3-D MESH COMPENSATED WAVELET LIFTING FOR 3-D+t MEDICAL CT DATA

Wolfgang Schnurrer, Thomas Richter, Jürgen Seiler, Christian Herglotz, and André Kaup

Multimedia Communications and Signal Processing
Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Cauerstr. 7, 91058 Erlangen, Germany
Email: {schnurrer, richter, seiler, herglotz, kaup}@lnt.de

ABSTRACT

For scalable coding, a high quality of the lowpass band of a wavelet transform is crucial when it is used as a downscaled version of the original signal. However, blur and motion can lead to disturbing artifacts. By incorporating feasible compensation methods directly into the wavelet transform, the quality of the lowpass band can be improved. The disadvantage in dynamic medical 3-D+t volumes from Computed Tomography is mainly given by expansion and compression of tissue over time and can be modeled well by mesh-based methods. We extend a 2-D mesh-based compensation method to three dimensions to obtain a volume compensation method that can additionally compensate deforming displacements in the third dimension. We show that a 3-D mesh can obtain a higher quality of the lowpass band by 0.28 dB with less than 40% of the model parameters of a comparable 2-D mesh. Results from lossless coding with JPEG 2000 3D and SPECK3D show that the compensated subbands using a 3-D mesh need about 6% less data compared to using a 2-D mesh.

Index Terms— Discrete Wavelet Transforms, Motion Compensation, Scalability, Computed Tomography, Signal Analysis

1. INTRODUCTION

Dynamical 3-D+t volumes can become very large in terms of storage so a scalable representation is desirable, e.g., for an efficient access and transmission. A downscaled version allows faster access, e.g., for browsing purposes. The lowpass band of a wavelet transform can be used as such a representation. By incorporating a compensation method directly into the wavelet transform, the transform can be adapted to the signal. A compensated wavelet transform in temporal direction is known as Motion Compensated Temporal Filtering (MCTF) [1]. The estimation of the grid point motion is a crucial task. We introduce a modification to the motion estimation for the 2-D mesh to avoid the degeneration of the mesh structure and thus obtain an improved inversion. Then, the extension to 3-D is presented. Simulation results are presented in Section 4.

Section 2 presents a brief review of the compensated wavelet lifting followed by a detailed description of the mesh compensation in Section 3.

2. COMPENSATED WAVELET LIFTING

Fig. 1 shows the lifting structure of the Haar wavelet transform extended by a compensation method. The transform is applied in temporal direction. The volumes \( f_t \) are indexed by the time step \( t \) in temporal direction. The lifting structure consists of a prediction step and an update step. The highpass coefficients \( H_{\text{P}} \) are computed in the prediction step according to

\[
H_{\text{P}} = f_{2t} - [W_{2t-1 \rightarrow 2t}(f_{2t-1})].
\]
Instead of the original volume $f_{2t-1}$, a predictor is subtracted from $f_{2t}$, denoted by the warping operator $W_{2t-1 ightarrow 2t}$ \[1\]. The compensation is called MC in Fig. 1. In the update step, the lowpass coefficients $LP_t$ are computed by

$$LP_t = f_{2t-1} + \frac{1}{2} W_{2t-2t-1}(HP_t). \tag{2}$$

As the index of $W$ in \[2\] shows, the compensation has to be inverted in the update step to achieve an equivalent wavelet transform. The inversion is denoted by IMC in Fig. 1. For a compensated wavelet transform, the mesh-based compensation method has the advantage that it is invertible \[6\]. To avoid rounding errors, rounding operations are applied to the fractional parts \[5, 7\]. The reconstruction of the original volume without any loss is very important, e.g., for medical image data.

3. MESH-BASED COMPENSATION

In 2-D, a mesh-based compensation is computed by putting a mesh over the reference image. The predictor is then computed by warping the reference image according to the deformation given by the motion vectors of the grid points \[5\]. A quadrilateral mesh topology leads to a bilinear transform \[8\] of the patches of the underlying image. Compared to an affine transform of a triangle mesh, the mesh-based compensation method has the advantage that it is invertible \[6\]. To avoid rounding errors, rounding operations are applied to the fractional parts \[5, 7\]. The reconstruction of the original volume without any loss is very important, e.g., for medical image data.

3.1. Motion Estimation in 2-D

The estimation of the grid point (GP) motion is a crucial task. Fig. 2(a) shows a detail of a quadrilateral mesh. The motion vectors between the GPs are computed from an interpolation. Thus, the choice of the motion vector of $P$ in the center is influenced by its eight surrounding neighbors. A modification of $P$ influences all surrounding neighbors as well. An optimum solution is a combination of all motion vectors of all GPs which is computationally too complex. Several iterative methods exist and we use an iterative refinement process as proposed in \[8, 10\]. We initialize the motion vectors with zeros. In every iteration, the current motion vector of each GP is updated in every direction, the corresponding warping of the underlying image is computed and the update with the smallest error metric is chosen. The nine update positions for $P$ are illustrated in Fig. 2(a). Independent GPs can be refined in parallel \[8\]. In order to avoid a degeneration of the mesh structure, the movement of the GPs can be limited to a specific search range \[11\]. This search range is limited by the direct neighbors, labeled by $A$, $B$, $C$, and $D$. It is marked light gray in Fig. 2(a). This prevents concave quadrilateral structures in the mesh that make a proper inversion very complicated or even impossible \[12\]. We introduce a further safety boundary $d$ to prevent the degradation to triangular structures. The resulting search area is illustrated in Fig. 2(b) in dark gray. Movements of $P$ have to stay inside the dark gray area. With this extension, a convex mesh structure is maintained. This is advantageous for the inversion of the compensation in the update step.

To avoid the occurrence of unconnected pixels \[13, 14, 15\], the movement of the boundary GPs of the complete mesh is further limited. The shape of the outer hull shall remain, so the four corner GPs are kept fix, i.e., no motion is allowed. The remaining GPs on the left and right boundary edge can only move up and down while the GPs on the top and bottom boundary edge can only move left and right. The 2-D mesh corresponds to a subsampling of a 2-D motion vector field where the motion vectors of the intermediate positions are obtained by bilinear interpolation.

3.2. Proposed Extension to 3-D

The mesh is extended by one dimension to obtain a 3-D compensation method. This corresponds to a further subsampling of the motion vector field in the third dimension. For the 3-D mesh, a motion vector of a GP has three components. The motion vectors of the intermediate positions are obtained by trilinear interpolation. A detail of a 3-D mesh is illustrated in Fig. 2(c). To estimate the motion vectors of the GPs, a similar method as for the 2-D mesh is applied. The refinement for the current GP is tested in the third dimension as well, resulting in 27 test positions. In this way, deformation in the third dimension is compensated. For every update position, the warping of the corresponding image cube is computed. The update for the motion vector with the smallest error metric is chosen.

Again, to avoid a degeneration of the mesh structure, the update positions are limited to a specific search area. In the 3-D case, this search area results in an octahedron marked by bold black lines in Fig. 2(c). For illustration, the octahedron of the search area in Fig. 2(c) is again shown in Fig. 2(d) in a different perspective. We add a safety boundary $d$ for the 3-D case as well. This is shown exemplarily for one plane of the octahedron. $P$ has to remain in the smaller dashed octahedron to maintain a proper mesh structure.

The movement of the GPs on the boundary is limited similar to the 2-D case to maintain the shape of the outer hull of the mesh. For instance, the GPs on the front and the back boundary face can only move up, down, left, and right.
is necessary to achieve better results than the 2-D mesh. On the one grid size in simulated compensation methods. For a compensated wavelet transformation method within the wavelet transform. The dashed green curve shows the result using a 2-D mesh. The comparison with the results from the 2-D mesh and the 3-D mesh shows again, that for a grid size of 4 in z-direction, the 3-D mesh yields better results. After 50 iterations, the 3-D mesh leads to a PSNR gain of 0.15 dB compared to the 2-D mesh.

Tab. 1 summarizes the results for the wavelet transform of all time steps of the 3-D+t volume. The results are given for 50 iterations, however, Fig. 3 and 4 show that a convergence of the refinement of the mesh motion is reached after about 20 iterations. The first column lists the compensation methods used for the compensated wavelet transform applied in temporal direction of the 3-D+t volume. The second column lists the number of parameters needed for the model of the compensation method. Although a motion vector of the 3-D mesh has three components instead of two, the overall number is smaller than for the 2-D mesh due to the subsampling in z-direction. For the 3-D mesh with grid size 16 × 16 × 4, less than 40% of the motion parameters are needed compared to the 2-D mesh with grid size 16 × 16. We assume that this leads to less side information for the 3-D case but we do not consider coding of the motion information in the following. The column MSE HP shows the mean energy of the highpass band and the column PSNR LP shows the quality of the lowpass band w.r.t. the corresponding reference volume f_{2t−1} after the wavelet transform. The overall results are consistent with the discussion of Fig. 3 and 4.

To evaluate the compressibility, the resulting subbands were coded losslessly using the standard wavelet coefficient coder partitioning embedded block (SPECK3D) [16]. We used the implementation available for the QccPack library [17]. For comparison, the volumes were also coded using the wavelet-based volume coder JPEG 2000 3D [18]. We used the OpenJPEG [19] implementation. For both subband volumes HP_{1} and LP_{1}, further 2 wavelet decomposition steps in slice direction and 5 wavelet decomposition steps in xy-direction were applied. For SPECK3D we applied the LeGall 5/3 wavelet as used in JPEG 2000 3D.

3 The CT volume data set was kindly provided by Siemens Healthcare.

4. SIMULATION RESULTS

For evaluating our compensation methods, we used a cardiac 3-D+t CT data set. This multidimensional 3-D+t volume has 10 time steps, 128 slices in z-direction and a resolution of 512 × 512 pixels in xy-direction at 12 bit per sample and shows a beating heart over time.

The safety boundary is chosen to d = 1 pixel. For the 2-D mesh, we used a grid size of 16 × 16 pixels. This size has proven to be reasonable in [5]. This corresponds to a subsampling of the motion vector field by 1 : 16^2 = 1 : 256 in xy-direction. This grid size is used for the 3-D mesh as well in xy-direction. For the grid size in z-direction, 16, 8 and 4 pixels were used to test different subsampling factors in z-direction.

Fig. 3 shows the mean energy of the highpass band HP_{1} for the simulated compensation methods. For a compensated wavelet transform, the highpass band can be regarded as the prediction error. The mean energy is plotted against the refinement iterations of the mesh-based compensation methods. The dotted blue curve results from a traditional wavelet transform when no compensation is used and thus it is independent of the iteration. The plot shows that the energy in the highpass band can be reduced significantly by using a compensation method within the wavelet transform. The dashed green curve shows the result using a 2-D mesh. The comparison with the results from the 3-D mesh method shows that a grid size of 4 in z-direction is necessary to achieve better results than the 2-D mesh. On the one hand, using a 3-D mesh has the advantage that a displacement in z-direction can be compensated. On the other hand, this leads to a subsampling of the motion information in z-direction.

Fig. 4 shows the resulting quality of the lowpass band in terms of PSNR (LP_{1}, f_{1}) in dB. Again, the metric is plotted against the iterations. The safety boundary avoids the degeneration of the mesh structure. Although an approximation is used for the inversion of the mesh-based methods, the quality of the lowpass band is increased by nearly 5 dB compared to the traditional transform. Further, the PSNR increases with a larger number of iterations. Comparing the results from the 2-D mesh and the 3-D mesh shows again, that for a grid size of 4 in z-direction, the 3-D mesh yields better results. After 50 iterations, the 3-D mesh leads to a PSNR gain of 0.15 dB compared to the 2-D mesh.

4.3. Inversion for the Update Step

In the update step of the compensated lifting [2], the compensation from the prediction step [1] has to be inverted to obtain an equivalent wavelet transform [5]. The mesh-based approach has the advantage that it is invertible [6]. We use an approximation from [6] for the inversion of the mesh-based compensation, thus accepting the induced error. Instead of calculating the inversion of the mesh warping, we take the negative values of the motion vectors at the GPs.

Fig. 3. The mean squared energy of the highpass band HP_{1} for no compensation, 2-D mesh and 3-D mesh is plotted against the refinement iteration used for the mesh-based compensation methods.

Fig. 4. The quality of the lowpass band in terms of PSNR (LP_{1}, f_{1}) in dB for no compensation, 2-D mesh and 3-D mesh is plotted against the refinement iterations used for the mesh-based compensation methods.
Table 1. The table lists summarized results for the considered compensation methods. \#mv params denotes the number of parameters needed for the model of the compensation method. The last row provides a delta between the 2-D mesh $16 \times 16$ and the 3-D mesh $16 \times 16 \times 4$.

| Compensation method       | \#mv params | Mean energy HP [dB] | Mean energy LP [dB] | SPECK3D [MByte] | JPEG 2000 3D [MByte] |
|---------------------------|-------------|---------------------|---------------------|-----------------|----------------------|
| none                      | 0           | 2976.2              | 44.30               | 80.3            | 82.9                 | 163.2               | 82.5            | 85.4 | 167.9 |
| 2-D mesh, $16 \times 16$  | 261 888     | 1280.5              | 47.95               | 91.2            | 89.4                 | 180.6               | 93.9            | 92.1 | 186.0 |
| 3-D mesh, $16 \times 16 \times 16$ | 26 037 | 1698.1              | 46.81               | 83.4            | 83.7                 | 167.1               | 85.8            | 86.3 | 172.1 |
| 3-D mesh, $16 \times 16 \times 8$ | 51 117 | 1422.7              | 47.58               | 84.0            | 84.1                 | 168.1               | 86.4            | 86.7 | 173.1 |
| 3-D mesh, $16 \times 16 \times 4$ | 101 277 | 1232.8              | 48.23               | 84.7            | 84.8                 | 169.5               | 87.1            | 87.5 | 174.6 |
| $\Delta$: 2-D $16 \times 16$ to 3-D $16 \times 16 \times 4$ | -61% | -47.7 | +0.28 | -7.1% | -5.1% | -6.1% | -7.2% | -5.0% | -6.1% |

Fig. 5. The first column shows temporal subsequent images of the original volume at one slice position. The white areas mark the origin of the details depicted in the remaining columns. Details from the highpass bands (HP, gray=0) are shown in the top row and details from the lowpass bands (LP) are shown in the bottom row. The respective compensation methods are listed between.

The columns entitled by SPECK3D and JPEG 2000 3D show the respective results in MByte from lossless coding of the lowpass band, the highpass band, and the sum of both. Using SPECK3D, the results are a little better, but the overall tendency is the same for both methods. When the quality of the lowpass image is of interest, no compensation method should be used when lossless coding of a CT volume is considered because the data contains a lot of correlated noisy structures that can be exploited by the traditional wavelet transform without a compensation method. In contrast to that, if a compensated wavelet transform is applied, it is not possible to exploit the structures of the noise any more.

When the lowpass band is used as a scalable representation, the quality is important. In this case, a compensation method can increase the quality, as shown by the 4th column of Tab. 1. The coding results for the compensated transform suggest that for a 3-D compensation it is better to use a 3-D mesh-based method than a 2-D mesh-based method. This might result from the fact that the 3-D mesh compensation is not computed independently slice-by-slice as for the 2-D mesh and thus leads to a smoother prediction.

The visual examples depicted in Fig. 5 support the results from Tab. 1. The lowpass band without compensation is blurred. The 3-D mesh can exploit motion in the z-direction over time and thus is able to further reduce the energy in the highpass band resulting in a sharper lowpass with less artifacts.

5. CONCLUSION

In this paper, we investigated mesh-based compensation methods for a compensated wavelet transform of medical 3-D+t volumes that contain deforming displacements over time. The proposed 3-D mesh compensation method is able to provide a prediction for a complete volume and thus is able to exploit deforming displacements in the third dimension as well. This is a huge advantage compared to a slice-wise applied 2-D mesh compensation as proposed in the literature.

The 3-D mesh-based compensation yields better compressible subbands. Within our simulation data, the filesize of the lossless coded subbands could be reduced by 6% using a 3-D mesh instead of a 2-D mesh. Further work aims at the investigation of a proper inversion of the mesh-based compensation and a compensated transform in z-direction within this framework.

Acknowledgment

We gratefully acknowledge that this work has been supported by the Deutsche Forschungsgemeinschaft (DFG) under contract number KA 926/4-2.
6. REFERENCES

[1] J.U. Garbas, B. Pesquet-Popescu, and A. Kaup, “Methods and Tools for Wavelet-Based Scalable Multiview Video Coding,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 2, pp. 113–126, Feb. 2011.

[2] V. Sanchez, R. Abugharbieh, and P. Nasiopoulos, “3-D Scalable Medical Image Compression With Optimized Volume of Interest Coding,” *IEEE Transactions on Medical Imaging*, vol. 29, no. 10, pp. 1808–1820, Oct. 2010.

[3] A. Cavallaro, F. Graf, H.-P. Kriegel, M. Schubert, and M. Thoma, “Region of Interest Queries in CT Scans,” in *Advances in Spatial and Temporal Databases*, pp. 56–73. Springer, 2011.

[4] W. Schnurrer, J. Seiler, and A. Kaup, “Analysis of Displacement Compensation Methods for Wavelet Lifting of Medical 3-D Thorax CT Volume Data,” in *Proceedings Visual Communications and Image Processing (VCIP)*, San Diego, CA, USA, Nov. 2012, pp. 1–6.

[5] W. Schnurrer, T. Richter, J. Seiler, and A. Kaup, “Analysis of Mesh-Based Motion Compensation in Wavelet Lifting of Dynamical 3-D+t CT Data,” in *Proceedings IEEE International Workshop on Multimedia Signal Processing (MMSP)*, Banff, Canada, Sept. 2012, pp. 152–157.

[6] A. Secker and D. Taubman, “Highly Scalable Video Compression Using a Lifting-Based 3D Wavelet Transform with Deformable Mesh Motion Compensation,” in *Proceedings International Conference on Image Processing (ICIP)*, Rochester, NY, USA, June 2002, vol. 3, pp. 749–752.

[7] A.R. Calderbank, I. Daubechies, W. Sweldens, and B.-L. Yeo, “Lossless Image Compression Using Integer to Integer Wavelet Transforms,” in *Proceedings International Conference on Image Processing (ICIP)*, Washington, DC, USA, Oct. 1997, pp. 596–599.

[8] Y. Nakaya and H. Harashima, “Motion Compensation Based on Spatial Transformations,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 4, no. 3, pp. 339–356, June 1994.

[9] A. Weinlich, P. Amon, A. Hutter, and A. Kaup, “Representation of Deformable Motion for Dynamic Cardiac Image Data Compression,” in *Proceedings SPIE Medical Imaging*, San Diego, CA, USA, Feb. 2012.

[10] G.J. Sullivan and R.L. Baker, “Motion Compensation for Video Compression Using Control Grid Interpolation,” in *Proceedings IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Toronto, Canada, Apr. 1991, pp. 2713–2716.

[11] J. Wang, M. Yu, Y. Xia, J. Chen, and Z. Xia, “A Novel Method of DT Mesh Motion Estimation and Compensation based on RDWT,” in *Proceedings IEEE International Conference on Communication Technology (ICCT)*, Chengdu, China, Nov. 2012, pp. 1291–1295.

[12] Y. Wang and O. Lee, “Use of two-dimensional deformable mesh structures for video coding. Part I - The synthesis problem: mesh-based function approximation and mapping,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 6, no. 6, pp. 636–646, Dec. 1996.

[13] J.-R. Ohm, M. Schaar, and J.W. Woods, “Interframe Wavelet Coding - Motion Picture Representation for Universal Scalability,” *Signal Processing: Image Communication*, vol. 19, no. 9, pp. 877–908, 2004.

[14] B. Girod and S. Han, “Optimum Update for Motion-Compensated Lifting,” *IEEE Signal Processing Letters*, vol. 12, no. 2, pp. 150–153, Feb. 2005.

[15] W. Schnurrer, J. Seiler, and A. Kaup, “Improving Block-Based Compensated Wavelet Lifting by Reconstructing Unconnected Pixels,” in *Proceedings International Symposium on Signals, Circuits and Systems (ISSCS)*, Iasi, Romania, July 2013, pp. 1–4.

[16] W. A. Pearlman and A. Islam, “Embedded and Efficient Low-complexity Hierarchical Image Coder and Corresponding Methods Therefor,” Dec. 2003.

[17] J.E. Fowler, “QccPack: An Open-Source Software Library for Quantization, Compression, and Coding,” in *Proceedings Applications of Digital Image Processing XXIII*, San Diego, CA, USA, Aug. 2000, vol. 4115, pp. 294–301.

[18] P. Schelkens, A. Munteanu, A. Tzannes, and C. Brislawn, “JPEG2000. Part 10. Volumetric data encoding,” in *Proceedings IEEE International Symposium on Circuits and Systems (ISCAS)*, Island of Kos, Greece, May 2006, pp. 3874–3877.

[19] A. Descampe, F. Devaux, H. Drolon, D. Janssens, and Y. Verschueren, “OpenJPEG 2.0.0,” http://www.openjpeg.org, Nov. 2012.