Motion Compensated JPEG2000 based video compression algorithms

T. Tuithung

Department of Computer Science and Engineering,
Indian Institute of Technology Kharagpur,
Kharagpur 721302, India
Fax: 91-3222-282206/282700
E-mail: themri@cse.iitkgp.ernet.in

S.K. Ghosh*

School of Information Technology,
Indian Institute of Technology Kharagpur,
Kharagpur 721302, India
Fax: 91-3222-282206/282700
E-mail: skg@iitkgp.ac.in
*Corresponding author

Jayanta Mukherjee

Department of Computer Science and Engineering,
Indian Institute of Technology Kharagpur,
Kharagpur 721302, India
Fax: 91-3222-282206/282700
E-mail: jay@cse.iitkgp.ernet.in

Abstract: A new approach of Motion-Compensated JPEG2000 (MCJ2K) video compression is proposed in this paper. It uses simplified MPEG-2 fundamentals along with JPEG2000 encoding for compression of the image frames. MCJ2K generates differential images using motion compensation algorithms and store the motion vectors in a customised header. The images are then compressed using JPEG2000 based scheme. MCJ2K have been studied on QCIF and CIF formated videos for various rate control mechanism. The proposed scheme produced encouraging results compared to standard MPEG-2 and MJPEG2000 schemes, especially at higher bit rates.

Keywords: video compression; JPEG2000; MJPEG2000; motion compensation; motion vectors; rate-control.

Reference to this paper should be made as follows: Tuithung, T., Ghosh, S.K. and Mukherjee, J. (2008) 'Motion Compensated JPEG2000 based video compression algorithms', *Int. J. Signal and Imaging Systems Engineering*, Vol. 1, Nos. 3/4, pp.197–212.

Biographical notes: T. Tuithung did her BTech in Computer Science and Engineering and MTech in Information Technology from North Eastern Regional Institute of Science and and Technology (NERIST), Itanagar, India. Currently she is doing her PhD in Computer Science and Engineering at Indian Institute of Technology, Kharagpur, India. Her areas of interest are image processing, video compression, algorithms.

S.K. Ghosh has done his PhD in the year 2002 from Indian Institute of Technology (IIT), Kharagpur, India from the Department of Computer Science and Engineering. He is presently working as an Associate Professor in the School of Information Technology, Indian Institute of Technology, Kharagpur. Prior to IIT Kharagpur, he worked for Indian Space Research Organisation (ISRO), Department of Space, Government of India, for about eight years in the field of Satellite Remote Sensing and GIS. His research interest includes image processing, computer networking, education technology, remote sensing and GIS.

Jayanta Mukherjee received his BTech, MTech, and PhD Degrees in Electronics and Electrical Communication Engineering from the Indian Institute of Technology (IIT), Kharagpur in 1985, 1987, and 1990, respectively. He is presently a Professor in the Department of Computer Science and Engineering, IIT Kharagpur. He was a Humboldt

Research Fellow at the Technical University of Munich in Germany for one year in 2002. He also has held short term visiting positions at the University of California, Santa Barbara, University of Southern California, and the National University of Singapore. His research interests are in image processing, pattern recognition, computer graphics, multimedia systems and medical informatics. He has published over 100 papers in international journals and conference proceedings in these areas. He served in various technical program committees of international conferences in related areas. He received the Young Scientist Award from the Indian National Science Academy in 1992. He is a Senior Member of the IEEE. He is a fellow of the Indian National Academy of Engineering (INAE).

1 Introduction

Over the years Discrete Cosine Transform (DCT) based encoding algorithms such as, MPEG-2 (Senda, 1995; Tudor, 1995), MPEG-4 (Sikora, 1997), H.264 (Wiegand et al., 2003; Wiegand and Sullivan, 2005) etc. have been proposed for video compression. The core compression algorithm used in such algorithms is the baseline JPEG (Wallace, 1991) encoding algorithm. A typical block diagram outlining these approaches are shown in Figure **??**.
All of these approaches have the following commonalities:

- Use of block DCT.
- The video stream is segmented into a Group Of Pictures (GOP), each starting with an Intra (I) frame, which is encoded by the usual static image compression algorithm such as block DCT based base-line JPEG compression algorithm.
- Block-wise (or Macroblock-wise) motion compensated error prediction. Prediction may be made from previous and future frames. However, in our simplified framework we have considered only predictions from the previous frame. This type of predicted error frame is termed as a 'P' frame. For each predictive frame, there is a set of motion vectors and corresponding error blocks, which are also further compressed using again the block-DCT based compression scheme, almost similar to baseline JPEG compression algorithm.
- In the above schemes, typical size of a DCT block is 8×8 and that of a macro-block for motion compensation is 16×16. Colour images are represented in $Y - C_r - C_b$ colour space and usually C_r and C_b components are subsampled at a ratio of $4:1:1$ before the application of DCT.

However one should note that there are variations on the motion compensated prediction algorithms, block and macro block sizes, intra-block prediction, object based video streaming, rate control strategies etc. among these approaches.

Other than DCT based approaches, wavelet based approaches are also increasingly being pursued in recent years. Specially, wavelet based static image compression schemes were found to be having better rate distortion performance over the DCT based JPEG (Wallace, 1991) compression scheme. Subsequently, the JPEG2000 (Taubman and Marcellin, 2002; Taubman, 1999; Skodras et al., 2000; Antonini et al., 1992) standard has incorporated wavelet based static image compression scheme. In wavelet based approaches usually 3-D wavelets (Xu et al., 1994; Taubman and Zakhor, 1994) are applied to a GOP. In some work (Leung and Taubman, 2005; Xu et al., 2002), motion compensation is used to improve the performance. However the usage of 3D wavelets over the GOP causes temporal artifacts (equivalent to jittering etc.) at its boundaries (Liang et al., 2005). In our proposed algorithm (Rath et al., 2006), we restrict ourselves to 2-D wavelets only. The scheme is similar to the DCT based schemes, where Intra frames (I-frame) and motion compensated Predicted frames (P-frame) are encoded using similar 2-D transform domain techniques. Like MPEG (LeGall, 1991), in our proposed algorithm we have used JPEG2000 (instead of JPEG as the core compression technique in MPEG) towards this. It may be noted that JPEG2000 has already been used in the MJPEG2000 (Dufaux and Ebrahimi, 2003) video compression standard, where all the frames are intra frames and no motion compensation is performed. The MJPEG2000 usually provides superior performance over MPEG-2 at very high bit rates. Our proposed scheme is a generalisation over MJPEG2000, as it allows motion compensation. In many cases, the proposed approach (Tuithung et al., 2006) is found to be having better rate distortion performance than the MJPEG2000 at very high bit rates. There are a few advantages on using JPEG2000 over the JPEG-like schemes for encoding prediction errors which have been adopted in the proposed algorithm:

1. Since the wavelet transform can be applied over the entire image, a JPEG2000 image does not exhibit the blocky artifacts common in highly compressed traditional JPEG images.
2. Scalability and multi-resolution representation make JPEG2000 as ideal choice for scalable video encoding scheme.
3. Precise frame wise rate control is possible while using JPEG2000. This is significant in the context

of errors being modelled as *Gaussian distribution* (Montgomery et al., 2001). It may be noted that MPEG-2 adopts macroblock wise rate control (for P frames), which makes error encoding algorithm little different than the standard baseline JPEG encoding scheme used for I frames. In our proposed approach, rate control mechanisms for both I and P frames remain same, which makes the encoder much simpler for implementation.

4 The other advantage of JPEG2000 scheme is that the scope of making the scheme lossless or semi-lossless by using reversible wavelets. This makes the proposed approach very attractive for the higher bit rate video transmission, specially the medical videos, where physicians demand very high quality of the decompressed video.

The proposed algorithm has the following distinct features:

1 It uses similar motion compensation algorithm that is used in MPEG-2. However, in the present work, we have demonstrated results using only full pixel motion compensation and also using only predictive frames (P frames).

2 To make the scheme semi-lossless, we have used the feed-back from the compressed video stream in the prediction of errors. We have also used reversible wavelets for encoding error frames.

3 Precise rate control is possible by allocating fixed bits for I and P frames. For efficient rate distortion performance, we have adopted dynamic rate control and distortion adjustment strategies.

The paper is organised in the following sections. In Section 2 we provide the basic outline of the proposed scheme and its performance on fixed bit allocation. We discuss about the rate and distortion control strategies in Section 3, which make the proposed scheme more efficient. Section 4 presents a brief overall discussion of the performance of the scheme. Finally, the conclusion is drawn in Section 5. Experimental results and graphs are provided in the respective sections. Further, we have considered the videos in YUV colour model for CIF and QCIF formats throughout this paper.

2 Base compression scheme

In our scheme, we have used JPEG2000 for compressing the intra frames or I frames as well as the error frames obtained through motion compensation. The I frame compression is the usual JPEG2000 still image compression. However, while compressing error frames (P or B), error values are computed for each macroblock of Y (16×16 pixels for each block), U and V (8×8 blocks for each of U and V) components by estimating motion vectors with respect to the previous (or/and next) frame(s). Motion compensation is made over the decompressed previous frames. Then the whole error frame is subjected to the JPEG2000 compression scheme. A typical block diagram of the scheme is shown in Figure **??**.

As JPEG2000 works with precise rate control (within a range of target bit rates), initially we have used fixed bit per pixel allocation for both I and P frames. A parameter α has been considered indicating the relative proportions (fractions) of bit investment on I and P frames. Let R be the target rate in bits per second (bps), G be the GOP (with single I frame and $(G-1)$ P frames) and f be the frame per second. Let h and w be the height and width of a frame. The bit per pixels for I frame (β_I) and for P frames (β_P) are given by:

$$\beta_I = \alpha \cdot \frac{R \cdot G}{f \cdot w \cdot h} \tag{1}$$

$$\beta_P = (1-\alpha) \cdot \frac{R \cdot G}{f \cdot w \cdot h \cdot (G-1)}. \tag{2}$$

For studying the performance of the compression schemes, we have used the Peak Signal to Noise Ratio (PSNR) measure. Let $I_s(x,y)$ denote the pixel value at (x,y) for the sth component ($s \in \{Y,U,V\}$) of the original frame and $I'_s(x,y)$ denote the corresponding recovered value after decompression. Then the PSNR is defined as:

$$\text{PSNR} = 20\log_{10}\left(\frac{255}{\sqrt{\frac{\Sigma_{\forall s \in \{Y,U,V\}}\Sigma_{\forall x}\Sigma_{\forall y}(I_s(x,y)-I'_s(x,y))^2}{\frac{3}{2}\cdot h \cdot w}}}\right). \tag{3}$$

For the entire video, the *average* PSNR value of all the frames has been considered. The effect of using reversible and irreversible wavelets has also been studied. The JPEG2000 compression has been implemented using the *Kakadu*[1] software. The number of levels of decomposition for I frame is taken as 5 and for P frame the value is 0. These values are chosen empirically. It can be observed that the use of reversible wavelets leads to higher PSNR values at higher bit rate. Experimental results have been shown in Figures **??** and **??**, with the proposed MCJ2K using reversible wavelets.

In Figures **??** and **??** we present the PSNR vs. *bit-rate* performances for the proposed schemes for different values of α using reversible and irreversible wavelets on QCIF and CIF videos respectively. The graph show that the MCJ2K using reversible wavelets provides higher PSNR values over MCJ2K with irreversible wavelets at high bit-rates. The PSNR values at typical bit-rates using reversible wavelets with different values of α for QCIF and CIF videos, are shown in Tables **??** and **??** respectively. It is observed that lower values of α performs better at high bit-rate. In Tables **??** and **??** the PSNR values are compared with those obtained by MPEG-2 and MJPEG2000 for QCIF and CIF videos respectively. One may observe that the use of feedback from the compressed stream improves the performance remarkably. The use of feedback makes the lossless reconstruction of P possible. In fact, the PSNR values of recovery of error frames and recovery of P frame become equivalent in such cases. This fact is stated in the Lemma 2.1.

Lemma 2.1: *The PSNR value of the error recovery remains as same as the PSNR value of the decompressed frame, when errors are computed from the previous decompressed frame.*

Proof: Let $\widehat{I}^{(d)}_{\text{prev}}$ be the motion compensated values from the decompressed previous frame. Let I_{curr} be the original current frame. Then the error values for the current frame are obtained as:

$$E = I_{\text{curr}} - \widehat{I}^{(d)}_{\text{prev}}. \tag{4}$$

After decompression, let the error values be $E^{(d)}$. Hence, the $I^{(d)}_{\text{curr}}$ after decompression could be expressed as:

$$I^{(d)}_{\text{curr}} = \widehat{I}^{(d)}_{\text{prev}} + E^{(d)}. \tag{5}$$

Hence,

$$\begin{aligned} |E - E^{(d)}| &= |(I_{\text{curr}} - \widehat{I}^{(d)}_{\text{prev}}) - E^{(d)}| \\ &= |(I_{\text{curr}} - \widehat{I}^{(d)}_{\text{prev}}) - (I^{(d)}_{\text{curr}} - \widehat{I}^{(d)}_{\text{prev}})| \\ &= |I_{\text{curr}} - I^{(d)}_{\text{curr}}|. \end{aligned} \tag{6}$$

Hence the PSNR values for error recovery and frame recovery are same. □

From the above lemma, we can observe that if the error recovery is *lossless*, the decompressed frame is also *lossless*. This motivates us to use the reversible wavelets. We have observed that usage of reversible wavelet slightly improves the scheme at the higher bit-rate. However, as the dynamic ranges of errors increase by two folds, we have scaled the values by $\frac{1}{2}$ before compression. For this reason, our scheme behaves as a semi-lossless scheme at higher bit rates.

2.1 Gaussian modelling of error frames

Lemma 2.1 is also significant as it converts the problem of compressing individual frames to the compression of error frames. It would be interesting to study how best these errors are modelled by a Gaussian distribution. In that case we would be able to use the well known rate distortion relationship of encoding of Gaussian distributed values, as described by the following lemma.

Lemma 2.2: *The values of a random variable following Gaussian distribution with the standard deviation σ would require $R(D)$ bits per symbol at a distortion (the Mean Square Error (MSE)) of D where $R(D)$ holds the following relationship with D*

$$R(D) = \begin{cases} \frac{1}{2}\log_2 \frac{\sigma^2}{D} & 0 \le D \le \sigma^2 \\ 0 & D > \sigma^2 \end{cases}. \tag{7}$$

Proof: Refer He and Mitra (2005). □

It was observed that in most cases the distributions of error values do not pass the significance levels (of 1%) of *Chi-Square* (χ^2) *Goodness of Fit test* (Montgomery et al., 2001) or the *Kolmogrov-Smirinov Test* (Montgomery et al., 2001). However, the distributions are observed as symmetric and highly peaked around its mean. Skewness and curtosis are measures of asymmetry and sharpness of the height relative to a normal distribution respectively. In Tables 1–3, we provide the average skewness and average curtosis measures (along with their standard deviations). They indicate that the Lemma 2.2 would provide an upper bound on the rate for a given distortion. We have made use of this fact in designing the rate control strategy.

2.2 Varying macroblock sizes

We have also studied the effect of different macroblock size on the performance of the proposed scheme. If errors are generated by larger macroblocks, it is expected the standard deviations of the error frames would be higher, on the other hand lower macroblocks should have smaller standard deviations. But there are increasing overheads on encoding the motion vectors for lower macroblock sizes. If we assume the bit assignment on encoding full pixel motion vectors is same (say mv_b for each motion vector, in our case $mv_b = 8$ bits), a macroblock of size $b_w \times b_h$ would require $\lceil \frac{w \cdot h}{b_w \cdot b_h} \rceil$ number of motion vectors. If σ_i, $i = 1, 2, \ldots, N$ is the standard deviation for the ith error frames the average bit requirement E_b for an error frame is given by:

$$E_b = \frac{1}{N}\sum_{i=1}^{N} \log_2(\sigma_i) + \frac{mv_b \cdot \lceil \frac{w \cdot h}{b_w \cdot b_h} \rceil}{w \cdot h}. \tag{8}$$

In Table 4, we have shown the theoretical average bit requirements for varying macroblock size. In Figure **??**, PSNR vs. rate curves are drawn for varying macroblock sizes. It is empirically observed that the macroblock size 16×16 yield better performance empirically. Theoretical bounds on the average bit requirement are also lower for different videos in such case. In Figure **??**(a), macroblock size 32×32 is showing better performance while in Figure **??**(b) and (c), macroblock size 16×16 shows better performance at lower bit-rates. Moreover MPEG-2 also uses a macroblock size of 16×16 and since our scheme uses the fundamentals of MPEG-2, thus the macroblock size 16×16 is chosen for implementing our scheme.

Table 1 Results on measures of skewness and curtosis of error frames for *Container. cif* video at 2449 Kbps

Frame component	*Avg_Skewness*	*Avg_Curtosis*	*σ Skewness*	*σ Curtosis*	*Min_Curtosis*
Y	0.45	41.82	1.53	42.29	9.53
U	−0.49	7.90	1.55	47.17	3.05
V	−0.40	6.88	0.51	6.53	4.01

Table 2 Results on measures of skewness and curtosis of error frames for *News.cif* video at 2640 Kbps

Frame component	*Avg_Skewness*	*Avg_Curtosis*	*σ Skewness*	*σ Curtosis*	*Min_Curtosis*
Y	0.39	96.08	1.98	46.56	5.90
U	−0.52	48.04	1.41	49.52	4.38
V	−0.01	57.19	1.42	54.97	4.31

Table 3 Results on measures of skewness and curtosis of error frames for *Foreman.cif* video at 3043 Kbps

Frame component	*Avg_Skewness*	*Avg_Curtosis*	*σ Skewness*	*σ Curtosis*	*Min_Curtosis*
Y	−0.03	17.20	0.70	10.42	4.75
U	−1.16	47.16	4.58	165.36	2.79
V	0.11	22.83	1.34	28.01	2.91

Table 4 Average bit requirement for Y-component with different macroblock size (8×8, 16×16 and 32×32) at higher bit-rate for CIF videos (Container, News, Foreman)

Video	*Vb_8*	*Vb_16*	*Vb_32*
Container	0.80	0.72	**0.70**
News	**0.95**	1.05	1.24
Foreman	**1.82**	1.85	1.99

2.3 Varying length of GOP

We have also observed the effect of variation of the length of GOP. In our scheme, as propagation of errors is arrested by the feedback from the compressed stream, the effect on variation of the GOP length is less (refer Figure **??**). In fact, the scheme works better for larger GOP length, as long as, there is a gain in rate for encoding a P frame compared to that of an Intra encoding. However, larger GOP will also cause, increasing values of standard deviations for the error frames at its trailers. This puts a restriction on the length of the GOP. This leads us to design a compression scheme for adaptive GOP. If the standard deviation of an error frame exceeds a threshold (called as σth in our work) the GOP is initialised. In Figure **??**, we have demonstrated the performances for adaptive GOPs. In Table 6, we have also provided the average lengths of GOP at varying σth.

3 Rate control

As discussed earlier, the use of JPEG2000 provides an opportunity for precise rate control. Moreover, the Gaussian nature of error frames motivates us to use the following lemma for the rate control of the MCJ2K scheme with fixed GOP.

Lemma 3.1: *Let $\overline{\beta_P}$ be the average bit per pixel for encoding a P frame. Let σ_i, $i = 1, 2, \ldots, (G-1)$ be the standard deviation of the ith P frame of a GOP of the length G. Then, the optimal bit per pixel at which the ith P frame would be encoded is given by:*

$$\beta_{P_i} = \overline{\beta_P} + \frac{1}{2}\log_2 \frac{\sigma^2}{{\sigma^*}^2} \tag{9}$$

where

$$\sigma^* = \left(\Pi_{i=1}^{(G-1)} \sigma_i\right)^{\frac{1}{(G-1)}} \tag{10}$$

with the optimum rate allocation the ith frame distortion D_i becomes,

$$D_i = {\sigma^*}^2 2^{-2\overline{\beta_P}}. \tag{11}$$

Proof: Refer Gray and Neuhoff (1998). □

3.1 Rate control of the MCJ2K

The above lemma motivates us to design the following rate control algorithm for the MCJ2K.

Algorithm *Rate_Control_MCJ2K*

Encoding Parameters: Length of GOP (G), relative bit investment for I frame (α), rate (R).

1 Determine bit per pixel for encoding I frame (β_I) and average bpp for encoding a P frame ($\overline{\beta_P}$) using equations (1) and (2).

2 Compute the geometric mean of the standard deviations (refer equation (10)) for the next *GOP*

from the current one and use it for the purpose of optimal bit assignment following the (refer equation (9)).

3 Compute the bpp for ith P frame (β_{P_i}) following the Lemma 3.1 and encode them accordingly.

End *Rate_Control_MCJ2K*

In this algorithm, P frames for the first GOP are encoded using the average bpp as computed from (refer equation (2)). The geometric mean of the standard deviations of all the P frames is also computed and used for encoding P frames of the next GOP using (refer equation (9)). The process is repeated for all the subsequent GOPs in the same way. The performance is shown in the Figure **??**. From the Table **??**, it is observed that $\alpha = 50\%$ gives better PSNR value. As such σ value at 5 and α at 50% are typically chosen for studying the performance of *Rate_Control_MCJ2K*.

At lower bit-rate, the achieved bit-rate is almost as same as the given target bit-rate but at higher bit-rate the achieved bit-rate gets saturated even if the target bit-rate is increased (refer Figure **??**).

3.2 Distortion control of the MCJ2K

Interestingly, the MCJ2K scheme is more suitable for controlling distortion or quality of the decompressed video. In this case, given a target PSNR Ω, the MCJ2K attempts to encode at the optimal rate for achieving the target rate. The distortion D is computed from Ω as follows:

$$D = 255^2 \times 10^{-\frac{\Omega}{10}}. \qquad (12)$$

Then following the Lemma 2.2, P frames are encoded. The performance of this strategy with different bpp's used for encoding I frames is shown in Figure **??**. The achieved PSNR values given the target PSNR values are also shown in the Figure **??**.

However, one may observe that there is a considerable gap between the achieved PSNR values and the target PSNR values. To reduce these gaps, we have adjusted the target PSNR values after encoding of each frame and measuring its PSNR with the original frame. Let the achieved PSNR value for the ith frame be $\Omega^{(d)}$. Then the target PSNR value for the $(i+1)$th frame is computed as:

$$\Omega^{(i+1)} = 2\Omega^{(i)} - \Omega^{(d)}. \qquad (13)$$

For the first P frame of a GOP, the target PSNR is initialised to Ω. The PSNR vs. rate curves with this distortion control strategy (referred as *Quality_Control_MCJ2K*) are shown in the Figure **??**. In Figure **??**, achieved PSNR values against the target PSNR values are also shown. It may be noted that due to the PSNR feedback, the achieved bit-rate gets saturated faster as compared to the scheme without using this feedback (refer Figure **??**).

4 Discussion

In the *base compression* scheme of MCJ2K (refer Section 2), we have experimented the bit-investment on I and P frames, using a parameter α. It was found empirically that lower value of α ($\alpha = 0.5$), produces better PSNR value at the higher bit-rate. We also experimented on using both reversible and irreversible wavelets and found that reversible wavelets outperforms irreversible wavelets (refer Lemma 2.1).

Using Gaussian modelling of error frames (refer Lemma 2.2), MCJ2K scheme is subjected to varying length of GOP base on a threshold value σ. We also found that GOP size increases with higher value of σ (refer Table 9).

Table 9 Average length of GOP for different σth for CIF videos

Video	$\sigma = 4.0$	$\sigma = 5.0$
Container	150	300
News	30	75
Foreman	5	10

In *rate control* of MCJ2K (refer Section 3), Gaussian modelling of error frames along with α (the relative bit-investment for I frame), has been implemented and found that it performs better than our base compression scheme at higher bit-rate.

In *distortion control* (quality control without feedback) of MCJ2K scheme (refer Section 3.2), given a target PSNR γ, the proposed MCJ2K algorithm attempts to encode at the optimal rate for achieving the target rate using equation (11) and Lemma 2.2. We also have quality control with feedback in which the target PSNR values are adjusted after encoding each frame (refer equation (12)). The effect of having feedback can be observed from Figures **??** and **??**.

From the Table **??**, it can be observed that quality control with feedback scheme of MCJ2K produces better PSNR values than that of quality control without feedback and rate control schemes of MCJ2K (refer Table **??**). However, it may be noted that quality control with feedback saturates faster than both quality-control without feedback and rate control schemes of MCJ2K (refer Table **??**). It is also worth mentioning that due to lower bound constraints, in both types of quality control schemes (with and without feedback), the scheme works from a certain high bit-rate only as compared to the rate control scheme of MCJ2K (refer Figures **??**, **??** and **??**).

Finally, in optimal bit allocation strategy, the MCJ2K scheme adjusts α for bit-investment for both the I and P frames. It behaves like MCJ2K rate control scheme at lower bit-rate but at the higher bit-rate it tends to behave like MJPEG2000, where the GOP size tends to become one. Hence this scheme works better than MPEG-2 at very high bit-rate (refer Figure **??**).

5 Conclusion

In this paper we have proposed a novel video compression algorithms which use JPEG2000 as the core compression engine for compressing *intra* frames as well as the motion compensated *prediction* error frames. The rate and quality control algorithms for the MCJ2K scheme have been developed and the MCJ2K is found to perform better than MPEG-2 and MJPEG2000 at higher bit-rates for many cases. Extensive experimentations have been carried out and the results are compared with standard schemes. It is also found that MCJ2K performs better for low motion videos as compared to high motion videos. The proposed MCJ2K scheme can find its applications like HDTV video conferencing, medical video transmission etc., where high quality video is required, even at the cost of high bit-rate.

References

Antonini, M., Barlaud, M., Mathieu, P. and Daubechies, I. (1992) 'Image coding using wavelet transform', *IEEE Transactions on Image Processing*, Vol. 1, No. 2, pp.205–222.

Dufaux, F. and Ebrahimi, T. (2003) 'Motion JPEG2000 for wireless applications', *Signal Processing Laboratory (EPLF), Proceedings of First International JPEG 2000 Workshop, Motion Analysis and Image Sequence Processing*, IEEE, Lugano, Switzerland, July, pp.1–5.

Gray, R.M. and Neuhoff, D.L. (1998) 'Quantization: invited paper', *IEEE Transactions on Information Theory*, Vol. 44, No. 6, pp.2325–2383.

He, Z. and Mitra, S.K. (2005) 'From rate-distortion analysis to resource-distortion analysis', *IEEE Transaction on Circuit and System for Video Technology*, Vol. 5, No. 3, pp.6–18.

LeGall, D. (1991) 'MPEG: a video compression standard for multimedia applications', *Communication of ACM*, Vol. 34, No. 4, pp.46–58.

Leung, R. and Taubman, D. (2005) 'Transform and embedded coding techniques for maximum efficiency and random accessibility in 3-D scalable compression', *IEEE Transactions on Image Processing*, Vol. 14, No. 14, pp.1632–1646.

Liang, J., Tu, C. and Tran, T.D. (2005) 'Optimal block boundary pre/postfiltering for wavelet-based image and video compression', *IEEE Transactions on Image Processing*, Vol. 14, No. 12, pp.2151–2158.

Montgomery, D.C., Runger, G.C. and Hubele, N.F. (2001) *Engineering Statistics*, 2nd ed., John Wiley, New York.

Rath, A., Tuithung, T., Ghosh, S.K. and Mukherjee, J. (2006) 'Compressing videos using JPEG2000', *Proceedings of the 3rd Workshop on Computer Vision, Graphics and Image Processing*, 12–13th January, Hyderabad, Andra Pradesh, India, pp.93–97.

Senda, Y. (1995) 'A simplified motion estimation using an approximation for the MPEG-2 real time encoder', *International Conference on Acoustics, Speech and Signal Processing*, Detroit, MI, USA, pp.3221–3224.

Sikora, T. (1997) 'The MPEG-4 video standard verification model', *IEEE Transactions on Circuit and System for Video Technology*, Vol. 7, pp.19–31.

Skodras, A.N., Christopoulos, C.A. and Ebrahimi, T. (2000) 'JPEG2000: the upcoming still image compression standard', *Proceeding of the 11th Portuguese Conf. on Pattern Recognition (RECPA00D 20: Invited paper)*, 11–12th May, Prto, Portugal, pp.359–366.

Tudor, P.N. (1995) 'MPEG-2 video compression', *Electronics and Communication Engineering Journal*, pp.257–264.

Tuithung, T., Sinha, D., Ghosh, S.K. and Mukherjee, J. (2006) 'Bit-investment policy of MCJ2K: a new video codec', *Proceedings of the 1st IEEE International Conference on Signal and Image Processing*, 7–9th December, Hubli, Karnataka, India, pp.648–652.

Taubman, D. and Marcellin, M.W. (2002) *JPEG 2000: Image Compression Fundamentals, Standards and Practice*, Kluwer, MA, Norwell.

Taubman, D. (1999) 'High performance scalable image compression with EBCOT', *Proceeding of IEEE International Conf. on Image Processing (ICIP99)*, Tokyo, Japan, Vol. 3, pp.344–348.

Taubman, D. and Zakhor, A. (1994) 'Multirate 3-D subband coding of video', *IEEE Transactions on Image Processing*, Vol. 3, No. 5, pp.572–589.

Wallace, G.K. (1991) 'The JPEG still picture compression standard', *Communication of ACM*, Vol. 34, No. 4, pp.30–44.

Wiegand, T., Sullivan, G.J., Bjontegaard, G. and Luthra, A. (2003) 'Overview of the H.264/AVC video coding standard', *IEEE Transaction on Circuit and System for Video Technology*, Vol. 13, pp.560–576.

Wiegand, T. and Sullivan, G.J. (2005) 'Video compression-from concepts to the H.264/AVC standard', *Proceeding of the IEEE*, Vol. 93, No. 1, pp.18–31.

Xu, J., Xiong, Z., Li, S. and Zhang, Y. (1994) '3-D embedded subband coding with optimal truncation (3-D Escot)', *J. Appl. Comput. Harmon. Anal.*, Vol. 10, pp.290–315.

Xu, J., Xiong, Z., Li, S. and Zhang, Y.Q. (2002) 'Memory-constrained 3-D wavelet transform for video coding without boundary effects', *IEEE Transaction on Circuit and System for Video Technology*, Vol. 12, No. 9, pp.812–818.

Note

[1]http://www.kakadusoftware.com