



## Technical Section

# Temporal wavelet-based compression for 3D animated models

Frédéric Payan\*, Marc Antonini

Laboratoire I3S - UMR 6070 CNRS/Université de Nice - Sophia Antipolis 2000, route des Lucioles - 06903 Sophia Antipolis, France

## Abstract

We present an efficient compression scheme for animated sequences of triangular meshes of the same connectivity. The proposed algorithm exploits the temporal coherence of the geometry component by using a temporal wavelet filtering. The quantization of the resulting wavelet coefficients is then optimized by a bit allocation process. This process dispatches the bit budget across the coefficient subbands according to their influence on the quality of the reconstructed sequence for one specific user-given bitrate. The proposed scheme is simple, fast, flexible, and scalable in frame rate and bitrate. Moreover, simulation results show that our approach is competitive for any kind of animated models, whatever the characteristics (parametrically coherent or not, fine/coarse meshes...), contrary to the related works.

© 2006 Elsevier Ltd. All rights reserved.

**Keywords:** Compression; Animation; Irregular meshes; Temporal wavelet; Lifting scheme; Bit allocation

## 1. Introduction

Today animated 3D objects are widely used in numerous domains (scientific applications, computer games, animation...) to represent realistic visual data. These animated 3D objects are most of the time represented by a sequence of irregular meshes sharing the same connectivity at any frame. The motion is obtained by modifying the positioning of the vertices, frame after frame (see Fig. 1). For complex animated mesh sequences, this involves a very large representation, and the compression is a relevant technique to ease storage or transmission of these data.

Compared to the compression of static meshes [1], the problem of producing compact representations for animated sequences has gone largely unaddressed. The overall scheme of the compression methods in this domain [2–10] is the following: an analysis step first transforms the original geometry to reduce the signal information, and then a predictive coding scheme is applied on the resulting details to exploit the temporal coherence between successive frames. The natural approach for analyzing time-varying data is to exploit the temporal coherence, like in

video processing [11–13]. Two of the most relevant approaches for animated sequences introduce for instance the principal component analysis (PCA) to exploit the temporal coherence of the geometry component [4,5]. Unfortunately this tool is efficient only for very specific sequences (with few global motion and a number of frames far higher than the number of vertices). Moreover, the time and memory complexity of this tool is high.

Although wavelets are very popular and particularly efficient in video compression for instance [11–13], they are poorly studied in animation compression. Moreover, the two only wavelet-based coders for animated mesh sequences introduce (*spatial*) wavelets for static meshes, and consequently their efficiency depends on the *parametrical coherence* of the mesh sequences [10,14].

Since we focus on sequences of irregular meshes sharing the same connectivity, one relevant solution to overcome the problem relating to the parametrical coherence is to apply the wavelet filtering along the time axis. Moreover, the fixed connectivity ensures the regularity of the geometry sampling along the time axis, and involves that the wavelet filtering is always valid for any kind of mesh sequences, parametrically coherent or not.

Therefore, we describe in this paper a *temporal wavelet-based coder* for animated sequences of meshes of the same connectivity. One key contribution is the use of a temporal

\*Corresponding author. Tel.: +33 4 92 94 27 22; fax: +33 4 92 94 28 98.  
E-mail addresses: [fpayan@i3s.unice.fr](mailto:fpayan@i3s.unice.fr), [am@i3s.unice.fr](mailto:am@i3s.unice.fr) (F. Payan).



Fig. 1. Several frames of the model CHICKEN.

wavelet transform (TWT) during the analysis step to exploit the temporal coherence of such data. The second key contribution of our compression scheme is the encoding of the resulting wavelet coefficients, optimized by a temporal model-based bit allocation initially proposed for the static semiregular meshes [15]. This coding scheme has the advantage to be valid whatever the wavelet used. We will experimentally show that the proposed coding scheme is simple, fast, flexible and competitive for any kind of animated models whatever the characteristics (parametrically coherent or not, fine/coarse meshes...), contrary to the state-of-the-art methods.

The rest of the paper is organized as follows. Section 2 presents the previous works relative to the compression of animated mesh sequences. Section 3 introduces the problem statement and the proposed approach. Section 4 describes the overall scheme of our compression algorithm. Section 5 gives a detailed description of the model-based bit allocation process. Section 6 deals with the temporal wavelet filtering. Simulation results of our coder are given in Section 7. Finally, we conclude and propose directions for future works in Section 8.

## 2. Related works

The first compression scheme for animated sequences of meshes, proposed by Lengyel, exploits the affine transformations [2]. The author proposes to split a mesh into several submeshes, and computes a rigid-body motion for each submesh. In this way, only a set of affine transformations are needed to represent a submesh, instead of all the displacements of the submesh vertices. Shamir and Pascucci proposed another approach based on the affine transformations, but combined to a multiresolution scheme [3]. Their technique first computes the best affine transformations of the mesh at the first frame to match the subsequent ones, and then encodes the temporal prediction errors. Other schemes based on predictive coding have been proposed to exploit the temporal and spatial correlations of mesh sequences [6,7,9]. The overall idea of all these

approaches is to predict the displacements of the vertices along the sequence and then to encode the residual errors.

In parallel, Alexa and Müller [4] proposed a coding scheme based on the PCA to represent the mesh sequences with only a small number of basis functions. Karni and Gotsman improved this method by further exploiting the temporal coherence and finally encode the PCA coefficients with a second-order linear predictive coding (LPC) [16].

In [8] Briceno et al. presented an original approach. The technique is to project each frame onto a 2D image, and then encode the resulting sequence of “2D images” with some well-known video techniques.

Recently, a wavelet-based compression method has been presented by Guskov and Khodakovsky [10]. They proposed to exploit the *parametric* coherence of some specific animated sequences of meshes, by applying a spatial multiresolution analysis on the frames. Each frame is consequently transformed into several sets of details, in other words the wavelet coefficients. In order to also exploit the temporal coherence, the subbands of coefficients are then progressively encoded by a predictive coding scheme based on a I-frames/P-frames approach (similar to some video compression methods).

In parallel, Yang et al. also proposed a wavelet-based compression algorithm, but for sequences of irregular meshes with changing connectivity [14]. They proposed to construct a semiregular mesh structure at the first frame and then match it to the other frames by using hierarchical motion estimation technique. They finally use a zerotree coding scheme to compress the motion compensation residuals between the successive frames.

## 3. Problem statement and proposed approach

Although they are widespread in video processing [11–13], wavelets are not fully exploited in animation compression. Actually, we observe that the works using wavelets in animation compression [10,14] propose an approach based on spatial wavelets (for static meshes). Such a 3D +  $t$  approach is evident when the input sequence

is a rigid-body motion. In that case, the wavelet coefficients are unchanged at any frame, and this strongly reduces the signal entropy of the moving shape [10]. But the interesting animated mesh sequences are not typically rigid, and consequently the coding performances strongly depend on the stability of the wavelet transform. For instance, the coder of Guskov and Khodakovskiy and Yang et al. [10] is irrelevant for the *Chicken* animation (Fig. 1), because of the lack of parametric coherence in the shape evolution [17]. As a result, the 3D +  $t$  approaches of Guskov and Khodakovskiy and Yang et al. [10,14] are efficient only for *parametrically coherent mesh sequences*. Moreover, a spatial decomposition is attractive only for fine meshes, in other words sequences with a high number of vertices (highlighted by the authors of [10]).

The objective of our work is to propose a wavelet coder for any kind of animated mesh sequences (parametrically coherent or not, fine or coarse meshes). We prefer using a TWT instead of a spatial one, for several reasons. First, the shared connectivity ensures that the geometry sampling is regular (in time). This implies that the analysis step is always valid, whatever the characteristics (manifold or not, sharp features, coarse meshes). Moreover, using a temporal wavelet overcomes the problem of the parametric coherence. In addition, using wavelets to exploit the temporal coherence ensures a compact representation of the transformed data, whatever the motion, and the number of vertices of the given animation, contrary to the PCA-based methods [4,5].

Then, to encode the different subbands of temporal details, we choose to expand the model-based coder of Payan and Antonini [15] to the animated mesh sequences (initially proposed for the semiregular meshes).

#### 4. Overview of our approach

Here is the outline of our coding scheme (Fig. 2).

- **Temporal wavelet transform (TWT):** a temporal wavelet filtering first transforms the original geometry component to reduce the signal information. The resulting one-level wavelet decomposition consists in a subband of detail frames (temporal *high-frequency*) and a subband

of approximation frames (temporal *low-frequency*), as shown in Fig. 7. Each vertex is represented by a 3D detail (or *wavelet coefficient*) at any frame.

- **Scalar quantization (SQ):** the resulting subbands of frames are then splitted into three 1D subbands according to the coordinates. These *coordinate* subbands are then encoded using uniform scalar quantizers SQ. The quantizers depend on the optimal quantization steps computed during the bit allocation process.
- **Bit allocation:** the allocation process optimizes the rate-distortion trade-off relative to the data quantization. The objective is to efficiently dispatch the bit budget across the subbands according to their influence on the quality of the reconstructed mesh sequence for one specific bitrate. A detailed description of this process is given in the next section.
- **Entropy coding:** the quantized details are then encoded with an entropy coder, to further compress the data and finally produce the bitstream. A simplified version of the context-based arithmetic coder of Payan and Antonini [18] is used.
- **Connectivity coding:** in parallel, the shared connectivity of the original sequence must be encoded and transmitted if we want to reconstruct the sequences after decompression. In this paper, we use the valence-based coder of Touma and Gotsman [19].

#### 5. Bit allocation

##### 5.1. Principle

The general purpose of a bit allocation process is to optimize the trade-off between the bitrate and the quality of the reconstructed data. In this paper, the allocation process allows to minimize the quantization losses of the geometry component for one user-given target bitrate.

Since the TWTs the original sequence in a multilevel structure, the objective of the allocation process is to dispatch the bit budget across the different subbands of the multilevel structure according to their influence on the quality of the reconstructed sequences.

More precisely, our allocation process computes the set of optimal quantizers  $\{q^*\}$  minimizing the reconstructed

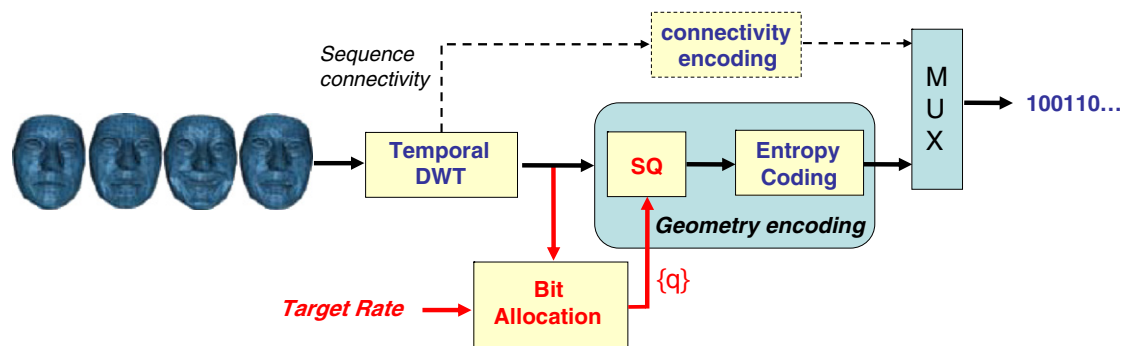


Fig. 2. Overall coding scheme.

mean square error  $D_T$  for one specific user-given target bitrate  $R_{target}$ . The solutions  $\{q^*\}$  are obtained by solving the problem

$$(\mathcal{P}) \begin{cases} \text{minimize} & D_T(\{q\}), \\ \text{with constraint} & R_T(\{q\}) = R_{target}, \end{cases} \quad (1)$$

with  $R_T$  the total bitrate.

### 5.2. Optimal solution

By using a Lagrangian approach, the constrained allocation problem  $\mathcal{P}$  can be solved by minimizing the criterion

$$J_\lambda(\{q\}) = D_T(\{q\}) + \lambda(R_T(\{q\}) - R_{target}), \quad (2)$$

with  $\lambda$  the Lagrangian operator.

The optimal quantization steps  $\{q^*\}$  are obtained by solving the following system:

$$\begin{cases} \frac{\partial J_\lambda(\{q\})}{\partial q} = 0, \\ \frac{\partial J_\lambda(\{q\})}{\partial \lambda} = 0. \end{cases} \quad (3)$$

As introduced in Section 4, the 3D subbands are splitted into three 1D (*coordinate*) subbands, and then encoded separately with three different scalar quantizers. So, the reconstructed mean square error can be defined by

$$D_T(\{q\}) = \sum_{i=0}^N w_i \sum_{j=1}^3 D_{i,j}(q_{i,j}), \quad (4)$$

where  $\{w_i\}$  are weights due to the wavelet non-orthogonality [20] ( $i$  is the resolution level),  $D_{i,j}$  the mean square error relative to the coordinate subband  $i,j$  ( $j = 1$  for the  $x$ -coordinates,  $j = 2$  for the  $y$ -coordinates, and  $j = 3$  for the  $z$ -coordinates), and  $q_{i,j}$  the associated quantization step. The weights  $\{w_i\}$  are defined by  $w_0 = (w_{lf})^N$  and  $w_i = w_{hf}(w_{lf})^{N-i} \forall i \neq 0$ , where  $w_{lf}$  and  $w_{hf}$  are two weights depending on the wavelet used [20]. Table 1 gives for instance the numerical values of these weights for different lifting schemes experimented during this work.

In parallel, the total bitrate  $R_T$  can be developed in

$$R_T(\{q\}) = \sum_{i=0}^N \sum_{j=1}^3 a_{i,j} R_{i,j}(q_{i,j}), \quad (5)$$

with  $R_{i,j}$  the bitrate relative to the coordinate subband  $i,j$ , and  $\{a_{i,j}\}$  the coefficients relative to the subsampling step (ratio between the size of the coordinate subband and the total number of samples). By merging (4) and (5) in (2),

Table 1  
Weights corresponding to different lifting schemes

| Filter   | [2,0] | [2,2]    | [4,2]        | [6,2]        |
|----------|-------|----------|--------------|--------------|
| $w_{lf}$ | 0.75  | 0.75     | 0.8203125    | 0.85299682   |
| $w_{hf}$ | 0.5   | 0.359375 | 0.3410644531 | 0.3359718323 |

system (3) becomes a system of  $(3N + 4)$  equations with  $(3N + 4)$  unknowns (the set  $\{q_{i,j}\}$  and  $\lambda$ ):

$$\frac{\partial D_{i,j}(q_{i,j})/\partial q_{i,j}}{\partial R_{i,j}(q_{i,j})/\partial q_{i,j}} = -\lambda \frac{a_{i,j}}{w_i}, \quad (6a)$$

$$\sum_{i=0}^N \sum_{j=1}^3 a_{i,j} R_{i,j}(q_{i,j}) = R_{target}. \quad (6b)$$

The solutions of (6a) may be obtained by inverting the equations. But this stage is a complex operation, and we overcome this problem by using an iterative algorithm depending on  $\lambda$  [15].

### 5.3. Overall algorithm

The optimal solutions of system (6) for the given bitrate  $R_{target}$  are computed thanks to the following overall algorithm:

- (1)  $\lambda$  is given. For each set  $(i,j)$ , compute  $q_{i,j}$  that verifies (6a);
- (2) while (6b) is not verified, calculate a new  $\lambda$  by dichotomy and return to step (1);
- (3) stop.

### 5.4. Distribution of the detail subbands

The computation of the quantization steps  $\{q_{i,j}\}$  as solutions of (6a) can be done according to different methods. Using a model-based algorithm is a relevant solution, provided that the processed data have specific statistical properties. Therefore, we study the distribution of the coordinate subbands at different decomposition levels.

Fig. 3 shows the probability density functions (pdf) of three typical coordinate subbands of temporal details (CHICKEN animation). We observe that these pdf can be modeled by a generalized Gaussian distribution (GGD). Hence, we can use the model-based algorithm presented in [15] to compute the solutions of (6a). This leads to a fast and low complex algorithm. The interested reader should refer to [15] for more details and explanations.

### 5.5. Encoding of the approximation frames

On the other hand, the typical pdf of coordinate subbands of approximation frames cannot be modeled by a GGD (see Fig. 4(a), (c) and (e)). To overcome this problem, we propose to use a *differential coding*. The overall idea of such a coding is to encode differences between samples instead of samples themselves [15].

So, we deal with the geometric differences between the coordinates of the approximation frames (in function of the vertex indexing), i.e., for each frame  $t$ :

$$[V(0, t), V(1, t) - V(0, t), V(2, t) - V(1, t), \dots, V(n_v, t) - V(n_v - 1, t)],$$



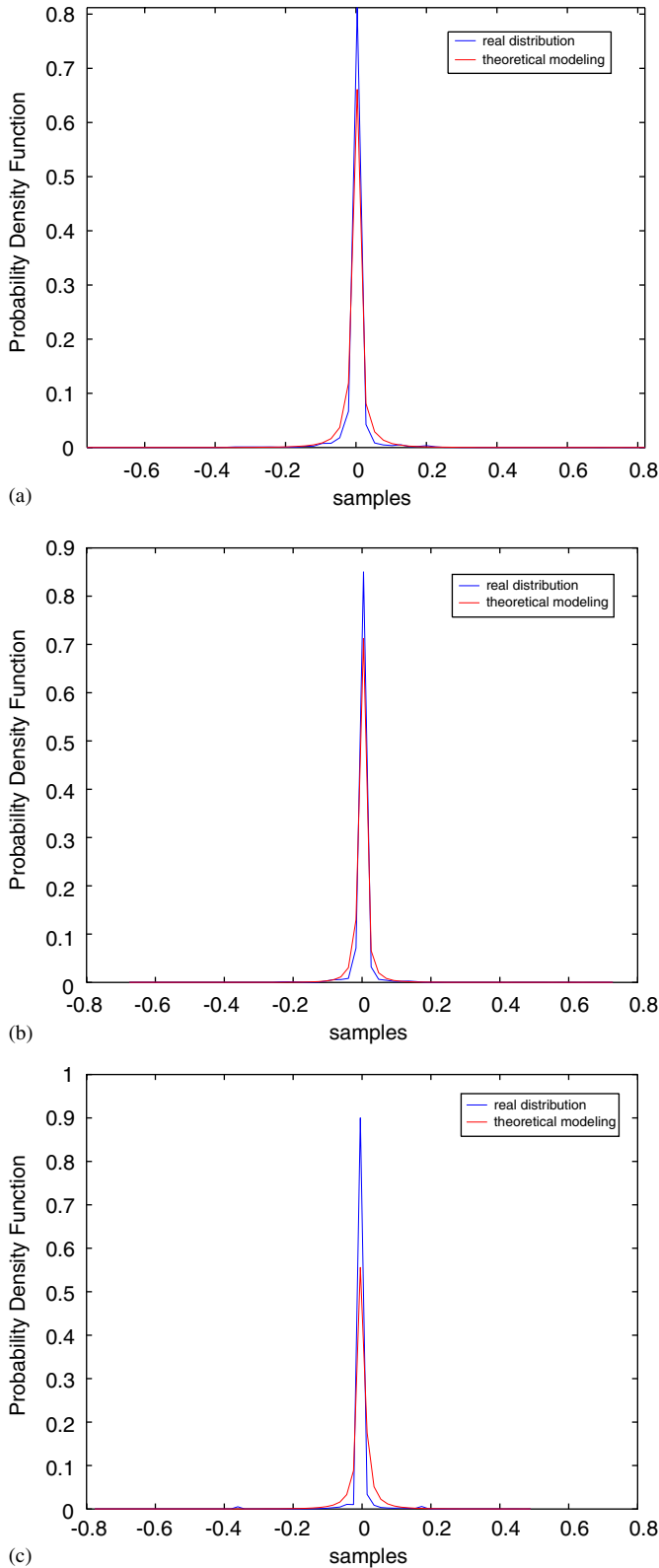


Fig. 3. Typical pdf of coordinate subbands of detail frames (CHICKEN animation). The blue and red curves represent, respectively, the real distribution and the approximated one modeled by a GGD. (a)  $X$ -coordinates (level 1); (b)  $Y$ -coordinates (level 1); (c)  $Z$ -coordinates (level 1).

with  $V(i, t)$  the positioning of the vertex  $i$  at the frame  $t$ , and  $n_v$  the number of vertices.

Fig. 4(b), (d) and (f) shows the typical pdf of three coordinate subbands of approximation frames (CHICKEN animation), when using the proposed differential coding. We observe that the distribution of the coefficients of the approximation frames finally can be modeled by a GGD as well. Moreover, this allows to exploit the spatial correlation between neighbor vertices, and further reduce the data information needed to represent the LF sequence.

This differential coding may appear to be a crude solution. Nevertheless, it is sufficient for our algorithm since the model-based algorithm presented in [15] can be finally applied on the LF data. This is partly due to the valence-based coder of Touma and Gotsman used to encode the connectivity. It provides a non-arbitrary indexing of the vertices sufficiently coherent [19]. Moreover, the low complexity of this approach leads to very fast encoding and decoding of the approximation frames. However, the improvement of this encoding is an interesting direction for future works (see Section 8).

## 6. Temporal wavelet transform for animated mesh sequences

### 6.1. Principle

Wavelets are powerful tools widely used in signal processing to provide a multiresolution representation of a given  $dD$  signal and enable decorrelation in space and frequency [21]. In case of time-varying signals ( $dD + t$ ), a wavelet filtering can be also applied along the temporal axis to exploit the temporal coherence (see Fig. 5). This discrete TWT leads to an efficient energy concentration on the low-pass temporal subbands.

As discussed in Sections 1 and 3, a TWT is a natural and relevant approach for the geometry component of animated mesh sequences sharing the same connectivity. The fixed connectivity implies a regular sampling in time and a fixed number of vertices at any frame. Thus, a discrete TWT can be easily applied on this kind of data.

An animated sequence of meshes sharing the same connectivity is defined by a set of  $T$  meshes  $\{f_1, f_2, \dots, f_T\}$ , where  $f_i$  represents the  $i$ th frame.  $f_i$  is defined by its geometry (i.e. the positioning of the vertices at the frame  $i$ ), and a list of triangles describing the connectivity of the vertices at any frame. As the number of vertices is fixed at any frame, we propose in this paper to apply a monodimensional TWT on the successive locations of each vertex (see Fig. 6) [22,23]. The principle is the following. Let us define the evolution of the vertex of index  $i$  along the time by

$$V(i) = \{V(i, 0), V(i, 1), \dots, V(i, T-1)\}, \quad (7)$$

with  $V(i, t)$  the positioning of this vertex at the frame  $t$ . For a given vertex  $i$ , such a temporal filtering provides two sets:

- the set of temporal details (or *high frequency* coefficients)  $h^{(1)}(i)$  relative to the set of odd samples

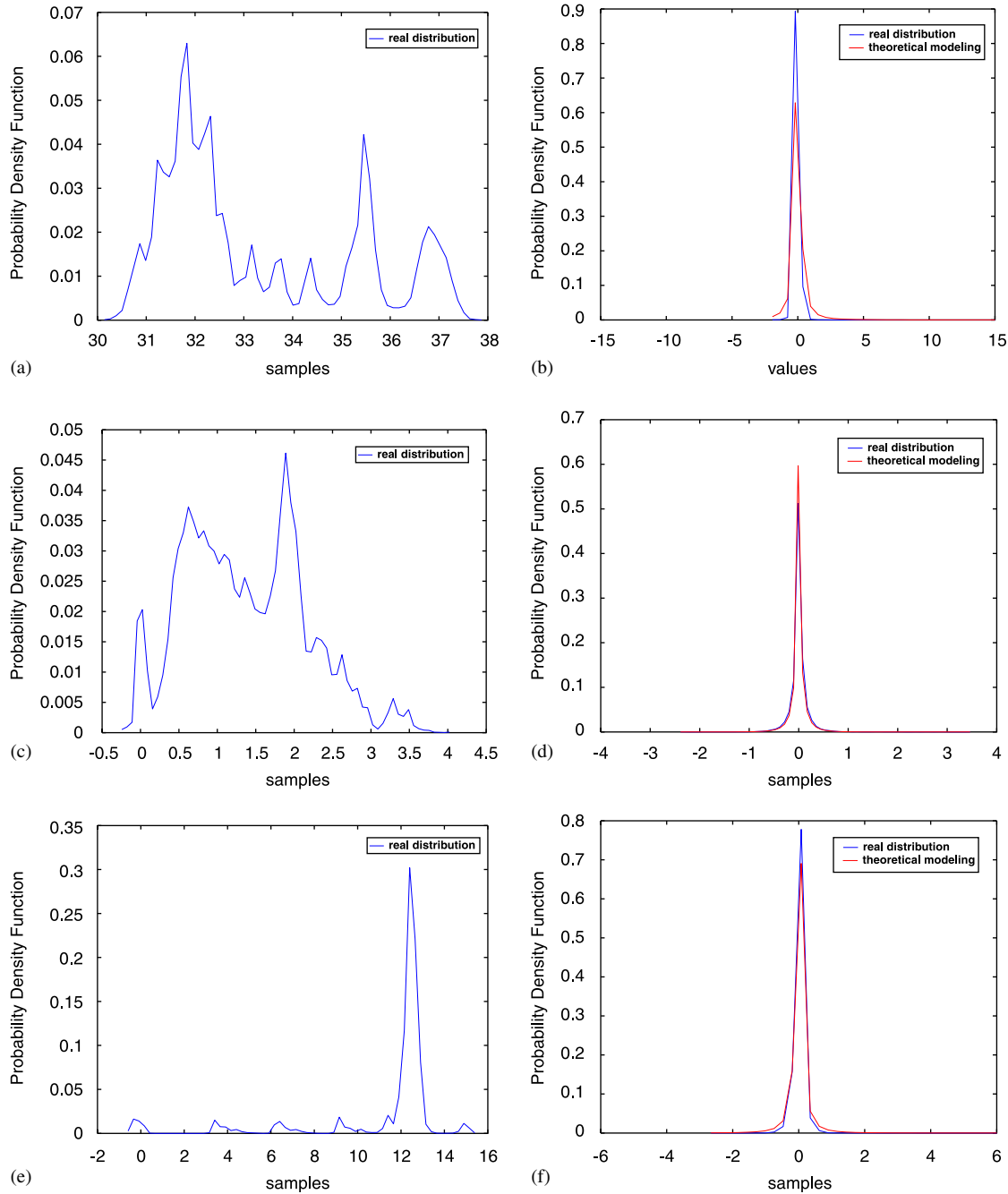


Fig. 4. Typical pdf of coordinate subbands of approximation frames, with and without differential coding (CHICKEN animation). The blue and red curves represent, respectively the real distribution and the approximated one modeled by a GGD. (a)  $X$ -coordinate subband without differential approach; (b)  $X$ -coordinate subband with differential approach; (c)  $Y$ -coordinate subband without differential approach; (d)  $Y$ -coordinate subband with differential approach; (e)  $Z$ -coordinate subband without differential approach; (f)  $Z$ -coordinate subband with differential approach.

$\{V(i, 2k + 1)\}$  (the upper index (1) represents the decomposition level);

- the set of *low frequency* coefficients  $l^{(1)}(i)$  relative to the set of even samples  $\{V(i, 2k)\}$ .

Applying this filtering in parallel on each vertex provides that the input sequence of  $T$  meshes is finally splitted into two subbands:

- the subband  $h^{(1)}$  defined by  $T/2$  detail frames;
- the subband  $l^{(1)}$  defined by  $T/2$  approximation frames.

One can obtain a multiresolution decomposition by subsequent filterings of the approximation subband. This decomposition consists in  $N$  subbands of detail frames  $\{h^{(r)}\}$  (with  $r$  the resolution level), and the approximation subband  $l^{(N)}$ . Fig. 7 shows the resulting 2-level decomposition of the FACE animation.

Numerous wavelet families exist depending on their mathematical properties, but also on their method of construction [24]. In this paper, we focus on the *lifting implementation* [25], because it has several advantages over

classical constructions: it allows easier and faster implementation (low computational cost), a fully in-place calculation, and the synthesis step can be directly deduced from the analysis step. Moreover, lifting-based implementation of any transversal wavelet filter can be designed [26].

### 6.2. Lifting implementation

If we consider a filter bank whose lifting implementation corresponds to a single lifting step (see Fig. 8), then the temporal details and the low frequency coefficients are obtained from the following equations:

$$h^{(1)}(i, k) = V(i, 2k + 1) - P(\{V(i, 2k)\}), \quad (8)$$

$$l^{(1)}(i, k) = V(i, 2k) + U(h^{(1)}(i)), \quad (9)$$

where  $P$  and  $U$  are, respectively, the predict and update operators. Fig. 9 shows for instance the principle of the lifting scheme [2, 0] with 2 levels of decomposition. Note that the synthesis step is easily obtained reversing the same lifting steps with opposite signs:

$$V(i, 2k) = l^{(1)}(i, k) - U(h^{(1)}(i)), \quad (10)$$

$$V(i, 2k + 1) = h^{(1)}(i, k) + P(\{V(i, 2k)\}). \quad (11)$$

## 7. Experimental results

First we compare the efficiency of several classical lifting schemes often used in video compression [13,27]: the schemes [2, 0], [2, 2], [4, 2] and [6, 2]. The corresponding

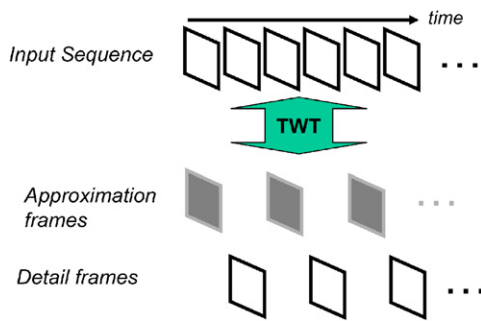


Fig. 5. Temporal wavelet filtering, one level of decomposition.

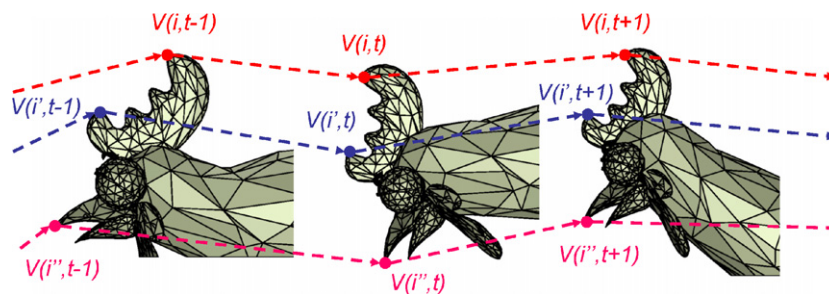


Fig. 6. Proposed temporal wavelet filtering for animated sequences of meshes sharing the same connectivity. The previous and next positionings of one vertex are used to compute the associated wavelet coefficient at the instant  $t$ .

coefficients of the predict and update operators are given in Table 2. Nevertheless, other wavelet families may be more relevant in animation compression, for instance the spline wavelets particularly useful for representing multiresolution curves [28]. In this way, a deep algebraic and comparative analysis of different wavelet families is one of the most interesting direction for future works.

We study three well-known animations with different characteristics (see Table 3): FACE, COW and CHICKEN. To evaluate the coding performances of our coder, we use the metric error introduced by Karni and Gotsman [5]. In the rest of the paper, this metric is called KG error, and is expressed in percent. This metric, corresponding to the relative discrete  $L_2$ -norm both in time and space, is given by

$$\text{KG error} = 100 \frac{\|G - \hat{G}\|}{\|G - E(G)\|}, \quad (12)$$

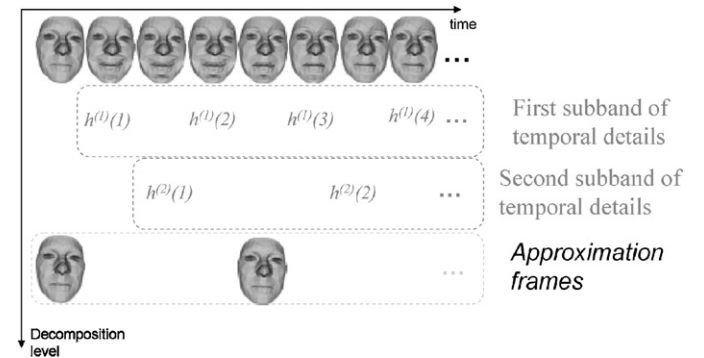


Fig. 7. The 2-level decomposition of the FACE animation.

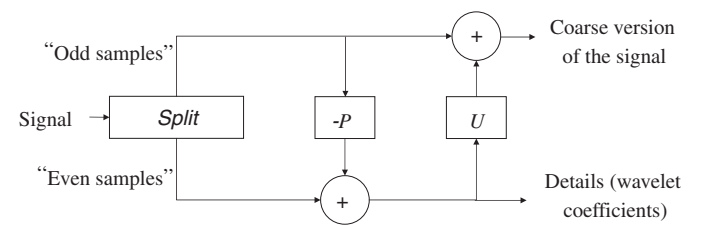


Fig. 8. Principle of the lifting implementation.

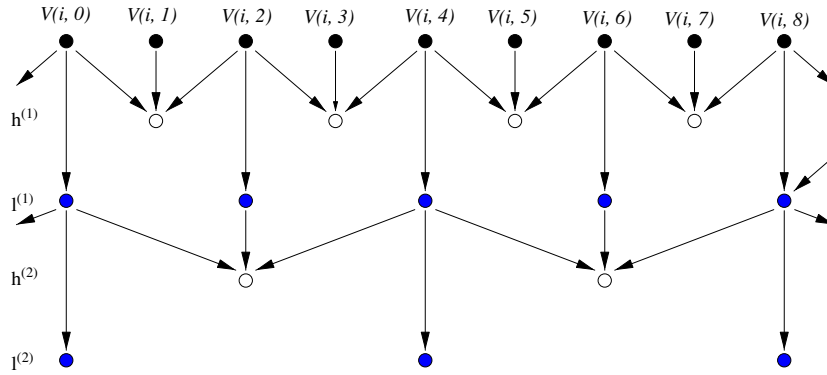


Fig. 9. Principle of the lifting scheme [2,0] with 2 levels of decomposition.  $V(i, t)$  is the position of the vertex  $i$  in the frame  $t$ .

Table 2 Predict and update operators of different lifting schemes

| Lifting scheme | Predict operator  |
|----------------|---|
| [2,0]          | $[-\frac{1}{2} \quad -\frac{1}{2}]$   |
| [2,2]          | $[-\frac{1}{2} \quad -\frac{1}{2}]$   |
| [4,2]          | $[\frac{1}{16} \quad -\frac{9}{16} \quad -\frac{9}{16} \quad \frac{1}{16}]$   |
| [6,2]          | $[-\frac{3}{256} \quad \frac{25}{256} \quad -\frac{75}{128} \quad -\frac{75}{128} \quad \frac{25}{256} \quad -\frac{3}{256}]$ |
| Lifting scheme | Update operator   |
| [2,0]          | —   |
| [x,2]          | $[\frac{1}{4} \quad \frac{1}{4}]$   |

Table 3 Characteristics of the animations

| Sequence | $T$    | $n_v$ |
|----------|--------|-------|
| CHICKEN  | 399    | 2916  |
| FACE     | 10,001 | 539   |
| COW      | 204    | 2904  |

where  $G$  is a matrix of dimension  $(3 \times n_v, T)$  containing the geometry of the original sequence,  $\hat{G}$  the quantized version of the geometry, and  $E(G)$  an average matrix in which the  $t$ th column is defined by

$$(\bar{X}_t(1 \dots 1), \bar{Y}_t(1 \dots 1), \bar{Z}_t(1 \dots 1))', \quad (13)$$

with  $\bar{X}_t$ ,  $\bar{Y}_t$ , and  $\bar{Z}_t$  the mean values of the coordinate sets of each frame  $t$ .

Figs. 10–12 show the curves KG error/bitrate of our coder for the three sequences according to the different lifting schemes (with three levels of decomposition). The bitrate is given in bits per vertex per frame. We observe that the scheme [2,0] provides the worst coding performances, whereas the schemes [4,2] or [6,2] provides the best performances. This is due to the fact that the schemes [4,2] and [6,2] take into account a larger neighborhood, and consequently leads to a better prediction.

Nevertheless, the schemes [2,2], [4,2] and [6,2] provide similar results at low bitrates for the FACE animation. The

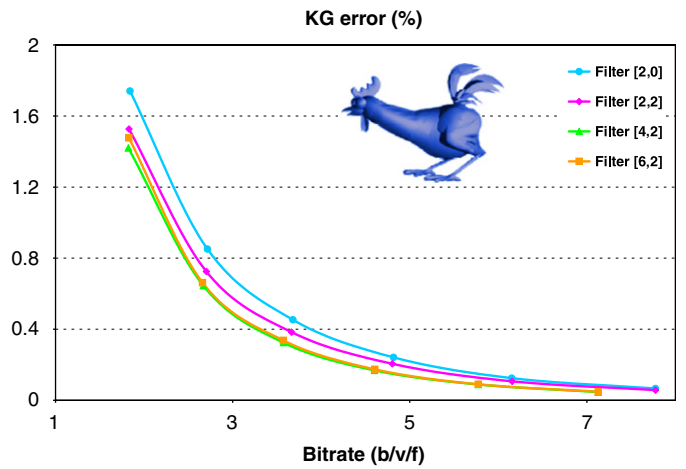


Fig. 10. Curve KG error/bitrate for the CHICKEN animation according to different lifting schemes.

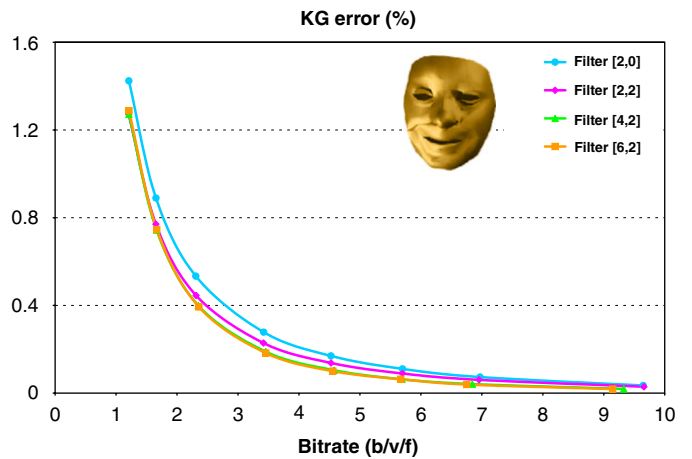


Fig. 11. Curve KG error/bitrate for the FACE animation according to different lifting schemes.

very small vertex displacements between subsequent frames of this animation involves that the prediction step provides similar results (same wavelet coefficients), whatever the complexity of the operator  $P$  used.



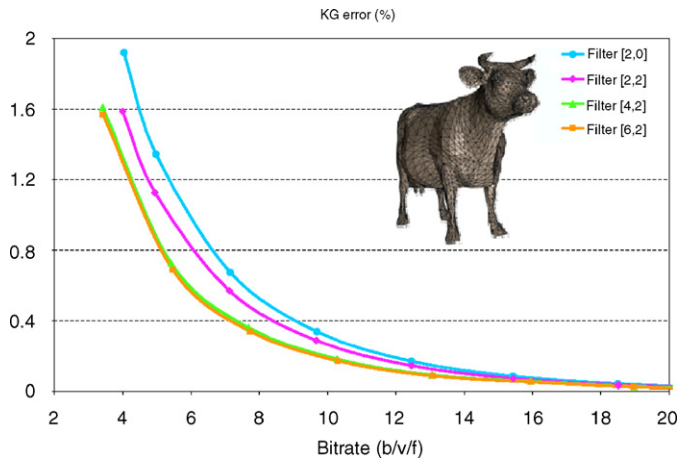


Fig. 12. Curve KG error/bitrate for the Cow animation according to different lifting schemes.

Table 4

Computational cost in operations per vertex per frame (ovf) according to different schemes

| Lifting scheme | [2,0] | [2,2]  | [4,2]  | [6,2]  |
|----------------|-------|--------|--------|--------|
| Complexity     | 6 ovf | 12 ovf | 18 ovf | 24 ovf |

As the filter [4,2] requires less computing resources in processing time and memory usage than the filter [6,2] (25% fewer operations, see Table 4) for similar results, we finally conclude that the filter [4,2] is the most interesting among the experimented filters.

We also study the influence of the number of decomposition levels. Figs. 13–15 show the curves KG error/bitrate for the three same sequences in function of the number of decomposition levels (when using the filter [4,2]). For the FACE, the best results are obtained with 7 decomposition levels. For the two other animated sequences, the best results are obtained with 4 decomposition levels. This difference is mainly due to the repetitive nature of the FACE, and the spatially restricted displacements of the vertices. These characteristics imply a larger temporal coherence (compared to the sequences CHICKEN and COW).

To show the interest of our approach, we also compare the performances of our coder with several state-of-the-art coders:

- the coder for static meshes of Touma and Gotsman [19] denoted by TG;
- the PCA-based coder for mesh sequences of Alexa and Müller [4] denoted by PCA;
- the coder for mesh sequences of Karni and Gotsman [5] denoted by KG. It combines the coder PCA to a linear prediction coding;
- the coder *Dynapack* of Ibarria and Rossignac [9];
- the spatial wavelet-based coder of Khodakovsky and Guskov [10], denoted by AWC.

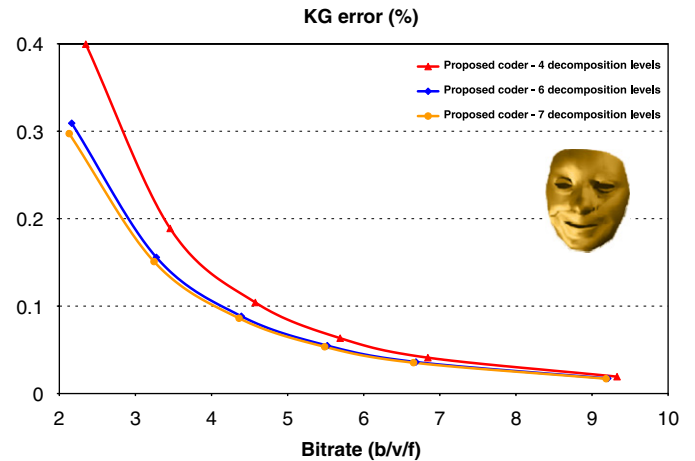


Fig. 13. Curve KG error/bitrate for FACE according to different numbers of decomposition levels (filter [4,2]).

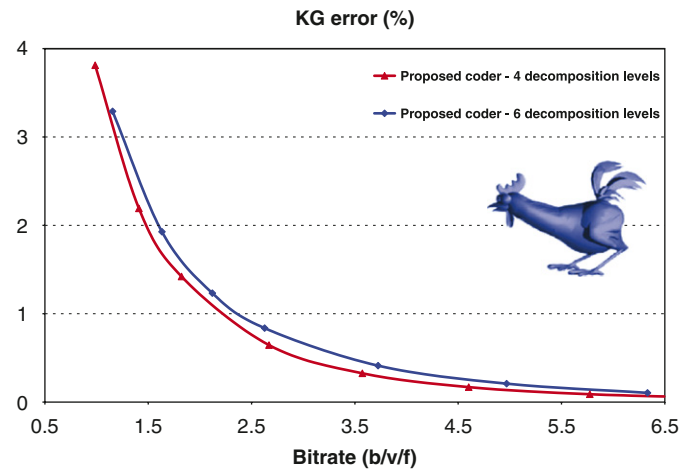


Fig. 14. Curve KG error/bitrate for CHICKEN according to different numbers of decomposition levels (filter [4,2]).

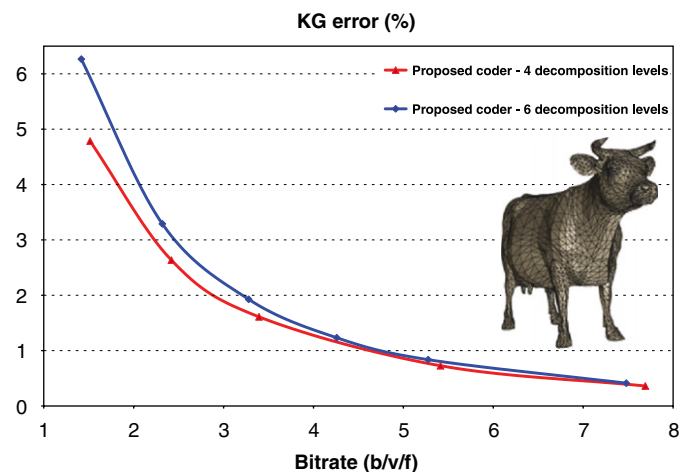


Fig. 15. Curve KG error/bitrate for Cow according to different numbers of decomposition levels (filter [4,2]).

All the results relative to these methods are extracted from [5,10], that is why some results are missing. This is also why we only use the KG error of Karni and Gotsman [5] as quality criterion. Notice that the bit budget needed to encode the connectivity of the animated sequences is included in the total bitrate.

First, Fig. 16(a) shows the curve KG error/bitrate for the sequence FACE. We observe that the PCA-based methods (PCA and KG) provide the best coding performances. This is not surprising given the characteristics of this sequence. It turns out that the PCA-based methods are particularly performant for long sequences of relatively coarse meshes [29]. On the other hand, our method is significantly more efficient than the other methods, and especially the coder AWC. Given the coarseness of this animated sequence (small number of vertices), the spatial multi-resolution nature of this coder cannot be fully exploited. As expected, the proposed temporal wavelet filtering does not depend on such characteristics. It shows one of the advantages of a temporal decomposition over a spatial decomposition.

Now, we observe the coding performances of our approach on the Cow (Fig. 16(b)), and on the CHICKEN (Fig. 16(c)). We recall that these animations are defined by short sequences of finer meshes and present much larger deformations than the FACE. First, we observe that for these two sequences our method clearly outperforms the PCA-based methods.<sup>1</sup> It turns out that the PCA-based methods are inefficient for sequences with fine meshes. This is due to the size of the “payload” archive of eigenmodes that becomes too large in case of short sequences [5]. Our simple temporal wavelet filtering does not provide such a side information during decompression. This shows the interest of our approach over the PCA-based methods.

Second, we observe our coding scheme provides better coding performances than AWC for the FACE, but not for the Cow. We assume this is due to the very small local deformations of the Cow, involving the parametric coherence, frame after frame. Nevertheless, compared to AWC, our method gives very competitive coding performances.

Concerning the sequence of the CHICKEN, we cannot compare our coder with AWC (Fig. 16(c)), because the results are not given in the corresponding paper [10]. Anyway, the method AWC is inefficient for such a sequence because of the lack of parametric coherence and the characteristics of this animated sequence (non-manifold and sharp features) [17].

From a compression point-of-view, we can conclude that our approach is competitive for any kind of animated sequences, and particularly relevant when the meshes are fine (more powerful than the PCA-based methods) and non

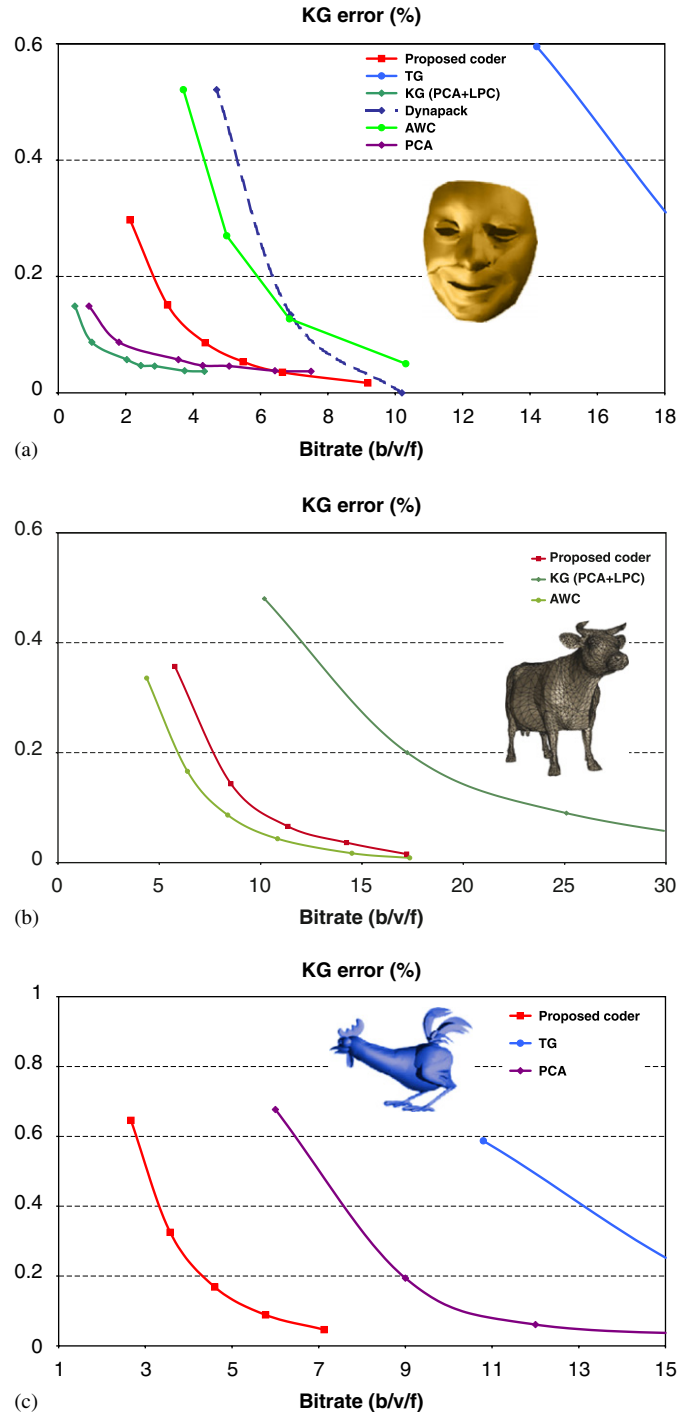


Fig. 16. Curves KG error/bitrate relative to different compression methods. (a) FACE (7 decomposition levels); (b) Cow (4 decomposition levels); (c) CHICKEN (4 decomposition levels).

parametrically coherent (more powerful than the method AWC).

## 8. Conclusions and future works

In this paper we have introduced a simple and efficient wavelet-based coder for any kind of animated sequences of

<sup>1</sup>The results of KG for the CHICKEN is missing because the authors claim that their method is not relevant for such a data. Consequently, the numerical values are not given in their paper [5].

irregular meshes sharing a same connectivity. One key contribution of this work is to propose a temporal wavelet filtering to process the geometry component of such data. Such a temporal analysis has the advantage to overcome the problem of the parametric coherence of a given animation assumed by previous wavelet coders in animation compression. The second key contribution is a model-based bit allocation that optimizes the quantization process. The resulting coder is simple, fast, and flexible (any 1D wavelet can be used). Experimentally, we show that our approach is always valid and competitive, whatever the characteristics of the animated mesh sequences (contrary to the state-of-the-art methods).

There are a lot of interesting directions for future works. The most important is a deep algebraic and comparative analysis of the numerous 1D wavelet families (orthogonal, spline...). Our future works should also concern the encoding of the approximation frames. Currently, the solution is crude but sufficient to exploit the model-based algorithm of Payan and Antonini [15] during the allocation process. Nevertheless, some more relevant solutions may improve the coding performances of our coder since the spatial coherence of the approximation frames is not fully exploited (using the works of Sorkine [30,31] for instance). Another interesting work should be a specific implementation to obtain a bitstream structure standing the temporal scalability, but also the random access across the time. This should be evident since our approach already allows the temporal scalability. Furthermore, this new implementation should not increase significantly the bitstream size.

## Acknowledgments

The sequence CHICKEN is the property of Microsoft Inc. and the sequence FACE was kindly generated by Demetri Terzopoulos. We are particularly grateful to Zach Karni for providing us with the sequences CHICKEN and FACE, and Igor Guskov for providing us with the sequence COW and the results of his method [10]. We are also particularly grateful to the PhD student Yasmine Boulfani for providing us with some simulation results of the proposed method. All the results of the state-of-the-art methods are extracted from the paper [5] or provided by Igor Guskov. We also want to acknowledge the anonymous reviewers for their advices, which permitted to improve the quality of this paper.

## References

- [1] Alliez P, Gotsman C. Recent advances in compression of 3D meshes. In: Proceedings of the symposium on multiresolution in geometric modeling, 2003.
- [2] Lengyel JE. Compression of time-dependent geometry. In: ACM symposium on interactive 3D graphics 1999. p. 89–96.
- [3] Shamir A, Pascucci V. Temporal and spatial level of details for dynamic meshes. In: Proceedings of virtual reality systems and techniques, 2001. p. 423–30.
- [4] Alexa M, Muller W. Representing animations by principal components. *Computer Graphics Forum* 2000;19:3.
- [5] Karni Z, Gotsman C. Compression of soft-body animation sequences. *Computers and Graphics* 2004;28:25–34.
- [6] Ahn J, Kim C, Kuo C, Ho Y. Motion compensated compression of 3D animation models. *IEEE Electronics Letters* 2001;37(24):1445–6.
- [7] Yang J, Kim C, Lee S-U. Compression of 3D triangle mesh sequences based on vertex-wise motion vector prediction. *IEEE Transactions on Circuits and Systems for Video Technology* 2002;12(12):1178–84.
- [8] Briceno H, Sander P, McMillan L, Gotler S, Hoppe H. Geometry videos: a new representation for 3D animations. *ACM symposium on computer animation*, 2003. p. 136–46.
- [9] Ibarria L, Rossignac J. Dynapack: space-time compression of the 3D animations of triangle meshes with fixed connectivity. *ACM symposium on computer animation*, 2003. p. 126–35.
- [10] Guskov I, Khodakovskiy A. Wavelet compression of parametrically coherent mesh sequences. In: *Eurographics/ACM SIGGRAPH symposium on computer animation*, 2004.
- [11] Turaga D, van der Schaar M, Pesquet-Popescu B. Complexity scalable motion compensated wavelet video encoding. *IEEE Transactions on Circuits and Systems for Video Technology* 2005;15(8):982–93.
- [12] Mehrseresh N, Taubman D. Spatial scalability and compression efficiency within a flexible motion compensated 3d-dwt. In: *Proceedings of the IEEE international conference on image processing (ICIP2004)*, Singapore; 2004.
- [13] Cagnazzo M, André T, Antonini M, Barlaud M. A model-based motion compensated video coder with jpeg2000 compatibility. In: *Proceedings of the IEEE international conference on image processing (ICIP2004)*, Singapore; 2004.
- [14] Yang J-H, Kim C-S, Lee S-U. Progressive compression of 3D dynamic sequences. In: *Proceedings of the ICIP 2004*, 2004.
- [15] Payan F, Antonini M. An efficient bit allocation for compressing normal meshes with an error-driven quantization. *Computer Aided Geometric Design. Special Issue on Geometric Mesh Processing* 2005;22:466–86.
- [16] Karni Z, Gotsman C. 3D mesh compression using fixed spectral basis. *Graphics interface conference proceedings*, 2001. p. 1–8.
- [17] Guskov I., Private communication.
- [18] Payan F, Antonini M. 3D multiresolution context-based coding for geometry compression. In: *Proceedings of IEEE international conference in image processing (ICIP)*. Spain: Barcelona; 2003.
- [19] Touma C, Gotsman C. Triangle mesh compression. *Graphics interface'98*, 1998. p. 26–34.
- [20] Payan F, Antonini M. Mean square error approximation for wavelet-based semiregular mesh compression. *IEEE Transactions on Visualization and Computer Graphics (TVCG)* 2006;649–57.
- [21] Mallat S. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 1989;11(7):674–93.
- [22] Payan F, Boulfani Y, Antonini M. Temporal lifting scheme for the compression of animated sequences of meshes. In: *Proceedings of IEEE international workshop VLBV*. Italy: Sardinia; 2005.
- [23] Payan F, Antonini M. Wavelet-based compression of 3d mesh sequences. In: *Proceedings of IEEE ACIDCA-ICMI'2005*. Tunisia: Tozeur; 2005.
- [24] Vetterli M, Herley C. Wavelet and filter banks: theory and design. *IEEE transactions on Acoustic Speech Signal Processing* 1992;40(9):2207–32.
- [25] Sweldens W. The lifting scheme: a custom-design construction of biorthogonal wavelets. *Applied and Computational Harmonic Analysis* 1996;3(2):186–200.
- [26] Daubechies I, Sweldens W. Factoring wavelet transforms into lifting steps. Technical report, 1996. URL ([citeseer.ist.psu.edu/daubechies96factoring.html](http://citeseer.ist.psu.edu/daubechies96factoring.html)).

- [27] Calderbank A, Daubechies I, Sweldens W, Yeo B. Wavelet transforms that map integers to integers. *Applied and Computational Harmonic Analysis* 1998;5(3):332–69.
- [28] Stollnitz EJ, DeRose T, Salesin DH. *Wavelets for computer graphics: theory and applications*. San Francisco, CA, USA: Morgan Kaufmann Publishers; 1996.
- [29] Karni Z, Gotsman C. Spectral compression of mesh geometry. In: *ACM SIGGRAPH conference proceedings*, 2000, p. 279–286.
- [30] Sorkine O, Cohen-Or D, Toledo S. High-pass quantization for mesh encoding. In: *Proceedings of the Eurographics/ACM SIGGRAPH symposium on geometry processing*, Eurographics Association, 2003. p. 42–51.
- [31] Chen D, Cohen-Or D, Sorkine O, Toledo S. Algebraic analysis of high-pass quantization. *ACM Transactions on Graphics* 2004;24(4): 1259–82.