International Journal of Engineering Research and Development
e-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com
Volume 7, Issue 2 (May 2013), PP. 58-62

Video Compression based on Singular Value Decomposition and Distributed Video Coding

S.Dhivya, M.Kannan
PG Scholar, Department of ECE, K.S.R. College of Engineering, Tiruchengode.
Assistant Professor, Department of ECE, K.S.R. College of Engineering, Tiruchengode.

ABSTRACT: Video compression with improved visual quality and with temporal redundancy reduction is a challenging task. Video coding standards such as H.261/3, MPEG-1/2/4 and H.264/AVC use motion estimation (ME) scheme to remove temporal redundancy which significantly improves the coding efficiency. However, the peak signal-to-noise ratio is less. A new video coding method is proposed based on Singular Value Decomposition (SVD) and Distributed Video Coding (DVC). The proposed scheme provides high peak signal-to-noise ratio(PSNR) with better visual quality and high compression efficiency compared to existing video codec's. The comparison of the proposed SVD coding plus distributed video coding scheme with the existing codec's shows its advantages and potential in applications. This method of video coding is suitable for scenarios such as mobile video phone, video surveillance systems, wireless video cameras, and disposable video cameras etc where high compression with enhanced visual quality are required.

Keywords – Compression efficiency, Distributed video coding, Peak-signal-to-noise ratio, Singular value decomposition, Video compression.

I. INTRODUCTION

Video is a sequence of still images representing scenes in motion. Video is characterized by huge amount of data. An uncompressed video (video conference) recorded with a resolution of 352x240 at 15 fps has a data rate of 30.41Mbps. As a result, transmission of raw video requires huge bandwidth. Also, the memory required to store this uncompressed video is enormous. These two drawbacks make it impractical to use raw video.

To reduce the transmission bandwidth and the storage requirements, video compression is done. During compression, the redundant information is removed. Compression algorithms typically exploit spatial and temporal redundancies present in a video. Temporal redundancy occurs when adjacent frames are highly correlated in a video. In MJPEG (Motion JPEG), each frame is coded independently using JPEG. Therefore MJPEG do not suffer from interframe dependence. However MJPEG exploits only spatial redundancy and does not exploit the temporal redundancies in the video and therefore it results in lower compression.

H.261 is the video compression standard and is a motion compression algorithm developed specifically for videoconferencing. Motion compensation involves working with groups of pixels to identify a group in the previous frame that best matches a group in the current frame. H.261 use prediction and motion estimation to reduce temporal redundancy. As the motion content of the images increases, the codec has to do more computations and usually has to give up on image quality to maintain frame rate.

H.264 offers higher compression rates and can be achieved by using both inter frame and intra frame coding. Spatial redundancy can be removed by intra frame coding and temporal redundancy can be removed by inter frame coding. The high compression efficiency is the high computational complexity due to the ME process used. In addition, with applications in which frames need to be edited and changed, all the frames affected by the changes have to be encoded again in an ME-based coder because of the inter frame dependence therefore, hybrid coding is not popular for situations requiring video editing and easy access to an individual frame.

Motion estimation and compensation is an efficient algorithm and a compression anywhere between 30:1 to 100:1 can be achieved. It is a challenge to remove temporal redundancy for getting high coding gain without significant increase in computational complexity. Temporal redundancy is exploited by the techniques such as block-based motion estimation (ME). Video coding standards like H.261/3, MPEG-1/2/4 and H.264/AVC implement this coding structure. Accordingly, coding efficiency is extensively improved.

However, due to the motion estimation process, computational complexity is high. The cost of high complexity are long computational time, high energy consumption, short battery life, and high cost of hardware implementation. In addition, the hybrid coding system causes inter-frame dependence in decompression. Therefore there is a need for algorithms, which achieve the necessary compression in real time with low power consumption. Hence in this paper, a new scheme is proposed for video coding based on Singular Value Decomposition and Distributed Video Coding which improves the visual quality of the video.

The rest of this paper is organized as follows. Section 2 gives the theoretical foundation of the proposed system and comparison with the existing relevant codecs. Section 3 describes the whole codec framework of the proposed SVD and DVC scheme. Section 4 presents the overall coding operation and experimental results with PSNR values, whereas Section 5 concludes this paper.

II. RELATED WORK

Asai proposed new video coding scheme optimized for high-resolution video sources[1] based on motion estimation and discrete cosine transform. Although coding efficiency is improved, computational complexity is high. Kyung-hee lee proposed a video coding method for H.264/AVC based on reference frame selection algorithm which reduced the complexity but frame loss is present.

Bahari proposed a low-power H.264 video compression[2] architectures for mobile communication based on pixel truncation by variable block-size where computational saving is achieved but increases the hardware design.

Fan Zhang proposed texture warping and synthesis based video compression[3]. A perspective motion model was employed to warp static textures, instead of encoding whole images or prediction residuals after translational motion estimation, and dynamic textures are created by utilizing texture synthesis. Using the features derived from complex wavelet transform, texture regions are segmented and further classified according to their spatial and temporal characteristics. In addition, to evaluate the quality of the reconstructed video, a compatible artifact-based video metric was proposed to prevent warping and synthesis artifacts. However this method had poor visual quality.

Gabriellini proposed an approach for intra-prediction in video coding[4] based on the combination of spatial closed-loop and open-loop predictions called as Combined Intra-Prediction, which enables better prediction of frame pixels that is desirable for efficient video compression and also addresses both the rate-distortion performance enhancement as well as low-complexity requirements that are imposed on codec's for targeted high-resolution content. The drawback of this method was that it does not provide better quality for complex video sequences.

Asai described a video coding scheme which used conventional block-based motion compensation plus discrete cosine transform coding approach that is suitable for hardware implementation of a high-resolution video codec. This scheme achieved significant development of coding efficiency, allowing the use of larger blocks for motion compensation and transform coding than conventional standards. However this method increased the computations.

III. VIDEO CODEC OPERATION

Video compression is done based on singular value decomposition plus distributed video coding method which improves the visual quality of the video sequences and also provide high peak signal-to-noise ratio (PSNR) compared with the existing video coding methods.

A. SVD process

Initially the given video sequence is divided into group of pictures(GOP) with n frames. Consider the ith frame denoted as F_i. The mean frame F_{mean} is found from n frames and the mean subtracted frames is obtained, i.e. $F_i' = F_i - F_{mean.}$ Then the mean-subtracted group of pictures(GOP) is divided into 8 × 8 group of blocks(GOBs) where each block in B is denoted as $B_{1,\ldots,}B_n$.

SVD is applied for the jth GOB to obtain the corresponding eigenvector matrices E_l and E_r and the group of 8 × 8 coefficient matrices $S_1 \ldots S_n$. The eigenvector matrix contains eight eigenvectors. The eigenvectors with small sub-indices are more important since they correspond to the top left coefficients which contain more information. These eigenvectors are often coded using vector quantization.

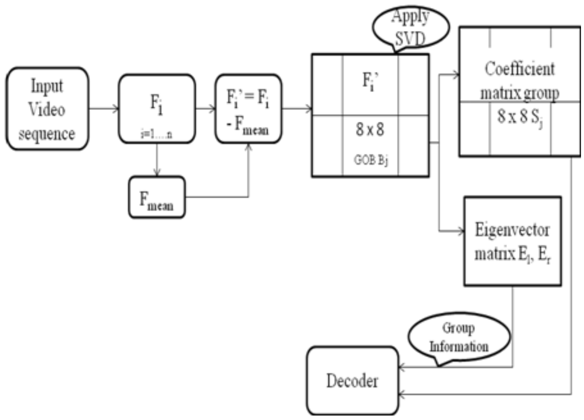

Fig 1. SVD encoder

After quantization, eigenvectors are transmitted as group information(GI) for each GOB. The longer the GOP, the less proportion the GI will occupy. With the increase in GOP size, there will be more high-magnitude SVD coefficients, which negatively impact the coding efficiency.

A rational GOP size is chosen, such that the portion of GI in each GOP and the number of high-magnitude coefficients are balanced. For video sequences, the optimal GOP size also ranges from 21 to 31.

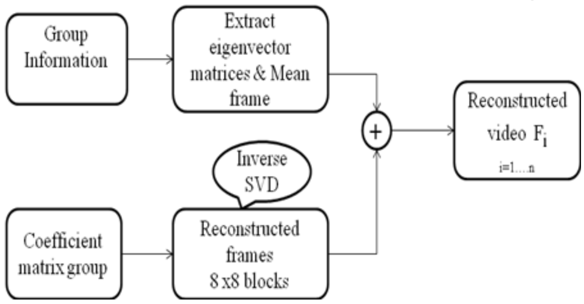

Fig 2. SVD decoder

In the decoder part, first, the group information is decoded; whereby the mean frame and eigenvector matrices are extracted. The reconstructed frames without the mean frame is obtained by applying inverse SVD. By adding the 8 × 8 blocks of frame with the reconstructed mean frame, we get the original frames F_i which forms the given video sequence. In this way, once the group information is successfully received, we are able to decode each frame independently.

B. DVC process

Distributed video coding (DVC) is a new video coding framework which is based on Wyner-Ziv (WZ) coding. In DVC, the complex task i.e., motion estimation is shifted to decoder, and decoder generates the side information, which is one type of input frame and the encoder only sends parity bits to correct errors or to improve its quality.

The incoming video sequence is first divided into group of pictures (GOPs). The first frame of each GOP is called key frame and conventional intra encoder is used to encode the key frame. The remaining frames in GOP are called as WZ frames and are encoded using distributed coding. Conventional intra encoder can be of H.264/AVC intra or JPEG2000. The WZ frames first go through transform such as discrete cosine transform (DCT) or discrete wavelet transform (DWT), depending on the conventional intra encoder chosen.

The transformed coefficients are then quantized (i.e., each coefficient is divided by an integer value) before splitting into bit planes. Typically, the result is a block in which most or all of the coefficients are zero, with a few non-zero coefficients. The video coding process produces a number of values that must be encoded to form the compressed bit stream.

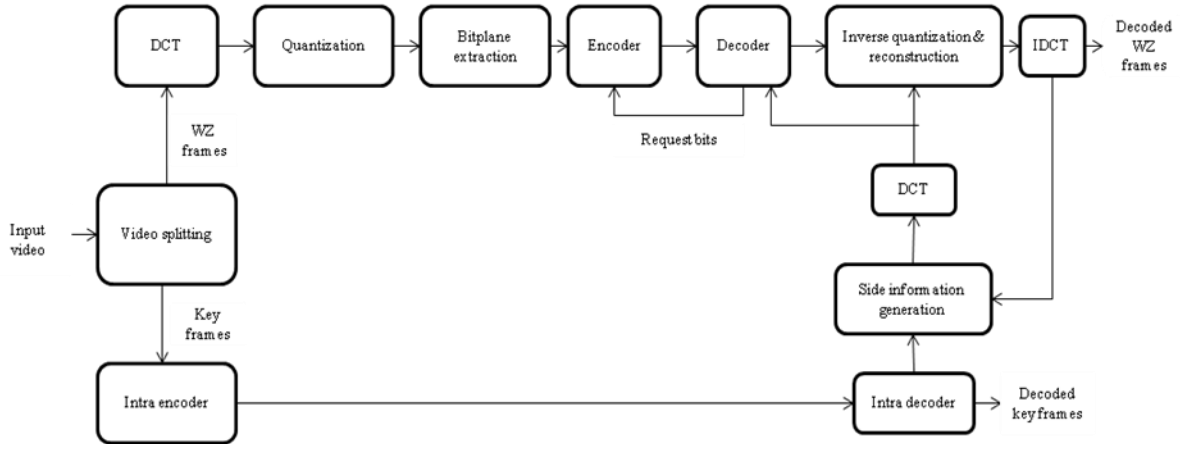

Fig 3. DVC codec

These values include quantized transform coefficients, information about the structure of the compressed data and the compression tools used during encoding, information about the complete video sequence. These values are converted into binary codes using any of the codes like Turbo, Trellis, or Arithmetic codes for encoding. After encoding, the bit stream can then be stored and/or transmitted. The codes are deployed here for error correction of frames, by sending the parity bits.

The parity bits stored in the buffer are sent to decoder based on the requests from the decoder. Therefore a feedback channel is needed between decoder and encoder. Side information (SI) is generated at the decoder by motion compensated interpolation or extrapolation of previously decoded frames. To decode the quantized bit planes, the side information is used along with the parity bits of the WZ frames requested via a feedback channel.

The reconstructed transform coefficients bands are obtained by the side information together with the decoded quantized bit planes. By applying inverse discrete cosine transform, the decoded WZ frames can be obtained and the key frames go through conventional intra encoder and decoder.

IV. PROPOSED TECHNIQUE

The proposed video coding scheme involves compressing the video sequence first by using SVD process. The output video obtained from the SVD process is further compressed by distributed video coding process which provides highly compressed video with better visual quality as described in Fig.4.

Also, the proposed system gives high peak signal-to-noise ratio when compared to existing video coding standards.

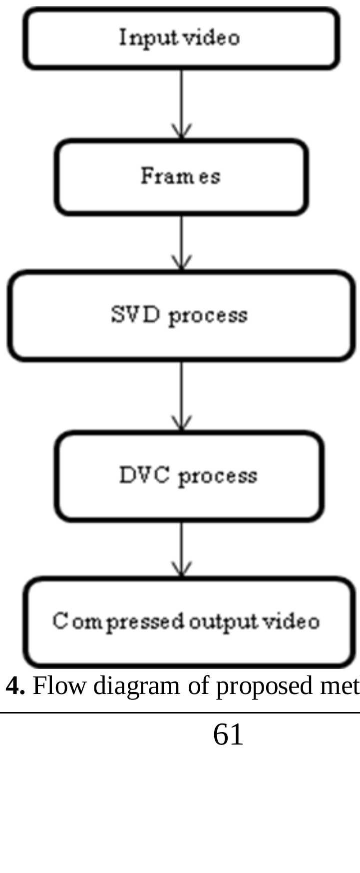

Fig 4. Flow diagram of proposed method

Initially, the input video is divided into group of frames. Then SVD process is applied to that frames which in turn produces compressed output video with certain PSNR value and compression ratio. Next the output video is again compressed using DVC method which produces highly compressed video with high PSNR improvement than the other codec's. Different video codec's are compared based on their performance in terms of peak-signal-to-noise ratio as shown in Fig. 5 where the proposed SVD and DVC video coding technique provides high PSNR value.

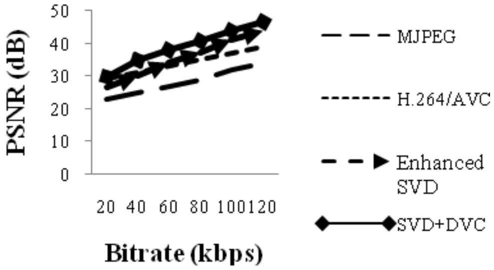

Fig 5. Performance comparison of different video codec's

V. CONCLUSION AND FUTURE WORKS

In this paper we have proposed a new scheme for video coding based upon Singular Value Decomposition (SVD) plus Distributed Video Coding (DVC). Compared with existing hybrid video codec's, our codec gives high coding efficiency in terms of peak signal-to-noise ratio(PSNR) by 3-8 dB for the given aero plane sequence compared with existing SVD method and hybrid video codec's. Therefore, our scheme provides better visual quality and this video coding method is suitable for applications such as mobile video phone, video surveillance systems, wireless video cameras, and disposable video cameras etc where high compression with enhanced visual quality are required. Hence developing this video coding method for complex videos with further complexity reduction will be a future challenge.

REFERENCES

[1]. Asai.K, Murakami.T, Yamagishi.S, Minezawa.A and Itani.Y, (2011), "New Video Coding Scheme Optimized for High-Resolution Video Sources," *IEEE Trans. Signal Process., vol. 5, no. 7,* pp. 1290–1297.

[2]. A. Bahari, T. Arslan, and A. T. Erdogan, (2009), "Low-power H.264 video compression architectures for mobile communication," *IEEE Trans. Circuits Syst. Video Technol., vol. 19, no. 9,* pp. 1251–1261.

[3]. Fan Zhang, and Bull.D.R, (2011), "A Parametric Framework for Video Compression Using Region-Based Texture Models," *IEEE Trans., Signal Process., vol. 5, no. 7,* pp. 1378–1392.

[4]. Gabriellini.A, Flynn.D, Mark.M and Davies.T., (2011), "Combined Intra- Prediction for High-Efficiency Video Coding," *IEEE Trans. Signal Process., vol. 5, no. 7,* pp. 1282–1289.

[5]. Zhouye Gu, Weisin Lin, Bu-sung Lee and Chiew Tong Lau, (2012), "Low-Complexity Video Coding Based on Two-Dimensional Singular Value Decomposition," *IEEE Trans. Image Process., vol.21, no.2,* pp. 674–687.

[6]. J. Liu, S. C. Chen, Z.-H. Zhou, and X. Y. Tan, (2010), "Generalized low-rank approximations of matrices revisited," *IEEE Trans. Neural Netw., vol. 21, no. 4,* pp. 621–632.

[7]. C. Brites, J. Ascenso, Q. Pedro, and F. Pereira, (2008), "Evaluating a feedback channel based transform domain Wyner-Ziv video codec", *IEEE Trans. Signal Process., vol. 23, no. 4,* pp. 269–297.

[8]. B. Girod, A. Aaron, S. Rane, and D.Rebollo-Monedero, (2005), "Distributed video coding", *IEEE, vol. 93, no. 1,* pp. 71-83.