Weighted bit allocation for multiresolution 3D mesh
generation compression

Frédéric Payan \textsuperscript{a} and Marc Antonini \textsuperscript{a}

\textsuperscript{a} I3S Laboratory, CNRS and University of Nice-Sophia Antipolis,
2000 route des Lucioles, 06903 Sophia-Antipolis, France

\textbf{ABSTRACT}

In this paper, we propose an efficient geometry compression method well-adapted to densely sampled semi-regular triangular meshes. Based on multiresolution analysis performed by wavelet transform, it includes a low complexity model-based bit allocation across wavelet subbands. The main contribution of this paper is the resolution of the sub-optimal bit allocation problem related to biorthogonal wavelet coders for 3D triangular meshes. Indeed, using biorthogonal filters weights the MSE distortion of the reconstructed quantized mesh. These weights can be computed from the wavelet filter bank. This permits to obtain a simple but efficient surface adapted distortion criterion for 3D wavelet coefficient coordinates, and to highly improve the performances of MSE-based codecs.

\textbf{Keywords:} Geometry coding, multiresolution analysis, bit allocation, weighted distortion, lifted Butterfly.

1. INTRODUCTION

In the 3D mesh coding setting, two approaches exist. On one hand, the oldest one exploits monoresolution representations. Generally, the meshes are irregular and compression techniques based on a specific topology information coding. Geometry information is generally quantized\textsuperscript{1} with a predictive method by exploiting already coded vertex positions to predict the following vertices.\textsuperscript{2,3} On the other hand, a more recent approach is the multiresolution method based on multiscale structures. Introduced by Lounsbery in 1994, the multiresolution analysis on 3D meshes permits progressive transmission, compression or display. Moreover, 3D Discrete Wavelet Transform (DWT), frequently used in image or video coding can be performed. This yields very performant methods like the PGC one.\textsuperscript{4} Despite, 3D DWT becomes more efficient on regular or semi-regular meshes and consequently remeshing methods\textsuperscript{5} are often processed before analysis. Even though multiresolution analysis has been introduced in mesh coding, the bit allocation methods, often related to multiresolution analysis and wavelet transforms, are few developed. These methods are frequent in image or video compression\textsuperscript{6} to fully exploit the multiresolution representations and are very performant. In this paper we propose a geometry compression scheme including a bit allocation technique well-adapted to 3D mesh geometry coding. The sub-optimal bit allocation problem related to the Mean Squared Error (MSE) distortion and non-orthogonal filters is solved for the triangular mesh case. We develop theoretically the computation of parameter values that weight the quantization error which appears on the reconstructed mesh. This optimal distortion criterion is included in a simple but efficient surface adapted bit allocation. The proposed compression scheme considerably improves the MSE-based codecs.

This paper is organized as follows. Section 2 introduces backgrounds. Section 3 presents the computation of the weights related to biorthogonal filters and Section 4 proposes our surface-adapted bit allocation process. Finally, we compare our algorithm to state-of-the-art methods and conclude in Section 5.

2. BACKGROUNDS

The overall scheme can be found in figure 1. The first step of the coder is a remeshing technique\textsuperscript{5} to obtain a multiscale semi-regular mesh from original irregular one. Then, a N-level multiresolution analysis is performed using a 3D DWT and results in:

This work is supported by a grant from the region PACA and Optaway Corporation in Sophia Antipolis, France.
• **N high frequency** wavelet coefficient subbands;

• **1 low frequency** subband corresponding to a coarse version of the original mesh;

• the **topology** information corresponding to the coarse mesh.

The LF topology coding is not treated in this paper. It can be coded with any topology coder like the Tounsi-Goisman method\(^1\) or the Edgebreaker technique.\(^3\)

Wavelet coefficients are tridimensional vectors \(X_i = \{x_{i,1}, x_{i,2}, x_{i,3}\}\), where \(i\) stands for the resolution index. These coefficients can be processed separately since they present low inter-correlation\(^7\) and can be modeled by a **Generalized Gaussian Distribution**\(^7,8\) (GGD). For LF coefficients we introduced predictive coding.\(^9\) Hence, each resolution level provides 3 subbands of scalar coefficients related to each coordinate axis. Thus, a N-level DWT provides \(3(N + 1)\) coefficient subbands that will be optimally scalar quantized by the bit allocation method. Finally, quantized coefficients will be coded with a 3D context-based arithmetic coder (3DCBAC) that we developed.\(^10\)

### 3. WEIGHTS FOR DISTORTION

#### 3.1. MSE Distortion of the reconstructed mesh

This section deals with the sub-optimal bit allocation problem related to the MSE and biorthogonal wavelet bases. Indeed, our bit allocation is based on the distortion measure arising from the wavelet coefficient quantization. Generally, this measure is processed with the MSE between original coefficients and the quantized ones. However, the minimization of the MSE of a reconstructed signal from biorthogonal wavelet coders depends on the filters.\(^11\) Thus, these filters must be taken into account in the bit allocation criterion.

The proposed coder includes a wavelet transform, that computes two subbands of coefficients, a **low frequency subband** and a **high frequency subband** in the case of a one-level decomposition (see figure 2). This wavelet transform is applied on a triangular edge lattice (see figure 3).

Kovacevic and Sweldens\(^12\) showed that filters applied on a triangular edge lattice (e.g., the lifted Butterfly scheme) can be deduced from a 4-channel interpolating filter bank (see figure 4 (a)). In that case, the input mesh across the analysis filters becomes 4 cosets \(x_0, x_1, x_2, x_3\). The matrix notation of the reconstruction error due to the quantization process across this synthesis 4-channel filter bank can be written as:

\[
e = H_0e_0 + H_1e_1 + H_2e_2 + H_3e_3,
\]
where $\epsilon_i$ is the quantization error vector on the $i^{th}$ coset. $H_i$ represents the matrix notation of the $i^{th}$ synthesis filter. The $i^{th}$ row of this matrix represents the synthesis filter coefficient values applied to the $i^{th}$ vertex. Moreover, for each row, the sum of values is constant. The corresponding MSE for a reconstructed mesh is:

$$D_T = \frac{1}{NbV} E(e^T e)$$

where $NbV$ is the number of input mesh vertices. Merging (1) in (2) results in

$$D_T = \frac{1}{NbV} E((H_0\epsilon_0 + H_1\epsilon_1 + H_2\epsilon_2 + H_3\epsilon_3)^T(H_0\epsilon_0 + H_1\epsilon_1 + H_2\epsilon_2 + H_3\epsilon_3))$$

Using the properties $x^T x = tr(x^T x)$ and $tr(x + y) = tr(x) + tr(y)$ gives:

$$D_T = \frac{1}{NbV} E[\sum_{i=0}^{3} \sum_{j=0}^{3} tr(H_i\epsilon_i e^T_j H_j^T)]$$

According to Usetvich’s paper, this formula can be written as

$$D_T = \frac{1}{NbV} \sum_{i=0}^{3} \sum_{j=0}^{3} tr(H_i^T H_j R_{ij})$$

where $R_{i,j} = E[e_i e_j^T]$ are the intercorrelation matrices. By assuming the quantization errors are white or mutually uncorrelated, we can write

$$\begin{cases} R_{i,j} = D_i I & \text{for } i = j, \\ R_{i,j} = 0 & \text{else}. \end{cases}$$

where $D_i$ is the distortion related to the quantization error of the $i^{th}$ coset. Then, MSE-based distortion results in

$$D_T = \frac{1}{NbV} [tr(H_0^T H_0)D_0 + tr(H_1^T H_1)D_1 + tr(H_2^T H_2)D_2 + tr(H_3^T H_3)D_3],$$

and can be simplified in

$$D_T = S_0 w_0 D_0 + S_1 w_1 D_1 + S_2 w_2 D_2 + S_3 w_3 D_3$$

with $S_i$ the ratio between the $i^{th}$ coset size and the input signal size and

$$w_i = \sum_{n} h_i^2(n) = \|h_i\|^2$$

the weights induced by the biorthogonal filters.

Hence, the weights depend on the synthesis filter bank coefficient values. As the proposed coder uses a lifting butterfly scheme, a good way to obtain these weights is using the 4-channel polyphase matrix decomposition related to the butterfly scheme that permits to link the filter bank with the predict and update operators.

### 3.2. The 4-channel lifting scheme and its polyphase matrix decomposition

As we said before, a filter applied on a triangular edge lattice (e.g., the lifted Butterfly scheme) can be deduced from a 4-channel interpolating filter bank (see figure 4 (a)). The corresponding polyphase matrix is then:

$$P = \begin{pmatrix} 1 & p_h & p_v & p_d \\ -u_h & 1 - u_h p_h & -u_v p_v & -u_d p_d \\ -u_v & -u_v p_h & 1 - u_v p_v & -u_d p_d \\ -u_d & -u_d p_h & -u_d p_v & 1 - u_d p_d \end{pmatrix}$$

Proc. of SPIE Vol. 5150 1949
where $p_h$, $p_u$ and $p_d$ (respectively $u_h$, $u_u$ and $u_d$) are the predictive operators (respectively the update operators) of a lifting scheme for the second, the third and the fourth channel (see figure 4 (b)). Moreover, Kovacevic and Sweldens gave the overall formulation of the synthesis filter bank\(^{12}\) for a 4-channel case:

$$h_i(z_x, z_y) = P_{0i}(z_x^2, z_y^2) + z_x P_{1i}(z_x^2, z_y^2) + z_y P_{2i}(z_x^2, z_y^2) + z_x z_y P_{3i}(z_x^2, z_y^2)$$  \(\text{(11)}\)

where $i$ is the channel index and $P_{ij}$ is the $ij^{th}$ element of the polyphase matrix $P$. Substituting the elements of (10) in (11) gives the following synthesis filters:

$$h_0(z_x, z_y) = 1 + z_x p_h(z_x^2, z_y^2) + z_y p_u(z_x^2, z_y^2) + z_x z_y p_d(z_x^2, z_y^2)$$  \(\text{(12)}\)

$$h_1(z_x, z_y) = -u_h(z_x^2, z_y^2) + z_x (1 - u_h(z_x^2, z_y^2)) p_h(z_x^2, z_y^2) - z_y u_h(z_x^2, z_y^2) p_u(z_x^2, z_y^2) + z_x z_y u_h(z_x^2, z_y^2) p_d(z_x^2, z_y^2)$$  \(\text{(13)}\)

$$h_2(z_x, z_y) = -u_u(z_x^2, z_y^2) - z_x u_u(z_x^2, z_y^2) p_h(z_x^2, z_y^2) + z_y (1 - u_u(z_x^2, z_y^2)) p_u(z_x^2, z_y^2) - z_x z_y u_u(z_x^2, z_y^2) p_d(z_x^2, z_y^2)$$  \(\text{(14)}\)

$$h_3(z_x, z_y) = -u_d(z_x^2, z_y^2) - z_x u_d(z_x^2, z_y^2) p_h(z_x^2, z_y^2) - z_y u_d(z_x^2, z_y^2) p_u(z_x^2, z_y^2) + z_x z_y (1 - u_d(z_x^2, z_y^2)) p_d(z_x^2, z_y^2)$$  \(\text{(15)}\)

$h_0$ is the low-pass filter and $h_1$, $h_2$ and $h_3$ represent the high-pass filters.

### 3.3. Computation of the weights

#### 3.3.1. 4-channel case

In section 3.1 we introduced the reconstruction MSE corresponding to a 4-channel scheme:

$$D_T = S_0 w_0 D_0 + S_1 w_1 D_1 + S_2 w_2 D_2 + S_3 w_3 D_3$$  \(\text{(16)}\)

with

$$w_i = \sum_n h_i^2(n) = |h_i|^2.$$  \(\text{(17)}\)
Figure 4. (a) the 4-channel synthesis filter bank. $x_i$ represents the $i^{th}$ coset of wavelet coefficients obtained by the analysis filter bank. (b) The corresponding lifting scheme. $p_i$ and $u_i$ are the predictive and update operators.

The 4-channel lifted scheme applied on triangular edge lattice provides 2 coefficient subbands (see figure 2). Indeed, the high frequency coefficients are obtained from the same predictive operator, its orientation is only rotated for each coset $x_1$, $x_2$ and $x_3$. Hence, the high frequency subband of the figure 2 includes these 3 cosets whereas the low frequency subband represents the coset $x_0$.

Hence, the one-level decomposition reconstruction MSE according to a 4-channel lifting scheme can be formulated in

$$D_T = S_{lf}w_{lf}D_{lf} + S_{hf}w_{hf}D_{hf},$$

where $w_{lf}$ and $D_{lf}$ (respectively $w_{hf}$ and $D_{hf}$) are the weight and the distortion corresponding to the LF (respectively HF) subband. $S_{lf}$ and $S_{hf}$ represent the ratio between subband sizes and original input size.

Comparing (16) with (18) provides 2 equations:

$$S_{lf}w_{lf}D_{lf} = S_0w_0D_0,$$
$$S_{hf}w_{hf}D_{hf} = S_1w_1D_1 + S_2w_2D_2 + S_3w_3D_3.\ (19)\ (20)$$

Assuming the stationarity of the 3 high frequency cosets, we can write $D_1 = D_2 = D_3 = \frac{D_0}{3}$. Moreover, $S_{lf} = S_0$ and $\frac{S_{hf}}{S_0} \simeq S_1 \simeq S_2 \simeq S_3$. Then, we can approximate the weights by:

$$\begin{cases} w_{lf} = \frac{w_0}{g} \\ w_{hf} \simeq \frac{w_0}{g} \end{cases} \ (21)$$

Considering the gains $K_{lf}$ and $K_{hf}$ introduced by the lifting scheme on the LF and HF subbands, the one-level decomposition weights become

$$\begin{cases} w_{lf} = \frac{1}{K_{lf}}w_0 = \frac{1}{K_{lf}}||h_0||^2 \\ w_{hf} = \frac{1}{K_{hf}}\frac{(w_0 - w_0 + w_0)}{9} = \frac{1}{K_{hf}}\left(\frac{||h_1||^2 + ||h_2||^2 + ||h_3||^2}{9}\right) \end{cases} \ (22)$$
3.3.2. The lifted butterfly scheme case

For the lifted Butterfly scheme, the predictive and updating coefficients in the $z^2$-transform space are given by\(^\text{13}\):

$$
\begin{align*}
\begin{cases}
p_h(x, z_y) &= \frac{1}{2}(1 + z_x^2) + \frac{1}{2}(z_x^2 z_y^2 + z_x^2) - \frac{1}{8}(z_x z_y^2 + z_x^2 z_y + z_x^2 y^2 + z_x y^2) \\
p_v(x, z_y) &= \frac{1}{2}(1 + z_y^2) + \frac{1}{2}(z_x^2 z_y^2 + z_x^2) - \frac{1}{8}(z_x z_y^2 + z_x^2 z_y + z_x^2 y^2 + z_x y^2) \\
p_d(x, z_y) &= \frac{1}{2}(z_x + z_y^2) + \frac{1}{2}(1 + z_x^2 z_y^2) - \frac{1}{8}(z_x^2 + z_y^2) + \frac{1}{8}(z_x + z_y^2)
\end{cases}
\end{align*}
$$

(23)

and the gains $K_{lf} = 2$ and $K_{hf} = 1$. Substituting (23) and (24) in equation (22), and computing the corresponding filter norms provides the weights for the lifted Butterfly scheme:

$$
\begin{align*}
\begin{cases}
w_{lf} &= \frac{169}{128} \simeq 1.3203125 \\
w_{hf} &= \frac{4450}{2296} \simeq 0.3628743490
\end{cases}
\end{align*}
$$

(25)

3.4. Multilevel Decomposition

Our coder includes a N-level Wavelet transform i.e. the lifting scheme is repeated N times. The optimal N-level 4-channel reconstruction distortion $D_T$ of the output signal is given by:

$$
D_T = \sum_{i=0}^{N} W_i D_i.
$$

(26)

with

$$
\begin{align*}
\begin{cases}
W_i = \frac{N C[i]}{N W} (w_{lf})^i & if \quad i = N; \\
W_i = \frac{N C[i]}{N W} (w_{hf})^i & else.
\end{cases}
\end{align*}
$$

(27)

where $N C[i]$ is the number of wavelet coefficients at the $i^{th}$ resolution level and $N W$ the number of semi-regular vertices.

4. OPTIMAL BIT ALLOCATION

4.1. Normal versus tangential information

As wavelet coefficients are computed in a local coordinate frame (induced by the surface tangent plane), the $z$ coordinates and the $x$ and $y$ coordinates depend respectively on the normal vectors and on the tangent plane at each vertex. Moreover, it can be easily verified that applying a noise on vertices according to their tangent plane does not affect the overall shape. But applying the same noise to vertices according to their normal direction considerably changes the surface geometry. Consequently, the metric error needs to be much less sensitive to quantization error of tangential (x and y axis) than normal coordinates (z axis). This implies that, from a rate distortion point of view, bits should be allocated preferentially to the local normal direction.

Taking into account these remarks, we introduce a weight $\Delta$ depending on the component direction in the bit allocation criterion and the distortion for the $i^{th}$ level wavelet coefficients becomes:

$$
d_p(X_i, \bar{x}_i) = (X_i - \bar{x}_i)^T \Delta (X_i - \bar{x}_i) \quad \text{with} \quad \Delta = \begin{pmatrix}
\delta_{i,1} & 0 & 0 \\
0 & \delta_{i,2} & 0 \\
0 & 0 & \delta_{i,3}
\end{pmatrix}
$$

1952 Proc. ofSPIE Vol. 5150
where $\hat{x}_i$ is the quantization of a wavelet coefficient $x_i \in \mathbb{R}^3$. Then, the weighted distortion for the resolution level $i$ is given by:

$$D_i^p = E[d_p(X_i, \hat{x}_i)] = \int_{x_i \in \mathbb{R}^3} d_p(x_i, \hat{x}_i)p_{X_i}(x_i)dx_i,$$

where $p_{X_i}(x_i)$ is the joint probability density function of vector $X_i$. Assuming that vector coordinates are independent, $p_{X_i}(x_i) = p_{X_{i,1}}(x_{i,1})p_{X_{i,2}}(x_{i,2})p_{X_{i,3}}(x_{i,3})$, and then,

$$D_i^p = \sum_{i=1}^{3} \int_{x_{i,j} \in \mathbb{R}} \delta_i \cdot d(x_{i,j}, \hat{x}_{i,j})p_{X_{i,j}}(x_{i,j})dx_{i,j}.$$

For a GGD, this formula can be rewritten as:

$$D_i^p = \sum_{j=1}^{3} \delta_i \cdot \sigma_{i,j}^2 \cdot D_i \left( \frac{q_{i,j}}{\sigma_{i,j}} \right),$$

where $D_i \left( \frac{q_{i,j}}{\sigma_{i,j}} \right)$ is given by Parisot et al.\(^4\) Then, taking into account the weights $W_i$ introduced by the non-orthogonal filters in section 3, the overall distortion of the reconstruction signal can be written as:

$$D_T = \sum_{i=0}^{N} \sum_{j=1}^{3} W_i \delta_i \cdot \sigma_{i,j}^2 \cdot D_i \left( \frac{q_{i,j}}{\sigma_{i,j}} \right).$$

### 4.2. Criterion

The general purpose of our bit allocation process is to determine the best set of quantization steps $\{q_{i,j}\}$ for each subband (in our case, sets of coordinates $\{x_{i,j}\}$ for $i$ and $j$ fixed) that minimizes the distortion $D_T$ at a given rate $R_{\text{target}}$.\(^4\)

By introducing Lagrangian operators, this constrained allocation problem can be written as:

$$J_a(\{q_{i,j}\}) = D_T + \lambda (R_T - R_{\text{target}}),$$

or equivalently

$$J_a(\{q_{i,j}\}) = \sum_{i=0}^{N} \sum_{j=1}^{3} W_i \delta_i \cdot \sigma_{i,j}^2 \cdot D_i \left( \frac{q_{i,j}}{\sigma_{i,j}} \right) + \lambda \left( \sum_{i=0}^{N} \sum_{j=1}^{3} a_{i,j} R_{i,j} \left( \frac{q_{i,j}}{\sigma_{i,j}} \right) - R_{\text{target}} \right), \quad (28)$$

where $\sigma_{i,j}$ is the variance of the subband $i,j$. The coefficients $a_{i,j}$ depend on the subsampling and correspond to $a_{i,j} = \text{size}(\{x_{i,j}\})/(3 \times \# \text{semi-regular vertices})$. $D_{i,j}$ and $R_{i,j}$ are respectively the distortion and the bitrate of each subband which depend on the quotients $\frac{q_{i,j}}{\sigma_{i,j}}$. The direction selection weights $\delta_{i,j}$ are introduced to permit an allocation preferentially to the local normal direction. Moreover, these weights are related to the resolution level: from coarser to finer levels, geometric information ratio between normal coordinates and tangential coordinates are more and more important. In this paper, we propose to choose these weights related to the resolution level. They are given by:

$$\begin{cases} \delta_{i,j} = \frac{\sigma_{i,j}}{\sigma_{3,j}} & \text{for } j = 3 \text{ (z axis)}; \\ \delta_{i,j} = 1 & \text{otherwise}. \end{cases}$$

By differentiating expression (28) with respect to $q_{i,j}$ and $\lambda$, and by solving the resulting system, we obtain the optimal relationships:

$$\begin{cases} h_{i,j} \left( \frac{q_{i,j}}{\sigma_{i,j}} \right) = \frac{\partial J_a(\{q_{i,j}\})}{\partial q_{i,j}} \\ \sum_{i=0}^{N} \sum_{j=1}^{3} a_{i,j} R_{i,j} \left( \frac{q_{i,j}}{\sigma_{i,j}} \right) - R_T = 0 \end{cases}, \quad (29)$$

$h_{i,j}$ is used in (29) to simplify the notations.\(^4\)

This allocation needs two functions depending on the distribution model: $ln(-h_{i,j}) = f_1(R_{i,j})$, $R_{i,j} = f_2(\frac{q_{i,j}}{\sigma_{i,j}})$. For low complexity purposes, we use pre-computed tables.\(^4\)
4.3. Algorithm

The bit allocation algorithm is the following:

1. \( \lambda \) is given. Compute \( -\lambda \frac{e_{ij}}{W_{ij}} \sigma_{ij} \sigma_{ij} = -h_{ij} \) and read the resulting bitrate \( R_{ij} \) from the first pre-computed function \( f_1 \).

2. While (29) is not verified (below a given threshold), calculate a new \( \lambda \) by dichotomy and return to step 1;

3. Compute \( \frac{h_{ij}}{e_{ij}} \) for each subband using the pre-computed function \( R_{ij} = f_2 \left( \frac{h_{ij}}{e_{ij}} \right) \).

5. EXPERIMENTAL RESULTS

5.1. Comparison with classical MSE

For these experiments, the Touma-Gotsman method\(^1\) is used to encode the topology information. Once quantized, coefficients are entropy coded with an original proposed 3D context-based coder.\(^10\) Measurements are processed with MESH tools.\(^15\) Figure 6 shows the global MSE (related to the bounding box) for two models, rabbit and venus. We compare results obtained with the proposed criterion and those obtained with classical MSE criterion.\(^7,9\) This figure shows a consequent improvement and the efficiency of the new bit allocation criterion. It proves that using a criterion only based on MSE is sub-efficient and introducing a weight on the normal information represents a simple method to improve the bit allocation.

5.2. Comparison with other coders

We compare the proposed method with the Touma-Gotsman coder,\(^1\) MPEG4\(^2\) and the PGC technique.\(^4\) Figure 6 shows that the proposed algorithm provides really better performances than mono-resolution methods and reaches the performances of PGC, one of the best multiresolution method. Visual results for Venus model can be found in figure 7 for different bitrates.

REFERENCES


---

![Figure 5](image1.png)

**Figure 5.** Proposed Criterion Vs. Classical MSE Criterion for models rabbit and venus.

![Figure 6](image2.png)

**Figure 6.** PSNR versus bitrate for models rabbit and venus.
Figure 7. Visual Results for Venus model.