	38.3	29.4	10.2	N/A	N/A
$3DWPM_2$ [CGEB07]	49.0	24.6	37.4	18.2	N/A	N/A
MSDM [LDD*06]	73.9	56.4	65.2	47.9	N/A	N/A
MSDM2 [Lav11]	80.4	66.2	89.6	76.2	86.7	79.6

Table 1: *Spearman and Pearson correlation (%) between Mean Opinion Scores and values from the metrics for the three publicly-available subjective databases. These data have been synthesised from [LC10] [Lav11].*

	chicken		dance		cloth		mocap	
	Pearson	*Spearman*	*Pearson*	*Spearman*	*Pearson*	*Spearman*	*Pearson*	*Spearman*
KGerror	-53	-23	-54	-24	-27	14	-34	-50
$D_a mean$	-49	-37	-53	-2	-24	13	-33	-49
$D_a peak$	-33	2	-60	-40	-29	20	-62	-53
Hausdorff	-32	2	-56	-36	-26	36	-53	-53
RMS	-69	-63	-57	-30	-28	20	-42	-50
STED	97	95	94	96	92	95	98	92

Table 2: *Spearman and Pearson correlation (%) between Mean Opinion Scores and values from the metrics for dynamic meshes. These data have been taken from [VS11].*

of static meshes, i.e. using the Spearman and Pearson co-efficient. Five metrics were compared (KG error, D_a error, average Hausdorff distance, average RMS error and STED error). The resulting Pearson coefficient was slightly negative for all the metrics except for STED. The results are summarised in Table 2. By using the STED algorithm and adjusting its parameters, the correlation with the results of the subjective experiment reached more than 0.9 in all the tests in terms of the Pearson coefficient.

6. Applications

6.1. Application to static mesh watermarking

In a static mesh watermarking [WLDB08a] algorithm, a piece of information, i.e. a *watermark*, is embedded into the functional part of a cover mesh. Applications of mesh watermarking include copyright protection (robust watermark), mesh authentication (fragile watermark) and content enhancement (high-payload watermark). In general, the embedding of a watermark will inevitably introduce some distortion to the original cover mesh. It is important to keep this distortion *imperceptible* to human eyes, so as to ensure that its insertion does not influence the intended use of the model and that the watermarked mesh does not look suspicious to an attacker.

In the literature, the mesh watermarking research has ben-

efited from the work on mesh perceptual quality assessment, or more generally from the work on human visual perception, in two different ways. First, the properties of the human visual system (mainly those of frequency sensitivity and visual masking) have been taken into account during the design of mesh watermarking algorithms, with the objective to achieve a better performance. Second, the emergence of objective mesh visual quality metrics has facilitated fair comparisons between different algorithms. In the following, we will present some details on these two points.

6.1.1. Use of HVS features for mesh watermarking

Use of frequency sensitivity. The geometry processing community has empirically noticed that in general high-frequency distortion on mesh surfaces is much more visible than low-frequency distortion. This observation provides insight regarding how to select watermark carriers in spectral mesh watermarking methods. However, compared to 2D images, performing a spectral transform on 3D triangle meshes is much more complicated. The standard solution is first to construct an $N \times N$ mesh Laplacian matrix (where N is the number of vertices) and then to use its eigenvectors as the transform basis [LZ10]. Different constructions of the mesh Laplacian matrix yield different transform bases. Existing spectral mesh watermarking methods may use different transforms, but they all embed a watermark, commonly

represented by a sequence of bits, in low-frequency coefficients of the mesh spectrum. For example, Ohbuchi et al. [OMT02] embed watermark bits only in the first 500 low-frequency coefficients obtained after a combinatorial Laplacian spectral analysis [KG00]. Lavoué et al. [LDD07] use even fewer coefficients, i.e. roughly the first 100, but for the very low-frequency coefficients the authors adaptively increase the embedding intensity because, empirically, more distortion can be introduced to these coefficients without being noticed. Liu et al. [LPG08] and Wang et al. [WLBD09] develop watermarking methods based on a new mesh spectral transform, namely the manifold harmonics transform [VL08], but the altered coefficients are always limited to the low-frequency end, i.e. approximately the lowest 100 frequencies.

Recently, an explanation has been proposed for the effect of different sensitivities of the HVS to the modifications of different mesh frequency components [Tor11]. Actually, by using results from Nodal sets theory [DF88] and some simple trigonometric computations, one can obtain the intrinsic frequency of each transform basis vector (i.e. each eigenvector of the mesh Laplacian matrix) and relate this frequency to the frequency as observed by human eyes under a regular viewing condition (the observed frequency is expressed in cpd, cycles per degree). Subsequently, we can relate the modification of a mesh spectral component to the HVS sensitivity as given by the contrast sensitivity function [MS74] (cf. Section 2.1), which is a function of the observed frequency. In Figure 8(b) and (c) are shown two deformed Venus models where low and high frequency coefficients have been altered, respectively. The observed frequencies of the two deformations are respectively around 0.1 and 4.0 cpd. According to the contrast sensitivity function, which has a peak at 4-6 cpd and drops very rapidly on either side of this peak, we know that the deformation in Figure 8(c) should be much more visible than that shown in Figure 8(b).

Frequency sensitivity of the HVS has also been utilised in other mesh watermarking methods that do not operate directly in a mesh spectral domain. For example, in the wavelet-based method of Wang et al. [WLDB08b], a robust watermark is embedded in the coarsest resolution level of a dense semi-regular mesh after it goes through a wavelet decomposition. The authors argue that the introduced distortion is of low frequency because after performing wavelet synthesis, the distortion would be spread over the surface and thus be smoothed. A similar strategy is employed in the subdivision surface watermarking scheme of Lavoué et al. [LDD07], where the watermark is embedded in the (coarse) control mesh of the subdivision surface. After subdivision, the watermark would be "diluted" on the mesh surface. Finally, in the moment-based method of Wang et al. [WLDB11], watermark bits are embedded through a low-frequency deformation of the cover mesh in which the surface is either globally pulled upwards or globally pushed
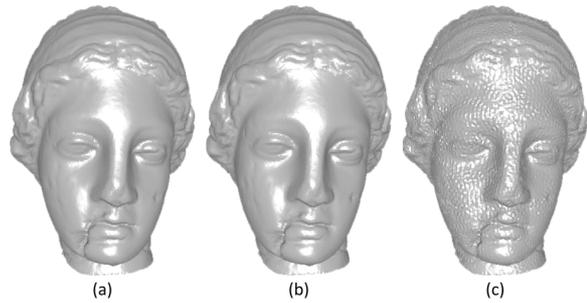

Figure 8: *From left to right are respectively the original Venus model, a deformed model after low-frequency modification in the spectral domain, and a model after high-frequency deformation. The induced MRMS errors of the two deformed meshes are exactly the same, but their visual impairments are quite different due to different frequency sensitivities of the HVS.*

downwards. In this way, the resulting deformation more or less follows the original shape of the surface and is of low frequency.

Use of visual masking effect. As mentioned in Section 2.1, the visual masking effect means that the existence of one visual signal may hide or reduce the visibility of another one. In the case of mesh visual quality assessment, this mainly implies that a local surface modification would be more visible in a smooth area than in a rough area.

The visual masking effect has been considered in some mesh watermarking methods, with the objective to achieve a better trade-off between watermark imperceptibility and the robustness/payload. The basic idea mainly consists in increasing the embedding intensity or the local embedding payload (i.e. the number of locally embedded bits) in rough areas of the cover mesh. Different roughness measures have been proposed in the literature of mesh watermarking, respectively based on the minimum [YIK03] or total [Bor06] length of the incident edges of a vertex, the dihedral angle between neighbouring facets [CW07] and the variance of the facet normal directions [YPSZ01]. Lavoué [Lav09] devises a roughness measure based on the curvature difference between the input mesh and a carefully smoothed version. When applying this measure for mesh watermarking, the author first computes the local roughness of a number of mesh patches, and then sets adaptive embedding intensity for each patch, i.e. the rougher the patch is, the stronger the intensity will be. This adaptive embedding can, to some extent, improve the robustness of the watermarking method of Ohbuchi et al. [OMT02], while ensuring the watermark imperceptibility thanks to the visual masking effect. Similarly, Kim et al. [KBT10] show that by using the mesh roughness measure of Corsini et al. [CGEB07] to adaptively determine the local embedding strength, they can achieve a better trade-

off between imperceptibility and robustness for the watermarking schemes proposed in [Ben99] and [CM03].

It can be seen that there does not exist a predominant roughness definition for polygonal meshes, at least in the context of mesh watermarking. We think that in order to find an adequate roughness definition, we may need to conduct psychovisual experiments to quantitatively and comparatively study the relationship between geometric quantities, e.g. those mentioned in the previous paragraph, and human visual perception and quality assessment.

6.1.2. Perceptually-based metric for mesh watermark benchmarking

When evaluating the imperceptibility of their mesh watermarking schemes, in most cases researchers simply use purely geometric measures such as Hausdorff distance and root mean squared error, or they just show some images of the original and watermarked meshes and leave the reader to judge the visual fidelity. This seems inadequate mainly for two reasons. First, as demonstrated by the results in Table 1, simple geometric measures do not correlate well with a subjective assessment. Second, the quality of the images illustrating the original and watermarked meshes highly depends on the viewpoint selection and the rendering technique used. Therefore, in order to ensure a fair and efficient comparison between different mesh watermarking algorithms, one reasonable solution is to use perceptually-based objective metrics. This was actually the motivation of the work of Corsini and Drelie Gelasca et al. [DGECB05, CGEB07] and Lavoué et al. [LDD*06].

Recently, MSDM has been integrated into a mesh watermarking benchmark that is publicly available online [WLD*10]. Different from the simple geometric measures, MSDM well reflects the visual fidelity between two meshes and thus is appropriate for the watermark imperceptibility assessment (see Figure 9 for an example). Like the widely used benchmarks for image watermarking [PAK98, STN*01], the basic idea of the mesh watermarking benchmark is to first fix thresholds of the introduced distortion and payload of the watermark, and afterwards assess the robustness against a series of common attacks. The authors of the benchmark consider that only fixing a threshold on MSDM is not sufficient, and it is also necessary to set a threshold of the geometric error, so as not to exaggerate the amount of low-frequency distortion introduced by watermark embedding that is rather imperceptible. Indeed, this control of geometric error is particularly important in applications such as computer-aided design and medical imaging. For more details on the benchmark and the evaluation results of several mesh watermarking schemes using the benchmark, readers can refer to [WLD*10] and [WLDB11].
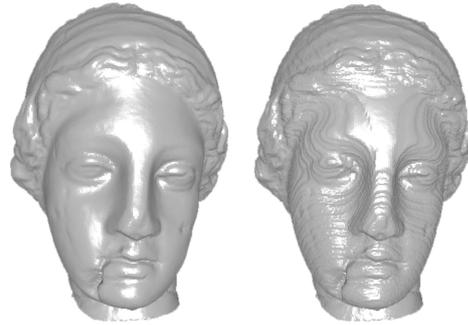

Figure 9: *On the left is the Venus watermarked using the method of Wang et al. [WLDB11] (MRMS= 1.69×10^{-3}, MSDM= 0.13), and on the right is the Venus watermarked using the method of Cho et al. [CPJ07] (MRMS= 0.936×10^{-3}, MSDM= 0.54). MSDM yields correct results that are consistent with a subjective assessment.*

6.2. Application to static mesh compression

The objective of compression is to reduce the size of the 3D data to improve storage and speed up transmission; it usually involves finding different ways of representing the mesh in order to remove data redundancy and maximise the amount of data implicitly encoded. Mesh compression techniques can be distinguished in two categories: *single-rate techniques*, where the mesh data is compressed as a whole and can only be decompressed after receiving the entire file, and *progressive techniques*, where the mesh is transmitted as a simple coarse mesh (low-resolution), and a refinement sequence allowing the viewer to update incrementally the level of detail of the mesh during the transmission. Both single-rate and progressive techniques introduce distortions on the object surface, which can be due to quantisation or simplification (in the case of progressive techniques); their main goal is to find the optimal compromise between the bitrate and the distortion, i.e. how to obtain a minimal file size while preserving high visual quality.

Compression has benefited from the work on mesh perceptual quality assessment and human visual perception, mostly in two ways. First, like for watermarking (see paragraphs above), the properties of the human visual system (frequency sensitivity and visual masking) have been integrated in the design of compression algorithms. Second, several compression frameworks have integrated perceptual metrics. In the following, these two points will be detailed.

6.2.1. Use of HVS features for mesh compression

Use of frequency sensitivity. Like for watermarking, several compression methods have used the fact that high-frequency distortion on mesh surfaces is much more visible than low-frequency. Sorkine et al. [SCOT03] conduct a Laplacian spectral analysis and then encode the transformed
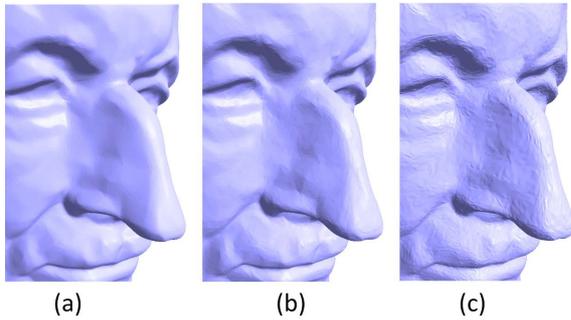
(a)   (b)   (c)

Figure 10: *(a) Original Max Planck, (b) results after high pass spectral quantisation, (c) result after uniform Cartesian quantisation. Models from (b) and (c) are associated with the same rate. The distortion from (c) (which is high-frequency) is clearly more visible than the distortion from (b) (which is low frequency). Images taken from [SCOT03].*

coordinates instead of the Cartesian coordinates; the strength of their method, called High-Pass Quantisation, is that it concentrates the quantisation error at the low-frequency end of the spectrum, which results in a quite imperceptible distortion since fine details are preserved. Figure 10 illustrates their results.

A side property of this frequency sensitivity is that the low-resolution level-of-details, in a progressive compression algorithm, do not need a high geometric precision; hence, they can be quantised roughly without impacting their visual appearance. Indeed, a geometric distortion applied on a low-resolution model (e.g. several hundreds of vertices) is very difficult to perceive because it causes a low-frequency geometric perturbation. Valette et al. [VCP09] and Lee et al. [LLD12] considered this principle and applied coarse-to-fine quantisation precision in their progressive compression algorithm to improve the rate-distortion trade-off.

Use of visual masking effect. The visual masking effect has been considered in mesh compression to hide quantisation artifacts: Lavoué et al. [Lav09] classify the vertices into two clusters (*rough* and *smooth*) according to their roughness value; *rough* vertices are then quantised with a lower precision than *smooth* ones, hence improving the rate-distortion trade-off. Roudet [Rou10] also considers roughness to segment the 3D mesh into patches and then consider an adaptive wavelet scheme for compression.

6.2.2. Use of perceptual metrics to drive mesh compression

The objective of compression is to optimise the rate-distortion trade-off; hence, one relevant method to solve this problem is to include a bit allocation process in the compression algorithm, driven by an error metric. For instance, Payan et al. [PA06] have proposed such a bit allocation pro-

cess for wavelet compression, based on the mean square error. Unfortunately, only few authors have used perceptual metrics for this purpose. Tian and AlRegib [TA08] proposed such a bit allocation framework for progressive compression of textured meshes; for a given bit budget, the optimal combination of mesh and texture resolution is calculated by optimising a simple perceptual metric previously defined in [TA04]. Cheng and Basu [CB07] considered a different scenario: they use the perceptual metric from Pan et al. [PCA05] to propose a transmission strategy for regular textured meshes optimising perceptual quality under packet loss in wireless networks; they break up the texture and mesh into packets by segmentation into overlapping components and subsampling. They then generalise this approach to irregular meshes [CYDB08]; the main idea is to distribute adjacent vertices into separate packets, so that packet loss does not result in a big gap, hence making possible a satisfactory interpolation.

6.3. Application to dynamic mesh compression

The most common use of dynamic mesh metrics is evaluation of distortion to the mesh caused by lossy compression. Being able to quantify the impact of the processing on various dynamic mesh datasets, it is possible to compare the performance of different processing methods to one another. Depending on the target application, a suitable metric can be chosen, e.g. a perceptually correlated one, if the mesh will be presented to human viewers, or a mean-square-error-related one for physical simulations or technical applications.

Compression of animated meshes has become an active field of research in recent years, building on the achievements of static mesh compression. The usual setting is that there is a series of static meshes, one for each frame of the animation, on the input, and the task is to encode the data into a data stream as short as possible, causing the smallest possible distortion. The meshes representing the frames are usually expected to share connectivity; thus, connectivity encoding is of little importance in this task, as the cost of connectivity code is spread over all the frames of the animation.

Impressive data rates of less than 0.5 bpfv (bits per frame and vertex) are achieved using the current state-of-the-art algorithms by exploiting both spatial and temporal coherence of the vertex positions. In search of further performance improvement, the research has recently focused on the problem of distortion evaluation, where some interesting results have already been obtained.

Until the proposal of the STED metric [VS11], there were no published comparisons utilising a perceptually correlated metric focused on dynamic meshes. To compare a new compression method with existing ones, the authors of the method need to know the performance of all the methods in a single metric. Obtaining such results requires evaluating all of the methods with a metric of choice. Since the

implementations of many of the methods are not publicly available, and some of them are difficult to obtain due to licensing restrictions, most researchers tend to take a short-cut and use the already published results of the methods. But that means they also have to adopt the metric these results were measured by. This creates a certain metric lock-in, where the same metrics, such as temporally-averaged RMSE, Hausdorff Distance, or KG error, are continually used, even though they have been proven not to correlate with human perception.
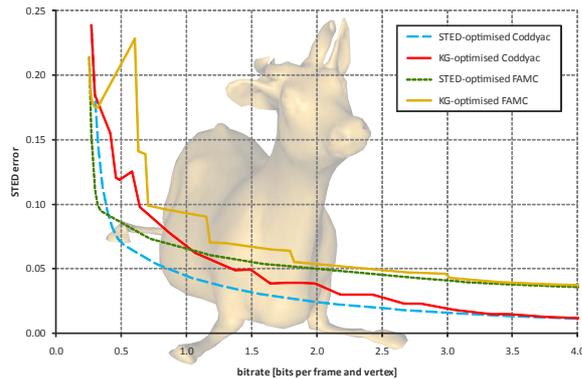

Figure 11: *Rate-distortion curves of the Coddyac [VS07] and FAMC [MSK*08] algorithms in the STED metric compressing "cow" animation. Solid lines show the performance when optimising for the KG error metric, dashed lines denote optimisation for the STED metric.*

Even when resulting meshes, processed by a particular method, are available, using a different metric than the authors did while creating the meshes to measure the distortion may not lead to a fair comparison. The parameters of the method were most probably set for optimal performance in the original metric, which does not generally mean they will be optimal for the new metric as well (see Figure 11). This is very important, since animated mesh compression algorithms usually require setting several configuration parameters, which significantly influence the character of distortion in the mesh. To perform a fair comparison using a particular distortion metric, all the compared compression algorithms need to be optimally configured with respect to that metric. Petřík and Váša proposed an optimisation technique for configuring a dynamic mesh compression algorithm optimally with respect to a chosen metric [PV10]. This technique has been subsequently used, together with the STED metric, in the development of the perception-driven compression algorithm published in [VP11]. The algorithm is based on Laplacian coordinates encoding, which has been previously used by Sorkine et al. [SCOT03] for the case of static meshes. The results confirm the previous finding that the kind of artifacts caused by quantisation of Laplacian coordinates is less perceptible than other kinds of artifacts.

Rus and Váša [RV12] proposed a method for smoothing discontinuities between clusters ("mesh deblocking") in meshes processed by clustering-based compression algorithms, such as FAMC [MSK*08] and Clustered Coddyac [RV10]. Such algorithms divide the mesh into several connected subsets (clusters) based on a certain criterion (similar motion of the vertices in a cluster, for example) and then encode these subsets separately. This procedure may cause the borders of the clusters to overlap or shift apart (see Figure 12). Human vision is much more sensitive to such unusual edges and discontinuities than, for example, RMSE-based metrics, and similarly sensitive are perceptual metrics. The deblocking method improves STED metric performance of existing algorithms that have been tuned to minimise a non-perception-correlated distortion metric, especially for low bitrates.
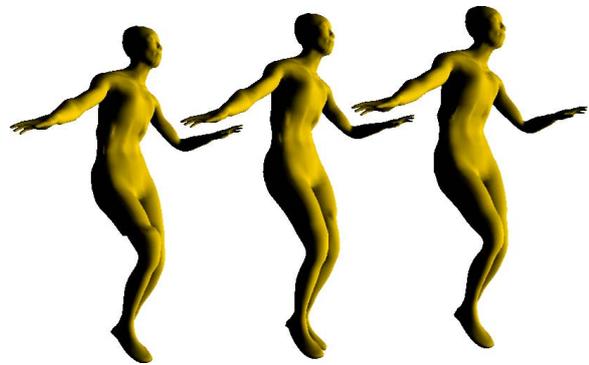

Figure 12: *A frame from the "dance" animation compressed by the Clustered Coddyac algorithm [RV10] (left) and then deblocked by the mesh deblocking method [RV12] (middle). STED value improved from* 0.055 *to* 0.047. *The original is shown on the right.*

7. Conclusions

Evaluating the perceptibility of distortion caused by mesh processing is currently still an open problem, and it seems unlikely that there is ever going to be a perfect solution that works in any circumstances. On the other hand, in recent years great advances have been made, and current metrics described in sections 3 and 4 provide a much better correlation with human perception of distortion.

Apart from maximising the statistical correlation with human perception, as described in section 5, there are many open questions in this area of research, some of which are probably never going to be answered completely. The main issues are:

- applicability of the metrics to various kinds of distortion that might be caused by newly proposed and possibly complex processing algorithms,

- the ability of the metrics to accurately predict perceived distortion in varying circumstances of the mesh usage, such as the characteristics of the display device, the rendering procedure, the viewing conditions and so on,
- applicability of the metrics to models of varying character and properties, such as model size, model roughness/smoothness, or the amount of detail represented in the model.

Even though it seems improbable that there is ever going to be a metric that works for any model under any circumstances and with any kind of distortion, it can be safely asserted that the current perceptual metrics, such as MSDM2 and STED, have made a significant first step in that direction. From this point, some initial efforts have already been made to continue the cycle of adopting the new metrics into processing algorithms and the hope for the future is that this process is going to provide feedback for construction of even better perception-correlated metrics.

8. Acknowledgements

References

[AK09] AYMEN KAMMOUN FRÉDÉRIC PAYAN M. A.: Bit allocation for spatio-temporal wavelet coding of animated semi-regular meshes. In *Proceedings of 15th International MultiMedia Modeling Conference (MMM)* (Sophia Antipolis, France, January 2009). 9

[AvMS10] AYDIN T. O., ČADÍK M., MYSZKOWSKI K., SEIDEL H.-P.: Video quality assessment for computer graphics applications. *ACM Transactions on Graphics 29*, 6 (Dec. 2010), 1. 7

[BAR∗91] BERNS R., ALMAN D., RENI L., SNYDER G., BALONON-ROSEN M.: Visual determination of suprathreshold color difference tolerances using probit analysis. *Color research and Application 16* (1991), 297–316. 5

[Bar99] BARTEN P. G. J.: *Contrast sensitivity of the human eye an its effects on image quality*. SPIE Press, Bellingham, Washington, 1999. 3

[BCLP11] BULBUL A., CAPIN T., LAVOUE G., PREDA M.: Measuring Visual Quality of 3D Polygonal Models. *IEEE Signal Processing Magazine 28*, 6 (2011), 80–90. 12, 13

[Ben99] BENEDENS O.: Geometry-based watermarking of 3D models. *IEEE Computer Graphics and Applications 19*, 1 (1999), 46–55. 16

[BHM08] BIAN Z., HU S.-M., MARTIN R.: Comparing small visual differences between conforming meshes. In *GMP* (2008), Chen F., Jüttler B., (Eds.), vol. 4975 of *Lecture Notes in Computer Science*, Springer, pp. 62–78. 8

[BHM09] BIAN Z., HU S., MARTIN R.: Evaluation for small visual difference between conforming meshes on strain field. *Journal of Computer Science and Technology 24*, 1 (Mar. 2009), 65–75. 8, 13, 14

[BKCG10] BULBUL A., KOCA C., CAPIN T., GÜDÜKBAY U.: Saliency for animated meshes with material properties. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (2010), ACM, pp. 81–88. 11

[BL12] BENSALMA R., LARABI M. C.: A perceptual metric for stereoscopic image quality assessment based on the binocular energy. *Springer Multidimensional Systems and Signal Processing 23*, 1-2 (2012). 4

[BM98] BOLIN M. R., MEYER G. W.: A perceptually based adaptive sampling algorithm. In *SIGGRAPH '98: Proceedings of the 25th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1998), ACM, pp. 299–309. 7

[Bor06] BORS A. G.: Watermarking mesh-based representations of 3-D objects using local moments. *IEEE Transactions on Image Processing 15*, 3 (2006), 687–701. 15

[BPL∗11] BOSC E., PEPION R., LE CALLET P., KOPPEL M., NDJIKI-NYA P., PRESSIGOUT M., MORIN L.: Towards a New Quality Metric for 3-D Synthesized View Assessment. *Selected Topics in Signal Processing, IEEE Journal of 5*, 7 (2011), 1332–1343. 7

[BS06] BARLAND R., SAADANE A.: Blind quality metric using a perceptual map for jpeg2000 compressed images. In *Proc. IEEE, ICIP* (Atlanta, GA, USA, 2006). 4

[BT06] BRUCE N., TSOTSOS J.: Saliency based on information maximization. *Advances in Neural Information Processing Systems 18* (2006), 155–162. 3

[CB07] CHENG I., BASU A.: Perceptually Optimized 3-D Transmission Over Wireless Networks. *IEEE Transactions on Multimedia 9*, 2 (Feb. 2007), 386–396. 17

[CDGE05] CORSINI M., DRELIE GELASCA E., EBRAHIMI T.: A Multi-Scale Roughness Metric for 3D Watermarking Quality Assessment. In *Workshop on Image Analysis for Multimedia Interactive Services 2005, April 13-15, Montreux, Switzerland.* (2005), ISCAS, SPIE. 8

[CFK09] CERF M., FRADY E. P., KOCH C.: Faces and text attract gaze independent of the task: Experimental data and computer model. *journal of vision 9*, 12 (2009), 1–15. 4

[CGEB07] CORSINI M., GELASCA E. D., EBRAHIMI T., BARNI M.: Watermarked 3-D Mesh Quality Assessment. *IEEE Transactions on Multimedia 9*, 2 (Feb. 2007), 247–256. 8, 12, 13, 14, 15, 16

[CH07] CHANDLER D. M., HEMAMI S. S.: Vsnr : A wavelet-based visual signal-to-noise ratio for natural images. *IEEE Transactions on Image Processing 16*, 9 (2007), 2284–2298. 5

[CM03] CAYRE F., MACQ B.: Data hiding on 3-D triangle meshes. *IEEE Transactions on Signal Processing 51*, 4 (2003), 939–949. 16

[CPJ07] CHO J. W., PROST R., JUNG H. Y.: An oblivious watermarking for 3-D polygonal meshes using distribution of vertex norms. *IEEE Transactions on Signal Processing 55*, 1 (2007), 142–155. 16

[CR68] CAMPBELL F. W., ROBSON J. G.: Application of fourier analysis to the visibility of gratings. *Journal of Physiology 197* (1968), 551–566. 2

[CS06] CLEJU I., SAUPE D.: Evaluation of supra-threshold perceptual metrics for 3D models. In *symposium on Applied percep-*

tion in graphics and visualization (July 2006), ACM Press, p. 41. 13

[CW07]   CHENG Y. M., WANG C. M.: An adaptive stegano-graphic algorithm for 3D polygonal meshes. *The Visual Computer 23*, 9 (2007), 721–732. 15

[CYDB08]   CHENG I., YING L., DANIILIDIS K., BASU A.: Robust and Scalable Transmission of Arbitrary 3D Models over Wireless Networks. *EURASIP Journal on Image and Video Processing 2008* (2008), 1–14. 17

[Dal93]   DALY S.: *Digital Images and Human Vision*. MIT press, 1993, ch. The Visible Differences Predictor : an Algorithm for the Assessment of Image Fidelity, pp. 197–206. 2, 3, 6

[Dal94]   DALY S.: A visual model for optimizing the design of image processing algorithms. In *International Conference on Image Processing* (1994), pp. 16–20. 6

[DF88]   DONNELLY H., FEFFERMAN C.: Nodal sets of eigen-functions on Riemannian manifolds. *Inventiones Mathematicae 93*, 1 (1988), 161–183. 15

[DGECB05]   DRELIE GELASCA E., EBRAHIMI T., CORSINI M., BARNI M.: Objective Evaluation of the Perceptual Quality of 3D Watermarking. In *IEEE International Conference on Image Processing (ICIP)* (2005), IEEE. 8, 16

[DPF03]   DUMONT R., PELLACINI F., FERWERDA J. A.: Perceptually-driven decision theory for interactive realistic rendering. *ACM Trans. Graph. 22*, 2 (2003), 152–181. 7

[Ebr09]   EBRAHIMI T.: Quality of Multimedia Experience : Past, Present and Future. In *ACM international conference on Multimedia* (2009), pp. 3–4. 12

[Fai05]   FAIRCHILD M. D.: *Color Appearance Models, 2nd ed.* Wiley-IS&T, Chichester, UK, 2005. 3, 4

[FSPG97]   FERWERDA J. A., SHIRLEY P., PATTANAIK S. N., GREENBERG D. P.: A model of visual masking for computer graphics. In *SIGGRAPH '97: Proceedings of the 24th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1997), ACM Press/Addison-Wesley Publishing Co., pp. 143–152. 3, 7

[HCA*12]   HERZOG R., CADIK M., AYDIN T. O., KIM K. I., MYSZKOWSKI K., SEIDEL H.-P.: NoRM : No-Reference Image Quality Metric for Realistic Image Synthesis. *Computer Graphics Forum 31*, 2 (2012). 7

[HHO04]   HOWLETT S., HAMILL J., O'SULLIVAN C.: An experimental approach to predicting saliency for simplified polygonal models. In *APGV '04: Proceedings of the 1st Symposium on Applied perception in graphics and visualization* (New York, NY, USA, 2004), ACM, pp. 57–64. 8

[HKL09]   HEU J., KIM C.-S., LEE S.-U.: Snr and temporal scalable coding of 3-d mesh sequences using singular value decomposition. *J. Visual Communication and Image Representation* (2009), 439–449. 9

[HZ08]   HOU X., ZHANG L.: Saliency detection: A spectral residual approach. In *IEEE Conference on Computer Vision and Pattern Recognition* (2008), pp. 1–8. 3

[IKN98]   ITTI L., KOCH C., NIEBUR E.: A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence 20* (1998), 1254–1259.

[IR03]   IBARRIA L., ROSSIGNAC J.: Dynapack: space-time compression of the 3d animations of triangle meshes with fixed connectivity. In *Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation* (Aire-la-Ville, Switzerland, Switzerland, 2003), SCA '03, Eurographics Association, pp. 126–135. 9

[ISO]   Iso/iec jtc1/sc29/wg11 ˛Sanimation framework extension core experiments descriptionŤ, doc. n6988, hong kong, china, 2005. 9

[JKJ*04]   JANG E. S., KIM J. D. K., JUNG S. Y., HAN M., WOO S. O.: Interpolator data compression for mpeg-4 animation. *IEEE Trans. Circuits Syst. Video Techn. 14*, 7 (2004), 989–1008. 9

[KBT10]   KIM K., BARNI M., TAN H. Z.: Roughness-adaptive 3-D watermarking based on masking effect of surface roughness. *IEEE Transactions on Information Forensics and Security 5*, 4 (2010), 721–733. 15

[Kee02]   KEELAN B. W.: *Handbook of Image Quality: Characterization and Prediction*. Marcel Dekker, New York, NY, 2002. 4

[Kel79]   KELLY D. H.: Motion and vision. ii. stabilized spatiotemporal threshold surface. *Journal of Optical Society of America 69* (1979), 1340–1349. 3

[KG00]   KARNI Z., GOTSMAN C.: Spectral compression of mesh geometry. In *SIGGRAPH '00: Proceedings of the 27th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 2000), ACM Press/Addison-Wesley Publishing Co., pp. 279–286. 8, 9, 13, 14, 15

[KG04]   KARNI Z., GOTSMAN C.: Compression of soft-body animation sequences. *Computer and Graphics 28* (2004), 25–34. 9

[KK93]   KARUNASEKERA S. A., KINGSBURY N. G.: A distorsion measure for blocking artifacts in images based on human visual sensitivity. In *SPIE, Visual Comm. and Image Processing* (1993), vol. 2094, pp. 474–486. 6

[KKK02]   KIM S., KIM S., KIM C.: Discrete differential error metric for surface simplification. In *Pacific Graphics* (2002), pp. 276–283. 8

[KTZC09]   KANAN C., TONG M., ZHANG L., COTTRELL G.: Sun: Top-down saliency using natural statistics. *Visual Cognition 17*, 6-7 (2009), 979–1003. 3

[KW02]   KUTTER M., WINKLER S.: A vision-based masking model for spread-spectrum image watermarking. *IEEE Transactions on Image Processing 11*, 1 (2002), 16–25. 3

[Lav09]   LAVOUÉ G.: A local roughness measure for 3D meshes and its application to visual masking. *ACM Transactions on Applied Perception (TAP) 5*, 4 (2009). 12, 13, 14, 15, 17

[Lav11]   LAVOUÉ G.: A Multiscale Metric for 3D Mesh Visual Quality Assessment. *Computer Graphics Forum 30*, 5 (2011), 1427–1437. 8, 13, 14

[LBF06]   LARABI M.-C., BRODBECK V., FERNANDEZ C.: A novel approach for constructing an achromatic contrast sensitivity function by matching. In *IEEE ICIP* (2006), pp. 441–441. 3

[LC10]   LAVOUÉ G., CORSINI M.: A comparison of perceptually-based metrics for objective evaluation of geometry processing. *IEEE Transactions on Multimedia 12*, 7 (2010), 636–649. 13, 14

[LDD*06]   LAVOUE G., DRELIE GELASCA E., DUPONT F., BASKURT A., EBRAHIMI T.: Perceptually driven 3D distance metrics with application to watermarking. In *SPIE* (Aug. 2006), vol. 6312, SPIE, pp. 63120L–63120L–12. 8, 12, 13, 14, 16

[LDD07]   LAVOUÉ G., DENIS F., DUPONT F.: Subdivision surface watermarking. *Computers & Graphics 31*, 3 (2007), 480–492. 15

[Len99]   LENGYEL J. E.: Compression of time-dependent geometry. In *In I3D Š99: Proceedings of the 1999 symposium on Interactive 3D graphics* (1999), ACM, pp. 89–95. 9

[LH01] LUEBKE D., HALLEN B.: Perceptually Driven Simplification for Interactive Rendering. In *Eurographics Workshop on Rendering Techniques* (2001), pp. 223–234. 6

[Lin00] LINDSTROM P.: *Model simplification using image and geometry-based metrics*. PhD thesis, Georgia Institute of Technology, Atlanta, GA, USA, 2000. 6

[LKT*07] LEE P.-F., KAO C.-K., TSENG J.-L., JONG B.-S., LIN T.-W.: 3d animation compression using affine transformation matrix and principal component analysis. *IEICE - Trans. Inf. Syst. E90-D* (July 2007), 1073–1084. 9

[LLD12] LEE H., LAVOUÉ G., DUPONT F.: Rate-distortion optimization for progressive compression of 3D mesh with color attributes. *The Visual Computer 28*, 2 (May 2012), 137–153. 17

[LO11] LARKIN M., O'SULLIVAN C.: Perception of simplification artifacts for animated characters. In *Proceedings of the ACM SIGGRAPH Symposium on Applied Perception in Graphics and Visualization* (New York, NY, USA, 2011), APGV '11, ACM, pp. 93–100. 10

[LPG08] LIU Y., PRABHAKARAN B., GUO X.: A robust spectral approach for blind watermarking of manifold surfaces. In *Proc. of the ACM Workshop on Multimedia and Security* (2008), pp. 43–52. 15

[LT00] LINDSTROM P., TURK G.: Image-driven simplification. *ACM Trans. Graph. 19*, 3 (2000), 204–241. 6, 10

[Lub93] LUBIN J.: *Digital Images and Human Vision*. MIT press, Cambridge, MA, 1993, ch. The use of psychophysical data and models in the analysis of display system performance, pp. 163–178. 6

[Lub95] LUBIN. J.: *A visual discrimination model for imaging system design and evaluation*. World Scientific Publishing Company, 1995, pp. 245–283. 6, 7

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. In *SIGGRAPH '05: ACM SIGGRAPH 2005 Papers* (New York, NY, USA, 2005), ACM, pp. 659–666. 8, 11

[LZ10] LÉVY B., ZHANG H.: Spectral mesh processing. In *Proc. of Siggraph Asia Course Notes* (2010), pp. 1–47. 14

[MK88] MOVSHON T., KIORPES L.: Analysis of the development of spatial sensitivity in monkey and human infants. *Journal of Optical Society of America JOSA A 5*, 12 (1988), 2166–2172. 2

[MK05] MUIJS R., KIRENKO I.: A no-reference blocking artifact measure for adaptive video processing. In *In European Signal Processing Conference* (2005). 4

[MNO07] MCDONNELL R., NEWELL F., O'SULLIVAN C.: Smooth movers: perceptually guided human motion simulation. In *Proceedings of the 2007 ACM SIGGRAPH/Eurographics Symposium on Computer animation* (Aire-la-Ville, Switzerland, Switzerland, 2007), SCA '07, Eurographics Association, pp. 259–269. 10

[MPG*09] MARAT S., PHUOC T., GRANJON L., GUYADER N., PELLERIN D., GUERIN-DUGUE A.: Modelling spatio-temporal saliency to predict gaze direction for short videos. *International Journal of Computer Vision 82*, 3 (2009), 231–243. 3

[MS74] MANNOS J. L., SAKRISON D. J.: The effects of visual fidelity criterion on the encoding of images. *IEEE Transactions on Information Theory 20*, 4 (1974), 525–536. 2, 5, 15

[MSK*05] MÜLLER K., SMOLIC A., KAUTZNER M., EISERT P., WIEGAND T.: Predictive compression of dynamic 3d meshes. In *Proc. ICIP2005, IEEE International Conference on Image Processing* (2005). 9

[MSK*06] MÜLLER K., SMOLIC A., KAUTZNER M., EISERT P., WIEGAND T.: Rate-distortion-optimized predictive compression of dynamic 3d mesh sequences. *Sig. Proc.: Image Comm. 21*, 9 (2006), 812–828. 9

[MSK*08] MAMOU K., STEFANOSKI N., KIRCHHOFFER H., MULLER K., ZAHARIA T., PRETEUX F., MARPE D., OSTERMANN J.: The new mpeg-4/famc standard for animated 3d mesh compression. In *3DTV Conference: The True Vision - Capture, Transmission and Display of 3D Video, 2008* (may 2008), pp. 97–100. 18

[MSKW06] MÜLLER K., SMOLIC A., KAUTZNER M., WIEGAND T.: Rate-distortion optimization in dynamic mesh compression. In *ICIP* (2006), pp. 533–536. 9

[Mys02] MYSZKOWSKI K.: Perception-based global illumination, rendering, and animation techniques. In *SCCG '02: Proceedings of the 18th spring conference on Computer graphics* (New York, NY, USA, 2002), ACM, pp. 13–24. 8

[MZP06] MAMOU K., ZAHARIA T., PRÊTEUX F.: A skinning approach for dynamic 3d mesh compression: Research articles. *Comput. Animat. Virtual Worlds 17* (July 2006), 337–346. 9

[NLF10] NAUGE M., LARABI M.-C., FERNANDEZ C.: A reduced-reference metric based on the interest points in color images. In *PCS2010* (Nagoya, Japan, 2010), pp. 610–613. 4

[NLF11] NAUGE M., LARABI M.-C., FERNANDEZ C.: Imqual: A web-service dedicated to image quality evaluation and metrics benchmark. In *IS&T Electronic imaging symposium: Image quality and System Performance* (Burlingame, CA, 2011). 5

[ODGK03] O'SULLIVAN C., DINGLIANA J., GIANG T., KAISER M. K.: Evaluating the visual fidelity of physically based animations. *ACM Transactions on Graphics 22* (July 2003), 527–536. 11

[OMT02] OHBUCHI R., MUKAIYAMA A., TAKAHASHI S.: A frequency-domain approach to watermarking 3D shapes. *Computer Graphics Forum (Proc. of Eurographics) 21*, 3 (2002), 373–382. 15

[PA06] PAYAN F., ANTONINI M.: Mean square error approximation for wavelet-based semiregular mesh compression. *IEEE transactions on visualization and computer graphics 12*, 4 (2006), 649–57. 17

[PAK98] PETITCOLAS F. A. P., ANDERSON R. J., KUHN M. G.: Attacks on copyright marking systems. In *Proc. of the International Workshop on Information Hiding* (1998), pp. 218–238. 16

[PCA05] PAN Y., CHENG I., ANUP BASU: Quality metric for approximating subjective evaluation of 3-D objects. *IEEE Transactions on Multimedia 7*, 2 (Apr. 2005), 269–279. 8, 12, 17

[PLR*04] PAN F., LIN X., RAHARDJA S., LIN W., ONG E., YAO S., LU Z., YANG X.: A locally adaptive algorithm for measuring blocking artifacts in images and videos. *Signal Processing Image Communication 19* (2004), 499–506. 4

[PS00] PAPPAS T. N., SAFRANEK R. J.: Perceptual criteria for image quality evaluation. In *Handbook of Image and video Processing*, Bovik A., (Ed.). Academic press, 2000, pp. 669–684. 5

[PV10] PETŘÍK O., VÁŠA L.: Finding optimal parameter configuration for a dynamic triangle mesh compressor. In *Proceedings of the 6th international conference on Articulated motion and deformable objects* (Berlin, Heidelberg, 2010), AMDO'10, Springer-Verlag, pp. 31–42. 18

[QM08] QU L., MEYER G. W.: Perceptually guided polygon reduction. *IEEE Transactions on Visualization and Computer Graphics 14*, 5 (2008), 1015–1029. 7

[Rec99] RECOMMENDATION ITU-T P.910 .: Subjective video quality assessment methods for multimedia applications. 12

[Rec02]   RECOMMENDATION ITU-R BT. 500-11 .: Methodology for subjective assessment of the quality of television pictures. 11, 12

[Rec07]   RECOMMENDATION ITU-R BT. 1788 .: Methodology for the subjective assessment of video quality in multimedia applications. 12

[Red97]   REDDY M.: *Perceptually Modulated Level of Detail for Virtual Environments*. PhD thesis, Dept. of Computer Science, University of Edinburgh, UK, 1997. 7

[RFWB07]   RAMANARAYANAN G., FERWERDA J., WALTER B., BALA K.: Visual equivalence: towards a new standard for image fidelity. In *SIGGRAPH '07: ACM SIGGRAPH 2007 papers* (New York, NY, USA, 2007), ACM, p. 76. 7

[RLFM08]   ROSSELLI V., LARABI M., FERNANDEZ-MALOIGNE C.: Perceptual color difference metric based on the perception threshold. In *IS&T Image Quality and System Performance V* (2008). 5

[Rou10]   ROUDET C.: A Region-Based Progressive Coding of Semi-Regular 3-D Meshes for View-Dependent Transmission. In *2010 Sixth International Conference on Signal-Image Technology and Internet Based Systems* (Dec. 2010), IEEE, pp. 51–59. 17

[RPG99]   RAMASUBRAMANIAN M., PATTANAIK S. N., GREENBERG D. P.: A perceptually based physical error metric for realistic image synthesis. In *SIGGRAPH '99: Proceedings of the 26th annual conference on Computer graphics and interactive techniques* (New York, NY, USA, 1999), ACM Press/Addison-Wesley Publishing Co., pp. 73–82. 7

[RR01]   ROGOWITZ B. E., RUSHMEIER H.: Are image quality metrics adequate to evaluate the quality of geometric objects? *Proceedings of SPIE* (2001), 340–348. 6, 7, 12

[RRP00]   RUSHMEIER H., ROGOWITZ B., PIATKO C.: Perceptual issues in substituting texture for geometry. In *SPIE* (2000), International Society for Optical Engineering; 1999, pp. 372–383. 12

[RV10]   RUS J., VÁŠA L.: Analysing the influence of vertex clustering on pca-based dynamic mesh compression. In *AMDO* (2010), pp. 55–66. 18

[RV12]   RUS J., VÁŠA L.: Deblocking for dynamic triangle meshes. In *To appear in proceedings of GRAPP 2012* (Feb. 2012). 18

[RvdLBC08]   RAJASHEKAR U., VAN DER LINDE I., BOVIK A. C., CORMACK L. K.: Gaffe: A gaze-attentive fixation finding engine. *IEEE Transcations on Image Processing 17*, 4 (2008), 564–573. 3

[SB06]   SHEIKH H., BOVIK A.: Image information and visual quality. *IEEE Transactions on Image Processing 15*, 2 (February 2006), 430–444. 5

[SBd05]   SHEIKH H. R., BOVIK A., DEVECIANA G.: An information fidelity criterion for image quality assessment using natural scene statistics. *IEEE Transactions on Image Processing 14*, 12 (December 2005), 2117–2128. 5

[Sch07]   SCHANDA J.: *Colorimetry: Understanding the CIE System*. John Wiley & Sons, Hoboken, NJ, USA, 2007. 4

[SCOT03]   SORKINE O., COHEN-OR D., TOLDEO S.: High-pass quantization for mesh encoding. In *Eurographics Symposium on Geometry Processing* (2003), pp. 42–51. 8, 13, 14, 16, 17, 18

[SPC04]   SAFRANEK R. J., PAPPAS T. N., CHEN J.: Perceptual criteria for image quality evaluation. In *In Handbook of Image and Video Processing*, Bovik A., (Ed.). Academic Press, 2004. 2, 3

[SSF07]   SILVA S., SANTOS B., FERREIRA C.: Comparison of methods for the simplification of mesh models using quality indices and an observer study. *SPIE* (2007), 64921L–64921L–12. 12, 13, 14

[SSFM09]   SILVA S., SANTOS B. S., FERREIRA C., MADEIRA J.: A Perceptual Data Repository for Polygonal Meshes. *2009 Second International Conference in Visualisation* (July 2009), 207–212. 12, 13

[STN*01]   SOLACHIDIS V., TEFAS A., NIKOLAIDIS N., TSEKERIDOU S., NIKOLAIDIS A., PITAS I.: A benchmarking protocol for watermarking methods. In *Proc. of the IEEE International Conference on Image Processing* (2001), vol. 3, pp. 1023–1026. 16

[TA04]   TIAN D., ALREGIB G.: FQM: A Fast Quality Measure for Efficient Transmission of Textured 3D Models. In *ACM Multimedia* (2004), pp. 684–691. 8, 17

[TA08]   TIAN D., ALREGIB G.: Batex3: Bit allocation for progressive transmission of textured 3-d models. *IEEE Transactions on Circuits and Systems for Video Technology 18*, 1 (2008), 23–35. 17

[Tau00]   TAUBIN G.: Geometric signal processing on polygonal meshes. *Eurographics State of the Art Reports 4* (2000). 2

[TLF08]   TULET O., LARABI M.-C., FERNANDEZ C.: Use of spatial adaptation for image rendering based on an extension of the ciecam02. In *International Conference on Computer Vision Theory and Applications* (Madera, Portugal, 2008), pp. 128–133. 4

[TOCH06]   TORRALBA A., OLIVA A., CASTELHANO M., HENDERSON J.: Contextual guidance of eye movements and attention in real-world scenes: The role of global features on object search. *Psychological Review 113* (2006), 766–786. 3

[Tor11]   TORKHANI F.: *Perceptual Quality Assessment of Dynamic Meshes*. Master thesis, Grenoble INP, France, 2011. 15

[VCP09]   VALETTE S., CHAINE R., PROST R.: Progressive lossless mesh compression via incremental parametric refinement. In *Proceedings of the Symposium on Geometry Processing* (July 2009), Eurographics Association, pp. 1301–1310. 17

[VL08]   VALLET B., LÉVY B.: Spectral geometry processing with manifold harmonics. *Computer Graphics Forum (Proc. of Eurographics) 27*, 2 (2008), 251–260. 15

[VP11]   VÁŠA L., PETŘÍK O.: Optimising perceived distortion in lossy encoding of dynamic meshes. *Comput. Graph. Forum 30*, 5 (2011), 1439–1449. 18

[VS06]   VÁŠA L., SKALA V.: A spatio-temporal metric for dynamic mesh comparison. In *AMDO* (2006), pp. 29–37. 10

[VS07]   VÁŠA L., SKALA V.: Coddyac: Connectivity driven dynamic mesh compression. In *3DTV-CON, The True Vision - Capture, Transmission and Display of 3D Video* (May 2007). 18

[VS11]   VÁŠA L., SKALA V.: A perception correlated comparison method for dynamic meshes. *IEEE Trans. Vis. Comput. Graph. 17*, 2 (2011), 220–230. 10, 13, 14, 17

[Wan95]   WANDELL B. A.: *Foundations of Vision*. Sunderland, MA, Sinauer Associates, 1995. 2, 3

[Wat87]   WATSON A.: The cortex transform: rapid computation of simulated neural images. *computer vision, graphics an image processing 39*, 3 (1987), 311–327. 6

[WBSS04]   WANG Z., BOVIK A., SHEIKH H., SIMONCELLI E.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing 13*, 4 (2004), 1–14. 5, 8

[WFM01] WATSON B., FRIEDMAN A., MCGAFFEY A.: Measuring and predicting visual fidelity. *ACM Siggraph* (2001), 213–220. 12, 13

[WHST01] WU J.-H., HU S.-M., SUN J.-G., TAI C.-L.: An effective feature-preserving mesh simplification scheme based on face constriction. *Computer Graphics and Applications, Pacific Conference on 0* (2001), 0012. 8

[Win02] WINKLER S.: *Vision models and quality metrics for image processing applications*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Lausanne, CH, 2002. 2

[WLBD09] WANG K., LUO M., BORS A. G., DENIS F.: Blind and robust mesh watermarking using manifold harmonics. In *Proc. of the IEEE International Conference on Image Processing* (2009), pp. 3657–3660. 15

[WLC*03] WILLIAMS N., LUEBKE D., COHEN J. D., KELLEY M., SCHUBERT B.: Perceptually guided simplification of lit, textured meshes. In *I3D '03: Proceedings of the 2003 symposium on Interactive 3D graphics* (New York, NY, USA, 2003), ACM, pp. 113–121. 7

[WLD*10] WANG K., LAVOUÉ G., DENIS F., BASKURT A., HE X.: A benchmark for 3D mesh watermarking. In *Proc. of the IEEE International Conference on Shape Modeling and Applications* (2010), pp. 231–235. 16

[WLDB08a] WANG K., LAVOUÉ G., DENIS F., BASKURT A.: A comprehensive survey on three-dimensional mesh watermarking. *IEEE Transactions on Multimedia 10*, 8 (2008), 1513–1527. 14

[WLDB08b] WANG K., LAVOUÉ G., DENIS F., BASKURT A.: Hierarchical watermarking of semi-regular meshes based on wavelet transform. *IEEE Transactions on Information Forensics and Security 3*, 4 (2008), 620–634. 15

[WLDB11] WANG K., LAVOUÉ G., DENIS F., BASKURT A.: Robust and blind mesh watermarking based on volume moments. *Computers & Graphics 35*, 1 (2011), 1–19. 2, 15, 16

[WS05] WANG Z., SIMONCELLI E. P.: Reduce-reference image quality assessment using a wavelet-domain natural image statistic model. In *Human Vision and Electronic Imaging X* (2005), vol. 5666, pp. 17–20. 4

[WSB03] WANG Z., SHEIKH H. R., BOVIK A. C.: Objective video quality assessment. In *The Handbook of Video Databases: Design and Applications*. CRC Press, 2003, pp. 1041–1078. 4

[Yar67] YARBUS A. L.: *Eye movements and vision*. Plenum Press, New York, 1967. 3, 4

[Yen98] YENDRIKHOVSKIJ S. N.: *Color reproduction and the naturalness constraint*. PhD thesis, Technische Universiteit Eindhoven, The Netherlands, 1998. 5

[YIK03] YU Z., IP H. H. S., KWOK L. F.: A robust watermarking scheme for 3D triangular mesh models. *Pattern Recognition 36*, 11 (2003), 2603–2614. 15

[YN04] YEE Y. H., NEWMAN A.: A perceptual metric for production testing. In *ACM SIGGRAPH 2004 Sketches* (New York, NY, USA, 2004), SIGGRAPH '04, ACM, p. 121. 11

[YPG01] YEE H., PATTANAIK S., GREENBERG D. P.: Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments. *ACM Trans. Graph. 20*, 1 (2001), 39–65. 8

[YPSZ01] YIN K., PAN Z., SHI J., ZHANG D.: Robust mesh watermarking based on multiresolution processing. *Computers & Graphics 25*, 3 (2001), 409–420. 15

[ZDL02] ZENG W., DALY S., LEI S.: An overview of the visual optimization tools in jpeg 2000. *Signal Processing: Image Communication 17*, 1 (2002), 85 – 104. 7

[ZW97] ZHANG X. M., WANDELL B. A.: A spatial extension of cielab for digital color image reproduction. *SID Journal 5*, 1 (1997), 61–63. 5

[ZZDZ10] ZHU Q., ZHAO J., DU Z., ZHANG Y.: Quantitative analysis of discrete 3D geometrical detail levels based on perceptual metric. *Computers & Graphics 34*, 1 (Feb. 2010), 55–65. 7