A Hybrid DCT-SVD Video Compression Technique (HDCTSVD)

Tong, Lin; Rao, K.R.

Abstract—*A new hybrid DCT-SVD (HDCTSVD) video compression technique is proposed in this paper. Discrete cosine transform (DCT) is widely used in video coding due to its high energy compaction and efficient computation complexity, singular value decomposition (SVD) is a transform that provides optimal energy compaction for any data. DCT and SVD are combined to achieve optimal performance of the transform part in the proposed video compression technique. SVD is used only for the blocks for which DCT cannot provide good compression. The decision criterion is set in the DCT domain. By dropping a certain number of coefficients in the DCT domain, the energy loss is calculated. Whether or not sending the block to SVD domain is based on the energy loss. The simulation shows that the proposed Hybrid DCT-SVD system provides good performance for both intra frame coding and inter frame coding.*

Index Terms—*adaptive vector quantization, DCT, inter frame, intra frame, SVD, video coding.*

1. INTRODUCTION

Transform coding is one of the core algorithms in video coding. By applying different linear transforms, the original data are decomposed into a set of coefficients in the corresponding transform domain. In the existing video coding standards such as H.264/AVC, H.263, MPEG, etc., discrete cosine transform (DCT) plays the role of transform coding due to its high energy compaction and efficient computation complexity [1]. Singular value decomposition (SVD) is a transform that provides optimal energy compaction for any data [2][4]. However, the high computational complexity associated with computing the eigenvectors and eigenvalues has limited its application. Besides, only modest compression can be achieved by using SVD because when eigenvalues are transmitted, the corresponding eigenvectors must also be transmitted [9]. Dapena and Ahalt proposed a hybrid DCTSVD algorithm for image compression [6]. This algorithm can reduce a lot computational cost that occurs by using only SVD and also improve the image quality when only using DCT.

Based on the characteristics of video data, we propose a hybrid DCT-SVD (HDCTSVD) technique for video coding. It will take advantages of DCT first, and use SVD only for the blocks that DCT does not compact energy well. Adaptive vector quantization (AVQ) technique is used to code the eigenvectors so that higher compression of SVD can be achieved [7].

This paper is organized as follows: Section 2 introduces SVD coding techniques. Section 3 briefly reviews the most popular video coding technique and explains the motivation of the proposed HDCTSVD technique based on the characteristics of data extracted from video compression. In Section 4, after a short review of the existing HDCTSVD algorithms in image coding, we propose our new video coding system. Section 5 shows the simulation results. Conclusions are presented in Section 6.

2. SVD CODING

SVD provides optimal energy compaction for any input data [8]. By taking only a few largest eigenvalues and corresponding eigenvectors will give good representation of the input data. Matrix $A(m \times n)$ has a singular value decomposition, which can be represented by

$$A = U\Lambda^{\frac{1}{2}}V^T \qquad (1)$$

where U is a $m \times m$ orthogonal matrix, V is a $n \times n$ orthogonal matrix, $\Lambda^{\frac{1}{2}}$ is a $m \times n$ matrix whose off-diagonal entries are all zeros. The diagonal elements of $\Lambda^{\frac{1}{2}}$ must satisfy $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq ... \geq \sqrt{\lambda_r} \geq \sqrt{\lambda_{r+1}} \geq ... \geq \sqrt{\lambda_n} \geq 0$, where r is the rank of A.

The columns of V can be calculated by solving

$$(A' - \lambda(n)I)v(n) = 0 \qquad n = 1,...,r \qquad (2)$$

where $A' = A^T A$. The columns of U are:

$$u(n) = \frac{1}{\sqrt{\lambda(n)}} Av(n) \qquad n = 1,...,r \qquad (3)$$

If q eigenvalues and corresponding eigenvectors are chosen to represent the original

Manuscript received May 31, 2006.

Tong, Lin (contact person), Electrical Engineering Department, University of Texas at Arlington, U.S., (email: lin_tong@hotmail.com)

Rao, K.R., Electrical Engineering Department, University of Texas at Arlington, U.S., (email: rao@uta.edu)

block, the reconstructed block is:

$$\hat{A} = \sum_{n=1}^{q} \sqrt{\lambda(n)}u(n)v^T(n) \qquad q \le r \qquad (4)$$

The square error is equal to the sum of the discarded eigenvalues:

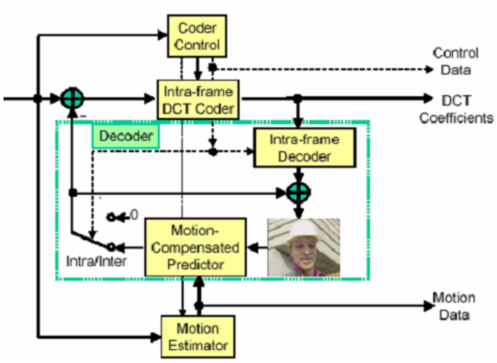
$$\sum_{m=1}^{r}\sum_{n=1}^{r}\left|A(m,n) - \hat{A}(m,n)\right|^2 = \sum_{n=q+1}^{r}\lambda(n) \qquad (5)$$

where $A(m,n)$ is the original pixel of the subblock and $\hat{A}(m,n)$ is the reconstructed pixel of the subblock. Thus, the energy of the reconstructed subblock is the energy in the q retained eigenvalues:

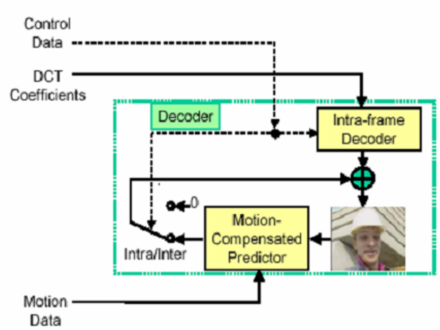
$$Energy = \sum_{n=1}^{q}\lambda(n) \qquad (6)$$

In conclusion, SVD decomposes the original data into two orthogonal matrices U and V^T, representing the eigenvectors, and a diagonal matrix $\Lambda^{\frac{1}{2}}$ representing the eigenvalues. Eigenvectors are usually encoded by vector quantization techniques, while eigenvalues are encoded by scalar quantization techniques.

3. Introduction of Video Coding and Characteristics of Video Data



(a)

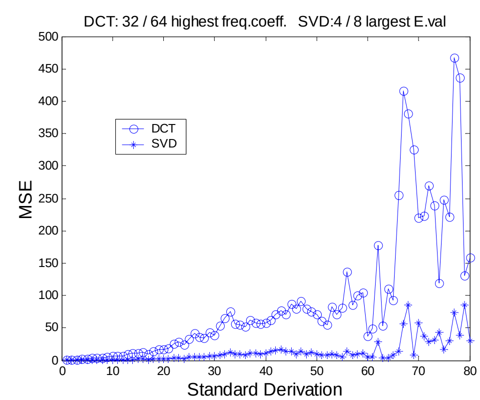


(b)

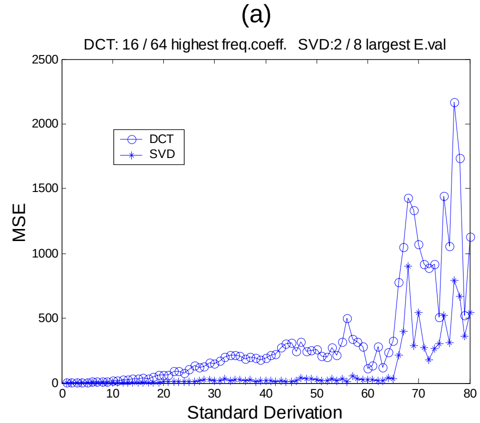
Fig.1. Block diagram of a video coding system. (a), Encoder (b), Decoder.

Video data are highly correlated data. In the most popular video coding techniques [3][13], the idea to compress video is to reduce both spatial correlation and temporal correlation of the original data. In the existing video coding standards (like H.263, H.264, MPEG2, MPEG4, etc), 2-dimentional DCT has been adopted to reduce spatial redundancy and inter-frame prediction is used to reduce temporal redundancy. When temporal prediction is applied, we call the coding mode INTER, as the compression techniques are applied to predict errors between adjacent frames; while if no temporal prediction is applied, the coding mode is called INTRA, as the redundancy reduction is applied to a single frame [3]. Fig.1 shows the block diagram of a typical video coding system.

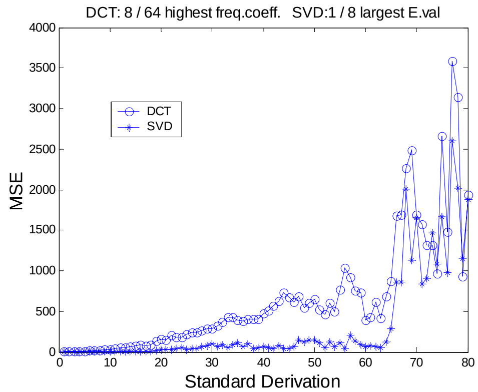
The standard deviation of a block describes the pixel activity for an INTRA block or movement level for an INTER block. High standard deviation indicates many abrupt changes in intensity in an INTRA block or heavy motion occurrence in an INTER block, while low standard deviation means nearly uniform intensity in an INTRA block or slight motion occurrence in an INTER block [11]. Applying DCT and SVD to blocks with different standard deviations will give us different results. DCT has the similar performance as SVD when the standard deviation of the block is low, but when the standard deviation is high, SVD performs much better than DCT [5].



(a)



(b)



(c)

Fig.2. Motivation analysis. (a),(b),(c), performance comparison between DCT and SVD by taking the same ratio of coefficients.

A series of tests have been done to illustrate this property. We take 45 frames from three different video sequences, Container, Foreman, and News [6]. They are QCIF sequences, with frame size 176x144, and 8 bpp. Each frame has been divided into 8x8 blocks. 17820 blocks are used in this testing, and the standard derivation of the testing blocks is concentrated in the range [0,80]. We applied DCT and SVD on these blocks. DCT coefficients are arranged in zigzag scan order, while SVD eigenvalues are arranged in descending order. In the three experiments, we dropped 50%, 25% and 12.5% of the lowest frequency coefficients in DCT or smallest eigenvalues in SVD domain respectively. There is no quantization and coding involved. Comparing the reconstructed block with the original block, we get the MSE for each block. Fig.2 shows the results in terms of mean square error versus standard deviation of the block, when different number of DCT coefficients or eigenvalues /eigenvectors retained in the reconstruction. It can be easily seen that for blocks with standard derivation less than 60, both transforms can achieve small MSE; while for the blocks with standard derivation greater than 60, the MSE caused by both transforms become high. Comparatively speaking, SVD performs better than DCT for high standard derivation blocks. Therefore, it can be concluded that SVD can be a compensating technique when DCT cannot provide good compression quality.

To achieve high compression, motion compensation technique with the search window 15x15 is applied. Generally speaking, INTER differential data will have relatively small standard deviation. While for the blocks with heavy motion, the standard deviation can be very high. Fig.3 shows the histogram of data in INTRA mode and differential data in temporal prediction INTER mode. By observing this distribution, it can be seen that both of them have high standard deviation blocks. Therefore both INTRA data and INTER differential data can benefit from the HDCTSVD technique.

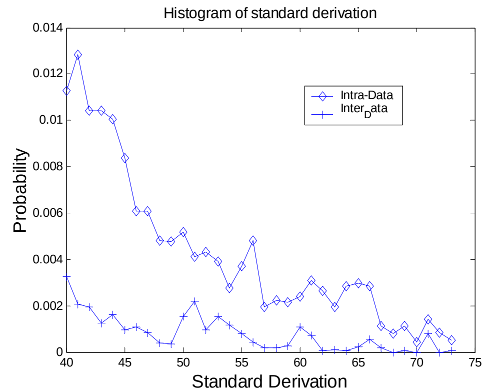


Fig.3. Distribution of INTRA data and INTER differential data at different standard derivations.

4. HDCTSVD VIDEO CODING SYSTEM

There are several approaches to combine DCT and SVD in image compression [6][12]. In these approaches, standard derivation is used to choose between SVD and DCT. Wongsawat extended the HDCTSVD [6] into color images and modified eigenvector quantization method by applying adaptive multistage vector quantization.

To design the HDCTSVD system for video compression, the approaches need to match the characteristics of video data well. The overall block diagram of the proposed HDCTSVD video coding system is shown in Fig 4. At the encoder side, both the INTRA data and the INTER differential data go through the HDCTSVD coder, either DCT or SVD coefficients will be transmitted. A HDCTSVD decoder is designed for the purpose of reconstruction. And the difference data between an inter frame and the previous reconstructed frame is processed by the HDCTSVD coder. The motion data and header information are transmitted directly. The decoder side is just the inverse process of the encoder side.

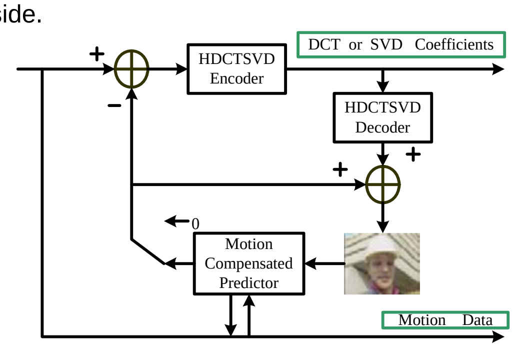


Fig.4. HDCTSVD encoder overview

The details of the proposed HDCTSVD encoder and decoder are illustrated in Fig.5.

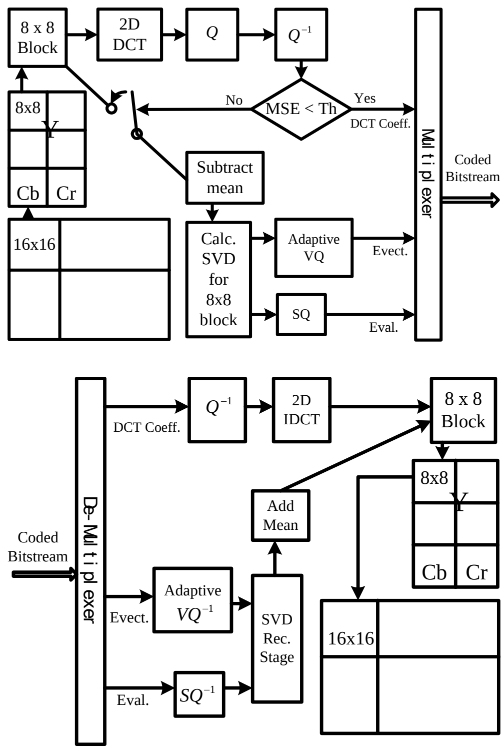
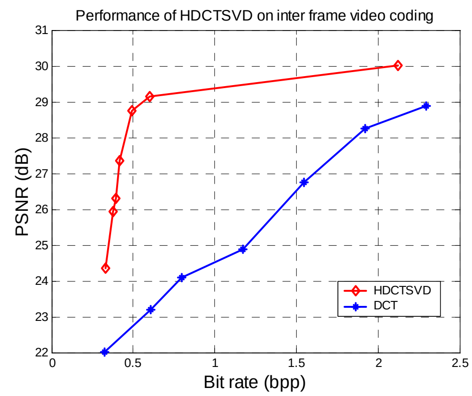


Fig.5. Block diagrams of HDCTSVD encoder and decoder.

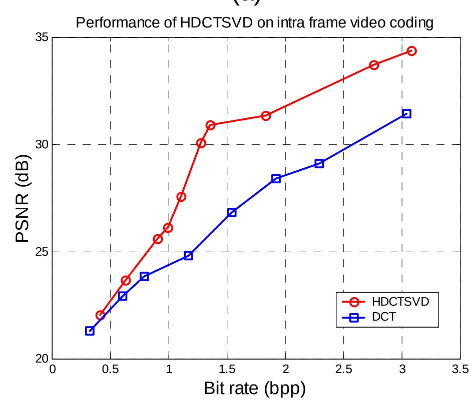
The system will nominally divide a video frame into 16x16 macro blocks. In each macro block, the data is organized as four luminance 8x8 blocks and two 8x8 chrominance blocks. 4:2:0 sampling pattern is used. The decision of whether DCT or SVD is chosen for a particular block is made in the DCT domain. We compare the DCT coefficients after quantization and energy dropping with the original DCT coefficients. If the MSE is smaller than our threshold, it means applying DCT gives good result at the given bit rate. In this case, the quantized DCT coefficients are transmitted. The first quantized DCT coefficient is coded using 8 bits, and the remaining quantized DCT coefficients are coded using 6 bits, each. If the MSE is greater than the threshold, the original block will go through the SVD branch. In the SVD branch, the mean of the block is subtracted and uniformly quantized using 8 bits. We calculate the eigenvalues and select a certain percentage of energy that is the same as the energy retained in the DCT domain. Then only a reduced number of eigenvectors are calculated. This decreases the computation cost. Scalar quantization is applied to the eigenvalues and adaptive vector quantization is applied to eigenvectors to achieve more compression [10].

At the decoder side, for each block, the first bit indicates which transform is used. If SVD is used, the eigenvalues will be recovered by inverse SQ. And matrices U and V^T are recovered by the inverse adaptive vector quantization. $U \sum V^T$ yields the reconstructed 8x8 block. If the flag indicating DCT is used, the quantizer DCT coefficients are retrieved.

5. SIMULATION RESULTS



(a)



(b)

Fig.6. Performance comparison of HDCTSVD and pure DCT for video sequence Carphone. (a) for INTER coding (b) for INTRA coding

Two sets of simulations have been performed. One is for INTRA frame coding, the other one is for INTER frame coding. The HDCTSVD video coding system is compared with a pure DCT video coding system, according to the rate-distortion performance. The formula to calculate distortion is:

$$PSNR = 10\log_{10}\frac{255^2}{MSE}dB \qquad (7)$$

$$MSE = \frac{1}{NM}\sum_{j=1}^{N}\sum_{i=1}^{M}[x(j,i)-\hat{x}(j,i)]^2 \qquad (8)$$

where $x(j,i)$ is the original value, and $\hat{x}(j,i)$ is the reconstructed value.

Figure 6 shows the results for INTRA frame coding and INTER frame coding respectively for Carphone sequence. The bit rate in DCT part is controlled by adjusting the percentage of energy that has been dropped. 8 bits are used to represent the first DCT coefficient in each block, while 6 bits are used for the rest of the transmitted DCT coefficients. The bit rate in the HDCTSVD system is controlled by applying SVD for a different number of blocks in a frame. Using 8x8 SVD requires high bit rate but provides good quality, while using 8x8 DCT requires lower bit rate but provides degraded quality. The computational cost varies at different bit rate since different percentage of blocks is processed by SVD. At very low bit rate, the computational cost is as low as a pure DCT system; while higher computational cost is needed when bit rate becomes higher as more SVD decomposition are applied. It can be seen that HDCTSVD system has better rate distortion performance than the pure DCT system for both intra frame coding and inter frame coding.

6. CONCLUSIONS

This paper proposes a new HDCTSVD video coding system. If DCT can compact the energy of the block well, DCT will be used; otherwise, SVD is applied. Adaptive VQ has been used to improve the compression in the SVD branch. Simulation results show that the proposed technique provides better performance than a pure DCT system for both INTRA frame coding and INTER frame coding. However, there is limitation using SVD for data compression because when eigenvalues are transmitted, the corresponding eigenvectors must also be transmitted. Although the proposed technique can outperform a pure DCT video coding system, it has limitation to outperform a more complicated DCT based video coding system. However, all the widely used video coding standards H.263, MPEG, H.264/AVC [13] integrate many efficient compression tools together with transform coding and entropy coding, the proposed HDCTSVD technique does not seem to have the potential to outperform those state of art video coding standards.

REFERENCES

[1] Jain, A.K., Fundamentals of Digital Image Processing, Englewood Cliffs, NJ: Prentice Hall, 1989.

[2] Ochoa, H., Rao, K.R.,"A low bit-rate hybrid DWT-SVD image coding system (HDWTSVD) for monochromatic images", IEEE, Symp. On Circuits and Systems, ISCAS, 2003, pp.472-475.

[3] Rao, K.R., Wang, J.J., Techniques and Standards for Image Video and Audio Coding, Upper Saddle River, NJ: Prentice Hall, 1996.

[4] Rao, K.R., Yip P.C., The Transform and Data Compression Handbook, Boca Raton: CRC Press, 2001.

[5] http://www.acticom.info/1494.html.

[6] Dapena, A., Ahalt, S., "A hybrid DCT-SVD image-coding algorithm", IEEE, Trans. on CSVT, 2002, pp.114-121.

[7] K. Sayood, Introduction to Data Compression. 3rd Edition, Mogan Kaufmann, 2006.

[8] Deprenttere, Ed. F., SVD and Signal Processing: Algorithms, Applications and Architectures, New York: North Holland, 1988.

[9] Aase, S.O., Husoy, J.H. and Waldemar, P., "A critique of SVD-based image coding systems", ISCAS'99, pp.13-16.

[10] Goldberg, M., Boucher, P.R., and Shlien, S., "Image Compression Using Adaptive Vector Quantization", IEEE Trans. on Communications, 1986, pp.180-187.

[11] Richardson, I.E.G., Video Codec Design, Chichester, West Sussex: Wiley, 2002.

[12] Wongsawat, Y, Ochoa,H, Rao, K.R., Oraintara, S., "A modified hybrid DCT-SVD image-coding system for color image", ISCIT, 2004, pp.766-769.

[13] Kwon, S., Tamhankar, A., Rao, K.R., "Overview of H.264/MPEG-4 Part 10", J. Visual Communication and Image Representation, vol. 17, pp.183-552, April 2006

Tong, Lin was born in Beijing, China in 1975. She received the B.S. degree from University of Science and Technology of China, in Automation Department, in 1999. From 1999 to 2000, she has been with the Legend Computer Company, Beijing, as a technical support engineer. She received the M.S. from Alfred University, NY, in Electrical Engineering in 2001. Since 2002, she has been working towards the Ph. D. at the University of Texas at Arlington. Her doctoral research is on rate control and HDCTSVD video coding. She worked as an intern in Philips Research Center in 2003 for five months.

Rao, K.R. received the Ph. D. degree in electrical engineering from The University of New Mexico, Albuquerque in 1966. Since 1966, he has been with the University of Texas at Arlington where he is currently a professor of electrical engineering. He, along with two other researchers, introduced the Discrete Cosine Transform in 1975 which has since become very popular in digital signal processing. He is the co-author of the books "Orthogonal Transforms for Digital Signal Processing" (Springer-Verlag, 1975), Also recorded for the blind in Braille by the Royal National Institute for the blind. "Fast Transforms: Analyses and Applications" (Academic Press, 1982), "Discrete Cosine Transform-Algorithms, Advantages, Applications" (Academic Press, 1990). He has edited a benchmark volume, "Discrete Transforms and Their Applications" (Van Nostrand Reinhold, 1985). He has coedited a benchmark volume, "Teleconferencing" (Van Nostrand Reinhold, 1985). He is co-author of the books, "Techniques and standards for Image/Video/Audio Coding" (Prentice Hall) 1996 "Packet video communications over ATM networks"(Prentice Hall) 2000 and "Multimedia communication systems" (Prentice Hall) 2002. He has coedited a handbook " The transform and data compression handbook," (CRC Press, 2001). Digital video image quality and perceptual coding, (with H.R. Wu),Taylor and Francis (2006). Introduction to multimedia communications: applications, middleware, networking, (with Z.S. Bojkovic and D.A. Milovanovic), Wiley, (2006). Some of his books have been translated into Japanese, Chinese, Korean and Russian and also published as Asian (paperback) editions. He has been an external examiner for graduate students from Universities in Australia, Canada, Hong Kong, India, Singapore, Thailand and Taiwan. He was a visiting professor in several Universities -3 weeks to 7 and1/2 months- (Australia, Japan, Korea, Singapore and Thailand). He has conducted workshops/tutorials on video/audio coding/standards worldwide. He has supervised several students at the Masters (59) and Doctoral (29) levels. He has published extensively in refereed journals and has been a consultant to industry, research institutes, law firms and academia. He is a Fellow of the IEEE.