Time-Varying Volume Geometry Compression with 4D Lifting Wavelet Transform

Yan Wang and Heba Hamza

NSF Center for e-Design, University of Central Florida, Orlando, FL 32816-2996, USA
{wangyan, hhamza}@mail.ucf.edu

Abstract. Geometry compression is an effective way to distribute high-volume geometry data within limited bandwidth and storage capacity. In this paper, a new time-varying 3D geometry compression method based on 4D lifted wavelet transform is presented. In this hybrid approach, geometric information and animation are compressed based volume grid values. Isosurfaces are reconstructed from decompressed grid values. With rescaling and integer-to-integer lifting, compression ratio is significantly increased without compromising quality of surfaces.

1 Introduction

There has been extensive research on 3D static geometry compression in the past decade [1,2]. In general, there are two approaches, mesh oriented and image oriented. In mesh oriented approach (e.g., [3]), both geometry (vertex coordinates in Euclidean space) and topology (connectivity among vertices) information is compressed. Accurate 3D meshes can be reconstructed, which is ideal for engineering applications. However, connectivity of mesh cannot be changed, which restrains it from general applications. In image oriented approaches (e.g., [4]), topology information is not considered. Volumetric geometry information is represented by voxels or points, and shapes can be compressed based on 2D image alike methods. The dynamics of topology can be captured easily, and the compression methods are general. However, to render a visually recognizable and appealing surface requires a large amount of data. A balanced approach considering these two ends will possibly introduce general solution with acceptable performance.

Visualization technologies tend to merge [5]. In the past three decades, digital signal processing has evolved from 1D to 3D data. However, only a few focus on dynamic geometry change over time. Similar to 2D video that complements 2D image, time-dependent 3D geometry can be looked as 3D video and has great potential in various applications, including communication, entertainment, scientific and medical computing, computer-aided design and engineering, as well as simulation and visualization. One can expect that 3D videos with compressed formats are standardized in the future and used as commonly as today's audio and video.

In general, there are two approaches in 3D animation compression. In the first approach, topology is assumed to be static, and there is no or small change in connectivity. Mesh compression is achieved by clustering and segmentation (e.g., Lengyel [6], Ahn et al. [7]), principal component analysis (e.g., Alexa and

M.-S. Kim and K. Shimada (Eds.): GMP 2006, LNCS 4077, pp. 670 – 676, 2006.

Müller [8], Sattler et al. [9]), motion predication (e.g., Ibarria and Rossignac [10], Karni and Gotsman [11]), parametric coding (e.g., Briceño et al. [12], Guskov and Khodakovsky [13]). In the second approach, topology may change arbitrarily between frames. Geometric data are encoded with voxel (e.g., Ma et al. [14]) and 3D wavelet coding (e.g., Guthe and Strasser [15], Sohn et al. [16]).

Different from the above methods, we propose a 3D animation compression scheme based on 4D lifted wavelet transform (LWT), which considers spatial and temporal coherence simultaneously. This method focuses on dynamic volume geometry compression with isosurface construction. Topology will not be coded as in mesh compression, as in many cases the connectivity exists purely for rendering purpose and it is not as essential as geometry information. Isosurface representation can reduce the data size for surface boundary reconstruction. With compressed geometry as well as related surface information such as normal directions and colors, it is sufficient to reconstruct surfaces for visualization with relatively low density volume information. In the rest of the paper, Section 2 gives the overview of the compression scheme. Sections 4 and 5 compare the floating point and integer lifting based on 4D LWT. Section 6 describes motion compensation issues in this scheme.

2 Overview of Proposed Compression Scheme

The volume data to be compressed is assumed to be regularly sampled, which is commonly used in scientific and medical visualization. Isosurface construction based on volume data can provide a realistic rendering and can also be applied in other visualization environments with artifacts and natural objects. The framework described here is based on the combination of volumetric data and isosurfaces, as illustrated in Fig. 1. 4D volumetric data is divided into groups of frames (GOFs) with no special requirement on GOF boundary. Each GOF is decomposed and compressed with 4D LWT at the server side. When received by client, the compressed data is decoded and decompressed. Isosurface is constructed based on the decompressed 4D volume data. The advantages of wavelet-based approaches include scalability and simultaneous redundancy reduction for spatial and time domains. With the inherent scalable representation, wavelet provides multi-resolution solution without extra costs. 4D LWT captures coherence locality of spatial and temporal domains. There is no need to divide intraframe data into blocks, as in MPEG-2 standards based on discrete cosine transform (DCT).

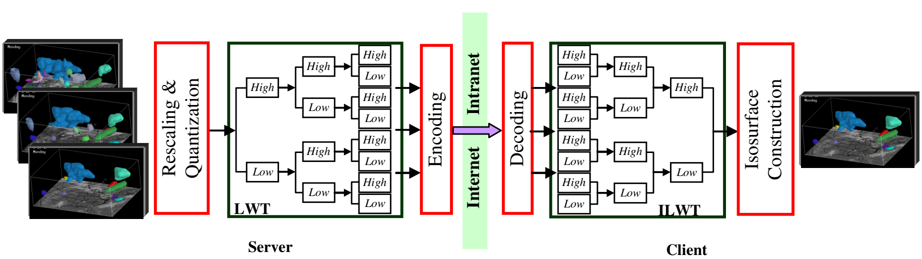


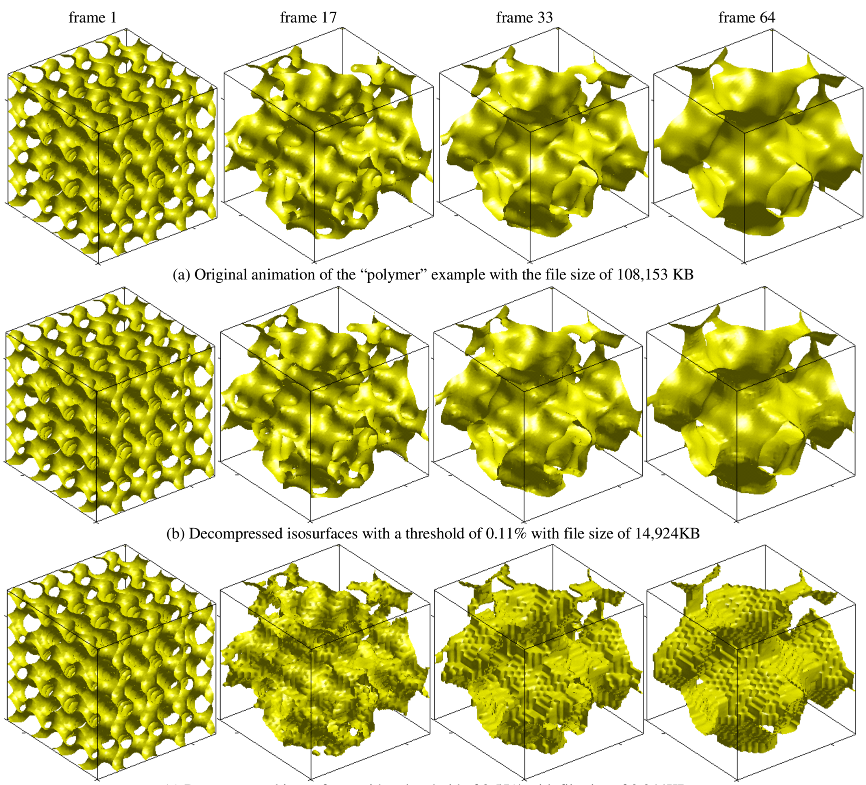
Fig. 1. Volume animation scheme based on 4D lifted wavelet transform

3 4D Lifted Wavelet Transform (LWT)

Lifting scheme [17] is also called second generation wavelet transform. It is a two-step filtering process: prediction and update. In the prediction step, even sequences are used to predict odd sequences. The prediction error forms the corresponding high-pass subband. In the update step, an approximation subband is obtained by updating even sequences with the scaled high-subband samples, generally described as

$$\begin{aligned} h_k[\mathbf{x}] &= f_{2k+1}[\mathbf{x}] + \sum_i p_i f_{2(k-i)}[\mathbf{x}] \\ l_k[\mathbf{x}] &= f_{2k}[\mathbf{x}] + \sum_j u_j h_{k-j}[\mathbf{x}] \end{aligned} \tag{1}$$

where $f_k[\mathbf{x}]$ is the input data to be processed at position $\mathbf{x}$ in frame k, h_k and l_k are resulting high-pass and low-pass sequences respectively, p_i and u_j are prediction and update coefficients of filters respectively. The main advantage of lifting is its memory efficiency in computation. Different from traditional traversal discrete wavelet transform (DWT), wavelet coefficient calculation in lifting can be embedded in-place.



(a) Original animation of the "polymer" example with the file size of 108,153 KB

(b) Decompressed isosurfaces with a threshold of 0.11% with file size of 14,924KB

(c) Decompressed isosurfaces with a threshold of 0.55% with file size of 3,944KB

Fig. 2. *Polymer* example with the grid size 64x64x64 in each frame

Another advantage is that backward transform is easy to find and has the same complexity as the forward transform.

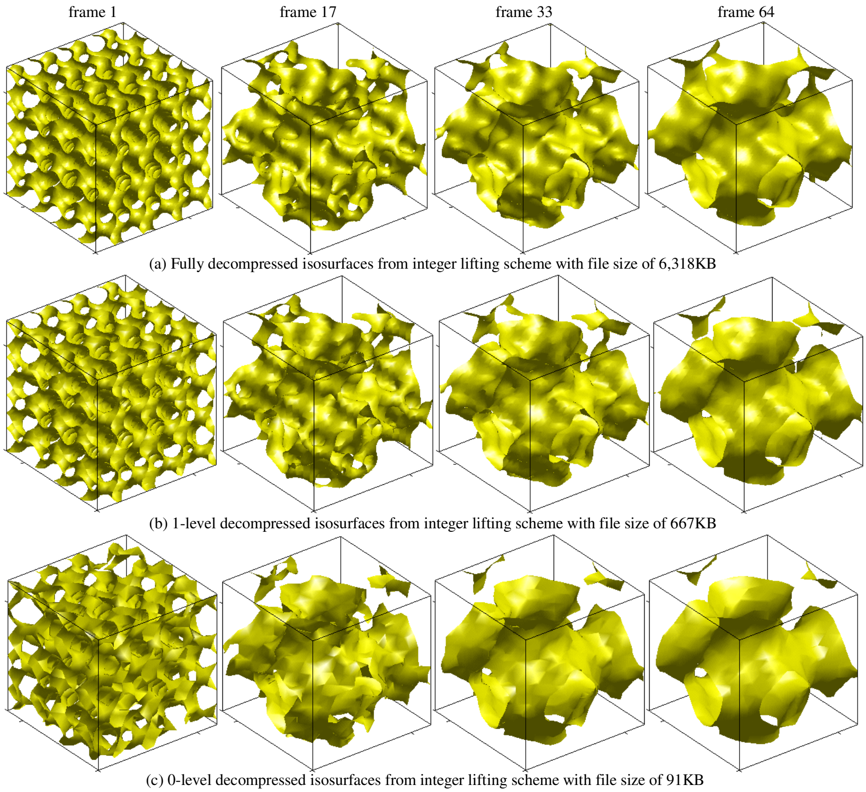
An important issue associated with spatio-temporal decomposition is the choice of filters. Different filters exhibit varied signal characteristics in terms of energy compactness in the transform domain and coding gain. Long filters tend to explore coherences of large regions or long period of time. However, they may blur boundary of occupation or movement. A dyadic decomposition approach is taken in this paper since it considers the coherence relation between time and space simultaneously.

An example is used to illustrate the visual effect. Fig. 2(a) shows four frames of the original 3D animation of polymer morphology, during which significant topology change is observed. A 2-level decomposition process is applied with Haar lifting scheme. Compression is achieved by setting the transformed coefficients to zero with the original values less than a threshold. Following the meta information including data type flags and dimensions of arrays, concatenated 4D arrays of coefficients form compressed files. No additional arithmetic coding is applied in the file size comparison. The highest coefficient magnitude in this example is M=89.944. If the threshold T is 0.11%, which is 0.11% of M, the decompressed surfaces are shown in Fig. 2(b) without significant visual difference. The classic marching cubes algorithm is applied in isosurface reconstruction. As T increases to 0.55%, the reconstructed surfaces are shown in Fig. 2(c). An interesting "voxel" effect occurs because the grid neighborhood with same isovalues expands. To avoid the voxel effect, resolution of grid needs to be reduced by either decreasing the density of grids at client side or taking advantage of inherent multi-resolution of wavelet transform thus transmitting only low resolution data from server side, which further increases compression ratio.

4 Rescaling and Integer-to-Integer Transform

Compared to classical wavelet transform, in which transformed wavelet coefficients are floating point numbers even if the original data are integers, lifting scheme supports lossless integer-to-integer transform [18]. It transforms integer data to integer coefficients. With inverse transform, original integer data can be reconstructed.

Another feature of the volume-based isosurface compression scheme is that isosurface construction is not sensitive to the number of bits used in coding if the range of grid values is large. As a result, floating-point grid values can be rounded to integers and integer-to-integer LWT can be used to increase the compression ratio without compromising quality. If the range of grid values is too small, the grid values can be rescaled before the rounding. A good rescaling strategy is to rescale the isosurface values to close to zero and the overall grid values to be evenly distributed between positive and negative sides. This reduces the number of bits to code values. In the example of Fig. 2, the maximum and minimum of grid values are ±7.8959. The grid values are multiplied by 100 before rounding. To reduce distortion, floor operation that rounds towards negative infinity is used. The result of integer lifting is depicted in Fig. 3, where the size of the file containing integer coefficients is 6,318KB compared to the floating coefficients of 61,900KB in Fig. 2(a) with full reconstruction. The quality of the surfaces is very close to the original ones.



(a) Fully decompressed isosurfaces from integer lifting scheme with file size of 6,318KB

(b) 1-level decompressed isosurfaces from integer lifting scheme with file size of 667KB

(c) 0-level decompressed isosurfaces from integer lifting scheme with file size of 91KB

Fig. 3. The *Polymer* example with integer-to-integer 4D LWT

5 Motion Compensation

Motion compensation (MC) is to remove temporal redundancy of video signal further in 4D subband coding. Block-based motion models are predominantly used in traditional motion-compensated MPEG coding. They can accurately represent very smooth motion fields but not complex ones. In contrast, deformable mesh motion model can improve motion compensation by tracking expansions and contractions while sustaining a continuous motion field. With the notation in (1),

$$
\begin{aligned}
h_k[\mathbf{x}] &= f_{2k+1}[\mathbf{x}] + \sum_i p_i f_{2(k-i)}[MC_{2k\to 2k+1}(\mathbf{x})] \\
l_k[\mathbf{x}] &= f_{2k}[\mathbf{x}] + \sum_j u_j h_{k-j}[MC_{2k+1\to 2k}(\mathbf{x})]
\end{aligned}
\qquad (2)
$$

is the MC scheme in spatial domain, where $MC_{i\to j}$ denotes motion-induced transformation from frame i to frame j. In traditional 2D video, MC increases signal-to-noise ratio thus video quality with the reduction of energy in high-pass temporal subbands. Yet it may introduce distortion because of motion inversion errors.

In the context of our approach, motion of surfaces is represented by value change of discrete grids, which is more resilient than direct pixel or voxel representation. We develop a Control Grid MC model for lifting transform to study the effectiveness on our 4D LWT scheme. In each frame, the 3D space is divided into small cells. The average value of within each cell is taken to be elements of motion vectors. For example, the Haar lifting with MC is shown in Fig. 4. Odd frames are predicted by even frames with motion vectors before the lifting process where high-pass temporal subbands are generated. Low-pass temporal subbands are generated with inverse motion vectors and lifting. Final high-pass and low-pass subbands are normalized.

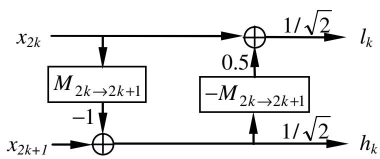


Fig. 4. MC in Haar lifting scheme

A simple "ball" example is used to test the Control Grid MC scheme. As shown in Fig. 5, shape distortion besides motion distortion is observed. As the cell size decreases, the distortion is reduced. The extreme case is that only one grid point is in each cell, which is indeed regular LWT. Compared to integer LWT, no compression gain is obtained with the motion compensation. The result indicates that block or grid based MC in our 3D video approach, in which grids are not as dense as pixel or voxel representation in 2D video, does not show significant advantages.

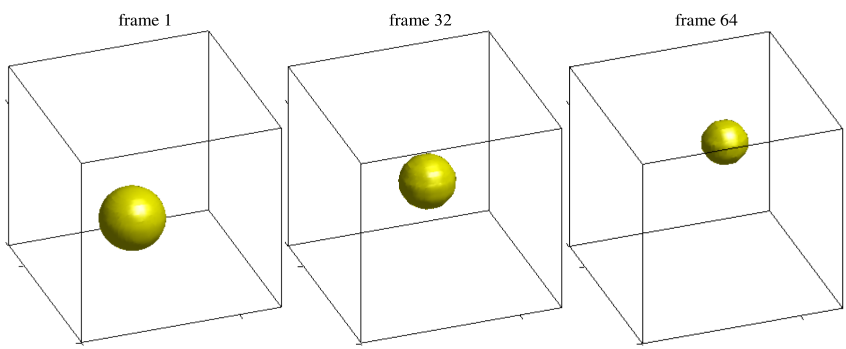


Fig. 5. Motion compensated *Ball* example shows distortion (MC cell size of 4x4x4, compressed file size is 6,238KB with motion vector included)

6 Concluding Remarks and Future Work

In this paper, a new time-varying 3D geometry compression scheme based on 4D lifted wavelet transform is presented. It is demonstrated that a hybrid approach with volume grid values and isosurfaces is feasible for 3D animation compression. Geometric information and animation are compressed based on volume grid values. Surfaces are reconstructed from grid values and isovalues. Rescaling and integer-to-integer LWT shows significant improvement on compression ratio without compromising quality of surfaces. Blocked based motion compensation appears unnecessary.

The proposed 4D geometry compression can be used in general applications such as scientific computing and visualization, collaborative engineering, modeling and simulation, teleconferencing, and entertainment. To enable application of 4D LWT to general 3D videos, future work may include volume data editing methods to support multiple isosurface construction, content-based motion compensation, and associated motion error, as well as filter selection and comparison.

References

1. Alliez, P. and Gotsman, C.: Recent advances in compression of 3D meshes. In: *Proc. Symp. on Mutiresolution in Geometric Modeling*. 2003
2. Peng, J., Kim, C.-S. and Kuo, C.-C. J.: Technologies for 3D mesh compression: A survey. *J. Visual Communication & Image Representation*. **16** (2005) 688-733
3. Taubin, G., Horn, W., Lazarus, F. and Rossignac, J.: Geometry coding and VRML. In: *Proc. the IEEE*. **96** (1998) 1228-1243
4. Muraki, S.: Volume data and wavelet transforms. *IEEE CG&A*. **13** (1993) 50-56
5. Nielson, G.M., Brunet, P., Gross, M., Hagen, H., Klimenko, S.V.: Research issues in data modeling for scientific visualization. *IEEE CG&A*. 14 (1994) 70-73
6. Lengyel, J.: Compression of time dependent geometry. In: *Proc. 1999 ACM Symp. on Interactive 3D Graphics*. (1999) 89-95
7. Ahn, J.-H., Kim, C.-S., Kuo, C.-C. J., and Ho, Y.-S.: Motion-compensated compression of 3D animation models. *IEE Elec. Lett.*, **37** (2001) 1445-1446
8. Alexa, M. and Müller, W.: Representing animations by principle components. *Computer Graphics Forum*. **19** (2000) 411-418
9. Sattler, M., Sarlette, R., and Klein, R.: Simple and efficient compression of animation sequences. In: *Proc. EUROGRAPHICS 2005*. (2005) 209-217
10. Ibarria, L. and Rossignac, J.: Dynapack: space-time compression of the 3D animations of triangle meshes with fixed connectivity. In: *Proc. EUROGRAPHICS 2003*. (2003) 126-135
11. Karni, Z. and Gotsman, C.: Compression of soft-body animation sequences. *Computers & Graphics*. **28** (2004) 25-34
12. Briceño, H.M., Sander, P.V., McMillan, L., Gortler, S., and Hoppe, H.: Geometry videos: A new representation for 3D animations. In: *Proc. EUROGRAPHICS 2003*. (2003) 136-146
13. Guskov, I. and Khodakovsky, A.: Wavelet compression of parametrically coherent mesh sequences. In: *Proc . EUROGRAPHICS 2004*. (2004) 183-192
14. Ma, K.-L., Smith, D., Shih, M.-Y., and Shen, H.-W.: Efficient encoding and rendering of time-varying volume data. *NASA/CR-1998-208424 ICASE Report No.98-22* (1998)
15. Guthe, S. and Strasser, W.: Real-time decompression and visualization of animated volume data. In: *Proc. IEEE Visualization 2001* (2001) 349-356
16. Sohn, B.-S., Bajaj, C., and Siddavanahalli, V.: Volumetric video compression for interactive playback. *Comp. Vis. & Image Understanding*. **96** (2004) 435-452
17. Sweldens, W.: The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. & Comp. Harmonic Analysis*. **3** (1996) 186-200
18. Calderbank, A.R., Daubechies, I., Sweldens, W., and Yeo, B.-L.: Lossless image compression using integer to integer wavelet transforms. In: *Proc. IEEE Int. Conf. Image Processing*. (1997) 596-599