Salih Burak Gokturk Anne Margot F. Aaron

APPLYING 3D METHODS TO VIDEO FOR COMPRESSION

EE 392J Final Report

Abstract

Video compression has been an active research area in the last decade. Amongst many proposed methods, motion compensated coding has taken the most attention and taken its place in many standards (Mpeg, H.26L, etc…) The main difference between motion compensated coding and image based coding schemes such as DCT is that motion compensated coding exploits temporal redundancies between consecutive frames, i.e., between two frames. In this paper, we investigate 3D methods that exploit the temporal redundancies on more than two frames. We first describe and evaluate transform based methods such as 3D DCT, and PCA. Next, we propose a new method called volume based motion compensated coding. Volume based motion compensated coding is an extension of regular motion compensated coding and works on stacks of images as opposed to images. Our results indicate that, 3D methods can provide better results than motion compensated coding in many of the video sequences.

1. Introduction

Motion compensated coding has been the most widely used video compression method amongst other compression schemes. The main difference of motion compensated coding from other transform and prediction based methods is its utilization of the motion characteristics in the images. While most other compression algorithms achieve compression by exploiting the redundancies only in spatial dimensions (in images), motion compensated coding exploits redundancies between two consecutive frames as well. In turn, motion compensated coding provides more efficient compression.

In this paper, we consider the video as a 3D signal and investigate ways of exploiting redundancies on more than two consecutive frames. Figure 1 gives an illustration of the redundancies on the temporal domain. Figure 1a gives a frame from the salesman sequence. Figure 1b and 1c shows the variation of row 80 and column 100 respectively over time. We can easily see that the variation of the 3D video signal is much less in the temporal dimension than the spatial dimensions. It is possible to achieve better compression by exploiting the redundancies in the temporal domain.

(a) (b) (c)

Figure 1. (a) An image from the salesman sequence. (b) Variation of a row over time. (c) Variation of the column over time.

For this purpose, we looked into various 3D methods. Among these are 3-D DCT, PCA and wavelet transformation. 3-D based methods work better than motion compensated coding especially when there is a smooth signal where most of the signal is concentrated in low frequencies. Next, we propose a new method called volume based motion compensated coding. This method is similar to regular motion compensated coding except that it works on stacks of images in stead of images directly.

The paper continues as follows: Section 2 describes previous work on 3D compression methods. Section III describes the 3D compression methods including our proposed method. Section IV describes our experiments with many sequences. Finally, Section V gives our conclusions and discusses possible future work.

2. Previous Work

3D transform based compression has been applied by some researchers. The 3D DCT has been proposed for both image [1,2] and video compression. Authors have argued that 3D DCT can be effective in compressing video sequences, especially those with little or no motion [3,4,5].

3D Wavelet transforms have also been investigated as a method for video compression [6-14]. Authors have shown that 3D wavelet transforms, using no motion compensation, provide better compression performance than current block-based motion compensated predictive methods. 3D wavelet decomposition has been combined with EZW [15] or SPHIT [16] coding to achieve good quality compression. A significant advantage of 3D wavelet transforms or 3D subband coding over other tranform methods is that the resulting encoded video is highly scalable, both in the spatial and temporal domains. The scalability and multiresolution properties of wavelet transforms have been utilized in previous work in designing scalable video codecs.

Although they have not yet been applied to video, vector quantization and principal component analysis (PCA) have been proposed in [17,18,19] for compression of images. In [18], PCA has been proposed as a better image processing approach compared to DCT since it exploits the intensity characteristics of data better.

In this paper, we would like to combine the results of previous 3D based methods and combine one of them, *i.e.* DCT, by a 3-D version of motion compensated coding. This way, not only do we exploit the motion characteristics better, but we also take advantage of the effectiveness of the 3-D methods.

3. Methods

In this section we describe the 3D compression methods we tried. 3D DCT is explained in Section 3.1. We explain how principal component analysis can be used for 3D data encoding in Section 3.2. We explain the wavelet transformation in Section 3.3. Finally, we describe our proposed method, *i.e.*, volume-based motion compensated coding in Section 3.4.

3.1. 3D Discrete Cosine Transformation

The Discrete Cosine Transform (DCT) is a real-valued, separable orthonormal transform whose basis vectors are composed of samples of cosine functions. The 3D DCT analysis is defined as follows:

Let $\boldsymbol{X}$ be a 3D signal of size M by N by T. Let $\boldsymbol{Y}$ be the 3D DCT of $\boldsymbol{X}$, also of size M by N by T. The elements of $\boldsymbol{Y}$ can be calculated as

$$Y_{rpq} = \alpha_r \alpha_p \alpha_q \sum_{t=0}^{T-1} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{tmn} \cos\frac{\pi(2t+1)r}{2T} \cos\frac{\pi(2m+1)p}{2M} \cos\frac{\pi(2n+1)q}{N} \qquad (1)$$

$$\alpha_r = \begin{cases} 1/\sqrt{T}, r = 0 \\ 2/\sqrt{T}, else \end{cases} \qquad \alpha_p = \begin{cases} 1/\sqrt{M}, p = 0 \\ 2/\sqrt{M}, else \end{cases} \qquad \alpha_q = \begin{cases} 1/\sqrt{N}, q = 0 \\ 2/\sqrt{N}, else \end{cases}$$

Since the DCT is separable, this 3D transform is the same as finding the 1 dimensional DCT along each of the three dimensions of $\boldsymbol{X}$.

The encoder block diagram of the 3D DCT compression algorithm is in Figure 2. The steps are summarized below.

1. Construct volumetric images by combining N frames into a stack. Here N is the depth of the constructed stack.
2. On each volumetric image:
   a. Divide each volumetric image into NxNxN volumetric blocks. (N is 8 in our experiments.)
   b. For each volume block, calculate the 3D DCT and quantize the coefficients.

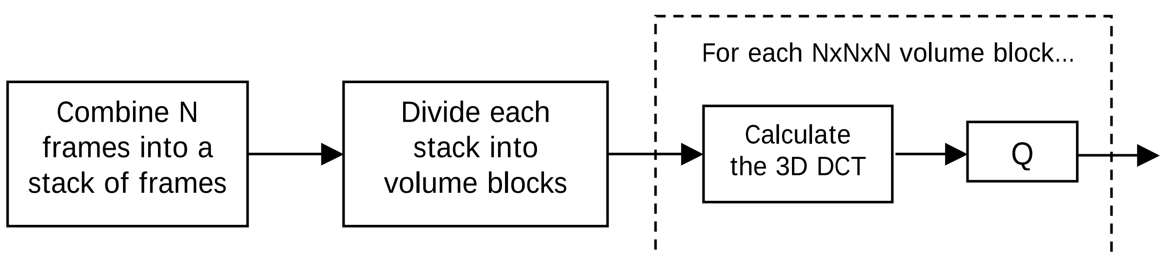


Figure 2. Block diagram for the encoder of the 3D DCT method.

When the 3D-DCT is applied to video, the transformation is also applied to the time dimension. It is therefore expected that the 3D-DCT will be more efficient than frame based 2D-DCT when there is correlation between frames.

3.2. 3D Principal Component Analysis

Principal Component Analysis, or singular value decomposition, is a way to make a low-rank estimation of data where the residual is minimized in the least squares manner. In this section, we describe the application of principal component analysis to representation of 3D intensity data (video). First, we have a training stage where the principal components are determined. The principal components are determined before the video is coded and sent to the decoder side. Once the principal components are determined, the 3D images are constructed in terms of them. This works as a low-rank prediction for the video signal as shown in the block diagram in Figure 3. The residual between the prediction and real video data is coded by a 3D DCT, and the quantized PCA and DCT coefficients are sent to the decoder.

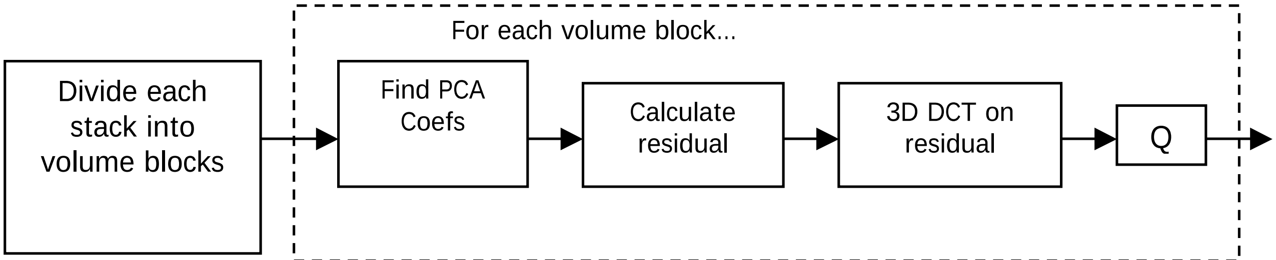


Figure 3. Block diagram of the encoder for the PCA method.

Similar to DCT, PCA works on stacks of images. Stacks of depth n are first divided into blocks of size nxnxn. (n is 8 in our experiments) As demonstrated on Figure 4., each of nxnxn block intensity values are listed into an nxnxn vector, called block vector. Each block vector is then listed into the block matrix B. This matrix is first normalized by subtracting the average block vector form every column and the normalized $\tilde{B}$ is constructed. Finally a singular value decomposition is performed on $\tilde{B}$. The left singular vectors of $\tilde{B}$ represent the orthogonal directions in nxnxn dimensional space where main intensity variations inside the blocks occur. We will refer to these directions as the principal components, P_i.

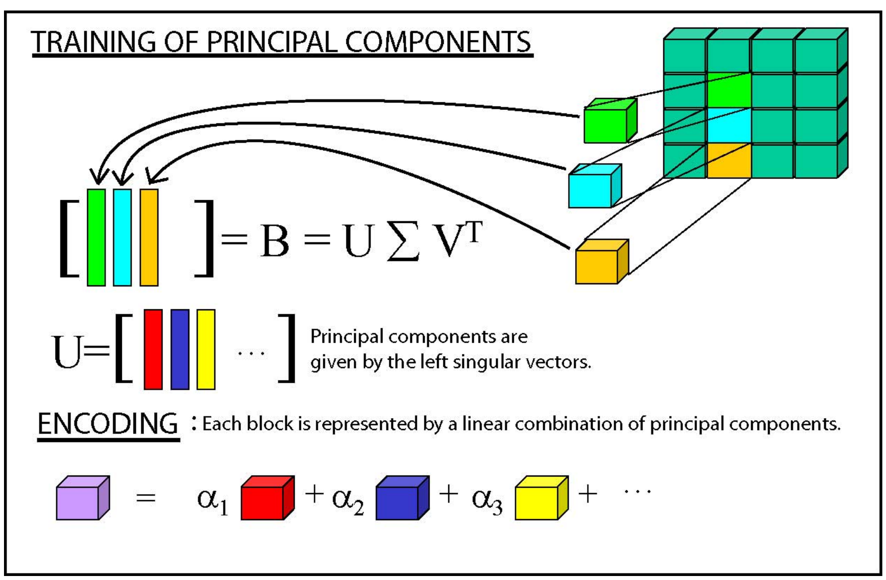


Figure 4. Illustration of the application of PCA on the 3D Data

Once the principal components are determined, each block is represented as a linear combination of the principal components:

$$\hat{X} = \sum_{i=1}^{k} \alpha_i P_i \text{ and } \alpha_i =< X, P_i >$$

where X is a block vector, $\hat{X}$ is a prediction for X, α_i is the PCA coefficient, <> is the inner vector product operation, and k the total number of used principal components. In our experiments the value of k was varied between 4-10. The PCA coefficients α_i are quantized and encoded. The residual between $\hat{X}$ and X is coded by 3D DCT. The DCT coefficients are quantized and encoded as well. In our experiments, we give the theoretical entropy values for the quantized coefficients.

3.3. 3D Discrete Wavelet Transformation

The Discrete Wavelet Transform (DWT) is a separable, dyadic tree-structured subband transform. The subband decomposition is performed by recursively passing the signal into a two-filter channel bank, where the successive decompositions are only done on the lowest subband. Since the 3D DWT is separable, a single step in the decomposition is composed of passing each dimension through the filter bank producing eight subbands per level. The block diagram of a single level 3D DWT is shown in Figure 5.

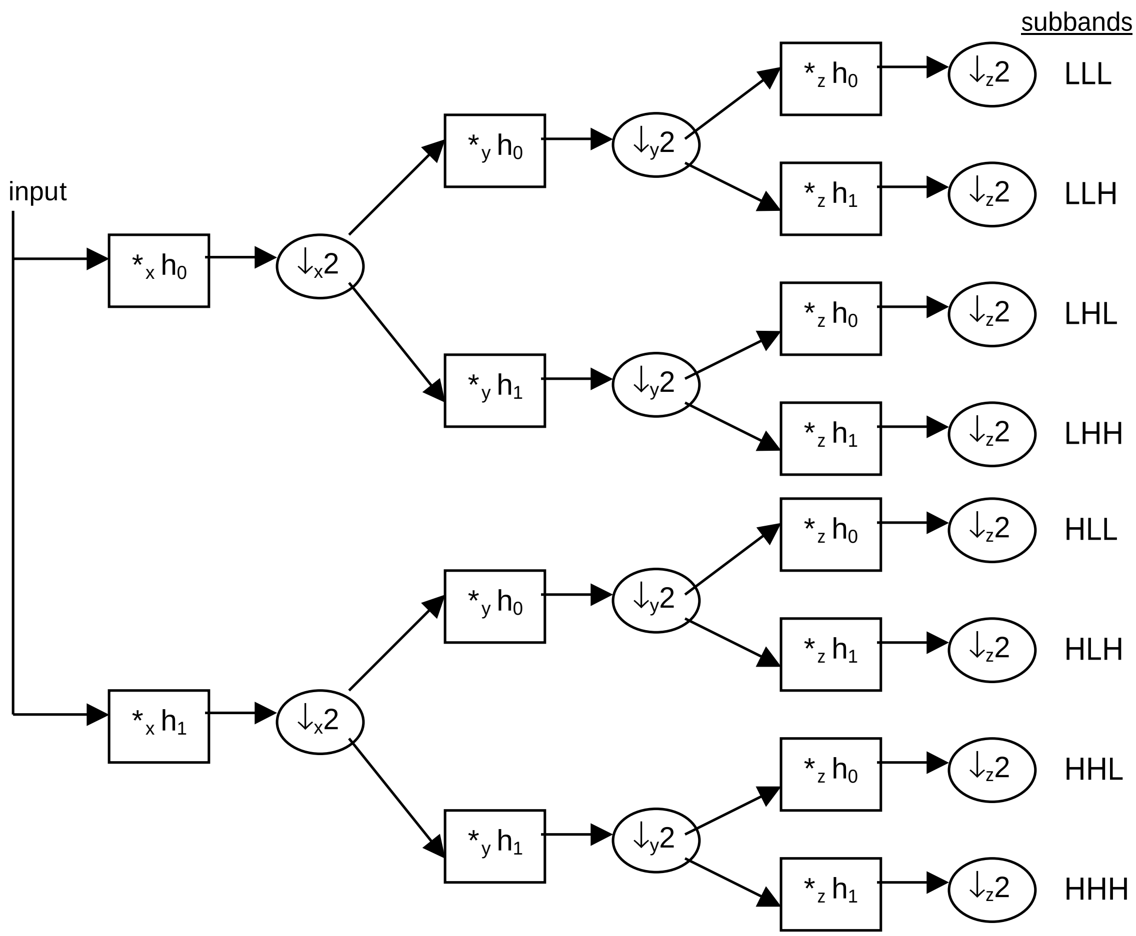


Figure 5. Block diagram of a single level decomposition for the 3D DWT

Like the 3D DCT and 3D PCA, the 3D DWT is applied to video by first dividing the sequence into stacks of N frames. Each stack is then divided into NxNxN volume blocks. The difference of this method from the previous two is that N must be a large number in order to take advantage of the multilevel decomposition. DWT recursions are then applied to each NxNxN volume block. Finally, for efficient compression of the wavelet coefficients, the EZW or the SPIHT algorithm can be applied. A block diagram of the proposed DWT encoding of video is shown in Figure 6.

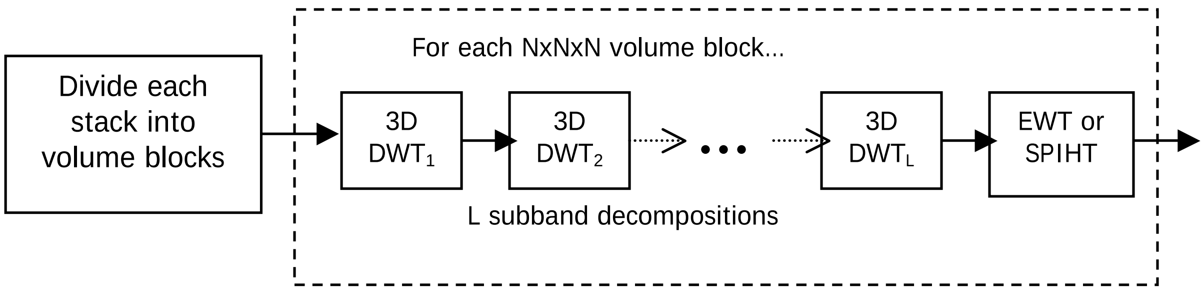


Figure 6. Block diagram of the encoder of the 3D DWT method

3.4. Volume Based Motion Compensated Coding

In order to exploit temporal redundancies more, we extended the motion compensated coding algorithm to 3D. Volume based motion compensated coding (3D Motion Compensated Coding (3D MC)) is quite similar to motion compensated coding, except that it works on stacks of images, as opposed to images in regular motion compensated coding.

The block diagram of the algorithm is given in Figure 7. The steps of the algorithm can be summarized as follows:

1. Construct volumetric images by combining each m image into a stack. Here m is the depth of the constructed stack.
2. On each volumetric image:
    a. Divide each volumetric image into NxNxM volumetric blocks. (N is 8 in our experiments. M is varied between 4 and 8)
    b. For each block of the image, find an NxNxM volumetric block in the previous (encoded) images that best resembles the current image. This prediction block is coded by a 3D motion vector. A prediction volumetric image is coded in this manner.
    c. The difference between the actual volumetric block and the prediction volumetric block is the error block. The error block is coded by 3D DCT. The coefficients of the DCT matrices are quantized with different methods and ways for various compression results.

Observe that, when M=1, the volume based motion compensated coding is exactly the same as regular motion compensated coding. In our implementation, we used full search to find the 3D motion vectors. We used two different measurement criterion: The first is the regular mean absolute difference distortion measure. The second is the standard deviation of the difference (residual) volumetric block. The latter is intended to obtain residual blocks with uniform distributions. The 3D DCT would then code the residual very efficiently. We observed in our experiments that the standard deviation measure always gave better results.

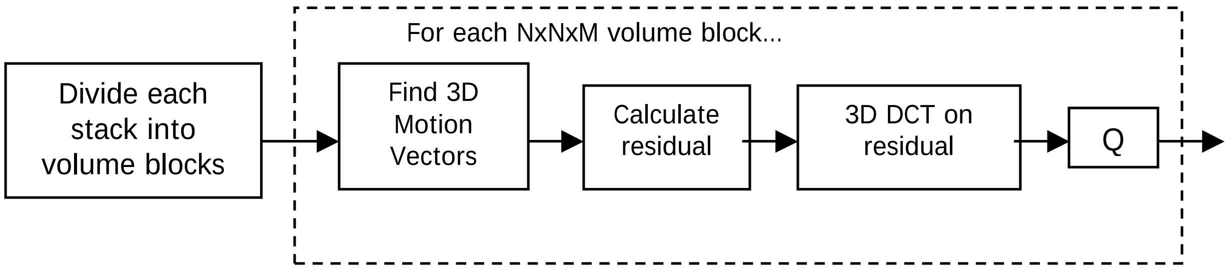


Figure 7. Block diagram of the volume based motion compensated coding.

3D methods usually would not work when there is non-uniform motion. Volume based motion compensated coding exposes even further constraints on the object motions in order to work well. Here are some of the cases that the volume based algorithm would not work:

1. When the motion is not smooth. This occurs when there is acceleration embedded in the motion. This problem would not occur in the regular motion compensated coding since each image is coded separately there.
2. Cases where motion compensated coding would not work. Among these are:
   a. The motion inside the block is not same for each pixel (voxel).
   b. Occlusion. This could potentially be dealt better by searching over several previous images. That is, the features could be occluded in the previous image I(t-1) but could be appearing in previous frames. The trade-off in searching over several previous images is the increased complexity.
   c. Sudden illumination changes

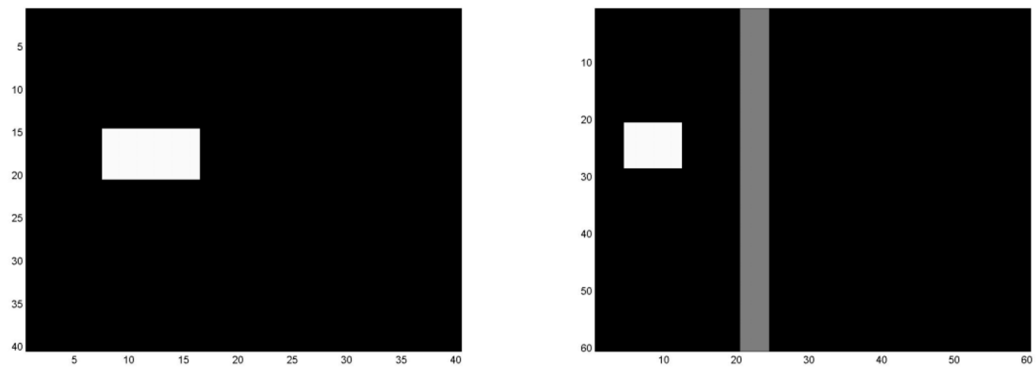
There are two main advantages of volume based motion compensated coding over regular motion compensated coding. First, more data is represented by the same number of motion vectors. This inherently provides a way for more efficient compression especially when there is no or little motion. Second, the residual of the volume based motion compensated coding is coded by a 3D DCT (as opposed to 2D DCT in motion compensated coding). We have discussed in the previous sections that 3D DCT provides better compression whenever there is correlation in the temporal domain (between frames). The residual block of the volume based motion compensated coding often has correlation in the temporal domain. Therefore, volume based motion compensated coding combined by 3D DCT potentially provides better results than motion compensated coding combined with 2D DCT.

In the current implementation, we use full search to find the best match for the motion vectors. We tried two different measurement criterions. The first is the mean absolute difference. The second is the standard deviation of the residual image. The main motivation behind the latter is that a uniformly distributed residual image is coded more efficiently by 3D DCT. In our experiments, we observed that the standard deviation measure usually gave better results compared to mean squared error.

We conducted synthetic experiments to demonstrate how the program works. In the first experiment, we had a rectangular box moving with constant velocity as shown in Figure 8(a). Since the motion was smooth, the video was perfectly reconstructed. In the second synthetic experiment, we added a non-moving rectangle to give an effect of occlusion. The initial frame is shown in Figure 8(b). Figure 9 shows one of the frames where we see the effect of occlusion. Figure 9(a) and (b) are the previous image and the current images respectively. Figure 10(a) shows the resulting predicted image by the regular motion compensated coding by looking only the previous image as the search region. Figure 10(b) is the resulting image with volume based motion compensated coding where we searched over several previous frames. We observed that volume based motion compensated coding dealt well with occlusion. Of course, regular motion compensation would also deal with occlusion when we search in more than one previous images. In our experiments, we will not look at previous images due to time complexity reasons.

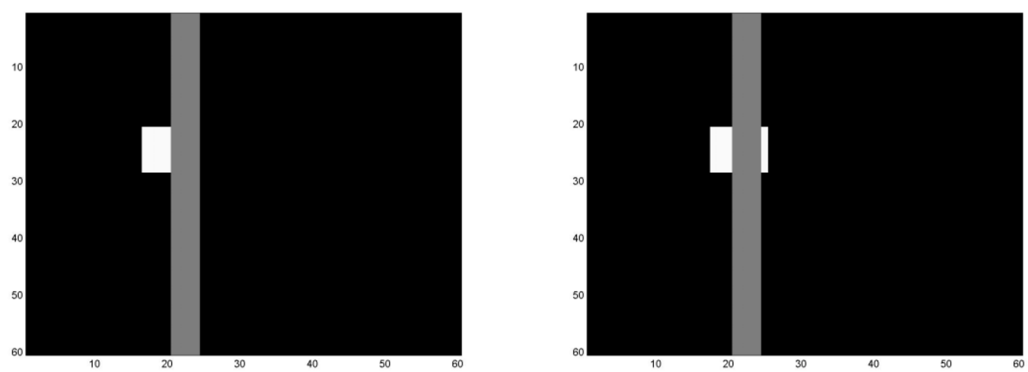
Volume based motion compensated coding matches volumetric blocks from the current stack of images with volumetric blocks from previous stack of images. In a such block match, we observe that the inner voxels of the blocks are usually matched better compared to the outlier voxels. We call this effect the ‘Centroid Effect’. Therefore, when we look at the constructed movies, we will see a pattern of good quality frames followed by worse quality frames. Figure 11 demonstrate four consecutive frames of a movie constructed by volume based coding. We can easily observe that the middle two frames have a better quality compared to the frames at the beginning and end. Although we currently do not handle the centroid effect, there could be methods for exploiting the centroid structure of the prediction, and should take its place in the future work.

In order to compare the predictions of motion compensated coding and volume based motion compensated coding, we coded the Miss. Am. sequence with both motion compensated coding and volume based motion compensated coding. The movements in this video are not very smooth, and we observed many artifacts in the resulting coded images. Figure 12(a) and 12(b) show the same encoded frame with regular motion compensated coding and volume based motion compensated coding (using 8x8x4 volume block size, and 0 search depth in time) respectively. Obviously, The regular motion compensated coding gives better prediction on each frame. The gain of volume based coding might come when the difference volume blocks are coded with 3-D DCT.



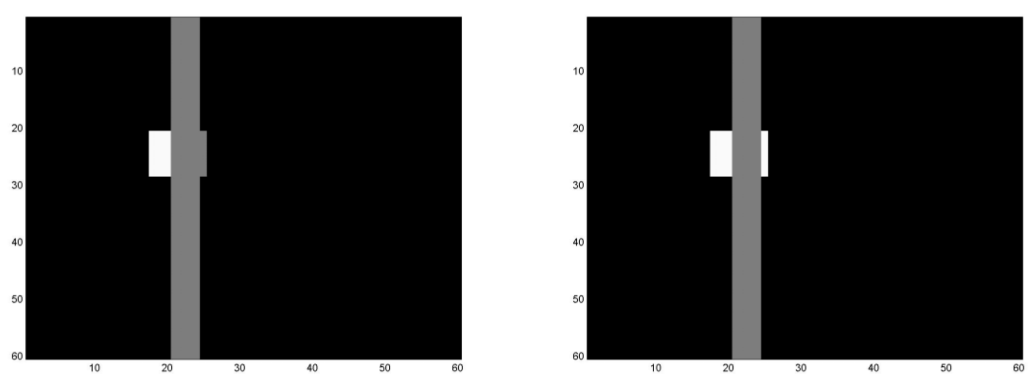
(a) (b)

Figure 8. Initial Frame of the (a) synthetic video with synthetic motion (b) synthetic video with occlusion



(a) (b)

Figure 9. Example of an occlusion. (a) Previous Image. (b) Current (to-be-coded) image.

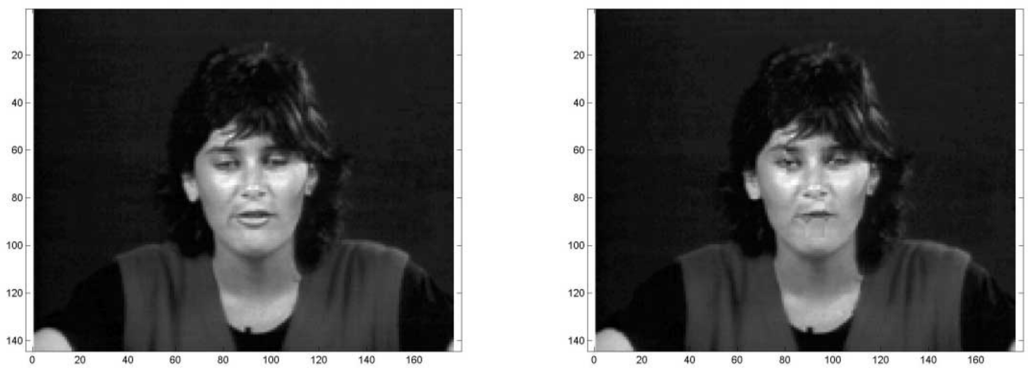


(a) (b)

Figure 10. The resulting encoded images for the image of Figure 5.b coded by (a) motion compensated coding. (b) Volume based motion compensated coding.

(a) (b) (c) (d)

Figure 11. Illustration of the centroid effect in volume based motion compensated coding. The central frames have a better prediction than the edge frames.



(a) (b)

Figure 12. An example image from the Miss.Am. sequence. (a) encoded coded by regular motion compensated coding (b) volume based motion compensated coding.

4. Experiments and Discussion

We performed experiments on the 3D DCT, 3D PCA and 3D MC methods. We compared these 3D algorithms to two 2D methods, namely, 2D DCT and 2D MC. For 2D DCT, each frame is coded independently by dividing the frame into square blocks and applying the 2D DCT on each block so no temporal redundancy is exploited at all. For the 2D MC, we applied the standard motion compensated predictive coding, with full search on a window of the previous frame and 2D DCT on the residual. The Matlab code for the experiments can be seen in Appendix B. Although we did not experiment with the 3D DWT we include our initial code of the 3D DWT implementation in this appendix.

To compare the algorithms we plotted the Bit Rate versus PSNR curves. For each data point in the plot, the bit rate was derived by quantizing the coefficients using a uniform quantizer and finding the entropy of the resulting coefficient symbols and motion vectors (if present). The quantized video sequence was then compared to the original and the PSNR was computed. PSNR was calculated over the same support region for all the experiments.

In Section 4.1, 4.2, 4.3 and 4.4 we discuss in detail the experimental results when the algorithms are applied to the Miss America, Salesman, Foreman and Bus Sequences, respectively. Section 4.5 contains Bit Rate versus PSNR plots for other sequences.

4.1. Miss America Sequence

The Miss America sequence is a low motion sequence with the motion confined to the person's lips and head. Since motion is low, temporal redundancy is high and it is expected that the 3D compression methods outperform the 2D methods for this sequence. Our experimental results, shown in Figure 13, verify this.

As it can be seen from the plot, the two 3D transforms – 3D PCA and 3D DCT – give the best PSNR for a given bit rate, with PCA slightly better than DCT. The 3D MC is also better than the 2D MC for bit rates lower than 0.55 bit per pixel. Specifically, at a bit rate of 0.14 bit per pixel, 3D MC is about 2 dB better in terms of PSNR compared to 2D MC. The 2D DCT method performs very badly with a PSNR 8 dB lower than the 3D transforms at 0.14 bit per pixel.

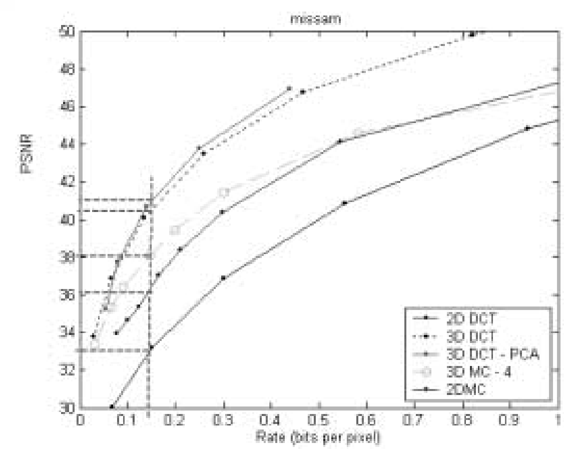


Figure 13. PSNR-Rate curve for the Miss Am sequence.

Figure 14 shows a comparison of a single frame from the compressed Miss America sequences. As it can be observed from the frame, the 2D DCT gives a very blocky and blurred sequences. For the 2D-MC compression, the distortion is manifested in rapidly varying noise which can be very striking to the viewer. For the 3D DCT and 3D PCA methods we observed less noise and smoother pictures. Although distortion is still observed, it does not vary rapidly from frame to frame so the noise is less evident. The absence of rapidly varying noise can be attributed to the fact that the same coefficients represent a stack of frames so the transition between those frames is smoother. Because of the centroid effect, the 3D MC sequence has a mixture of noisy frames and smooth frames.



(a) (b) (c)


(d) **(e)**

Figure 16. A reconstructed frame of Miss Americaq at 0.14 bit/pixel by (a) 2D DCT, (b) 2D MC, (c) 3D MC, (d) 3D DCT and (e) 3D PCA

4.2. Salesman Sequence

Salesman sequence seems to involve more motion compared to the Miss America sequence, yet the motion takes place only in a very concentrated area. Due to the little amount of motion taking place on the overall image, we observed that 3D based methods get better results compared to motion compensated coding.

Figure 15 shows the PSNR-Rate curve for the salesman sequence among many methods. First of all, we observe that all of the methods outperform the 2D image based DCT coding scheme. The other methods seem to perform similar to each other except for the low bit range. We observe that volume based motion compensated coding performs better than 3D DCT and motion compensated coding at bit rates lower than 0.25 bits/pixel. More specifically, around 0.11 bits/pixel rate, the PSNR of volume based motion compensated coding is 1(1.5) dB more than motion compensated coding (3D DCT).

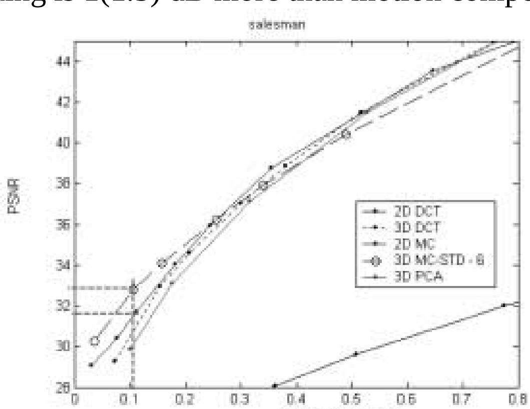


Figure 15. PSNR-Rate curve for the salesman sequence.

Figure 16(a) and (b) gives a frame constructed at 0.11 bits/pixel by volume based motion compensated coding and motion compensated coding respectively. We observe that volume based motion compensated coding provides a less noisy image in this case. Similarly, Figure 17(a) and (b) provide a frame from the reconstructed videos of volume based motion compensated coding and 3D DCT respectively. Here, we observe different artifacts, *i.e.*, very blurred image in the 3D DCT versus noisier image in volume based motion compensated coding. Since the artifacts are different in this case, it is subjective to choose between the two methods.

(a) (b)

Figure 16. A reconstructed frame at 0.11 bits/pixel by (a) volume based motion compensated coding (b) motion compensated coding.

(a) (b)

Figure 17. A reconstructed frame at 0.14 bits/pixel by (a) volume based motion compensated coding (b) 3D DCT.

4.3. Foreman Sequence

The Foreman sequence contains more motion than the previous sequences. In this sequence, non-uniform motion is caused by the camera as well as the man's face and head movement.

The Bit Rate versus PSNR curves below (see Figure 18) show that the 3D DCT performs the best among the methods. Even if there is a large amount of motion, the wide planar surfaces (the foreman's hat and the background building), still provide for a lot of temporal redundancy at a given spatial location. It can also be observed that the PCA comparative performance deteriorates as we go to lower bit rates. One reason for this is that the video is roughly divided into half smooth temporal signal and half high frequency temporal signals. So at lower bit rates, it is harder to represent both types accurately using a small number of principal components. Furthermore, since the motion is non-uniform from frame to frame, the 3D-MC method is less appropriate and the motion is better captured by the 2D-MC technique. Figure 19 gives a comparison of a single frame from the foreman sequence compressed using the difference methods.

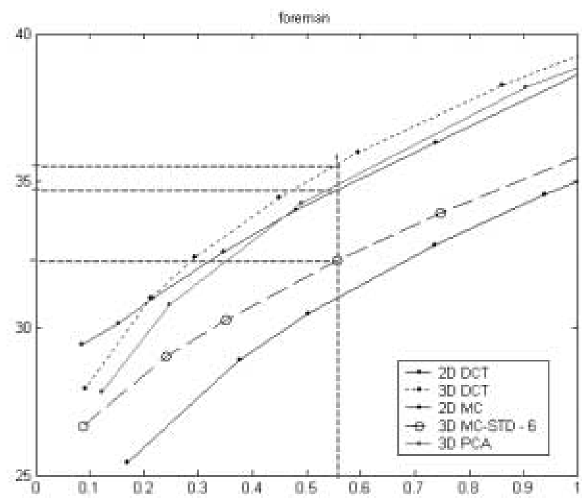
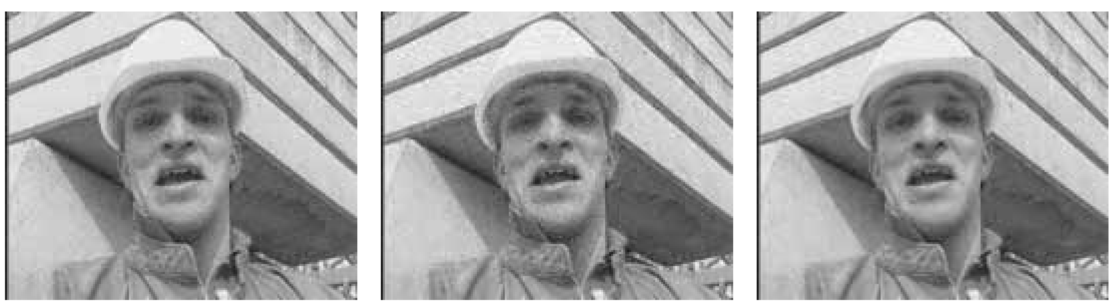


Figure 18. PSNR-Rate curve for the foreman sequence.

Based on the plots, we can see that at 0.56 bit per pixel, the 3D-DCT method perform 1 dB better than 2D-MC while 3D-MC performs 2.5 DB worse than 2D-MC.



(a) (b) (c)

Figure 19. A reconstructed frame of Foreman at 0.56 bit/pixel by (a) 2D MC, (b) 3D MC, (c) 3D DCT

4.4. Bus Sequence

The bus sequence has a lot of motion going on in the scene. Since the video was captured by a hand-held camera, there was also a lot of motion even in the non-moving parts of the scene. Due to the movements of the camera, there was a global motion in the video, which was captured well with regular motion compensated coding. This global motion, however, is very non-uniform over consecutive sets of frames, and as a result each volumetric block of volume based motion compensated coding involved objects moving with instantly changing accelerations, which is a big drawback for volume based motion compensated coding. Similarly, the correlation between consecutive frames of an 8x8x8 DCT block was nearly 0. As a result, 3D DCT does not perform any better than 2D DCT.

Figure 20 gives the PSNR-rate curve for the bus sequence. As expected, regular motion compensated coding provides the best results amongst all of the methods.

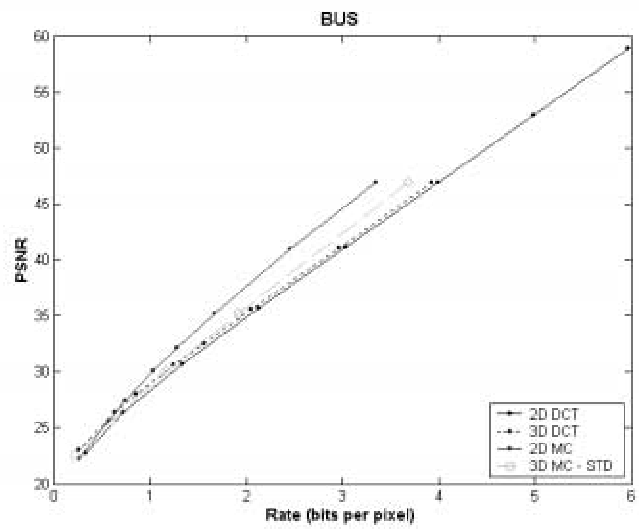
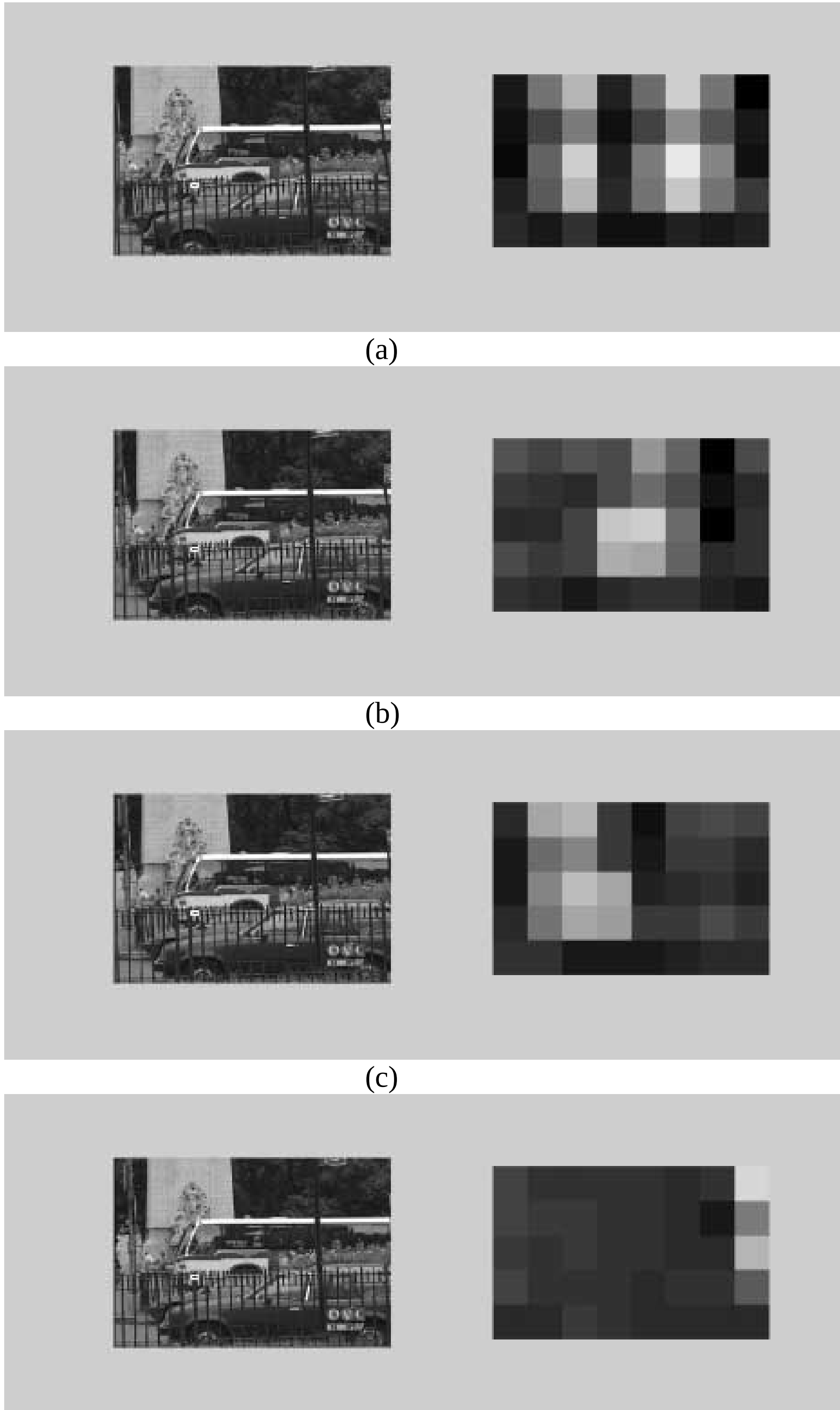


Figure 20. PSNR-Rate curve for the bus sequence.

In order to illustrate why 3D methods do not work that well, we would like to show 4 consecutive frames from the bus sequence. In Figure 21(a) – (d), we give the four frames on the left, and a zoomed region from these sequences. Observe that, the zoomed region constantly changes in these four frames. This is exactly what is happening everywhere throughout the sequence due to the non-uniform global motion embedded in the sequence. Due to this high frequency behavior in the temporal domain, the 3D methods failed in this case.



(a)

(b)

(c)

(d)

Figure 21. Illustration of high frequency characteristics in the temporal domain of the bus sequence. Observe how the zoomed window in the right various across consecutive frames.

4.5. Other Sequences

In this Section we provide the PSNR-Rate curves for the other sequences that we experimented with. Figure 22-24 show the results for Claire, mother daughter, and car phone respectively. In general, we observe that either of 3D DCT or volume based motion compensated coding provided better results in terms of PSNR versus Rate.

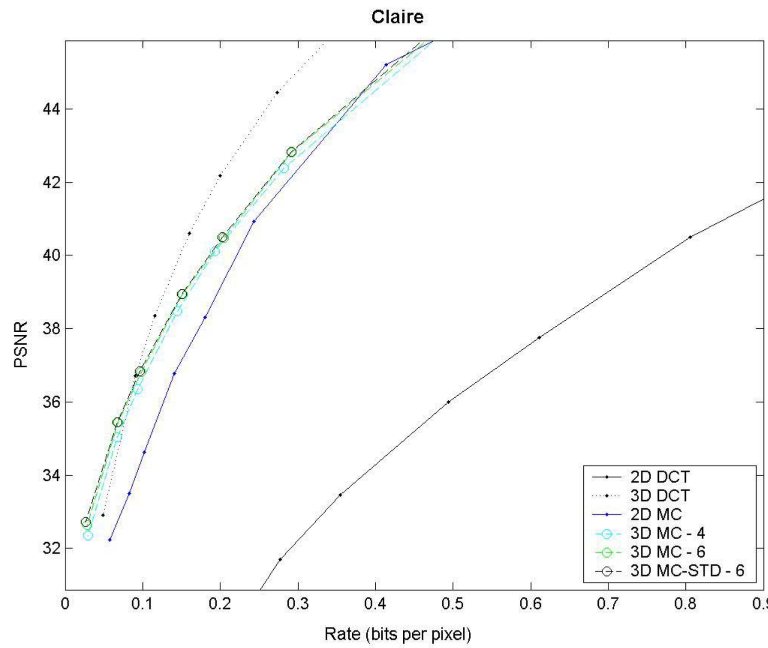


Figure 22. PSNR-Rate curve for the Claire sequence.

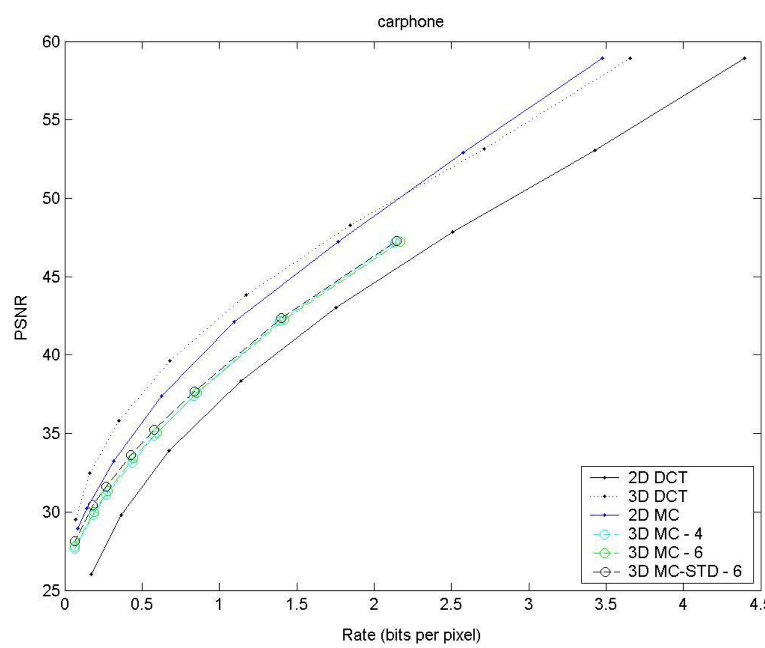


Figure 23. PSNR-Rate curve for the car-phone sequence.

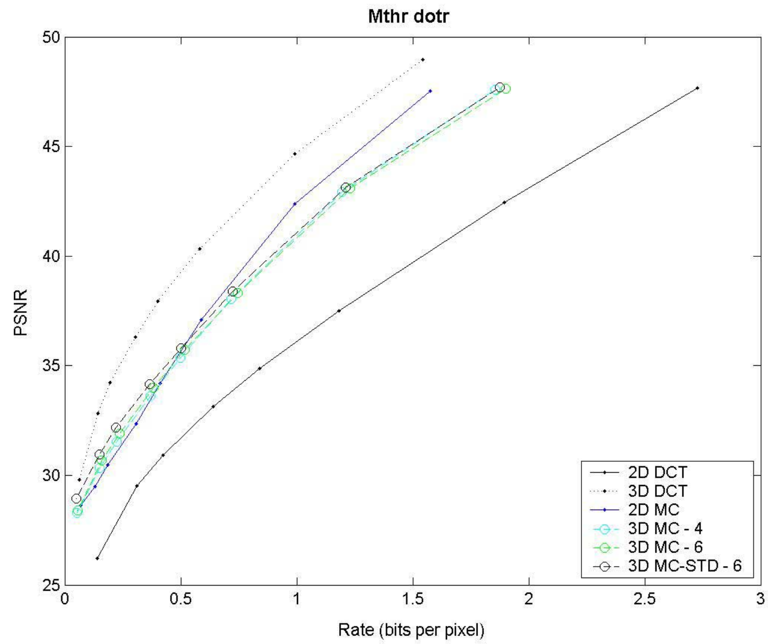


Figure 24. PSNR-Rate curve for the Mother-Daughter sequence.

5. Conclusions and Future Work

The video signal has high temporal redundancies between a number of frames and this redundancy has not been exploited enough by current video compression techniques. In this research, we showed that 3D methods, such as 3D DCT, 3D PCA and 3D MC, which make use of the temporal redundancy between several frames, can perform better than standard 2D MC especially for low motion sequences. Although 3D methods require more memory for frame storage and possibly more complex processing, continuous advancements in processing and storage technology will overcome this disadvantage. With the apparent gains in compression efficiency we foresee that 3D methods will be the future of video compression.

There are various directions for future investigations. First of all, we would like to explore methods that would exploit the centroid effect that is observed in volume based motion compensated coding. Currently we use 3D DCT for coding the residual, and 3D DCT does not exploit the centroid structure at all. Another direction could be to combine volume based motion compensation with other 3D transformations such as wavelet transformation. Another extension could be to use only the temporal domain redundancies, i.e. by a run length coding of each pixel or collection of pixels along time.

References

[1] Forman, M.C., Aggoun, A., McCormick, M., "*A novel coding scheme for full parallax 3D-TV pictures*," Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing, 1997, pp. 2945 – 2947, vol.4, April 1997, Munich, Germany.

[2] Abousleman, G.P., Marcellin, M.W., and Hunt, B.R., "*Compression of hyperspectral imagery using the 3-D DCT and hybrid DPCM/DCT*, "IEEE Transactions on Geoscience and Remote Sensing, Volume 33, Issue 1, Jan. 1995, pp. 26 –34.

[3] Bauer, M and Sayood, K., "*Video coding using 3 dimensional DCT and dynamic code selection*," . Proc. IEEE Data Compression Conference, 1995. p. 451, March 1995, Snowbird, UT, USA

[4] Chan, R.K.W.and Lee, M.C, "*3D-DCT quantization as a compression technique for video sequences*," Proc. International Conference on Virtual Systems and MultiMedia, 1997. VSMM '97. Proceedings, pp. 188 – 196, Sept. 1997, Geneva, Switzerland.

[5] Jingmin Song, Zhang Xiong, Xudong Liu and Yun Liu, "*PVH-3DDCT: an algorithm for layered video coding and transmission*," Proc. The Fourth High Performance Computing in the Asia-Pacific Region, 2000, pp. 700 – 703, vol.2, May 2000, Beijing, China.

[6]. D.G.Sampson, E.A.B. de Silva, and M. Ghanbari, "*Low bit-rate video coding using Wavelet quantization*", Proc. Inst. Elect. Eng., 1995, vol 142, Page:141-148.

[7] Gi Hun Lee; Joon Ho Song; Rae-Hong Park, "*Three-dimensional DCT/WT compression using motion vector segmentation for low bit-rate video coding*," Proceedings of International Conference on Image Processing, 1997, Volume: 3, Page(s): 456 -459 vol.3, 1997.

[8] Beong-Jo Kim and Pearlman, W.A., "*An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT*)", Proc. Data Compression Conference, 1997, pp. 251 - 260, March 1997, Snowbird, UT, USA

[9]Vass, J.; Bing-Bing Chai; Xinhua Zhuang, "*3DSLCCA-a highly scalable very low bit rate software-only wavelet video codec* ",Proc. IEEE Second Workshop on Multimedia Signal Processing, 1998, pp. 474 – 479, Dec. 1998, Redondo Beach, CA, USA.

[10] Gang Lin and Zemin Liu, "*3D wavelet video codec and its rate control in ATM network*," Proc. IEEE International Symposium on Circuits and Systems, 1999, pp. 447 – 450, vol.4, June 1999, Orlando, FL.

[11]. Bernabe, G.; Gonzalez, J.; Garcia, J.M.; Duato, J., "*A new lossy 3-D wavelet transform for high-quality compression of medical video"*, IEEE EMBS International Conference on Information Technology Applications in Biomedicine, 2000, Page(s): 226 –231.

[13]. Lazar, D.; Averbuch, A., "*Wavelet-based video coder via bit allocation*," IEEE Transactions on Circuits and Systems for Video Technology, Volume: 11 Issue: 7, July 2001, Page(s): 815 –832.

[14] Seigneurbieux, P. and Zixiang Xiong, "*3-D wavelet video coding with rate-distortion optimization*," Proc. International Conference on Information Technology: Coding and Computing, 2001. pp. 322 – 326, April 2001, Las Vegas, NV, USA

[15] J.M. Shapiro, "*Embedded image coding using zerotrees of wavelet coefficients*," IEEE Transactions on Signal Processing, vol. 41, no. 12, pp. 3445-3462, December 1993.

[16] A. Said and W. Pearlman, "*A new, fast, and efficient image codec based on set partitioning in hierarchal trees*", IEEE Transacations on Circuits and Systems for Video Technology, vol. 63, pp. 243-250, June 1996.

[17] Gokturk SB, Tomasi C, Girod B, Beaulieu CF, "*Medical Image Compression Based on Region of Interest, with Application to Colon Ct Images*," in 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Istanbul, Turkey, 2001.

[18] J.S. Taur, and C.W. Tao, "*Medical Image Compression using Principal Component Anlyasis*," International Conference on Image Processing, Volume: 1 , 1996 , Page(s): 903 -906 vol.2.

[19] M. Yoshioka, and S. Omatu, "*Image Compression by nonlinear principal component analysis*," IEEE Conference on Emerging Technologies and Factory Automation, EFTA '96,Volume: 2 , 1996 , Page(s): 704 -706 vol.2.

APPENDIX A - Division of Work

Algorithms
The algorithms were designed through discussions of both members of the group.

Implementation
2D and 3D Block Matching and Motion Compensation – Burak
3D DCT – Anne
3D PCA – Burak
Quantization and Entropy Calculations of 2D and 3D DCT Coefficients – Anne
Quantization and Entropy Calculations of PCA Coefficients – Burak
Entropy Calculations of Motion Vectors – Burak
PSNR Calculations – Anne
Synthetic Experiments – Burak
Main scripts – Anne and Burak
Initial 3D DWT Code – Anne
Movie Creation script for the presentation - Burak

Experiments
Both members of the group repeated all the experiments on at least five different sequences.

Report
Abstract - Burak
Introduction – Burak
Previous Work – Burak and Anne
3D DCT – Anne
3D DWT – Anne
3D PCA – Burak
3D MC – Burak
Experiments and Analysis
- Experiment Set-up – Anne
- Miss America – Anne
- Salesman – Burak
- Foreman – Anne
- Bus – Burak
- Other Sequences – Burak

Conclusions and Future Work – Burak and Anne

APPENDIX B – Matlab Code

Please refer to attached source files in source.zip.