DOI: 10.1111/cgf.12001

COMPUTER GRAPHICS *forum*
Volume 32 (2013), number 1 pp. 101–125

Perceptual Metrics for Static and Dynamic Triangle Meshes

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

[1]Visual Computing Laboratory, ISTI-CNR, Italy
massimiliano.corsini@isti.cnr.it
[2]Université de Poitiers, CNRS, XLIM-SIC UMR 7252, France
chaker.larabi@univ-poitiers.fr
[3]Université de Lyon, CNRS, Insa-Lyon, LIRIS UMR 5205, France
guillaume.lavoue@insa-lyon.fr
[4]Centre of Computer Graphics and Visualization, University of West Bohemia, Pilsen, Czech Republic
{opetrik, lvasa}@kiv.zcu.cz
[5]GIPSA-lab, CNRS UMR5216, France
kai.wang@gipsa-lab.grenoble-inp.fr

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.

Keywords: metric, distortion, perception, mesh, surface, evaluation, watermarking, compression, simplification, human vision, animation, comparison

ACM CCS: 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

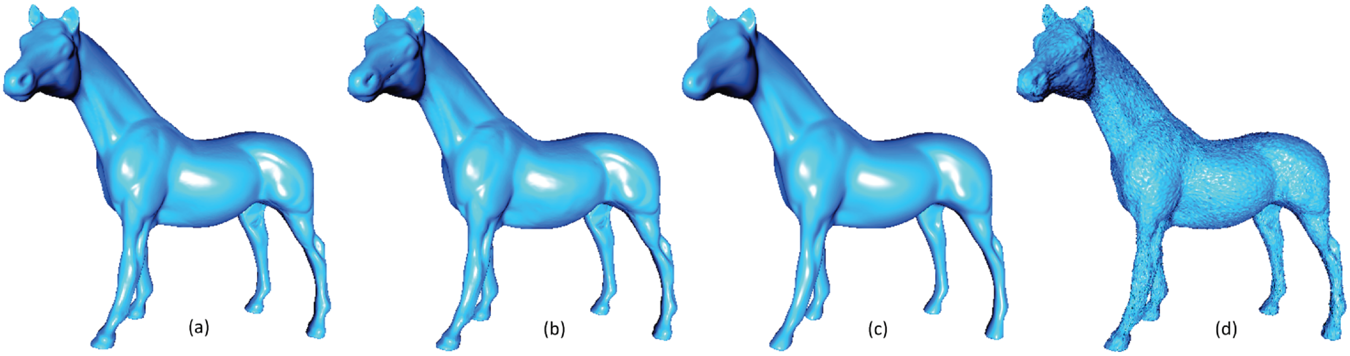

Figure 1: *Original and distorted versions of the Horse model, all associated with the same maximum root mean square error* ($MRMS = 1.05 \times 10{-}3$)*. (a) Original model. Results after (b) watermarking from Wang et al. [WLDB11], (c) Laplacian smoothing [Tau00], (d) Gaussian noise addition.*

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 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.

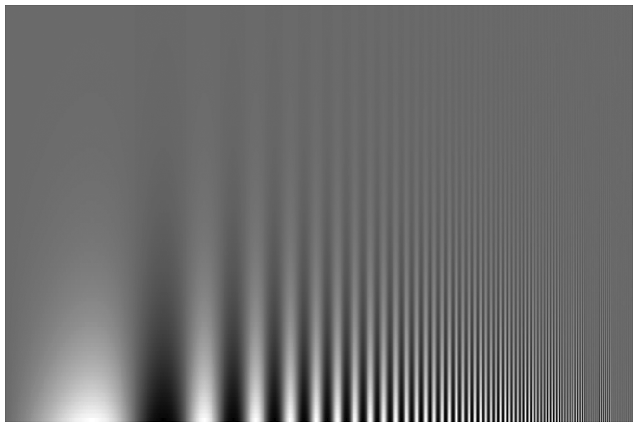

Figure 2: *Campbell and Robson chart [CR68].*

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 modelling the HVS behaviour is an essential step for developing image-based applications [Wan95]. This allows one to take advantage of the end-user perception to hide or highlight specific details and thus evaluate the perceived quality of an image or an image sequence [SPC04, Win00].

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 imperceptible 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 bandpass filter characterises the luminance CSF with a peak frequency between 4 and 6 cpd (cycles by degree) and a cut-off around 30 cpd. Assuming that an observer can be in a range of viewing distances, the CSF can be considered as the envelope of functions each describing the behaviour for a given range of spatial frequencies.

One of the most popular analytical models was introduced by Mannos and Sakrison [MS74] in the 1970s 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 bandpass CSF giving more sensitivity to horizontal and vertical spatial frequencies in comparison to diagonal frequencies. Another complex model has been proposed 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(FR) evaluation) or between the provided stimulus and an implicit reference (for a no-reference (NR) 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 behaviour with cut-off frequencies lower than those of luminance. This behaviour is partially explained by the fact that edge detection/enhancement does not occur in the chromatic dimension [Fai05].

Finally, the spatio-temporal CSF aims at defining the detection threshold of a dynamic stimulus with regards to its spatial and temporal frequencies. This has been demonstrated by Kelly's [Kel79] pioneering work through experiments measuring threshold contrast while viewing travelling sine-waves instead of using flickering spatial gratings. Figure 3 gives the variation of the CSF function of the velocity, showing that the constant velocity CSF curves are regular from the shape point of view for any velocity higher than 0.1 degree per second. A widely used formula has been derived from the results based on analytical approximations giving a spatio-velocity CSF that is convertible to spatio-temporal CSF by an appropriate equation. Daly has extended Kelly's model to comply with characteristics of the CRT displays [Dal98].

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, luminance adaptation is 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. Secondly, texture masking points out the maximised visibility of a stimulus on homogeneous regions rather than on textured ones [Wan95].

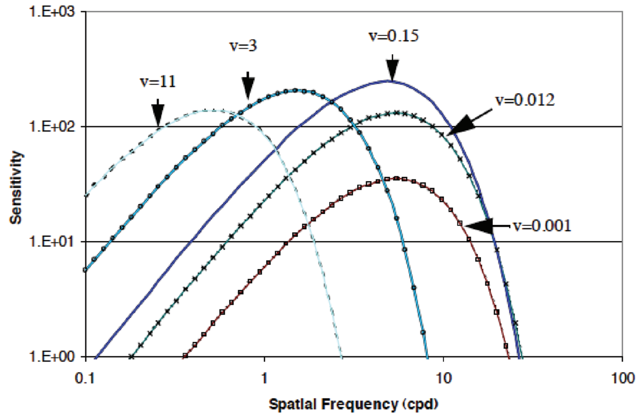

Figure 3: *Velocity-dependent contrast sensitivity function as defined empirically by Kelly [Kel79].*

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 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 study and understand this property. One of the most famous and often-cited studies was performed by Yarbus in the 1960s [Yar67], where he 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. One can note that the visual path corresponding to a free exploration is different from the path obtained when subjects were asked to judge a given situation.

Visual saliency models try to mimic the HVS 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], [PKPH12] 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

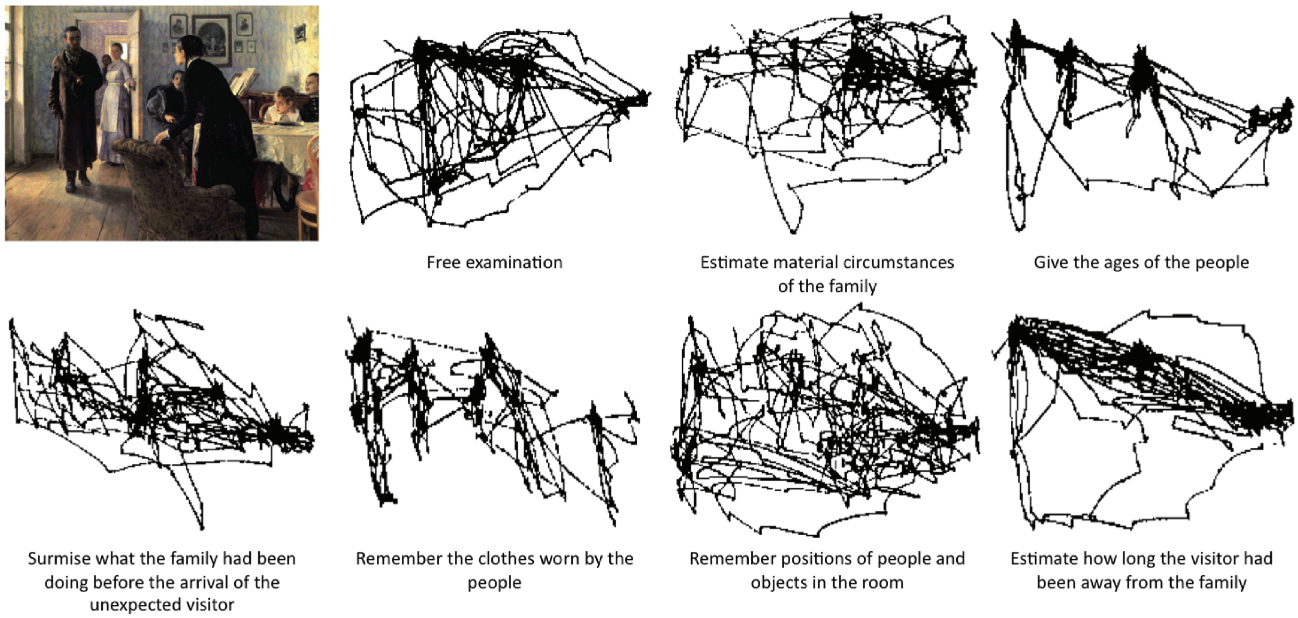

Figure 4: *Experiments performed by Yarbus [Yar67] on how the task given to a person influences the eye movement.*

demonstrated 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. The perception of a colour stimulus is partly dependent on the environment's properties, such as background colour and lighting conditions. 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.

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 FR, reduced reference (RR) and 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. Several metrics have been proposed in recent years and are described in [NLF10] and [WS05]. NR metrics are distortion-based; the specialised metric looks for a specific artefact 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] and [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 to study the performance of these metrics with regard to human judgment. 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]. Another web-application dedicated to quality assessment of images and video with any dynamic range has been developed (http://drim.mpi-sb.mpg.de/).

The signal-oriented metrics do not take into account any comprehensive HVS model with regard to quality evaluation. 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 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 visual information fidelity [SB06] has been developed with the aim to quantify the loss of image information because of the distortion process.

The second type of FR metrics uses a single-channel modelling of the HVS. In this context, the HVS is seen as a spatial filter whose characteristics are given by the 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

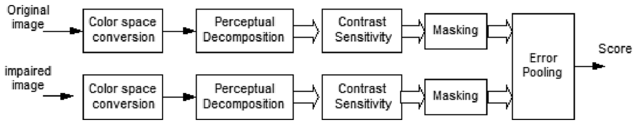

Figure 5: *Block diagram of perceptual metrics.*

the detection of sinusoidal gratings. An efficient metric called visual signal-to-noise ratio has been developed in [CH07] for quantifying the visual fidelity of natural images. 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 modelled 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 5.

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 processing of the chrominance channels. Then, a perceptual decomposition (multi-channel decomposition) is applied 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 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, by grouping them according to their main principle. Table 1 provides a summary of these metrics and their main characteristics.

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 *model-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 [Rog01] 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 analysing 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, for example the animation of an object or the user interaction. This last aspect, i.e. the feedback between the actions of the user dealing with an interactive applications and the corresponding movements of the objects, is in general not taken into account by the existing metrics.

3.2. Signal-oriented and HVS-based approaches

As raised in Section 2., there are two different approaches to developing perceptual metrics: *HVS-based* and *signal-oriented*. The HVS-based approach takes into account the complex mathematical models of the psychophysical and physiological mechanisms of the HVS to develop the perceptual metric although the signal-oriented 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 artefacts will be perceived by a human observer. The signal-oriented approach is preferable when the visual task to model is particularly complex. In the development of perceptual metrics for static meshes both the HVS-based and the signal-oriented approaches have been employed; the HVS-based approach has been mainly used for the development of image-based perceptual metrics (see Section 3.3.), instead, for the model-based perceptual metrics (see section 3.4.) the signal-oriented approach is usually followed because of the complexity of the perceptual

Table 1: *Properties of existing metrics for static meshes. The column 'Constraints' indicates if the metric require the meshes to share the same connectivity or the same level of details. The method [CDGEB07] does not require similar connectivity but similar vertex density. The correlation data have been synthesised from [LC10], [Lav11], [VR12], [WTM12] and [TWC12].*

	Image/model	HVS/signal	Principle	Constraints	Correlation
[Red97],[LH01] [WLC*03]	Image	HVS	CSF model	No	Not evaluated
[BM98], [Lin00], [QM08]	Image	HVS	Sarnoff VDM	No	Not evaluated
[Mys98], [RPG99] [DPF03], [ZZDZ10]	Image	HVS	Daly VDP	No	Not evaluated
[MG10]	Image/model	HVS	Spatio-temporal CSF [BK80]	No	Not evaluated
GL_1 [KG00]	Model	Signal	Geometric Laplacian	Yes	Poor
GL_2 [SCOT03]	Model	Signal	Geometric Laplacian	Yes	Poor
SF [BHM09]	Model	Signal	Triangle deformation	Yes	Poor
$3DWPM$ [CDGEB07]	Model	Signal	Roughness	No*	Moderate
$MSDM$ [LDD*06]	Model	Signal	Local curvature statistics	Yes	Moderate|High
$MSDM2$ [Lav11]	Model	Signal	Local curvature statistics	No	High
$TPDM$ [TWC12]	Model	Signal	Curvature directions	No	High
$FMPD$ [WTM12]	Model	Signal	Roughness	No	High
$DAME$ [VR12]	Model	Signal	Dihedral angles	Yes	High

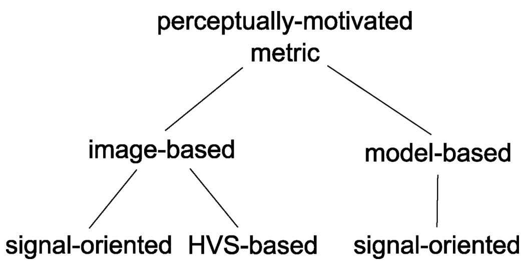

Figure 6: *Relationships between the application domain of the metric (image/model) and the approach followed to develop the metric (HVS-based/signal-oriented).*

mechanisms involved. Figure 6 summarizes the relationships between these approaches and our categorization.

3.3. Image-based perceptual metrics

These metrics are mainly used in perceptually driven rendering and in perceptually based mesh simplification. 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. The goal of *perceptually based simplification* is to reduce the number of polygons of a given mesh making the visual differences between the simplified model and the original one imperceptible.

3.3.1. *CSF*

One of the first studies for perceptually driven rendering 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.

Luebke and Hallen [LH01] developed a perceptually based simplification algorithm based on a simplified version of the CSF. 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.

Menzel and Guthe [MG10] propose a perceptual simplification algorithm based on contrast and spatial frequency computation [BK80] which is able to handle different materials through image-based bi-directional reflectance distribution function analysis. This algorithm is able to perform almost all the calculation directly on vertices instead of rendered images. However, it still uses the rendered views to evaluate the masking effect, thus it can be classified as an hybrid image-/model-based method.

3.3.2. *VDP*

Ferwerda *et al.* [FSPG97] were the first to propose a perception model for Computer Graphics with particular attention to visual masking, extending the original Daly *VDP* metric. In this work, they demonstrated how surface texture can hide some visual artefacts of the geometry in given shading conditions, in particular polygonal tessellation.

In the context of perceptually based rendering, Myszkowski *et al.* [Mys98] exploited the VDP to steer adaptive mesh subdivision and generate a simplified mesh for radiosity computation that produces the same perceived rendering of the original one. In the same context, Ramasubramanian *et al.* [RPG99] proposed a framework to reduce the computational burden of a global illumination (direct plus indirect lighting) rendering system. First, they evaluated a perceptual *threshold map* taking into account the direct illumination of the scene and then they use this map to drive the indirect illumination computation, which is usually the most computationally expensive task in a global illumination rendering system.

Another interesting approach is the one 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] because of its accuracy and computational efficiency.

Zhu *et al.* [ZZDZ10] studied the relationship between the viewing distance and the perceptibility of model details using 2D metrics (VDP and SSIM) for the optimal design of discrete LOD for the visualization of complex 3D building facades.

3.3.3. *Visual discrimination model (VDM)*

Bolin and Meyer [BM98] used a perceptual model to optimise the sampling for ray-tracing algorithms. The visual differences metric 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 metric is used to drive adaptively of the sampling in a ray-tracing framework.

In the ambit of perceptually driven mesh simplification, Lindstrom and Turk [LT00] proposed to render the model being simplified from several viewpoints and use a fast image quality metric (RMS error) to evaluate the impact of the simplification. Lindstrom [Lin00] proposed also to replace the 2D RMS error with a perceptual image metric based on a simplified version of the Sarnoff VDM. This algorithm is particularly effective for textured 3D models.

Qu and Meyer [QM08] considered the visual masking properties of 2D texture maps to drive simplification and re-meshing 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 re-meshing 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 re-meshing 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.

3.3.4. *Visual equivalence*

Recently, perceptual evaluation has been moved to a higher level of investigation concerning visual mechanisms. Ramanarayanan *et al.* [RFWB07] for first time, 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. A subsequent work about visual equivalence has been also presented by Ferwerda *et al.* [FRBW08]. Visual equivalence has been recently used to investigate the perception of rendering artefacts caused by global illumination approximations on different materials and shape [KFB10].

3.3.5. *Spatio-temporal (video)*

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. Recently, Aydin *et al.* [ACMS10] proposed another perceptual metric for the quality evaluation of computer-generated video based on the high dynamic range (HDR) VDP [MMS04], i.e. the extension of the VDP to HDR images. The application of these spatio-temporal perceptual metrics in the context of 3D model visual fidelity evaluation has, to our knowledge, never been investigated.

3.3.6. *Other*

Some other interesting methods have recently been proposed: Guthe *et al.* [GMSK09] introduce a perceptual metric based on spatio-temporal CSF dedicated to bidirectional texture functions (BTFs). This metric is proposed to measure the visual quality of the various compressed representations of BTF data. Bosc *et al.* [BPL*11] conducted a perceptual study about quality evaluation of synthesized views generated for 3D TV application, i.e. depth image-based rendering. Herzog *et al.* [HCA*12] introduce a NR perceptual quality metric to identify visual artefacts in images synthetically generated with rendering techniques.

3.4. Model-based perceptual metrics

The main limitation of the image-based metrics in the context of Computer Graphics applications is that, in general, as demonstrated by the experiments conducted by Rogowitz and Rushmeier [Rog01], 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.

3.4.1. *Geometric laplacian*

Several researchers have investigated the use of signal-oriented perceptual metrics for the evaluation of specific artefacts. 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. This type of metrics usually does not correlate well with the human perception.

3.4.2. *Curvature-based*

Model-based metrics are used in different applications. One of these is to control mesh simplification algorithms, to reduce the number of vertices although 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 multi-resolution curvature maps.

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 $MSDM2$, 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, Torkhani *et al.* [TWC12] have proposed a metric called $TPDM$ (Tensor-based Perceptual Distance Measure). Different from $MSDM$ and $MSDM2$ that only make use of the mesh curvature amplitude, $TPDM$ also takes into account the mesh principal curvature directions (therefore, the full information of the curvature tensor [ACSD*03]) for the purpose of mesh visual quality assessment. The motivation is that the directions of the two principal curvatures are those along which surface normals vary the slowest (e.g. along creases) and the fastest (e.g. across creases), which represent structural features of the surface and thus should be visually important. Moreover, $TPDM$ accounts for the visual masking effect by introducing a weighting factor in the computation of a perceptually oriented distance of mesh curvature tensors. Preliminary results [TWC12] show that it would be beneficial to use both amplitudes and directions of the curvatures for the assessment of mesh visual quality.

3.4.3. *Roughness-based*

Following the idea that a measure of the visual artefacts produced by watermarking should be based on the amount of *roughness* introduced on the surface, Corsini *et al.* [CDGEB07] 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 (at multiple scales) about the dihedral angles, i.e. the angle between the normals of two adjacent faces. The second method by Drelie Gelasca *et al.* [DGECB05] is based on the consideration that visual artefacts should be better perceived on smooth surfaces: 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.

Wang *et al.* [WTM12] developed a simple and fast metric for comparing the visual difference between a pair of meshes. As in [DGECB05] and [CDGEB07], the proposed Fast Mesh Perceptual Distance ($FMPD$) is based on the local and global roughness modification of the mesh surface. First, the mesh local roughness is computed as the Laplacian of the discrete Gaussian curvature. The local roughness is then modulated by using some simple mathematical functions, so as to account for the visual masking effect. The basic idea is to induce a large perceptual distance where a smooth region becomes rough, but to induce a small distance where a rough region still remains rough after distortion. The global roughness of a mesh is computed as the normalised surface integral of the modulated local roughness, and the visual difference between two meshes is simply the difference between their global roughness values. Besides, its low computing time and its good coherence with subjective scores, $FMPD$ also has the potential to be used as a *RR* metric. A Matlab implementation of $FMPD$ is available at http://www.gipsa-lab.inpg.fr/ kai.wang/publications_en.html.

3.4.4. *Strain energy*

Bian *et al.* [BHM08][BHM09] developed a model-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. Bian *et al.* tested some variants of this metric by choosing different weighting strategy of the local contributions, but from their experimental results they concluded that the unweighted version gave results similar to the tested variants; hence, it is preferable because of its simplicity.

3.4.5. *Texture geometry*

Tian and AlRegib [TA04] and Pan *et al.* [PCB05] 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 [PCB05]. Their results underline the fact that the perceptual contribution of image texture is, in general, more important than the model's geometry, i.e. the reduction of the texture resolution is perceived more degraded than the reduction of model's polygons (geometry resolution).

3.4.6. *Oriented dihedral angles*

Motivated by the promising results based on comparing dihedral angles of [CDGE05], Váša and Rus [VR12] have proposed a metric based on measuring distortion of corresponding oriented dihedral angles. The metric only works for comparing meshes of shared connectivity, but in this field it provides a very high correlation with the results of subjective testing and for such case it is also significantly faster to evaluate than competing perceptual metrics, which makes it a suitable candidate for iterative optimization algorithms. In addition, the masking effect is addressed by using a weighting term that reflects the original dihedral angle. An implementation of the dihedral angle mesh error (DAME) mesh comparer can be downloaded at http://compression.kiv.zcu.cz.

4. Metrics for Dynamic 3D Meshes

The approaches to dynamic mesh distortion evaluation (for overview see Table 2) 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 perception in the static case will most likely also fail when applied to the dynamic case. Moreover, although CSF has a bandpass characteristic, its spatio-temporal counterpart ST-CSF is inherently a low-pass function. This means that some low-frequency mesh distortions, that are not visible in static case, might become readily visible in the context of animation. We call such cases *temporal artefacts*, and a proper metric probably should detect this kind of artefact as well. However, temporal artefacts of course cannot be detected by a static mesh metric applied in per-frame fashion.

A typical example of such an artefact 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 unnotable on a 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, 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] and [KPA09]—show temporal development of root mean squared error (RMSE) or its average, although 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 artefacts 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 $KG(M, M') = f(M, AMSE(M, M'))$. Because of this fact, one might expect that the KG error metric will show the same insufficiencies as any other averaged static metric based on mean squared error (MSE).

4.3. D_a error

Another metric designed specifically for animated meshes has been proposed by Jang *et al.* [JKJ*04]. This metric 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 error vectors projected into coordinate axes, 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 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

Table 2: *Properties of existing metrics for dynamic meshes (the column 'Constraints' indicates if the metric require the meshes to share the same connectivity).*

	Principle	Based on static mesh metric	Constraints	Handling of temporal artifacts	Correlation
KG [KG04]	Normalized average MSE	Yes	Yes	No	Poor
D_a [JKJ*04]	Error vectors	No	Yes	No	Poor
4D Hausdorff [VS06]	Hausdorff distance	No	No	No	Poor
STED [VS11]	Edge difference	No	Yes	Yes	High
PDiff [LO11]	Image comparison	Yes	No	No	Not evaluated

vertices and all pairs of subsequent frames of the animation, and finally normalises the result.

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 of oscillating vertices. The metric uses a different formula for straight and twisted vertices to correctly compute the ribbon area. However, this leads to a smaller contribution from vertices oscillating around a central position (an obvious temporal artefact) than from vertices that are constantly dislocated in time.
(2) Lack of rotation invariance. Because of 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 papers dealing with dynamic mesh compression [MSK*05, MSKW06, MSK*06], and it became part of the MPEG-4 standard in the form of Animation Framework eXtension Core Experiments Description [ISO11].

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. 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. Spatio-temporal edge difference (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 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 artefacts 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' artefacts are more notable 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 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

Although 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 artefacts. The authors identified three types of artefacts 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 artefacts in static and dynamic cases on the perception of the simplified mesh. The results of the study show that the silhouette artefacts are the most easily identified by human observers, although 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 artefacts, 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. Because 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, although 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, although 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:

$$P(A) = f(P_{\text{angular}}(A), P_{\text{momentum}}(A), P_{\text{spatio-temporal}}(A)), \qquad (1)$$

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 $P_{\text{spatio-temporal}}$ 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 in Sections 3. and 4. 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 MOSs and the objective scores computed by the metrics provides an excellent indicator of the performance of these metrics and a very good way to evaluate them quantitatively.

This section presents the protocols usually used in subjective tests (see 5.1.), the existing tests and the available MOS databases (corpus of models with MOSs; see 5.2.) and some evaluation and comparison results of the existing metrics on these databases (see 5.3.).

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) The subjective experiment is conducted where human observers give their opinion or some ratings about the perceived distortions of the database objects. A 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 MOS 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) Because 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. The reliability of the MOS may also be checked by computing the 95% confidence intervals or the intra-class correlation coefficient.
(4) The correlation can then be computed between the MOSs of the objects and their associated metric's values; usually two correlation coefficients are considered: the Spearman Rank Order Correlation and the Pearson Linear Correlation Coefficient. 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 optimises the matching between the values given by the objective metric and the subjective opinion scores provided by the subjects.

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 corpus should contain different kinds of models and different types of distortions and not focus on a specific scenario. The range of the visual impacts of the distortions have to be correctly balanced. It is also 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. 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. $\in [0, 100]$).
- 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 [CDGEB07]: 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.

5.2. Existing subjective databases

This subsection presents the subjective tests conducted by the scientific community related to 3D (3D + t) quality assessment and details more particularly some publicly available databases that provide the corpus of models and the MOSs.

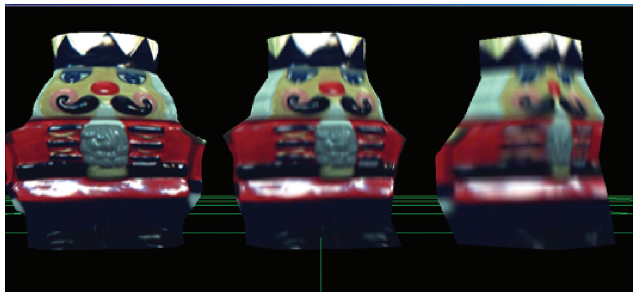

Figure 7: *Evaluation interface for the subjective test of Pan et al. [PCB05]. The observers were asked to compare the target stimulus (centre) with the two referential stimuli (left and right) and assign it one of the following ratings: very poor (1), poor (2), fair (3), good (4), very good (5). Reprinted from [PCB05].*

5.2.1. *3D static mesh*

Several authors have made subjective tests involving 3D static models [RRP00, WFM01, Rog01, PCB05, CS06, LDD*06, CDGEB07, SSF07, SSFM09, Lav09, VR12]. Their experiments had different purposes and used 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 [Rog01] focused on a mesh simplification scenario; their test databases were created by applying different simplification algorithms at different ratios on 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 [Rog01].

Rushmeier *et al.* [RRP00] and Pan *et al.* [PCB05] 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. Figure 7 illustrates the evaluation interface from [PCB05], the stimulus to rate is presented with two referential stimuli corresponding to the best and worst cases to help the observer to calibrate his rating; Pan *et al.* [PCB05] also provided a perceptual metric predicting this visual quality and evaluated it quantitatively by studying the correlation with subjective MOS from their experiment.

Corsini *et al.* [CDGEB07] 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 artefacts. The authors then used the MOSs 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 details the subjective experiments from [LDD*06], [Lav09], [SSF07] and [VR12], as 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 RockerArm). 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). Note that, 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. Note that, 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). Note that, 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/.

The UWB Compression Database [VR12] was created at the University of West Bohemia, Czech Republic. It contains 68 models created from 5 reference meshes. For each reference mesh, all the model versions share the connectivity of the original. The main purpose of the database is to evaluate different kinds of artefacts introduced by different compression algorithms. In contrast to previous experiments, instead of scoring, a binary forced choice paradigm has been adopted when collecting the user opinions. That means that for each of the 69 users in the test, triplets of meshes were presented, with one mesh being designated as original, and two randomly chosen distorted versions. The users were only asked to select the preferred version of the two distorted ones. The collected data is available both in terms of user choices and in scores computed from the choices in a manner described in [VR12]. The data can be obtained at http://compression.kiv.zcu.cz/.

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 data sets, i.e. for each of the five data sets 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 MOSs 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 3 summarises these correlation results, together with recent results presented in [VR12] and [WTM12]. 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], [LDD*06], [BHM09], [VR12]—or the same LODs—[CDGEB07].

As the table shows, the MSDM2 and FMPD metrics provide very good results on all the databases, although the DAME metric works well for databases of meshes with shared connectivity. 3DWPM metrics also have a correct behaviour. In contrast to that, 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 [CDGEB07].

Table 3: *Spearman and Pearson correlation (%) between mean opinion scores and values from the metrics for the four publicly available subjective databases. These data have been synthesised from [LC10], [Lav11][VR12] and [WTM12]. The correlations are computed for whole databases, with the exception of the compression database, where per-model averages were used, because the data acquiring procedure does not capture inter-model relations.*

	General purpose [LDD*06]		Masking [Lav09]		Simplification [SSF07]		Compression [VR12]	
	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson	Spearman	Pearson
$Hausdorff$	13.8	1.3	26.6	4.1	49.4	25.5	24.5	14.0
RMS	26.8	7.9	48.8	17.0	64.3	34.4	52.0	49.0
GL_1 [KG00]	33.1	12.6	42.0	15.7	N/A	N/A	66.9	70.6
GL_2 [SCOT03]	39.3	18.0	40.1	14.7	N/A	N/A	73.9	76.1
SF [BHM09]	15.7	0.5	38.6	2.4	N/A	N/A	57.4	34.8
$3DWPM_1$ [CDGEB07]	69.3	38.3	29.4	10.2	N/A	N/A	81.9	84.1
$3DWPM_2$ [CDGEB07]	49.0	24.6	37.4	18.2	N/A	N/A	80.9	82.3
$MSDM$ [LDD*06]	73.9	56.4	65.2	47.9	N/A	N/A	83.1	91.5
$MSDM2$ [Lav11]	80.4	66.2	89.6	76.2	86.7	79.6	78.0	89.3
$FMPD$ [WTM12]	81.9	83.5	80.2	80.8	87.2	89.3	81.8	88.8
$DAME$ [VR12]	76.6	75.2	68.1	58.6	N/A	N/A	85.6	93.5

Unfortunately, image-based metrics have not been quantitatively tested on these public databases, whereas several authors [WFM01] and [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 of static meshes, i.e. using the Spearman and Pearson coefficient. 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 4. 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

In this section, we focus on the applications of various perceptual and non-perceptual distortion metrics for static and dynamic meshes to performance improvement and evaluation of mesh watermarking, evaluation and optimisation of lossy mesh compression, and also mesh simplification.

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 benefited 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 HVS (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. Secondly, 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 noted 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.*

Table 4: *Spearman and Pearson correlation (%) between Mean Opinion Scores and values from the metrics for dynamic meshes. These data have been taken from [VS11]. Since psychometric fitting has not been used in this work, negative correlations may appear, indicating an inverse correlation on the dataset used.*

	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

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 because of different frequency sensitivities of the HVS.*

[WLBD09] develop watermarking methods based on a new spectral mesh 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 [DF8] 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). Subsequently, we can relate the modification of a mesh spectral component to the HVS sensitivity as given by the CSF [MS74] (*cf.* Section 2.1.), which is a function of the observed frequency. In Figures 8(b) and (c) we show two deformed Venus models after respectively low- and high-frequency deformation. The observed frequencies of the two deformations are respectively around 0.1 and 4.0 cpd. According to the CSF, 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 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 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], although 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.* [CDGEB07] to adaptively determine the local embedding strength, they can achieve a better trade-off between imperceptibility and robustness for two previously proposed mesh watermarking schemes.

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 to find an adequate roughness definition, we may need to conduct psycho-visual 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 3, simple geometric measures do not correlate well with a subjective assessment. Secondly, the quality of the images illustrating the original and watermarked meshes highly depends on the viewpoint selection and the rendering technique used. Therefore, 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 Drelie Gelasca *et al.* [DGECB05], Corsini *et al.* [CDGEB07] 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 benchmark for image watermarking [PAK98], 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].

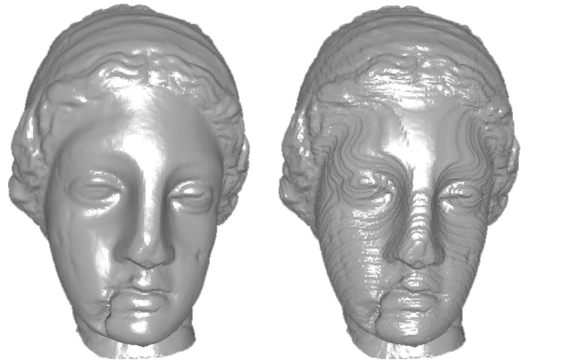

Figure 9: *On the left is the Venus watermarked using the method of Wang et al. [WLDB11] (MRMS=* 1.69×10^{-3}, MSDM= 0.13*), and on the right is the Venus watermarked using the method of Cho et al. [CPJ07] (MRMS=* 0.936×10^{-3}, MSDM= 0.54*). MSDM yields correct results that are consistent with a subjective 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 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 LOD 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 HVS (frequency sensitivity and visual masking) have been integrated in the design of compression algorithms. Secondly, 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 use 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 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 as fine details are preserved (see Figure 10).

A side property of this frequency sensitivity is that the low-resolution LODs, in a progressive compression algorithm, do not need a high geometric precision; hence, they can be quantised

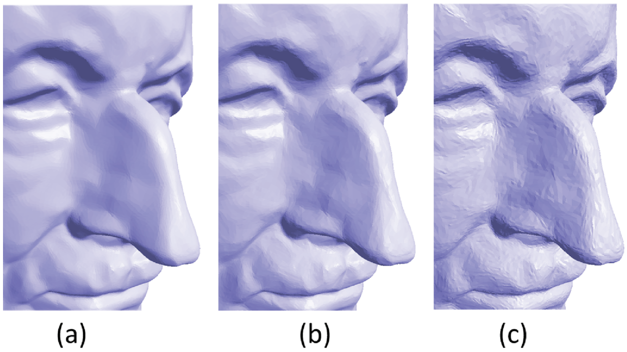

Figure 10: *(a) Original Max Planck, (b) results after high pass spectral quantisation, (c) result after uniform Cartesian quantisation. Models from (b) and (c) are associated with the same rate. The distortion from (c) (which is high-frequency) is clearly more visible than the distortion from (b) (which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

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 artefacts: 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 process 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.* [PCB05] 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 sub-sampling. 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 data sets, 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 became 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 [VP11, PV11], 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. Because the implementations of many of the methods are not publicly available, and some of them are difficult to obtain because of licensing restrictions, most researchers tend to take a shortcut 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.

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, because 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

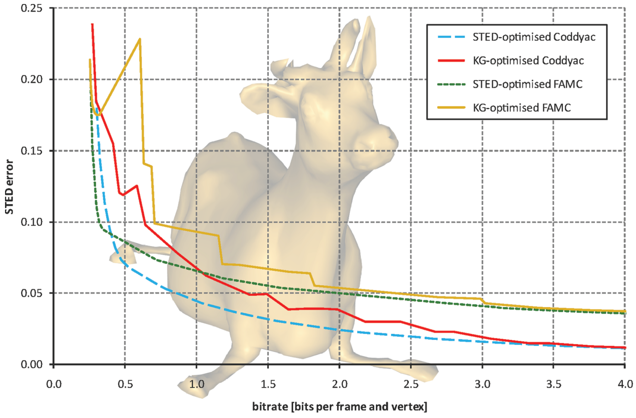

Figure 11: *Rate-distortion curves of the Coddyac [VS07] and FAMC [MSK*08] algorithms in the STED metric compressing cow animation. Solid lines show the performance when optimising for the KG error metric, dashed lines denote optimisation for the STED metric.*

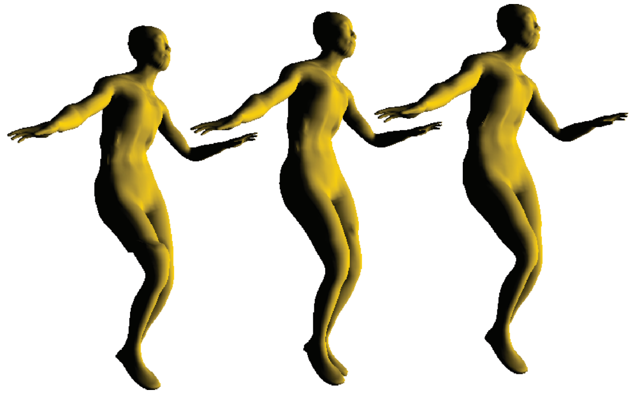

Figure 12: *A frame from the "dance" animation compressed by the Clustered Coddyac algorithm[RV10] (left) and then deblocked by the mesh deblocking method[RV12] (middle). STED value improved from* 0.055 to 0.047. *The original is shown on the right.*

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 artefacts caused by quantisation of Laplacian coordinates is less perceptible than other kinds of artefacts.

Rus and Váša [RV12] proposed a method for smoothing discontinuities between clusters ('mesh deblocking') in meshes processed by clustering-based compression algorithms, such as frame-based animated mesh compression (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 causes artefacts at cluster borders (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 de-blocking method improves STED metric performance of existing algorithms that have been tuned to minimise a non-perception-correlated distortion metric, especially for low bit rates.

6.4. Application to mesh simplification

Mesh simplification is a process that reduces the number of mesh vertices to a desired degree using edge collapses, or vertex removal and hole-filling. Such technique may be required, e.g for fast display, or preview, of very detailed meshes, for the creation of multiple levels of detail of a mesh, or in progressive mesh storage and transmission. In most cases, the simplified mesh is required to retain as much of the features or details of the original as possible. This can be achieved by using a distortion metric to drive the simplification. Such approaches use the metric to determine a *simplification cost*. To each potential simplification step (i.e. each vertex to be removed or each edge to be collapsed), a cost is assigned based on the difference between the original mesh and the simplified version including this step. The simplification step with the lowest cost is then performed and the affected costs are updated. This greedy approach attempts to achieve the smallest difference between the result and the original. However, it has not been proven that such method will ensure an optimal result. The metric used for calculating the costs has to be connectivity independent, although it can assume a close proximity of the compared surfaces.

Several metrics have been used for evaluating the difference in the simplification process. Klein *et al.* [KLS96] and Borouchaki and Frey [BF05] use the Hausdorff distance to assign costs to the simplification steps. Garland and Heckbert [GH97] proposed employing a quadric error metric based on the distance from vertex tangent planes. Lindstrom and Turk [LT98] use a similar metric with the addition of volume and boundary preservation. Garland and Zhou [GZ05] extended the quadric approach to spaces of arbitrary dimensions. Cohen *et al.* [CVM*96] proposed creating an inner and an outer envelope by offsetting the mesh surface and then finding triangles between the original mesh vertices that cover as much original vertices as possible while not intersecting the two envelopes.

Another group of algorithms use a render-based approach implementing error metrics for images. Lindstrom and Turk [LT00] utilise a simple mean squared error to compare the renders, whereas Luebke and Hallen, in their view-dependent approach [LH01], rather employ a CSF-based metric. Lindstrom performed a thorough comparison of geometry- and render-based simplification methods in his PhD thesis [Lin00], including the use of perceptual image metrics for the render-based approach.

An important aspect of mesh simplification is the evaluation of the final simplified mesh. Although it may seem best to use the same metric in the process as for the evaluation, authors often use simpler and computationally less expensive means for driving the simplification process than for the final comparison, to decrease computation times, although still attempting to optimise their simplification for a more complex metric. Most simplification papers rely the Hausdorff distance or its variant to compare the results [KLS96, GH97, BF05, GZ05]. In [LT98] and [LT00], the Metro tool was used for results evaluation. In [LT00], the average root mean squared error between renders (non-perceptual image metric) of the simplified and the original mesh was also measured. Luebke and Hallen [LH01] performed a subjective experiment with 4 users to compare their method to others. Although a few authors use a perceptual measure in the simplification process, to our knowledge, no simplification algorithm was specifically designed to minimise an object-space perceptual metric.

7. Conclusions

Evaluating the perceptibility of distortion caused by mesh processing is currently still an open problem with many unresolved issues. 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, and
- 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, FMPD, DAME and STED, have made a significant first steps 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.

Acknowledgements

This work has been partly supported by the Ministry of Education, Youth and Sports of the Czech Republic under the research programme LC-06008 (Center for Computer Graphics). It has also been partly supported by the French National Research Agency (ANR) through the COSINUS programme (project COLLAVIZ ANR-08-COSI-003) and the VERSO programme (project CAIMAN ANR-08-TO-002). A partial support has been received from the project UE-FEDER NAVII.

References

[ACMS10] AYDIN T. O., CADIK M., MYSZKOWSKI K., SEIDEL H.-P.: Video quality assessment for computer graphics applications. *ACM Transactions on Graphics (Proceedings of SIGGRAPH Asia) 29*, 6 (Dec. 2010), Article No. 161:1–161:12.

[ACSD*03] ALLIEZ P., COHEN-STEINER D., DEVILLERS O., LÉVY B., DESBRUN M.: Anisotropic polygonal remeshing. *ACM Transactions on Graphics (Proceedings of SIGGRAPH) 22*, 3 (July 2003), 485–493.

[Bar99] BARTEN P. G. J. : *Contrast Sensitivity of the Human Eye and its Effects on Image Quality*. SPIE Press, Bellingham, WA, USA, 1999.

[BCLP11] BULBUL A., CAPIN T., LAVOUÉ G., PREDA M.: Assessing visual quality of 3D polygonal models. *IEEE Signal Processing Magazine 28*, 6 (November 2011), 80–90.

[BF05] BOROUCHAKI H., FREY P. J.: Simplification of surface mesh using Hausdorff envelope. *Computer Methods in Applied Mechanics and Engineering 194*, 48–49 (Nov. 2005), 4864–4884.

[BHM08] BIAN Z., HU S.-M., MARTIN R. R.: Comparing small visual differences between conforming meshes. In *Proceedings of Geometric Modeling and Processing Conference* (Tokyo, Japan, 2008), vol. 4975, *Lecture Notes in Computer Science*, Springer, pp. 62–78.

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

[BK80] BURBECK C. A., KELLY D. H.: Spatiotemporal characteristics of visual mechanisms: excitatory-inhibitory model. *Journal of Optical Society of America 70*, 9 (Sep. 1980), 1121–1126.

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

[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* (2012), (in press).

[BM98] BOLIN M. R., MEYER G. W.: A perceptually based adaptive sampling algorithm. In *Proceedings of SIGGRAPH* (Orlando, FL, USA, 1998), ACM, pp. 299–309.

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

[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. *IEEE Journal of Selected Topics in Signal Processing 5*, 7 (Nov. 2011), 1332–1343.

[BS06] Barland R., Saadane A.: Blind quality metric using a perceptual map for JPEG2000 compressed images. In *Proceedings of International Conference on Image Processing* (Atlanta, GA, USA, 2006), IEEE, pp. 2941–2944.

[BT06] Bruce N., Tsotsos J.: Saliency based on information maximization. In *Proceedings of Neural Information Processing Systems Conference* (Vancouver, Canada, 2006), MIT Press, pp. 155–162.

[CB07] Cheng I., Basu A.: Perceptually optimized 3-D transmission over wireless networks. *IEEE Transactions on Multimedia 9*, 2 (Feb. 2007), 386–396.

[CDGE05] Corsini M., Drelie Gelasca E., Ebrahimi T.: A multi-scale roughness metric for 3D watermarking quality assessment. In *Proceedings of Workshop on Image Analysis for Multimedia Interactive Services* (Montreux, Switzerland, 2005), SPIE, pp. 1–4.

[CDGEB07] Corsini M., Drelie Gelasca E., Ebrahimi T., Barni M.: Watermarked 3-D mesh quality assessment. *IEEE Transactions on Multimedia 9*, 2 (Feb. 2007), 247–256.

[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 (Nov. 2009), 1–15.

[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 (Sep. 2007), 2284–2298.

[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 (Jan. 2007), 142–155.

[CR68] Campbell F. W., Robson J. G.: Application of Fourier analysis to the visibility of gratings. *Journal of Physiology 197*, 3 (Aug. 1968), 551–566.

[CS06] Cleju I., Saupe D.: Evaluation of supra-threshold perceptual metrics for 3D models. In *Proceedings of Symposium on Applied Perception in Graphics and Visualization* (Boston, MA, USA, 2006), ACM, pp. 41–44.

[CVM*96] Cohen J., Varshney A., Manocha D., Turk G., Weber H., Agarwal P., Brooks F., Wright W.: Simplification envelopes. In *Proceedings of SIGGRAPH* (New Orleans, LA, USA, 1996), ACM, pp. 119–128.

[CW07] Cheng Y. M., Wang C. M.: An adaptive steganographic algorithm for 3D polygonal meshes. *The Visual Computer 23*, 9–11 (Sep. 2007), 721–732.

[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* (Jan.2008), 1–14.

[Dal93] Daly S.: The visible differences predictor: An algorithm for the assessment of image fidelity. In *Digital Images and Human Vision*. A. b. Watson (Ed.). MIT Press, Cambridge, MA, USA, (1993), pp. 197–206.

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

[Dal98] Daly S.: Engineering observations from spatiovelocity and spatiotemporal visual models. In *Proceedings of Conference on Human Vision and Electronic Imaging* (San Jose, CA, USA, 1998), vol. 3299, SPIE, pp. 180–191.

[DF88] Donnelly H., Fefferman C.: Nodal sets of eigenfunctions on Riemannian manifolds. *Inventiones Mathematicae 93*, 1 (Feb. 1988), 161–183.

[DGECB05] Drelie Gelasca E., Ebrahimi T., Corsini M., Barni M.: Objective evaluation of the perceptual quality of 3D watermarking. In *Proceedings of International Conference on Image Processing* (Genoa, Italy, 2005), IEEE, pp. 241–244.

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

[Ebr09] Ebrahimi T.: Quality of multimedia experience: Past, present and future. In *Proceedings of ACM Multimedia* (Beijing, China, 2009), ACM, pp. 3–4.

[Fai05] Fairchild M. D.: *Color Appearance Models* (2nd edition). Wiley IS&T, Chichester, UK, 2005.

[FRBW08] Ferwerda J. A., Ramanarayanan G., Bala K., Walter B. J.: Visual equivalence: An object-based approach to image quality. In *Proceedings of Color Imaging Conference* (Portland, OR, USA, 2008), IS&T, pp. 347–354.

[FSPG97] Ferwerda J. A., Shirley P., Pattanaik S. N., Greenberg d. P.: A model of visual masking for computer graphics. In *Proceedings of SIGGRAPH* (Los Angeles, CA, USA, 1997), ACM, pp. 143–152.

[GH97] Garland M., Heckbert P. S.: Surface simplification using quadric error metrics. In *Proceedings of SIGGRAPH* (Los Angeles, CA, USA, 1997), ACM, pp. 209–216.

[GMSK09] Guthe M., Müller G., Schneider M., Klein R.: BTF-CIELab: A perceptual difference measure for quality assessment and compression of BTFs. *Computer Graphics Forum 28*, 1 (Mar. 2009), 101–113.

[GZ05] GARLAND M., ZHOU Y.: Quadric-based simplification in any dimension. *ACM Transactions on Graphics 24*, 2 (Apr. 2005), 209–239.

[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 (Proceedings of Eurographics) 31*, 2 (May 2012), 545–554.

[HHO04] HOWLETT S., HAMILL J., O'SULLIVAN C.: An experimental approach to predicting saliency for simplified polygonal models. In *Proceedings of Symposium on Applied Perception in Graphics and Visualization* (Los Angeles, CA, USA, 2004), ACM, pp. 57–64.

[HKL09] HEU J., KIM C.-S., LEE S.-U.: SNR and temporal scalable coding of 3-D mesh sequences using singular value decomposition. *Journal of Visual Communication and Image Representation 20*, 7 (Oct. 2009), 439–449.

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

[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*, 11 (Nov. 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 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (San Diego, CA, USA, 2003), Eurographics Association, pp. 126–135.

[ISO11] ISO/IEC JTC1/SC29/WG11: Animation Framework eXtension (AFX). 12419. Core Experiments Description of AFX. Doc. N6988, Hong Kong, China, 2011.

[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 Transactions on Circuits and Systems for Video Technology 14*, 7 (July 2004), 989–1008.

[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 (Dec. 2010), 721–733.

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

[Kel79] KELLY D. H.: Motion and vision. II. Stabilized spatiotemporal threshold surface. *Journal of Optical Society of America 69*, 10 (Oct. 1979), 1340–1349.

[KFB10] KŘIVÁNEK J., FERWERDA J. A., BALA K.: Effects of global illumination approximations on material appearance. *ACM Transactions on Graphics 29*, 4 (July 2010), 112:1–112:10.

[KG00] KARNI Z., GOTSMAN C.: Spectral compression of mesh geometry. In *Proceedings of SIGGRAPH* (New Orleans, LA, USA, 2000), ACM, pp. 279–286.

[KG04] KARNI Z., GOTSMAN C.: Compression of soft-body animation sequences. *Computers & Graphics 28*, 1 (Feb. 2004), 25–34.

[KK93] KARUNASEKERA S. A., KINGSBURY N. G.: A distortion measure for blocking artifacts in images based on human visual sensitivity. In *Proceedings of Conference on Visual Communications and Image Processing* (Cambridge, MA, USA, 1993), vol. 2094, SPIE, pp. 474–486.

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

[KLS96] KLEIN R., LIEBICH G., STRASSER W.: Mesh reduction with error control. In *Proceedings of IEEE/ACM Visualization Conference* (San Francisco, CA, USA, 1996), IEEE, pp. 311–318.

[KPA09] KAMMOUN A., PAYAN F., ANTONINI M.: Bit allocation for spatio-temporal wavelet coding of animated semi-regular meshes. In *Proceedings of International Multimedia Modeling Conference* (Sophia Antipolis, France, 2009), vol. 5371, *Lecture Notes in Computer Science*, Springer, pp. 128–139.

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

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

[Lav09] LAVOUÉ G.: A local roughness measure for 3D meshes and its application to visual masking. *ACM Trans. on Applied Perception 5*, 4 (Jan. 2009), Article No.21:1–21:23.

[Lav11] LAVOUÉ G.: A multiscale metric for 3D mesh visual quality assessment. *Computer Graphics Forum (Proceedings of Symposium on Geometry Processing) 30*, 5 (Aug. 2011), 1427–1437.

[LBF06] LARABI M.-C., BRODBECK V., FERNANDEZ C.: A novel approach for constructing an achromatic contrast sensitivity function by matching. In *Proceedings of International Conference on Image Processing* (Atlanta, GA, USA, 2006), IEEE, pp. 441–441.

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

[LDD*06] LAVOUÉ G., DRELIE GELASCA E., DUPONT F., BASKURT A., EBRAHIMI T.: Perceptually driven 3D distance metrics with application to watermarking. In *Proceedings of Conference on Applications of Digital Image Processing* (San Jose, CA, USA, 2006), vol. 6312, SPIE, p. 63120L.

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

[Len99] LENGYEL J. E.: Compression of time-dependent geometry. In *Proceedings of Symposium on Interactive 3D Graphics* (Atlanta, GA, USA, 1999), ACM, pp. 89–95.

[LH01] LUEBKE D. P., HALLEN B.: Perceptually-driven simplification for interactive rendering. In *Proceedings of Eurographics Workshop on Rendering Techniques* (London, UK, 2001), Springer-Verlag, pp. 223–234.

[Lin00] LINDSTROM P.: *Model Simplification Using Image and Geometry-Based Metrics*. PhD thesis, Georgia Institute of Technology, USA, Nov. 2000.

[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 Transactions on Information and Systems E90-D*, 7 (Jul. 2007), 1073–1084.

[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.

[LO11] LARKIN M., O'SULLIVAN C.: Perception of simplification artifacts for animated characters. In *Proceedings of Symposium on Applied Perception in Graphics and Visualization* (Toulouse, France, 2011), ACM, pp. 93–100.

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

[LT98] LINDSTROM P., TURK G.: Fast and memory efficient polygonal simplification. In *Proceedings of IEEE Visualization Conference* (Research Triangle Park, NC, USA, 1998), IEEE, pp. 279–286.

[LT00] LINDSTROM P., TURK G.: Image-driven simplification. *ACM Transactions on Graphics 19*, 3 (Jul. 2000), 204–241.

[Lub93] LUBIN J.: The use of psychophysical data and models in the analysis of display system performance. In *Digital Images and Human Vision*, A. B. Watson (Ed.). MIT Press, Cambridge, MA, USA, (1993), pp. 163–178.

[Lub95] LUBIN J.: A visual discrimination model for imaging system design and evaluation. In *Vision Models for Target Detection and Recognition*, E. Peli (Ed.). World Scientific Publishing Company, Singapore, (1995), pp. 245–283.

[LVJ05] LEE C. H., VARSHNEY A., JACOBS D. W.: Mesh saliency. *ACM Transactions on Graphics (Proceedings of SIGGRAPH) 24*, 3 (Jul. 2005), 659–666.

[LZ10] LÉVY B., ZHANG H.: Spectral mesh processing. In *Proceedings of SIGGRAPH Asia Course Notes* (Seoul, Korea, 2010), ACM, pp. 1–47.

[MG10] MENZEL N., GUTHE M.: Towards perceptual simplification of models with arbitrary materials. *Computer Graphics Forum 29*, 7 (Sep. 2010), 2261–2270.

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

[MK05] MUIJS R., KIRENKO I.: A no-reference blocking artifact measure for adaptive video processing. In *Proceedings of European Signal Processing Conference* (Antalya, Turkey, 2005), pp. 1–4.

[MMS04] MANTIUK R., MYSZKOWSKI K., SEIDEL H.-P.: Visible difference predicator for high dynamic range images. In *Proceedings of International Conference on Systems, Man and Cybernetics* (The Hague, The Netherlands, 2004), IEEE, pp. 2763–2769.

[MNO07] MCDONNELL R., NEWELL F., O'SULLIVAN C.: Smooth movers: Perceptually guided human motion simulation. In *Proceedings of ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (San Diego, CA, USA, 2007), Eurographics Association, pp. 259–269.

[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 (Apr. 2009), 231–243.

[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 (Jul. 1974), 525–536.

[MSK*05] MÜLLER K., SMOLIC A., KAUTZNER M., EISERT P., WIEGAND T.: Predictive compression of dynamic 3D meshes. In *Proceedings of International Conference on Image Processing* (Genoa, Italy, 2005), vol. 1, IEEE, pp. 621–624.

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

[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 *Proceedings of 3DTV-CON: The True Vision - Capture, Transmission and Display of 3D Video* (Istanbul, Turkey, 2008), IEEE, pp. 97–100.

[MSKW06] MÜLLER K., SMOLIC A., KAUTZNER M., WIEGAND T.: Rate-distortion optimization in dynamic mesh compression. In *Proceedings of International Conference on Image Processing* (Atlanta, GA, USA, 2006), IEEE, pp. 533–536.

[Mys98] MYSZKOWSKI K.: The visible differences predictor: Applications to global illumination problems. In *Eurographics Workshop on Rendering* (1998), vol. 98, pp. 223–236.

[Mys02] MYSZKOWSKI K.: Perception-based global illumination, rendering, and animation techniques. In *Proceedings of Spring Conference on Computer Graphics* (Budmerice, Slovakia, 2002), ACM, pp. 13–24.

[MZP06] MAMOU K., ZAHARIA T., PRÊTEUX F.: A skinning approach for dynamic 3D mesh compression. *Computer Animation and Virtual Worlds 17*, 3–4 (Jul. 2006), 337–346.

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

[NLF11] NAUGE M., LARABI M.-C., FERNANDEZ C.: ImQual: A web-service dedicated to image quality evaluation and metrics benchmark. In *Proceedings of Conference on Image Quality and System Performance* (Burlingame, CA, USA, 2011), vol. 7867, SPIE, p. 786714.

[ODGK03] O'SULLIVAN C., DINGLIANA J., GIANG T., KAISER M. K.: Evaluating the visual fidelity of physically based animations. *ACM Transactions on Graphics 22*, 3 (Jul. 2003), 527–536.

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

[PA06] PAYAN F., ANTONINI M.: Mean square error approximation for wavelet-based semiregular mesh compression. *IEEE Transactions on Visualization and Computer Graphics 12*, 4 (Jul.–Aug. 2006), 649–657.

[PAK98] PETITCOLAS F. A. P., ANDERSON R. J., KUHN M. G.: Attacks on copyright marking systems. In *Proceedings of International Workshop on Information Hiding* (Portland, OR, USA, 1998), vol. 1525, *Lecture Notes in Computer Science*, Springer, Heidelberg, Germany, pp. 218–238.

[PCB05] PAN Y., CHENG I., BASU A.: Quality metric for approximating subjective evaluation of 3-D objects. *IEEE Transactions on Multimedia 7*, 2 (Apr. 2005), 269–279.

[PKPH12] PERAZZI F., KRÄHENBÜHL P., PRITCH Y., HORNUNG A.: Saliency filters: Contrast based filtering for salient region detection. In *Proceedings of International Conference on Computer Vision and Pattern Recognition* (Providence, RI, USA, 2012), IEEE, pp. 733–740.

[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*, 6 (Jul. 2004), 499–506.

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

[PV10] PETŘÍK O., VÁŠA L.: Finding optimal parameter configuration for a dynamic triangle mesh compressor. In *Proceedings of International Conference on Articulated Motion and Deformable Objects* (Port d'Andratx, Spain, 2010), vol. 6169, *Lecture Notes in Computer Science*, Springer, pp. 31–42.

[PV11] PETŘÍK O., VÁŠA L.: Improvements of MPEG-4 standard FAMC for efficient 3D animation compression. In *Proceedings of 3DTV-CON: The True Vision - Capture, Transmission and Display of 3D Video* (Antalya, Turkey, 2011), IEEE, pp. 1–4.

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

[Rec99] Recommendation ITU-T P.910: Subjective video quality assessment methods for multimedia applications, 1999.

[Rec02] Recommendation ITU-R BT.500-11: Methodology for subjective assessment of the quality of television pictures, 2002.

[Rec07] Recommendation ITU-R BT.1788: Methodology for the subjective assessment of video quality in multimedia applications, 2007.

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

[RFWB07] RAMANARAYANAN G., FERWERDA J. A., WALTER B., BALA K.: Visual equivalence: Towards a new standard for image fidelity. *ACM Transactions on Graphics (Proceedings of SIGGRAPH) 26*, 3 (Jul. 2007), Article No. 76:1–76:11.

[Rog01] ROGOWITZ B. E.: Are image quality metrics adequate to evaluate the quality of geometric objects? In *Proceedings of Conference on Human Vision and Electronic Imaging* (San Jose, CA, USA, 2001), vol. 4299, SPIE, pp. 340–348.

[Rou10] ROUDET C.: A region-based progressive coding of semi-regular 3-D meshes for view-dependent transmission. In *Proceedings of International Conference on Signal Image Technology and Internet-Based Systems* (Kuala Lump, Malaysia, 2010), IEEE, pp. 51–59.

[RPG99] RAMASUBRAMANIAN M., PATTANAIK S. N., GREENBERG D. P.: A perceptually based physical error metric for realistic image synthesis. In *Proceedings of SIGGRAPH* (Los Angeles, CA, USA, 1999), ACM, pp. 73–82.

[RRP00] RUSHMEIER H. E., ROGOWITZ B. E., PIATKO C.: Perceptual issues in substituting texture for geometry. In *Proceedings of Conference on Human Vision and Electronic Imaging* (San Jose, CA, USA, 2000), vol. 3959, SPIE, pp. 372–383.

[RV10] RUS J., VÁŠA L.: Analysing the influence of vertex clustering on PCA-based dynamic mesh compression. In *Proceedings of International Conference on Articulated Motion and Deformable Objects* (Port d'Andratx, Spain, 2010), vol. 6169, *Lecture Notes in Computer Science*, Springer, pp. 55–66.

[RV12] RUS J., VÁŠA L.: Deblocking for dynamic triangle meshes. In *Proceedings of International Conference on Computer Graphics Theory and Applications* (Rome, Italy, 2012), SciTePress, pp. 48–57.

[RvdLBC08] Rajashekar U., van der Linde I., Bovik A. C., Cormack l. K.: GAFFE: A gaze-attentive fixation finding engine. *IEEE Transactions on Image Processing 17*, 4 (Apr. 2008), 564–573.

[SB06] Sheikh H. R., Bovik A. C.: Image information and visual quality. *IEEE Transactions on Image Processing 15*, 2 (Feb. 2006), 430–444.

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

[SCOT03] Sorkine O., Cohen-or D., Toldeo S.: High-pass quantization for mesh encoding. In *Proceedings of Eurographics Symposium on Geometry Processing* (Aachen, Germany, 2003), Eurographics Association, pp. 42–51.

[SPC04] Safranek R. J., Pappas T. N., Chen J.: Perceptual criteria for image quality evaluation. In *Handbook of Image and Video Processing*, A. C. Bovik (Ed.). Academic Press, Burlington, MA, USA, (2004), pp. 669–684.

[SSF07] Silva S., Sousa Santos B., Ferreira C.: Comparison of methods for the simplification of mesh models using quality indices and an observer study. In *Proceedings of Conference on Human Vision and Electronic Imaging* (San Jose, CA, USA, 2007), vol. 6492, SPIE, p. 64921L.

[SSFM09] Silva S., Sousa Santos B., Ferreira C., Madeira J.: A perceptual data repository for polygonal meshes. In *Proceedings of International Conference in Visualisation* (Barcelona, Spain, 2009), IEEE, pp. 207–212.

[TA04] Tian D., Alregib G.: FQM: A fast quality measure for efficient transmission of textured 3D models. In *Proceedings of ACM Multimedia* (New York, NY, USA, 2004), ACM, pp. 684–691.

[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 (Jan. 2008), 23–35.

[Tau00] Taubin G.: Geometric signal processing on polygonal meshes. In *Proceedings of Eurographics State-of-the-Art Reports* (Interlaken, Switzerland, 2000), Eurographics Association, pp. 1–11.

[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*, 4 (Oct. 2006), 766–786.

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

[TWC12] Torkhani F., Wang K., Chassery J.-M.: A curvature tensor distance for mesh visual quality assessment. In *Proceedings of International Conference on Computer Vision and Graphics* (Warsaw, Poland, 2012), vol. 7594, *Lecture Notes in Computer Science*, Springer, pp. 253–263.

[VCP09] Valette S., Chaine R., Prost R.: Progressive lossless mesh compression via incremental parametric refinement. *Computer Graphics Forum (Proceedings of Symposium on Geometry Processing) 28*, 5 (Jul. 2009), 1301–1310.

[VL08] Vallet B., Lévy B.: Spectral geometry processing with manifold harmonics. *Computer Graphics Forum (Proceedings of Eurographics) 27*, 2 (Apr. 2008), 251–260.

[VP11] Váša L., Petřík O.: Optimising perceived distortion in lossy encoding of dynamic meshes. *Computer Graphics Forum (Proceedings of Symposium on Geometry Processing) 30*, 5 (Aug. 2011), 1439–1449.

[VR12] Váša L., Rus J.: Dihedral Angle Mesh Error: A fast perception correlated distortion measure for fixed connectivity triangle meshes. *Computer Graphics Forum (Proceedings of Symposium on Geometry Processing) 31*, 5 (Aug. 2012), 1715–1724.

[VS06] Váša L., Skala V.: A spatio-temporal metric for dynamic mesh comparison. In *Proceedings of International Conference on Articulated Motion and Deformable Objects* (Port d'Andratx, Spain, 2006), vol. 4069, *Lecture Notes in Computer Science*, Springer, pp. 29–37.

[VS07] Váša L., Skala V.: CoDDyAC: Connectivity driven dynamic mesh compression. In *Proceedings of 3DTV-CON: The True Vision - Capture, Transmission and Display of 3D Video* (Kos, Greece, 2007), IEEE, pp. 1–4.

[VS11] Váša L., Skala V.: A perception correlated comparison method for dynamic meshes. *IEEE Transactions on Visualization and Computer Graphics 17*, 2 (Feb. 2011), 220–230.

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

[Wat87] Watson A. B.: The cortex transform: Rapid computation of simulated neural images. *Computer Vision, Graphics and Image Processing 39*, 3 (Sep. 1987), 311–327.

[WBSS04] Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E.: Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing 13*, 4 (Apr. 2004), 1–14.

[WFM01] Watson A. B., Friedman A., Mcgaffey A.: Measuring and predicting visual fidelity. In *Proceedings of SIGGRAPH* (Los Angeles, CA, USA, 2001), ACM, pp. 213–220.

[WHST01] Wu J.-H., Hu S.-M., Sun J.-G., Tai C.-L.: An effective feature-preserving mesh simplification scheme based on face constriction. In *Proceedings of Pacific Graphics* (Tokyo, Japan, 2001), IEEE, pp. 12–21.

[Win00] Winkler S.: *Vision Models and Quality Metrics for Image Processing Applications*. PhD thesis, Ecole Polytechnique Fédérale de Lausanne, Switzerland, Dec. 2000.

[WLBD09] Wang K., Luo M., Bors A. G., Denis F.: Blind and robust mesh watermarking using manifold harmonics. In

Proceedings of International Conference on Image Processing (Cairo, Egypt, 2009), IEEE, pp. 3657–3660.

[WLC*03] Williams N., Luebke D., Cohen j. D., Kelley M., Schubert B.: Perceptually guided simplification of lit, textured meshes. In *Proceedings of Symposium on Interactive 3D Graphics* (Monterey, CA, USA, 2003), ACM, pp. 113–121.

[WLD*10] Wang K., Lavoué G., Denis F., Baskurt A., He X.: A benchmark for 3D mesh watermarking. In *Proceedings of International Conference on Shape Modeling and Applications* (Aix-en-Provence, France, 2010), IEEE, pp. 231–235.

[WLDB08a] Wang K., Lavoué G., Denis F., Baskurt A.: A comprehensive survey on three-dimensional mesh watermarking. *IEEE Transactions on Multimedia 10*, 8 (Dec. 2008), 1513–1527.

[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 (Dec. 2008), 620–634.

[WLDB11] Wang K., Lavoué G., Denis F., Baskurt A.: Robust and blind mesh watermarking based on volume moments. *Computers & Graphics 35*, 1 (Feb. 2011), 1–19.

[WS05] Wang Z., Simoncelli E. P.: Reduce-reference image quality assessment using a wavelet-domain natural image statistic model. In *Proceedings of Conference on Human Vision and Electronic Imaging* (San Jose, CA, USA, 2005), vol. 5666, SPIE, pp. 17–20.

[WSB03] Wang Z., Sheikh H. R., Bovik A. C.: Objective video quality assessment. In *Handbook of Video Databases: Design and Applications*. B. Furht and O. Marqure (Eds.). CRC Press, Boca Raton, FL, USA, (2003), pp. 1041–1078.

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

[Yar67] Yarbus A. L.: *Eye Movements and Vision*. Plenum Press, New York, NY, USA, 1967.

[YIK03] Yu Z., Ip h. h. S., Kwok l. F.: A robust watermarking scheme for 3D triangular mesh models. *Pattern Recognition 36*, 11 (Nov. 2003), 2603–2614.

[YN04] Yee y. H., Newman A.: A perceptual metric for production testing. In *Proceedings of SIGGRAPH Sketches* (Los Angeles, CA, USA, 2004), ACM, p. 121.

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

[YPSZ01] Yin K., Pan Z., Shi J., Zhang D.: Robust mesh watermarking based on multiresolution processing. *Computers & Graphics 25*, 3 (Jun. 2001), 409–420.

[ZDL02] Zeng W., Daly S., Lei S.: An overview of the visual optimization tools in JPEG 2000. *Signal Processing: Image Communication 17*, 1 (Jan. 2002), 85–104.

[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.