Bit Allocation For Spatio-temporal Wavelet Coding Of Animated Semi-regular Meshes

Aymen Kammoun, Frédéric Payan, and Marc Antonini

Laboratoire I3S (UMR 6070 CNRS-Université de Nice - Sophia Antipolis)
2000, Route des Lucioles - 06903 Sophia Antipolis - France
{kammoun,fpayan,am}@i3S.unice.fr
www.i3s.unice.fr/~creative/

Abstract. In this paper, we propose a compression scheme for animated semi-regular meshes. This scheme includes a spatio-temporal wavelet filtering to exploit the coherence both in time and space. In order to optimize the quantization of both spatial and temporal wavelet coefficients, the proposed compression scheme also includes a model-based bit allocation. The experimental results show that this approach significantly improves the compression performances, when comparing with previous similar approaches.

Key words: Spatio-temporal decomposition, Wavelet filtering, Animation, Semi-regular meshes, Bit Allocation, Coding.

1 Introduction

Typically, an animation is represented by a sequence of triangular meshes sharing the same connectivity at any frame. Even if the static triangulation has the advantage to strongly reduce the amount of information needed to represent such data, compression is useful when the number of vertices increases, because of the amount of geometric information.

During the last decade, most of works focused on the compression for sequences of irregular meshes. When observing the state of the art for such data, we can note techniques using spatial and/or temporal prediction [1, 2]; principal component analysis (PCA) [3–5]; motion-based segmentation [6, 7]; wavelet-based analysis tool [8–11].

Among the wavelet-based methods, the pioneering coder for sequences of irregular meshes was based only on temporal wavelet filtering [8]. The reason is that the geometry of an animation has a regular temporal sampling (same connectivity at any frame), while having an irregular "spatial" sampling (the frames are irregular meshes). And it is well-known that wavelets are more efficient on regular grids than irregular ones (in image coding for instance). Recently, another coder using temporal wavelet filtering but implemented in lifting scheme was proposed [10]. In parallel, a motion-compensated temporal wavelet filtering has been also developed in [6]. The main idea of all these coders is to apply a one-dimensional wavelet filtering on the successive positions of each vertex.

Despite the challenge of applying wavelets on irregular grids, some coders using irregular spatial wavelet was proposed. For instance, Guskov and Khodakovsky have developed in [9] a coder for parametrically coherent mesh sequences, using an irregular wavelet-based decomposition. Unfortunately, from an analysis point-of-view, all these methods tend to be suboptimal since wavelets are not fully exploited both in spatial and temporal dimension. With this outlook, Cho et al. recently proposed in [11] to expand the wavelet-based coder for static irregular meshes of [12] to the time-varying surfaces.

In geometry compression, it has been shown that combining remesher and wavelets based on semi-regular sampling leads to efficient techniques for compressing static shapes [13–15]. The semi-regular meshes have the advantage to present a multiresolution structure which facilitates wavelet filtering (to exploit the spatial coherence), and also visualization, levels of details, progressive transmission and so on. In order to have the same properties for animations, similar approaches have been lately proposed for animations defined by sequences of meshes with static or independent connectivity [16, 17]. The resulting semi-regular animations finally have the same relevant properties as the static meshes and also a temporal coherence which can be exploited to perform efficient compression.

We propose in this paper a spatio-temporal wavelet-based coding scheme for sequences of semi-regular meshes sharing the same connectivity. The proposed coder includes an improved version of the bit allocation process presented in [16], which optimizes the quantization of the temporal wavelet coefficients but also the spatial ones (contrary to [16]). The key contribution of this paper is a formulation of the distortion criterion which takes into account the quantization errors of all the data (temporal wavelet coefficients, spatial wavelet coefficients and low frequency signal).

The rest of this paper is organized as follows. Section 2 gives an overview of the proposed method. Section 3 shows the principle of a spatio-temporal wavelet filtering. Section 4 shows the bit allocation for a spatio-temporal decomposition. Section 5 shows some results for several animations, and we conclude in section 6.

2 Overview of the proposed approach

Let us consider an animated semi-regular mesh $\mathcal{F}$, represented by a sequence of T meshes sharing the same connectivity: $\mathcal{F} = \{f_1, f_2, ..., f_t, ...f_T\}$. Each mesh f_i corresponds to one *frame* of the animation. Semi-regular animations can be obtained by using a remeshing technique such as [16, 17]. Here is the outline of our method, illustrated by figure 1.

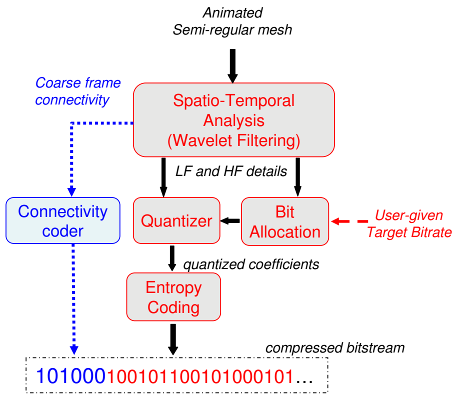


Fig. 1. Overview of the complete proposed coding scheme.

- **Spatio-temporal Analysis**: The animation $\mathcal{F}$ has a regular sampling in time but also A semi-regular one in "space". So, a temporal wavelet-based filtering is first applied, followed by the spatial butterfly-based lifting scheme [13]. Hence, we obtain a spatio-temporal decomposition of the input animation (see figure 2): several sets of temporal and spatio-temporal high-frequency details (or *wavelet coefficients*), and a short sequence of coarse meshes (*low frequency signal*). The list of triangles of one sole coarse mesh represents the whole connectivity information.
- **Bit Allocation**: this process optimizes the rate-distortion trade-off relative to the data quantization. The objective is to efficiently dispatch the bits across the subbands of wavelet coefficients, according to their influence on the quality of the reconstructed mesh sequence. With this outlook, this process will compute the optimal quantization step for each subband, for one user-given total bitrate.
- **Entropy coding**: once quantized, the coefficients are entropy coded with the context-based arithmetic coder of [10].
- **Connectivity coder**: to reconstruct the compressed data after storage and/or transmission, the connectivity information must be preserved. For this purpose, the list of triangles is encoded in parallel with [18] and included in the compressed bitstream.

3 Spatio-temporal decomposition

To create a spatio-temporal decomposition of a semi-regular animation, we choose to apply successively a temporal wavelet filtering and a spatial one. These two filterings are based on lifting scheme implementations [19].

Temporal wavelet filtering

The principle of our temporal wavelet filtering is to consider each vertex trajectory as a one-dimensional (1D) signal [10]. Therefore, a classical 1D lifting scheme is applied on the successive positions of each vertex in parallel. After this step, we obtain two subsets: a high frequency subband H_1^t defined by $T/2$ detail frames (*temporal wavelet coefficients*): $H_1^t = \{h_{1,0}^t, h_{1,1}^t, ..., h_{1,T/2}^t\}$ and a low frequency subband L_1^t defined by $T/2$ approximation frames (*temporal low frequency signal*): $L_1^t = \{l_{1,0}^t, l_{1,1}^t, ..., l_{1,T/2}^t\}$. One can obtain a multiresolution decomposition by subsequent filterings of the low frequency subbands. This decomposition finally consists in N_t subbands of detail frames $H_{r^t}^t$ (where r^t represents the temporal resolution), and the low frequency sequence $L_{N_t}^t$.

Spatial wavelet filtering

After applying the temporal wavelet filtering, we apply the spatial wavelet filtering based on the butterfly lifting scheme [13] on all the frames of the low frequency sequence $L_{N_t}^t$. For each frame $l_{N_t,k}^t$ of this sequence, this step gives N_s subbands of spatial wavelet coefficients $h^s l_{r^s,k}^t$ (where r^s represents the spatial resolution level) and a coarse version $l^s l_{N_s,k}^t$ of the given frame.

The principle of this decomposition is illustrated by figure 2. In this figure, we apply 2 levels of temporal decomposition, and 2 levels of spatial decomposition on the animated DOLPHIN.

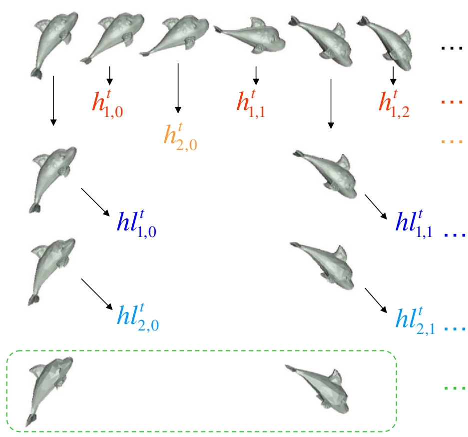


Fig. 2. Spatio-temporal decomposition based on wavelet filtering: 2 levels of temporal decomposition giving the temporal wavelet coefficients (in red and orange), and 2 levels of spatial decomposition giving the spatial ones (in blue). The meshes in the dotted rounded box (in green) represent the coarse sequence (spatio-temporal low frequency signal).

Given the resulting output multiresolution structure of this analysis stage, we now have to deal with the necessary bit allocation process which will optimize the quantization of the different subbands.

4 Bit allocation

Given a multiresolution structure, we can not use the same quantizer for all the subbands since the energy is not uniformly dispatched between them. We recall that once transformed, the relevant data (from a coding point of view) is most of times concentrated in the lower frequency subbands, while having the fine and perhaps negligible details in the highest frequency subbands. Therefore we generally use a bit allocation process, to compute the best quantizer for each subband, in order to obtain the best trade-off between rate (compressed file size) and quality of the reconstructed data (distortion).

In our case, the general purpose of the bit allocation process is precisely to determine the best set of quantization steps $\{q^s_{i,j}, q^t_{i,j}\}$ -where $q^s_{i,j}$ and $q^t_{i,j}$ are respectively the quantization steps for the spatial subbands and for the temporal subbands- that minimizes the reconstruction error $D_T(\{q^s_{i,j}, q^t_{i,j}\})$ for a given rate R_{target}. This can be formulated by the following problem:

$$(\mathcal{P}) \begin{cases} \text{minimize} & D_T\left(\{q^s_{i,j}, q^t_{i,j}\}\right) \\ \text{with constraint} & R_T\left(\{q^s_{i,j}, q^t_{i,j}\}\right) = R_{target}. \end{cases} \quad (1)$$

Before solving this problem, we first must formulate the reconstruction error D_T.

Note: To know how solving such an optimization problem and have more details, please refer to [15].

4.1 Distortion criterion

Temporal decomposition Let us consider a temporal wavelet coder, using one level of decomposition (see figure 3). In this case, the reconstruction error

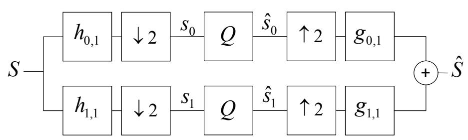


Fig. 3. Model of a temporal wavelet coder used to link the quantization effects with the reconstruction error after decoding. $(h_{0,1}, g_{0,1})$ and $(h_{1,1}, g_{1,1})$ are respectively the low-pass analysis/synthesis filters and the high-pass ones. Q represents the quantization process. $\downarrow 2$ and $\uparrow 2$ are respectively the downsampling and upsampling stages.

D_T can be written as the weighted sum of the MSE $(\sigma^t_{01})^2$ and $(\sigma^t_{11})^2$ relative to the quantization of the low frequency signal and of the wavelet coefficients [20]:

$$D_T = \frac{1}{2} w^t_{01} (\sigma^t_{01})^2 + \frac{1}{2} w^t_{11} (\sigma^t_{11})^2, \tag{2}$$

where w^t_{ij} are the weights due to the non-orthogonality of the temporal wavelet filters ($i = 0$ gives a low frequency subband, $i = 1$ a high frequency subband and j is the resolution level). Those weights are given by:

$$w^t_{01} = \frac{2}{N} tr\left((G^t_{01})^T G^t_{01}\right) \tag{3}$$

and

$$w^t_{11} = \frac{2}{N} tr\left((G^t_{11})^T G^t_{11}\right), \tag{4}$$

where N is the number of vertices and G^t_{ij} is the matrix of reconstruction relative to the synthesis filter g^t_{ij}.

Let us consider now a temporal wavelet coder with N_t levels of decomposition. By using (2), we find that the reconstruction error at any level of decomposition $(n-1)$ $(2 \leq n \leq N_t)$ can be given by:

$$(\sigma^t_{0(n-1)})^2 = \frac{1}{2} w^t_{0n} (\sigma^t_{0n})^2 + \frac{1}{2} w^t_{1n} (\sigma^t_{1n})^2. \tag{5}$$

Since we use the same synthesis filter for each level of decomposition,

$$g_{0i} = g_{01} \textit{ and } g_{1i} = g_{11} \ \forall i \in [2, N_t], \tag{6}$$

and consequently, the weights are also similar whatever the level:

$$w^t_{0i} = w^t_{01} = w^t_{lf} \textit{ and } w^t_{1i} = w_{11} = w^t_{hf} \ \forall i \in [2, N_t]. \tag{7}$$

By combining the new weights of (7), (2) and (5), we finally obtain:

$$D_T = \frac{1}{2^{N_t}} (w^t_{lf})^{N_t} (\sigma^t_{0N_t})^2 + \sum_{j=1}^{j=N_t} \left(\frac{1}{2^j} (w^t_{lf})^{j-1} w^t_{hf} (\sigma^t_{1j})^2 \right). \tag{8}$$

Spatial decomposition The distortion for N_s levels of spatial decomposition of a given triangular mesh can be estimated by [15]:

$$(\sigma_S)^2 = (w^s_{lf})^{N_s} (\sigma^s_{0N_s})^2 + \sum_{j=1}^{j=N_s} \left((w^s_{lf})^{j-1} w^s_{hf} (\sigma^s_{1j})^2 \right), \tag{9}$$

where w^s_{lf} and w^s_{hf} are the weights due to the non-orthogonality of the spatial wavelet transform for the low and high frequency subbands, respectively.

Spatio-temporal decomposition Now, we can generalize the distortion criterion for a spatio-temporal decomposition with N_t levels of temporal decomposition and N_s levels of spatial decomposition. For each frame of the temporal low frequency subband, we obtain N_s spatial subbands. So, the number of spatial subbands is $N_{BF}N_s$, where N_{BF} is the number of frames in the temporal low frequency subband. By replacing $(\sigma^t_{0N_t})^2$ by (9) in (8), we finally obtain:

$$\begin{aligned} D_T = \tfrac{1}{2^{N_t}}(w^t_{lf})^{N_t} \textstyle\sum_{f=1}^{f=N_{BF}} \Big(\tfrac{1}{N_{BF}} \Big((w^s_{lf})^{N_s} (\sigma^s_{0N_s})^2 \\ + \textstyle\sum_{j=1}^{j=N_s} \Big((w^s_{lf})^{j-1} w^s_{11} (\sigma^s_{1j})^2 \Big) \Big) \Big) \\ + \textstyle\sum_{j=1}^{j=N_t} \Big(\tfrac{1}{2^j} (w^t_{lf})^{j-1} w^t_{11} (\sigma^t_{1j})^2 \Big) \end{aligned} \quad (10)$$

4.2 Model based bit allocation

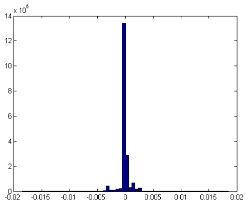
As explained in [15], a bit allocation is generally processed with an iterative algorithm. At each iteration, we need to evaluate the distortion (and the associated bitrate) of each subband for one given set of quantization steps to check the convergence of the algorithm. In this way, one can achieve a real prequantization and then compute real distortion and bitrate for each subband. But such a method is time-consuming, particularly when dealing with large animations. One relevant method to overcome this problem is using statistical models to estimate distortion and bitrate of each subband without quantizing data. To perform such a model-based process, we must study the distribution of the processed wavelet coefficients. Figure 4 shows some results for the animated VENUS. We observe that the data follow a Gaussian distribution whatever the level of decomposition, and consequently we can use the model-based bit allocation proposed in [15].

5 Results

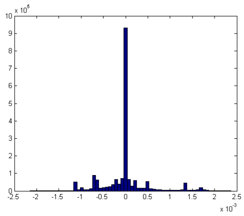
In this section we compare the proposed compression scheme to the following coders:

- The temporal wavelet coder proposed in [10] applied to the original irregular animations;
- The previous spatio-temporal wavelet coder without an optimal bit allocation proposed in [16].

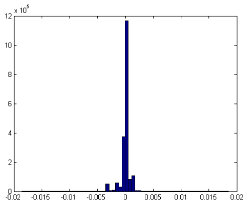
To evaluate the quality of compressed animations, we use the RMSE metric based on Hausdorff distance given in [21]. First, figures 5, 6, and 7 show the average reconstruction error (computed on all the frames) of these coders in function of different bitrates, for the animated DOLPHIN, VENUS and HORSE. We observe that the proposed compression scheme based on a spatio-temporal decomposition and an optimal model-based bit allocation gives the best results. It clearly proves the interest of using a spatio-temporal decomposition and an



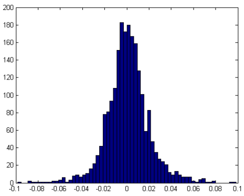
(a) X-coordinates (first level of temporal decomposition).



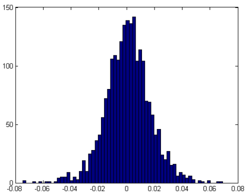
(b) Y-coordinates (first level of temporal decomposition).



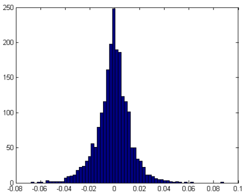
(c) Z-coordinates (first level of temporal decomposition).



(d) X-coordinates (second level of spatial decomposition).



(e) Y-coordinates (second level of spatial decomposition).



(f) Z-coordinates (second level of spatial decomposition).

Fig. 4. Typical distribution of the wavelet coefficients of the animated semi-regular VENUS with the spatio-temporal decomposition used.

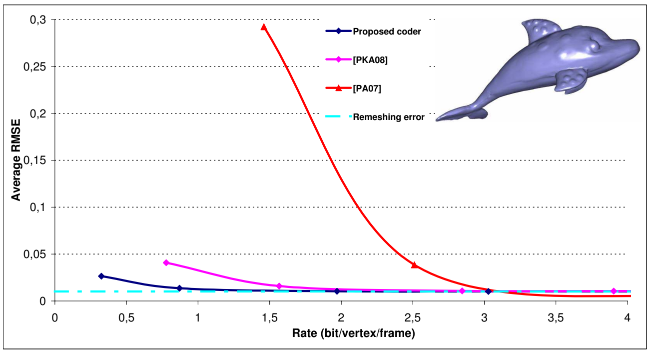


Fig. 5. Rate-Distortion curve for the animated DOLPHIN.

optimal bit allocation process for such data instead of the previous allocation process proposed in [16]

In addition, figures 8, 9 and 10 show the evolution of the RMSE frame after frame when using the different coders, at given bitrates. We can observe that the proposed coder significantly improves the coding performances. With a smaller bitrate, the proposed coder gives a smaller reconstruction error for the majority

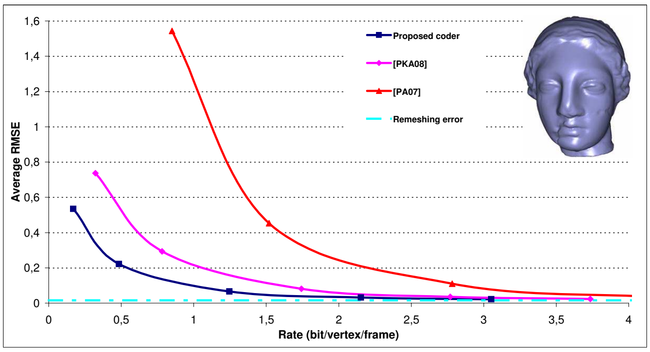


Fig. 6. Rate-Distortion curve for the animated VENUS.

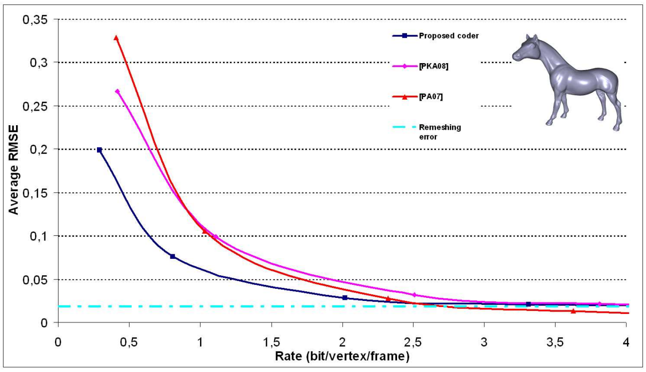


Fig. 7. Rate-Distortion curve for the animated HORSE.

of frames. Furthermore, it tends to produce a more regular quality over the frames (no peak of distortion, unlike the two other coders).

6 Conclusion and future works

In this paper, we have presented a new framework for compressing animations sharing the same connectivity at any frame. Inspired by works in geometry processing for static meshes, we have developed a spatio-temporal wavelet-based coding scheme for animated semi-regular meshes. Semi-regular animations offers the advantage of making the spatial wavelet-based analysis easier, since wavelets are more efficient on (semi-)regular grids than on irregular ones. The key contribution of this paper is the bit allocation process which optimizes the quantization of all the subbands of temporal and spatial wavelet coefficients created by the analysis tool. Furthermore, the statistics of the wavelet coefficients allows us to use a model-based algorithm, which makes the allocation process very fast. Experimentally, we show that the proposed method improves significantly the compression performances of the previous coders [10] and [16], which makes our new coder particularly interesting. For future works we plan to improve the spatio-temporal filtering and add a motion estimation/compensation technique to further improve the performances of our coder.

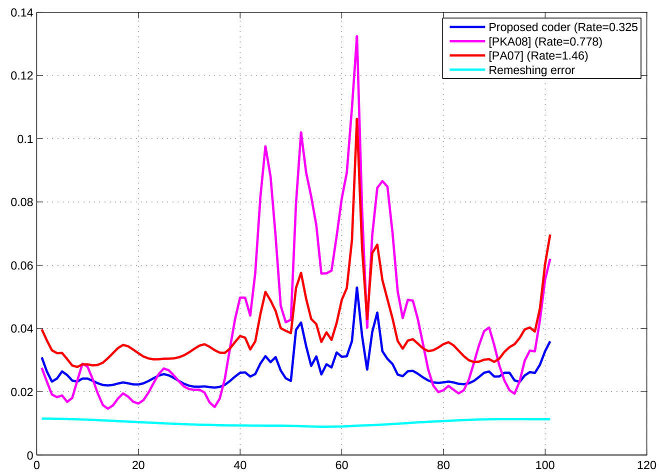


Fig. 8. RMSE-frame curve for the animated DOLPHIN.

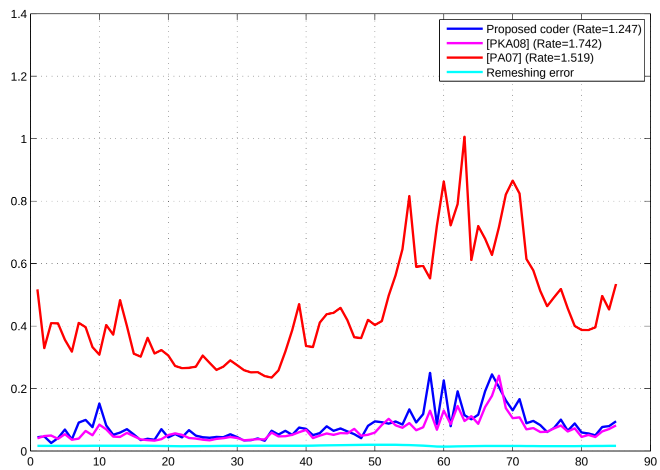


Fig. 9. RMSE-frame curve for the animated VENUS.

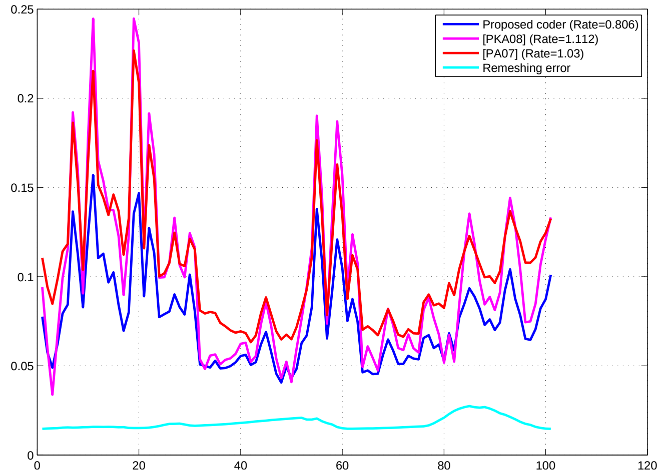


Fig. 10. RMSE-frame curve for the animated HORSE.

Acknowledgements

Original VENUS HEAD and HORSE models is courtesy of Cyberware. We are particularly grateful to J-H. Yang for providing us the original DOLPHIN model. The animated models DOLPHIN, VENUS and HORSE can be found at: http://www.i3s.unice.fr/˜fpayan/

References

1. Ibarria, L., Rossignac, J.: Dynapack: space-time compression of the 3D animations of triangle meshes with fixed connectivity. ACM Symp. Computer Animation (2003) 126–135
2. 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 **12**(12) (Dec. 2002) 1178–1184
3. Alexa, M., Muller, W.: Representing animations by principal components. Computer Graphics Forum 19, **3** (2000)
4. Karni, Z., Gotsman, C.: Compression of soft-body animation sequences. Computers and Graphics **28** (2004) 25–34
5. Vasa, L., Skala, V.: Coddyac: Connectivity driven dynamic mesh compression. In: 3DTV Conference 2007. (2007)
6. Boulfani, Y., Antonini, M., Payan, F.: Motion-based mesh clustering for mcdwt compression of 3d animated meshes. In: Proceedings of EUSIPCO'2007, Poland (September 2007)
7. Mamou, K., Zaharia, T., Prêteux, F., Kamoun, A., Payan, F., Antonini, M.: Two optimizations of the mpeg-4 famc standard for enhanced compression of animated 3d meshes. Proceedings of IEEE International Conference in Image Processing (September 2008)

8. Lopes, A., Gamito, M.: Wavelet compression and transmission of deformable surfaces over networks. In: Proceedings of the 10th Portuguese Computer Graphics Meeting. (2001) 107–114
9. Guskov, I., Khodakovsky, A.: Wavelet compression of parametrically coherent mesh sequences. In: Eurographics/ACM SIGGRAPH Symposium on Computer Animation. (August 2004)
10. Payan, F., Antonini, M.: Temporal wavelet-based geometry coder for 3d animations. Computer & Graphics **31** (january 2007) 77–88
11. Cho, J., Kim, M., Valette, S., Jung, H., Prost, R.: 3d dynamic mesh compression using wavelet-based multiresolution analysis. In: IEEE International Conference on Image Processing (ICIP'06). (October 2006)
12. Valette, S., Prost, R.: A wavelet-based progressive compression scheme for triangle meshes : Wavemesh. IEEE Transactions on Visualization and Computer Graphics **10**(2) (mars/avril 2004)
13. Khodakovsky, A., Schröder, P., Sweldens, W.: Progressive geometry compression. In Akeley, K., ed.: Siggraph 2000, Computer Graphics Proceedings, ACM Press / ACM SIGGRAPH / Addison Wesley Longman (2000) 271–278
14. Khodakovsky, A., Guskov, I.: Normal mesh compression. Geometric Modeling for Scientific Visualization, Springer-Verlag (2002)
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 **22** (July 2005) 466–486
16. Payan, F., Kamoun, A., Antonini, M.: Remeshing and spatio-temporal wavelet filtering for 3d animations. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Las Vegas, US (March-April 2008)
17. Yang, J., Kim, C., , Lee, S.: Semi-regular representation and progressive compression of 3d dynamic mesh sequences. IEEE Transactions on Image Processing **15**(9) (2006) 2531–2544
18. Touma, C., Gotsman, C.: Triangle mesh compression. Graphics Interface'98 (1998) 26–34
19. Sweldens, W.: The lifting scheme: A custom-design construction of biorthogonal wavelets. Applied and Computational Harmonic Analysis **3**(2) (1pril 1996) 186–200
20. Usevitch, B.: Optimal bit allocation for biorthogonal wavelet coding. In: DCC '96: Proceedings of the Conference on Data Compression, Washington, DC, USA, IEEE Computer Society (1996) 387
21. Aspert, N., Santa-Cruz, D., Ebrahimi, T.: Mesh: Measuring errors between surfaces using the hausdorff distance. In: IEEE International Conference in Multimedia and Expo (ICME). (2002)