THEORETICAL DISTORTION ESTIMATION IN L-INFINITE WAVELET-BASED CODING OF SEMI-REGULAR MESHES

Ruxandra Florea[1,2], Leon Denis[1,2], Jan Lievens[1,2], Peter Schelkens[1,2], Adrian Munteanu[1,2]

[1] Vrije Universiteit Brussel, Department of Electronics and Informatics (ETRO), Pleinlaan 2, 1050 Brussels, Belgium
[2] Interdisciplinary Institute for Broadband Technology (IBBT), Gaston Crommenlaan 8 b102, 9050 Ghent, Belgium

ABSTRACT

Subdivision-based wavelet coding techniques yield state-of-the-art performance in scalable compression of semi-regular meshes. However, all these codecs make use of the L-2 distortion metric, which gives only a good approximation of the global error produced by lossy coding of the wavelet coefficients. The L-infinite metric has been proven to be a suitable metric for applications where controlling the local, maximum error on each vertex is of critical importance. In this context, an upper bound formulation for the L-infinite distortion for a wavelet-based coding scheme operating on semi-regular meshes is derived. In addition, we propose a rate-distortion optimization algorithm that minimizes the rate for any target L-infinite distortion. It is shown that our L-infinite coding system outperforms the state-of-the-art and that an L-2 driven coding approach for semi-regular meshes loses ground to its L-infinite driven version when the goal is to have a tight control on the local reconstruction error.

Index Terms— L-infinite coding, semi-regular mesh compression, subdivision-based wavelets.

1. INTRODUCTION

Polygonal mesh representations have evolved from the relatively low resolution ones common in the early years of the last decade to virtual objects consisting of millions of vertices that better fit the new age of high-definition graphics and fast microprocessors. The list of domains that make use of mesh representations is quite extensive. One can start with the popular gaming industry and continue with medical imaging, coding of topographic landscapes, mesh geometry watermarking or 3D CAD.

Regardless of the application environment, handling such highly detailed 3D objects imposes considerable demands in storage, transmission, computational and display resources. Scalable mesh coding techniques are therefore particularly important when the aim is to provide a quality and resolution-scalable representation of the 3D object as well as region-of-interest coding and client-view adaptation. This allows for coding, transmitting and rendering such meshes on a broad range of end-user terminals, transmission environments and computational capabilities.

The literature provides several scalable wavelet-based mesh coding technologies that satisfy these requirements. Among them is our MeshGrid representation method [1], which employs a volumetric wavelet transform and efficient volumetric coding technologies. The state-of-the-art in scalable mesh coding systems that deploy wavelets on surfaces is the popular Progressive Geometry Compression (PGC) technique of [2], which relies on the well-known zerotree coding paradigm [3, 4] to encode the wavelet coefficients. Recently, we have designed scalable intraband and composite mesh coding techniques for semi-regular meshes [5] that outperform the PGC compression system.

A common denominator of existing mesh coding techniques, including the aforementioned ones, is the choice of the L-2 distortion metric [6-10] which provides a good approximation of the global error produced by lossy coding of the wavelet coefficients. Though very popular, the L-2 metric is not appropriate for applications where geometry accuracy is critical, such as mesh geometry watermarking, industrial applications (3D CAD, architectural design etc.) or coding of topographic surfaces. The L-infinite distortion overcomes this short-coming by imposing a tight bound on the local error.

To our knowledge, the only existent L-infinite coding approach for meshes is that of MeshGrid proposed in [11], whereas all subdivision-based wavelet techniques, including those of [2] and [5], employ the L-2 distortion metric. Since coding techniques that deploy wavelets on surfaces have demonstrated a wide use in practical applications and coding literature, the goal of this paper is to design a coding methodology ensuring an L-infinite upper bound when employing subdivision-based wavelet coding systems. Our approach is based on our previously designed Scalable Intraband Mesh (SIM) coding system of [5], for which an efficient estimate for the L-infinite distortion is derived.

The paper is structured as follows. Section 2 describes

This work was supported by the Flemish agency for Innovation by Science and Technology (IWT) (ICocoon project and PhD grant R. Florea) and the Fund for Scientific Research-Flanders (FWO project G.0177.12). Datasets are courtesy of Cyberware and Headus. Also, we would like to thank Igor Guskov for providing us with normal meshes and the Trireme remesher.

the employed coding architecture. The L-infinite coding problem is formulated in Section 3. Section 4 describes the L-infinite estimator for the considered wavelet-based coding system, whereas Section 5 reports the experimental results obtained with our L-infinite coding approach. Finally, Section 6 draws the conclusions of our work.

2. SYSTEM OVERVIEW

In our work we employ our recently developed Scalable Intraband Mesh (SIM) coding technique [5]. The unlifted Butterfly [12] wavelet transform is used to construct lower resolution approximations of the semi-regular input mesh M^J. Recursively applying the transform J times generates a coarse approximation of the input mesh, i.e. the base mesh S_J^A, and J high-frequency detail subbands $S_{j=1,\ldots,J}$. The SIM codec losslessly stores the base mesh S_J^A in the final bit-stream. Commonly though, the base mesh is stored using a single rate coder, for instance the TG coder [13]. The wavelet subbands are quantized using Successive Approximation Quantization (SAQ). The coding scheme then exploits the intraband statistical dependencies between the wavelet coefficients through octree decomposition and encoding. To improve the performance even further arithmetic entropy coding is adopted [5].

3. PROBLEM FORMULATION

In general terms, the scalable coding system decomposes the input 3D mesh into J independent sources of information (not including the base mesh), each of them being then progressively encoded. For wavelet-based coding techniques such as MeshGrid [1] or PGC [2], these sources correspond to the wavelet subbands and are encoded in a bitplane-by-bitplane manner. The problem to be solved is determining the layers of information that need to be coded from each source such that the estimated distortion at the decoding side is minimized subject to a total target rate. Alternatively, one can impose a bound on the estimated distortion and minimize the required rate by optimizing the rate allocation.

Let D_{tot} denote the total reconstruction error in the spatial domain. Each source of information j, with $1 \le j \le J$, has a contribution D_j to the total distortion. For additive distortion metrics, one can express the total distortion in the spatial domain as a linear combination of distortion contributions from each source of information:

$$D_{tot} = \sum_{j=1}^{J} w_j D_j(R_j) \qquad (1)$$

where R_j is the rate associated with source j, and w_j are weights reflecting the contribution of the distortion in subband j, with respect to the total distortion. For subdivision-based wavelet coding systems, such as [2] or [5], these weights depend only on the wavelet transform of choice and on the type of distortion metric. Recall that the SIM codec employs the unlifted Butterfly wavelet transform. In this paper, we employ the L-infinite distortion.

The L-infinite norm, defined as $\left\|V-\tilde{V}\right\|_\infty = \max_i \left|v_i - \tilde{v}_i\right|$, represents the MAXimum Absolute Difference (MAXAD) between the original position of the vertices $v_i \in V$ and their decoded versions, $\tilde{v}_i \in \tilde{V}$ [11]. We note that, in addition to ensuring a bound on the local error [11, 14], the L-infinite metric can also provide an upper bound to the Hausdorff distance [11], when the resolution of the original mesh and that of the decoded mesh are identical.

4. SOLUTION METHODOLOGY

In the following, we introduce an estimator for the L-infinite distortion in the context of our SIM coding system and discuss a rate-distortion optimization algorithm that minimizes the rate for a given MAXAD bound.

4.1. Data-independent L-infinite estimator

Intuitively, a quantization error produced in a certain wavelet subband will be translated (via wavelet synthesis) into a contribution to the total reconstruction error occurring in the spatial domain. Due to the linear nature of the transform, it is possible to define a linear relation that combines the various quantization errors produced in the detail subbands into corresponding errors occurring in the spatial domain.

We start from the classical lifting scheme and, for one level of decomposition, we derive the errors on the even samples, $\varepsilon(2m) = \varepsilon_{S_1^A}(m) - \varepsilon_{S_1}(m) * U(m)$, and on the odd samples $\varepsilon(2m+1) = \varepsilon_{S_1^A}(m) * P(m) + \varepsilon_{S_1}(m)$. In line with the notations of Section 2, $\varepsilon_{S_1^A}(m)$ and $\varepsilon_{S_1}(m)$ refer to quantization errors in the approximation and the detail subband, respectively. $P(m)$ and $U(m)$ are the predict and update functions, respectively, and $*$ denotes the convolution operator. One notes that the SIM codec relies on the unlifted Butterfly transform [12], which excludes the update step. Under worst-case scenario assumptions, it is now possible to maximize the different error-contributions from the different wavelet subbands, and determine the smallest upper-bound of the MAXAD:

$$\begin{cases} \sup\left|\varepsilon(2m)\right| = \sup\left|\varepsilon_{S_1^A}(m)\right| \\ \sup\left|\varepsilon(2m+1)\right| = \sum_{i=0}^{7}\left|P(i)\right| \cdot \sup\left|\varepsilon_{S_1^A}(m-i)\right| + 1\cdot \sup\left|\varepsilon_{S_1}(m)\right| \end{cases} \qquad (2)$$

where $P(i)$, $i = \overline{0,7}$ are the coefficients of the Butterfly transform. Let $K_{S_1^A} = \sum_i \left|P(i)\right|$ and K_{S_1} denote the weights with which the approximation and the detail subband, respectively, contribute to the reconstruction error. For the unlifted Butterfly transform $K_{S_1^A} = 1.5$, $K_{S_1} = 1$ and one can see that $\sup\left|\varepsilon(2m+1)\right| > \sup\left|\varepsilon(2m)\right|$, meaning that the

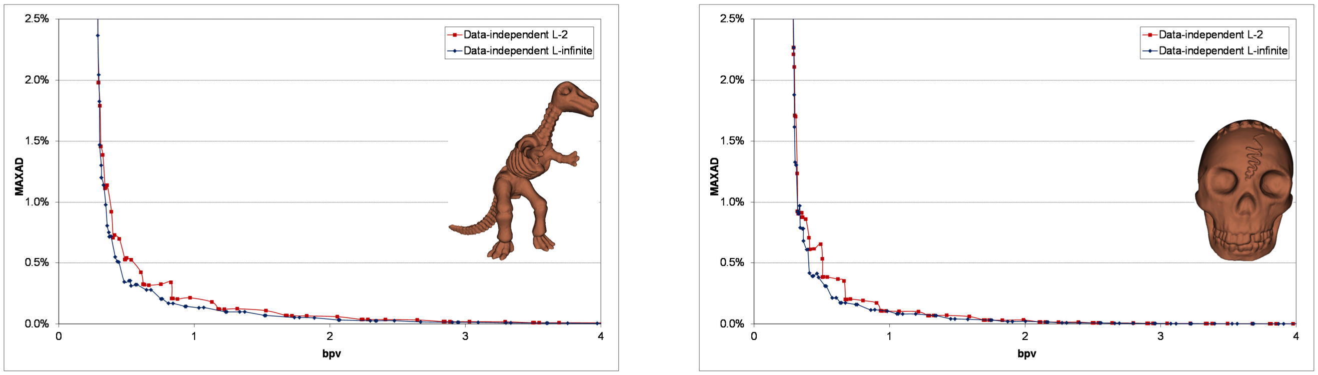

Figure 1. Actual MAXAD (as a percentage of the bounding box diagonal) versus rate (bpv) for the *Dino* (left) and *Skull* (right) mesh models for the (data-independent) L-2 and L-infinite driven SIM codec.

largest reconstruction error occurs on the odd samples. Proceeding recursively, for a J-level decomposition we get:

$$\sup\left|\varepsilon(2m+1)\right| = (K_{S_J^A})^J \cdot \sup\left|\varepsilon_{S_J^A}(m)\right| + \sum_{j=1}^{J}(K_{S_J^A})^{j-1}\cdot K_{S_j}\cdot\sup\left|\varepsilon_{S_j}(m)\right| \tag{3}$$

where $K_{S_J^A}=1.5$ and $K_{S_j}=1$, $1\le j\le J$.

The SIM codec [5] employs the classical SAQ, which is an embedded scalar quantizer for which the deadzone size is twice the size of the other quantizer bins. Let $\Delta_{j,0}$ be the bin-size of the quantizer at the finest quantization level applied on the subband $S_j, 1\le j\le J$ and $b_j\in\mathbb{Z}_+$ be the number of discarded bitplanes for subband S_j. One notes that for any S_j and b_j, the dead-zone bin-size for the quantizer employed by SIM at level b_j is expressed as $2^{b_j+1}\cdot\Delta_{j,0}$, while the other bin-sizes are given by $2^{b_j}\cdot\Delta_{j,0}$. Consequently, the maximum quantization error that can be introduced is half the size of the deadzone and the supremum of the quantization errors in subband S_j is given by $2^{b_j}\cdot\Delta_{j,0}$. As mentioned in Section 2, the coarsest resolution approximation is never quantized. Hence, we can formulate the smallest upper bound M_{tot} of the MAXAD D_{tot}, as:

$$D_{tot}\le M_{tot}=\sum_{j=1}^{J}(K_{S_J^A})^{j-1}\cdot K_{S_j}\cdot\Delta_{j,0}\cdot 2^{b_j} \tag{4}$$

where J is the total number of decomposition levels, S_J^A represents the approximation band after J decomposition levels and $S_j, 1\le j\le J$ are the detail subbands.

In the transmission scheme employed by the original SIM codec [5], the bin-sizes at the finest quantization level are identical across all subbands. Moreover, each decoding point ensures that a complete bitplane is received, meaning that the same number of bitplanes is discarded across all wavelet subbands [5]. However, the subband transmission order employed in the original SIM codec is not necessarily optimal when using the L-infinite distortion metric.

In the following, we will investigate rate-distortion optimization in L-infinite sense. In our computations, we employ the generic expression of the L-infinite estimator given by (4); the problem to be solved is to identify the number of bitplanes to be transmitted from each subband such that the rate is minimal for any given MAXAD bound.

4.2. Rate-distortion optimization algorithm

After J decomposition levels, the intraband bit-plane coding scheme generates an embedded bit-stream Π_j for each subband S_j, $1\le j\le J$. The bit-streams Π_j can be truncated at a predefined set of points τ_{b_j,Π_j}, for b_j discarded bit-planes in subband S_j. Each truncation point τ_{b_j,Π_j} is associated with a certain bitrate $R_j(b_j)$ and a distortion $D_j(b_j)$. The problem at hand is to ensure an optimized performance in rate-distortion sense. This is formulated as a constrained optimization problem, by which the optimal truncation points τ_{b_j,Π_j} need to be determined such that the rate is minimal subject to a constraint on the total distortion.

In JPEG2000 [15] the global optimization problem is split in optimization problems per subband. However, the non-additive nature of the L-infinite metric prevents us from taking this route. In addition, the problem is not necessarily convex and the upper bound formulation of (4) does not guarantee a convex optimization problem either. In order to obtain a convex problem we consider as eligible truncation points only those points that lie on the convex hull of the distortion-rate (D-R) function in each subband. So, the solution is to obtain the convex-hull of the D-R function in each band; the problem is then convex and the solution can be determined using the method of Lagrange multipliers – see e.g. [11]. The optimal distortion-rate slopes $\lambda_j(b_j)$ are calculated using the bisection-method as follows:

$$\lambda_j(b_j)=w_j\frac{D_j(b_j+1)-D_j(b_j)}{R_j(b_j)-R_j(b_j+1)}=w_j\frac{\Delta D_{j,b_j}}{\Delta R_{j,b_j}} \tag{5}$$

where w_j are the weighting factors mentioned in (1). When computing the increase in rate $\Delta R_{j,b_j}$ when an additional quantization level b_j is encoded, we have considered the actual number of bits per vertex required for each subband at each decoding point. $\Delta D_{j,b_j}$ denotes the decrease in

Figure 2. Visual comparison between the original *Skull* mesh model (left), the decoded version when employing the L-infinite driven SIM codec (center) and the L-2 driven SIM codec (right) at a target MAXAD of 2.1% of the bounding box diagonal corresponding to a rate of 0.607 bpv.

distortion between two successive truncation points (i.e. subband bitplanes). $\Delta D_{j,b_j}$ is estimated and we have considered both the L-2 and the L-infinite distortion.

In the L-2 case, the distortion is modeled similar to [14] and [8]. Unlike [8], where the optimal quantization steps are derived for a target rate, we focus on SAQ. Moreover, we employ the classical high-rate approximation $\Delta D_{j,b_j} = 2^{2b_j} \Delta_{j,0}^2 / 12$. Similar to (4), this is a data-independent estimation and corresponds to a classical transmission of subband bitplanes used in wavelet coding of images [3, 4, 16] and in the standardized mesh coding technology MeshGrid [1]. The weights w_j are numerically obtained such that the transform is approximately unitary, giving $w_j = 2^{j-J}$, $1 < j \le J$ and $w_1 = 1.25 \cdot 2^{1-J}$.

In the L-infinite case, the distortion at every possible decoding point is estimated with the data-independent estimator of (4). As previously mentioned, the SIM codec employs Successive Approximation Quantization, for which the decrease in distortion between two decoding points is given by $\Delta D_{j,b_j} = 2^{b_j} \Delta_{j,0}$. In this case, the employed weights are given by $w_j = (K_{S_j^A})^{j-1} \cdot K_{S_j}$ with $K_{S_j^A}$ and K_{S_j} having the same values as in (4).

The slopes $\lambda_j(b_j)$ from all decomposition levels j and all bitplanes are sorted in a monotonically decreasing order, indicating the order in which the subbands have to be transmitted. This corresponds to a global distortion-rate curve for which the slopes are monotonically decreasing. That is, the global distortion-rate curve ensures that the rate is minimal for each target MAXAD and it indicates the appropriate number of bitplanes to be sent from each subband. Similar to [11], the proposed approach offers scalability in L-infinite sense.

5. EXPERIMENTAL RESULTS

In a first set of experiments, we compare the SIM codec when driven by the data-independent L-2 and the data-independent L-infinite distortion estimators. Figure 1 depicts experimental results in terms of actual MAXAD versus bitrate. Each dot on the rate-distortion curves corresponds to a valid decoding point, where the local error is bounded and guaranteed. Irrespective of the model, one observes that the L-2 driven SIM codec is frequently prone to large vertex errors ("error spikes"). This stems from the very nature of the L-2 metric, which does not set bounds on the local errors, quantifying the global error instead.

Visually, this phenomenon is depicted in Figure 2. The *Skull* model is compressed at a user-specified MAXAD bound using the L-infinite codec, employing the data-independent estimator of (4). The L-2 driven codec then compresses the geometry of the model at the same rate as the L-infinite codec, but in this case the L-2 distortion is minimized. The green color indicates that the imposed MAXAD bound is not exceeded. Whenever the distortion value is higher than the imposed bound, the corresponding vertex is represented in red. The conclusion to be drawn is that the L-2 distortion can lay no claim on minimizing the local error on the vertices and that an L-infinite approach should be followed in order to offer this functionality.

To our knowledge, the only wavelet-based L-infinite-oriented mesh codec in the literature is the MeshGrid-based system proposed in [11]. We therefore consider this coding system to be the state-of-the-art in L-infinite coding of meshes. In the last set of experiments, we compare the proposed technique against the state-of-the-art. Figure 3 depicts experimental results for two MeshGrid models (*Humanoid* and *Heart*). Both systems make use of their corresponding data-independent L-infinite estimators. The comparison is made in terms of the actual obtained L-infinite distortion versus required bitrate. Both models have been originally created for the MeshGrid system and were not compatible with the SIM codec. For this reason, a remeshing procedure was required, using the remesher of [17]. One notes that at high rates, for the *Heart* model, the L-infinite driven SIM codec reaches the remeshing error, while MeshGrid further lowers the MAXAD as the rate is increased above 17 bpv. At low-to-medium rates, the performance differences are substantial, indicating that for L-infinite coding, a compression system deploying wavelets on surfaces, as proposed in this paper, should be favored

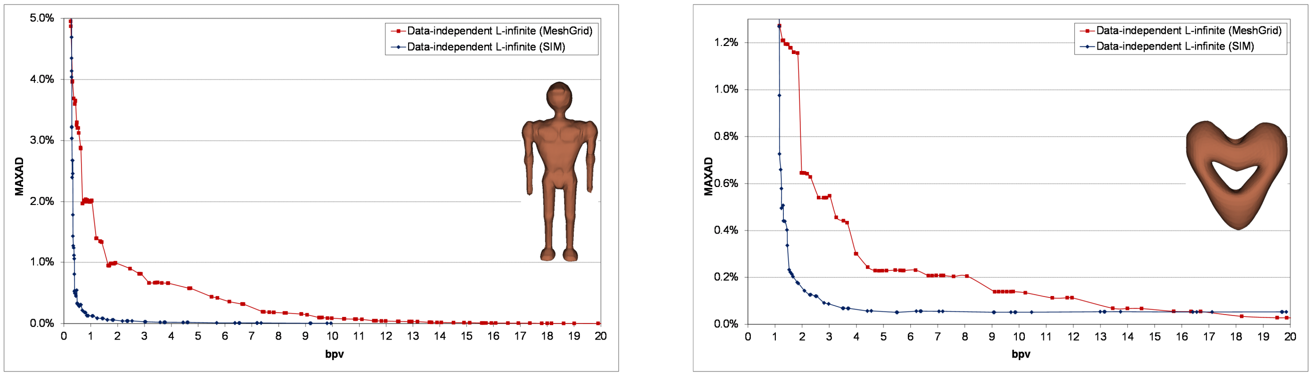

Figure 3. Actual MAXAD (as a percentage of the bounding box diagonal) versus rate (bpv) for the *Humanoid* (left) and *Heart* (right) mesh models when employing the L-infinite driven versions of the MeshGrid and SIM codecs.

over a coding architecture that employs a volumetric wavelet transform, as in the case of MeshGrid [11].

6. CONCLUSIONS

In this paper, we propose a data-independent estimator for the L-infinite distortion in the context of subdivision-based wavelet compression of semi-regular meshes. A rate-distortion optimization algorithm is also proposed that minimizes the rate for a given distortion bound. The results show that an L-2 driven subdivision-based wavelet codec lays no claim on ensuring a bound on the local error, being frequently prone to error spikes, and that an L-infinite driven codec is a viable solution to this problem.

In addition, the proposed system brings substantial performance improvements in L-infinite coding over the state-of-the-art L-infinite MeshGrid-based system of [11]. We note that, although very fast in computational terms, a data-independent L-infinite estimator relies solely on high-rate assumptions. Future work will revolve around the derivation of a more accurate, data-dependent estimator.

7. REFERENCES

[1] I. A. Salomie, A. Munteanu, A. Gavrilescu, G. Lafruit, P. Schelkens, R. Deklerck, and J. Cornelis, "MeshGrid - a compact, multi-scalable and animation-friendly surface representation," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 14, pp. 950-966, 2004.

[2] A. Khodakovsky, P. Schroder, and W. Sweldens, "Progressive geometry compression," *Proc. of SIGGRAPH,* 2000.

[3] J. M. Shapiro, "Embedded image coding using zerotrees of wavelet coefficients," *IEEE Trans. on Signal Processing,* vol. 41, pp. 3445-3462, 1993.

[4] A. Said and A. W. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 6, pp. 243-250, 1996.

[5] L. Denis, S.-M. Satti, A. Munteanu, J. Cornelis, and P. Schelkens, "Scalable intraband and composite wavelet-based coding of meshes," *IEEE Trans. on Multimedia,* vol. 12, pp. 773-789, 2010.

[6] G. AlRegib, Y. Altunbasak, and J. Rossignac, "Error-resilient transmission of 3-D models," *ACM Trans. on Graphics,* vol. 24, pp. 182-208, 2005.

[7] F. Moran and N. Garcia, "Comparison of wavelet-based 3-D model coding techniques," *IEEE Trans. on Circuits and Systems for Video Technology,* vol. 14, pp. 937-949, 2004.

[8] F. Payan and M. Antonini, "Mean square error approximation for wavelet-based semiregular mesh compression," *IEEE Trans. on Visualisation and Computer Graphics,* vol. 12, pp. 649-657, 2006.

[9] S.-B. Park, C.-S. Kim, and S.-U. Lee, "Error resilient 3-D mesh compression," *IEEE Trans. on Multimedia,* vol. 8, pp. 885-895, 2006.

[10] D. Tian and G. AlRegib, "Multistreaming of 3-D scenes with optimized transmission and rendering scalability," *IEEE Trans. on Multimedia,* vol. 9, 2007.

[11] A. Munteanu, D. C. Cernea, A. Alecu, J. Cornelis, and P. Schelkens, "Scalable L-Infinite coding of meshes," *IEEE Trans. on Visualization and Computer Graphics,* vol. 16, pp. 513-528, 2010.

[12] N. Dyn, D. Levine, and J. A. Gregory, "A butterfly subdivision scheme for surface interpolation with tension control," *ACM Trans. on Graphics,* vol. 9, pp. 160-169, 1990.

[13] C. Touma and C. Gotsman, "Triangle mesh compression," in *Proc. of Graphics Interface '98*, Vancouver, 1998.

[14] A. Alecu, A. Munteanu, J. Cornelis, and P. Schelkens, "Wavelet-based scalable L-infinity-oriented compression," *IEEE Trans. on Image Processing,* vol. 15, 2006.

[15] D. S. Taubman and M. W. Marcellin, *JPEG2000: Image compression fundamental, standards and practice*. Dordrecht: Kluwer Academic Publishers, 2001.

[16] A. Munteanu, J. Cornelis, G. Van der Auwera, and P. Cristea, "Wavelet image compression - the quadtree coding approach," *IEEE Trans. on Information Technology in Biomedicine,* vol. 3, pp. 176-185, 1999.

[17] I. Guskov, "Manifold-based approach to semi-regular remeshing," *Graphical Models,* vol. 69, pp. 1-18, 2007.