Three Dimensional Motion Vectorless Compression

Rohini Nagapadma and Narasimha Kaulgud*

Department of E & C, National Institute of Engineering, Mysore
* Talent Transformation Group, Wipro electronics city, Bangalore

Summary
Video Compression can be performed by Motion Compensation and Three Dimensional Motion Vectorless method. In Motion Compensation [6], [8], [9], motion is estimated to get a predictor block in a past reference frame and the error between two blocks with the relative displacement is sent, for each block in the current frame. This reduces the number of bytes needed to represent the video. The Motion Vectorless video compression using the Three Dimensional Discrete Cosine Transform (3D-DCT) algorithm developed exploits spatial and temporal redundancies and does not use the computational intensive Motion Compensation algorithm to remove interframe correlation. Since all DCT multiplications are real it is easy to realize. Quantization and Run-length coding follow 3D-DCT. Actual timing analysis is obtained to show that the Three Dimensional scheme generates near real time video compression.
Key words:
Video Compression, Motion Vectorless, Motion Compensation, DCT, Conditional Replenishment

1. Introduction

We are witnessing a phenomenal increase in the use of images in many different applications. This is mainly due to technological advances impacting several image operations, the availability of sophisticated software tools for manipulation and management and the World Wide Web providing easy access to a wide range of users. Typical applications using huge amounts of images are medical imaging, remote sensing, entertainment, digital libraries, distance learning, training and multimedia. Digital images require huge amounts of space for storage and large bandwidths for transmission. For example, a single 640 x 480 pixel color image using 24 bits/pixel requires close to one Megabyte space. Despite the technological advances in storage and transmission, the demands placed on the storage capacities and on the bandwidth of communication exceed the availability. Image compression has proved to be a viable technique as one solution response with various standards [3].

2. Introduction to Motion Vectorless Method

The problem of real time video compression is difficult and an important one. This has inspired a great deal of research activity. Moreover the problem encountered in transmission of video signals is the inordinate amount of data involved. Amount of data can be reduced by exploiting the spatial & temporal redundancies within and across video frames.

DCT plays a vital role in exploiting the spatial redundancy & compressing the data without much visible distortion. DCT based image coding is the basis for all the image & video compression standards. The basic computation in a DCT based system is the transformation of an N x N image block from the spectral domain to the DCT domain [1], [3], [7]. For image compression standards, N is equal to 8. In normal compression scheme a single frame is subdivided into 8 x 8 blocks, each of which is independently processed as shown in Figure 1. Each block is transformed into 2D-Discrete Cosine Transform resulting in 8 x 8 block of DCT coefficients. These coefficients are quantized to produce minimal visual artefacts, while maximally reducing the representational entropy of the coefficients. Quantized coefficients are then encoded into a compressed data stream. The reduced entropy of the quantized coefficients is reflected in the higher compression ratio of data.

3. Methodology of Motion Vectorless Method

The Motion Vectorless algorithm developed in our work is based on the 3D-Discrete Cosine Transform [4], [12], [13]. This algorithm takes a full motion digital video stream and divides it into groups of eight frames. Each group of eight frames is considered as a three dimensional image, where X and Y are spatial components and Z is the temporal component. Each frame in the image is divided into 8 x 8 blocks, forming 8 x 8 x 8 cubes [10] as illustrated in the Figure 1.

The transformed 512-point discrete signal is a function in three dimensions and contains both spatial and temporal information. Most of the energy is contained in low frequency co-efficients while the majority of the high frequency co-efficients have zero or near zero values [5], [7]. In the next step, all 512 DCT co-efficient are quantized using a 512-element quantization table.

Manuscript received April 5, 2009
Manuscript revised April 20, 2009

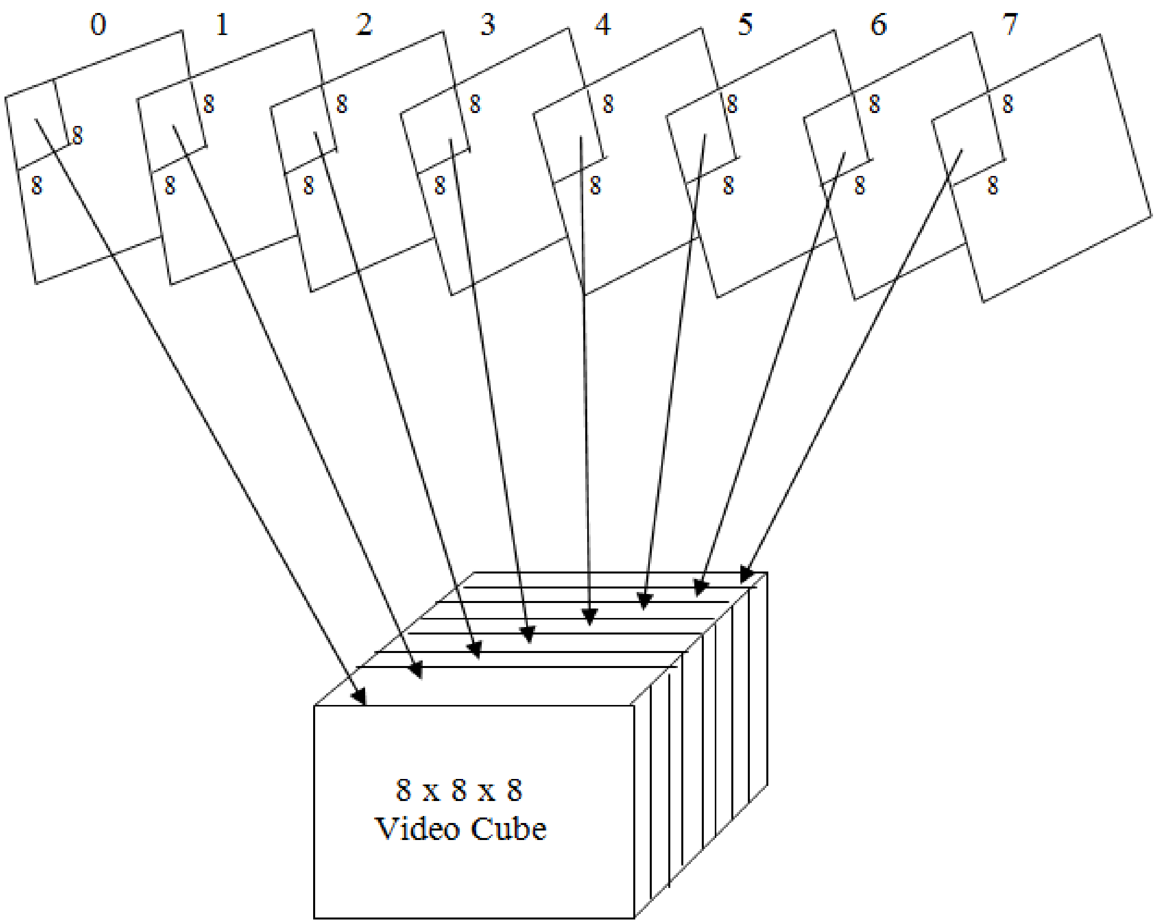


Figure 1: 8 x 8 x 8 video cube

Quantization introduces minimum error while increasing the number of zero value co-efficient. The result of the quantization operation is a collection of smaller-valued co-efficients, a large number of which are zero. These co-efficients are then converted into compact binary sequence using an entropy coder. The entropy coding operation starts with re-ordering the co-efficients in descending order of expected value. This sequence has the benefit of collecting sequentially the largest number of zero-valued co-efficients. The runlength of zero co-efficients is computed & the alphabet of symbols to be encoded becomes the Runlength of zeroes appended to the length of the non-zero co-efficient. This binary sequence represents the compressed 8 x 8 x 8 block. The steps of the above are indicated in Figure2.

The 3D-DCT encoder and decoder are symmetrical and each has only three parts [10], [11], [12] as shown in Figure 2. They are the forward or inverse 3D-DCT calculations, the quantization or dequantization and the Run-Length Coding (RLC) or decoding. If the video signal digitalization and the video cubes creation are skipped over, the 3D-DCT is the first part of the encoder.

The 3D Forward Discrete Cosine Transformation used in the 3D-DCT image coder is based on the following formula:

$$S(w,v,u)=\alpha_{3D}(w,v,u)\sum_{z=0}^{N_F-1}\sum_{y=0}^{N_R-1}\sum_{x=0}^{N_C-1}\{s(z,y,x)\cos(t_1)\cos(t_2)\cos(t_3)\} \quad (1)$$

Where $S(w,v,u)$ represents the transformed 3D coefficient and *w*, *v* and *u* are the spatial indices.

The proposed 3D-DCT image decoder employs the following 3D Inverse Discrete Cosine Transformation [1]:

$$s(z,y,x)=\sum_{w=0}^{N_F-1}\sum_{v=0}^{N_R-1}\sum_{u=0}^{N_C-1}\alpha_{3D}(w,v,u)S(w,v,u)\cos(t_1)\cos(t_2)\cos(t_3) \quad (2)$$

Where $s(z,y,x)$ represents the 3D inverse transformed coefficient and *x, y* and z are the time indices.

Also

$$t_1 = \frac{(2x+1)u\pi}{2N_C},$$

$$t_2 = \frac{(2y+1)v\pi}{2N_R},$$

$$t_3 = \frac{(2z+1)w\pi}{2N_F},$$

$$\alpha_{3D}(w,v,u) = \sqrt{\frac{2}{N_R}\frac{2}{N_C}\frac{2}{N_F}}C(v)C(u)C(w)$$

Here,

$$C(u), C(v), C(w) = \begin{cases} \frac{1}{\sqrt{2}} & u,v,w = 0 \\ 1 & otherwise \end{cases}$$

The second part of the encoder is the quantization. The purpose of this operation is to reduce number of the insignificant DCT coefficients, and hence to increase the compression ratio. All DCT coefficients in the same data cube are quantized using the following formula:

$$F_q(u,v,w) = round\left(\frac{F(u,v,w)}{Q(u,v,w)}\right) \qquad (3)$$

where *u, v,* and *w* are the spatial indices; *F(u, v, w)* refers to the coefficient value before the quantization; *Q(u ,v, w)* denotes the element in the quantization table; $F_q(u,v,w)$ represents the quantized co-efficient.

DCT-based image coding is the basis for all the image and video compression standards. The basic computation in a DCT based system is the transformation of an N x N image block from the spectral domain to the DCT domain. The cubes are reordered in a way that maximizes the number of successive zeros in the stream. The high correlation of the zero, probability of a coefficient with its position in the cube makes this possible. Run-length encoding is subsequently applied to the stream.

4. Implementation of Motion Vectorless Method

Following are the steps for implementing the Motion Vectorless method.

4.1a Reading The Input Video

The standard QCIF image is read and divided into 8 x 8 blocks. Thus we get 22 blocks of 8 x 8 dimensions in a row and 18 blocks in a column.

4.1b Compressing The Input Video

Eight frames are read together which form a group. Further each 8 x 8 x 8 cube of 512 pixels is transformed into the frequency domain by applying the forward 3D-DCT. Quantization is done by dividing the block from standard quantization matrix. RLC [5] is used to code the quantized components.

4.1c Decompressing The Video

At the receiver end, the Runlength coded sequence is decoded. The result obtained is very similar to the vectorised sequence. The sequence is then restored in the matrix format for dequantization. The matrix obtained is dequantized where the elements are multiplied by the respective elements of the quantization table. The dequantized arguments are passed to the IDCT function. Applying the 3D- IDCT transform the data is converted back from frequency domain to spatial domain. Inverse DCT matrix closely resembles the original matrix. The above steps are repeated for every block to get complete image compression and decompression.

5. Experimental Results and Conclusions

The results for timing analysis and compression ratio are given below.

QCIF video	Motion Vectorless method
Claire.qcif	**8.99**
Suzie.qcif	**8.83**
Stefan.qcif	**8.73**
Soccer.qcif	**8.96**

Table 1: Timing analysis (in seconds) for Motion Vectorless method

Timing analysis for Motion vectorless method is as shown in Table 1 for various video clippings. We have evaluated this technique on video sequences that vary in their characteristics.

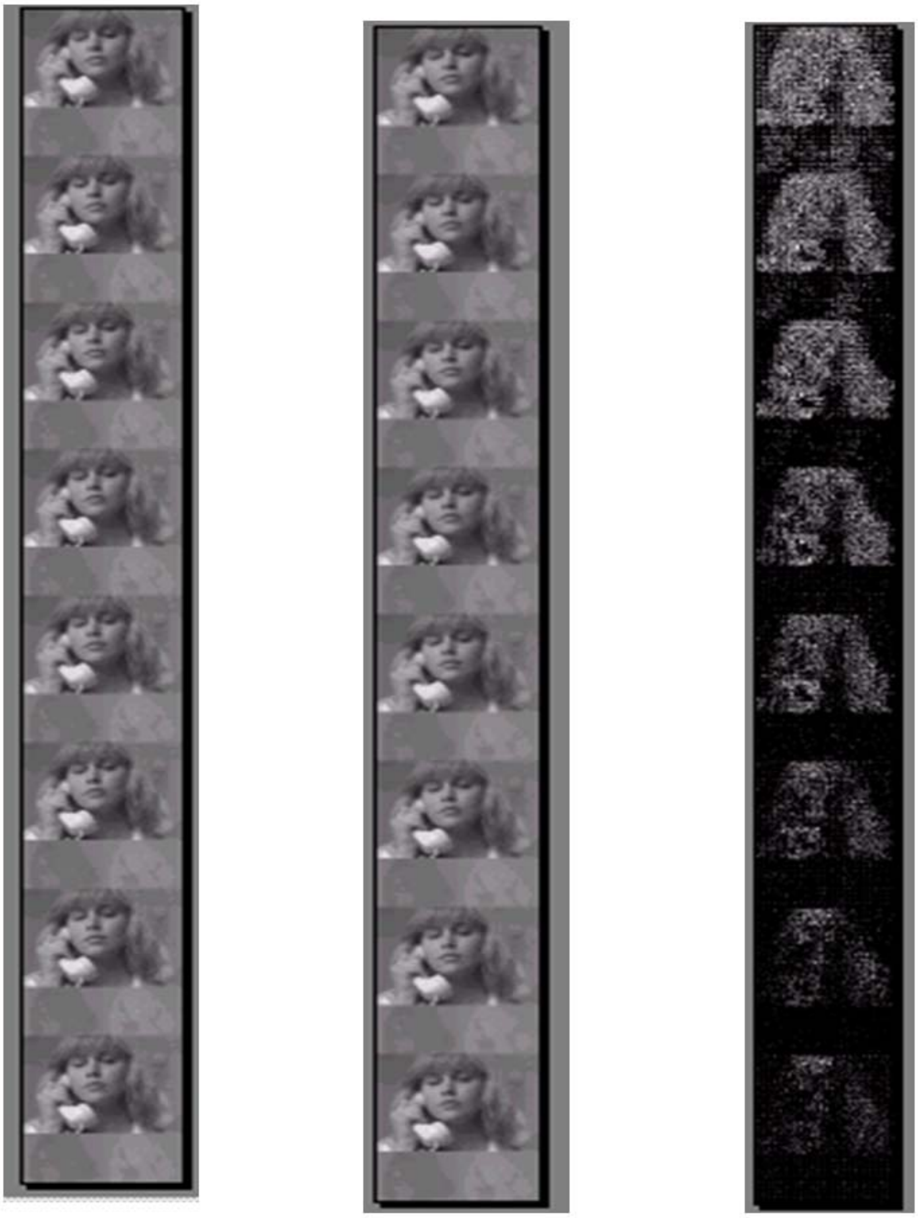
Method	Compression Ratio

3D-DCT	**15.00**

Table 2: Compression ratio for Motion Vectorless method

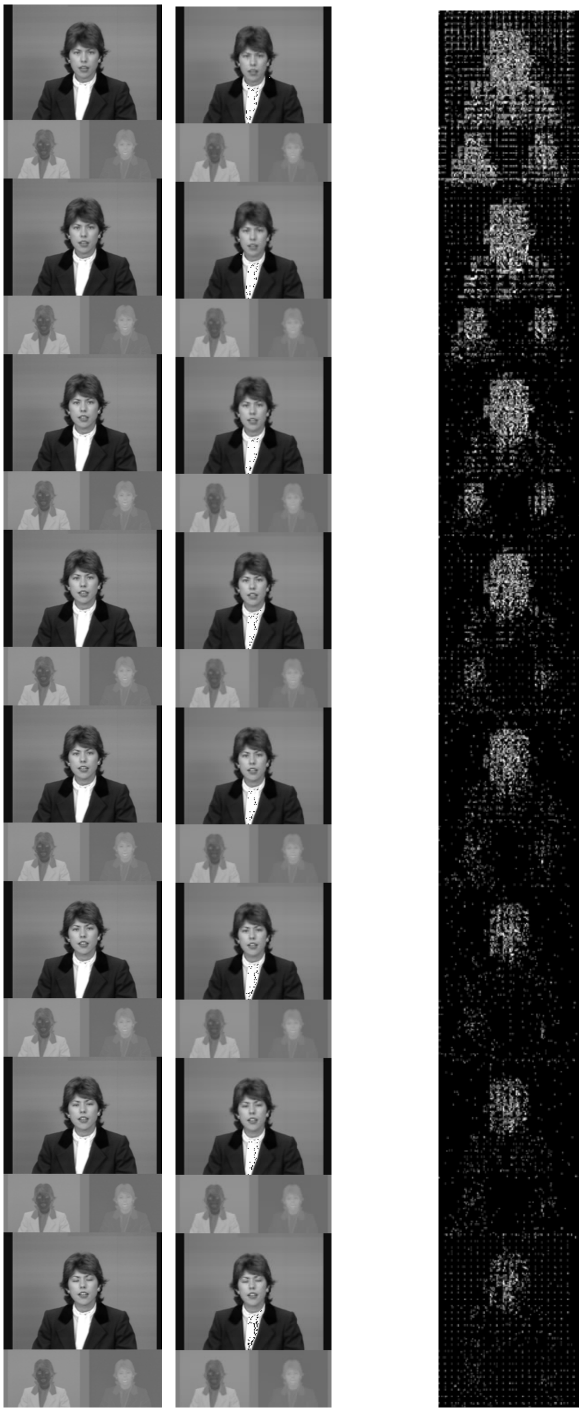
Due to simplicity 3D compression is a good choice for use in many applications. It ensures inexpensive equipment at the cost of slightly higher bandwidth consumption. Sophisticated quantization and coding techniques can be incorporated to achieve even higher compression ratios. Compression ratio achieved in this method is satisfactory. This is shown in Table 2.

DCT has the beneficial properties of forward/inverse 3D-symmetry, fast separable implementation and excellent energy compaction but at the cost of slightly higher bandwidth consumption.



(a) (b) (c)

Figure 3:Results for Motion Vectorless Method
a) Original Frames of Suzie.qcif b) Reconstructed Frames c) DCT Frames



(a) (b) (c)

Figure 4: Motion Vectorless method
(a) Original frames of Claire.qcif (b) Reconstructed frames (c) DCT frames.

In Figures 4 and 5, reconstructed images show little distortion at the center of the image i.e., on the white tie area in the form of black dots. Except for this, the reconstructed image is of very good quality in both the methods. This minor distortion may be removed using appropriate filtering techniques.

References

[1] Ahmed N, *et al*, Jan.1974, *Discrete cosine transform*, IEEE Transactions on Computers, Vol. C-23, pp. 90-93.

[2] Chan Y.L and Siu W.C, May 1997, *Variable Temporal-Length 3-D Discrete Cosine Transform Coding*, IEEE Transactions on Image Processing, Vol. 6, No. 5, pp. 758-763.

[3] Furht B., 1999, *Video Presentation and Compression*, in Handbook of Multimedia Computing, CRC Press, Boca Raton, FL.

[4] Furht B, *et al*, 2003, *An Adaptive Three-Dimensional DCT Compression Based on Motion Analysis*, Proceedings of the ACM Symposium on Applied Computing.

[5] Jain A.K, 1994, *Fundamentals of Image Processing*, Prentice-Hall India.

[6] Netravali A.N. and Stuller J.A, Sept. 1979, *Motion-compensation transform coding*, Bell Syst. Tech J., Vol. 58, No. 7, , pp. 1703-1719.

[7] M. Nelson, 1992 *The Data Compression Book*, San Mateo, CA, M&T Books.

[8] Rohini Nagapadma, *et al*, Jan 2004, *Motionless Video Compression*, Proceedings of National conference on Control Communication & Information Systems, Goa College of Engineering .

[9] Rohini Nagapadma, *et al*, Dec 2004, *Algorithm For Motion Compensation & Conditional Replenishment* , Proceedings of International conference on Human Machine Interfaces – 2004, JN Tata auditorium, IISc. Bangalore.

[10] Rohini Nagapadma, *et al*, Dec 2004, *Vectorless 3D Video Compression*, Proceedings of International conference on Human Machine Interfaces – 2004, JN Tata auditorium, IISc. Bangalore.

[11]Rohini Nagapadma, *et al*, March 2005, Proceedings of *Video Compression For Telemedicine*, National conference on Image Processing, M.S. Ramaiah school of advanced studies.

[12]Westwater R. and Furht B, February 1996, *The XYZ Algorithm for Real-Time Compression of Full-Motion Video*, Journal of Real-Time Imaging, Vol. 2, No. 1, , pp.19-34.

[13]Xiuqi Li and Borko Furht, *An Approach to Image Compression Using Three-Dimensional DCT*, Department of Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL 33431.

AUTHORS

Rohini Nagapadma received the B.E degree in Electronics and Communication Engineering from Mysore University, Mysore in 1983. She received her M.Tech degree in 1991 from IIT- Madras in the area of Communication and High Frequency. Presently she is serving as Assistant Professor in the department of E & C, National Institute of Engineering, Mysore. She is presently persuing Ph.D in the area of Image Processing. She has presented 12 papers in both National and International conferences. Her areas of interest are Multimedia communications and Computer networking.

Narasimha Kaulgud received his M.Tech degree in 1988 from Mysore University in Industrial Electronics. He obtained Ph.D in the area of Image and Video Processing from IIT-Bombay. He has 15 years of teaching experience in the field of signal processing, artificial neural networks and others Presently, he is in the Talent Transformation Group in Wipro Technologies, Bangalore. He has presented 41 papers in both National and International conferences.