A New Unequal Error Protection Technique Based on the Mutual Information of the MPEG-4 Video Frames over Wireless Networks

H. Ghanei Yakhdan*, M. Khademi** and J. Chitizadeh**

Abstract: The performance of video transmission over wireless channels is limited by the channel noise. Thus many error resilience tools have been incorporated into the MPEG-4 video compression method. In addition to these tools, the unequal error protection (UEP) technique has been proposed to protect the different parts in an MPEG-4 video packet with different channel coding rates based on the rate compatible punctured convolutional (RCPC) codes. However, it is still not powerful enough for the noisy channels. To provide more robust MPEG-4 video transmission, this paper proposes a modified unequal error protection technique based on the mutual information of two video frames. In the proposed technique, the dynamic channel coder rates are determined online based on the mutual information of two consecutive video frames. With this technique, irregular and high motion areas that are more sensitive to errors can get more protection. Simulation results show that the proposed technique enhances both subjective visual quality and average peak signal to noise ratio (PSNR) about 2.5 dB, comparing to the traditional UEP method.

Keywords: MPEG-4 video, error resilient, UEP, mutual information.

1 Introduction

With the rapid development of mobile communications, robust transmission of video over wireless networks is becoming an increasingly important application requirement. On the one hand, because of the vast amount of data necessary to represent digital video, the source video data must be compressed before storing or transmitting. On the other hand, the highly compressed data is very sensitive to channel errors. Therefore, using error resilience tools to protect the compressed video data from the channel errors becomes very important.

During the last decade, many video compression standards have been proposed by ITU-T and ISO for different applications, such as, MPEG-1, MPEG-2, MPEG-4, H.26X and so on. From among these standards, MPEG-4 [1] is the most suitable for wireless links [2-4]. There are three ways to encode a video frame in an MPEG-4 video codec [5]. Those are intra-coded (I-frame), predictive-coded (P-frame) and bi-directionally predictive–coded (B-frame). An I-frame is a video frame that has been encoded without reference to any other frame of video. P-frames use the motion prediction from the I-frame or the previous P-frames. B-frames encoding accept the bidirectional prediction between the neighboring I-frame and P-frames, or two P-frames. To make the compressed bitstream more robust to channel errors, the MPEG-4 video compression standard has incorporated several error resilience tools including resynchronization markers, data partitioning, reversible variable length coding and header extension codes. However, these tools are not powerful enough for typical wireless channels with high bit error rates, so channel coding must be introduced [6-9].

Recently, many researchers have focused on channel coding for MPEG-4 video transmission over wireless channels and many techniques have been proposed. The most popular of these techniques is the traditional unequal error protection (TUEP) technique [10], which can protect the different parts in an MPEG-4 video packet with different channel coding rates based on the RCPC codes [11]. However, this technique is not powerful enough for this application due to performance issues [12] and it must be modified.

Iranian Journal of Electrical & Electronic Engineering, 2009.
Paper first received 4 Oct. 2008 and in revised form 24 Jan. 2009.
* The Author is with the Department of Electrical Engineering, Yazd University, Yazd, Iran.
E-mail: ho_gh58@stu-mail.um.ac.ir
** The Authors are with the Department of Electrical Engineering, Ferdowsi University of Mashhad, Mashhad, Iran.
E-mails: khademi@um.ac.ir, chitizadeh@hotmail.com.

In order to modify the TUEP technique many works have been done [13-23]. In [13], an adaptive constrained optimization is implemented by dynamically varying the UEP level in a video packet based on the channel conditions. The potential advantage of this approach is that it attempts to achieve maximum perceptual quality for the given channel conditions and gives better results for higher loss percentages at relatively low cost during run time.

Similar approaches have been evaluated in [14-16]. While [15] uses unconstrained optimization, a constraint on system probability of failure rather than on channel conditions is applied in [14]. In doing so, Grangetto et al. have enforced a tight hound on minimum achievable PSNR but have not prevented any channel inflicted quality loss. It is argued here that channel constraints have to be implemented adaptively and the packet structure changed dynamically for the model in [14] to work effectively. In [16] however, an unconstrained optimization is done on the overall rate distortion performance of a joint source-channel coding system.

In [17], an UEP scheme for scalable video has been proposed. However, this scheme did not consider time-varying channel conditions or handling packet loss simultaneously.

A proportional unequal error protection scheme has been presented in [18] for the protection of MPEG-4 video bitstream. In this approach, channel coder rates are determined based on the video packet length. Such a scheme introduces a delay of one packet.

In [19], an improved unequal error protection has been presented. The proposed scheme reorganizes the components in the video packet and partitions it into four parts. Thus the coded macroblock indication or COD bits can be more protected with lower code rate than other bits in the motion part. This approach does not increase the total bit rate for transmission in most cases, but increases the computation complexity.

An error-resilient MPEG-4 video system has been proposed in [20]. System ensures different sections of a bitstream of different importance are protected by different channel coding algorithms with different data protection capability. In the proposed system, the bitstream gets optimal protection without incurring much overhead, but requires a feedback between transmitter and receiver.

In [21], an adaptive UEP approach based on the video motion has been proposed. The proposed technique uses the motion information for the adaptive FEC coding, but it does not differentiate between different sections of the video packets.

When the coded packet size is fixed, a heuristic optimization method which has low complexity and obtains performance approaching to the optimal solutions for various channel conditions and transmission rates has been proposed in [22]. However, in this method, the motion information has not been used to the coding process.

In the reported works for control of the channel coding rate the amount of instantaneous temporal information of the video sequence has not been taken into consideration. In our previous work, a reformed unequal error protection approach based on the number of bits required for coding the motion vector was proposed. However, if a scene change exists in two consecutive video frames, then this method will product poor performance. In order to cope with this problem, a modified unequal error protection (MUEP) technique based on the mutual information of two consecutive video frames is proposed. This means that, the dynamic channel coder rates are determined online based on the content of the video scene. With this technique, the bits in the motion parts for P-frames can get more protection. Since these bits are much more important for motion vector (MV) decoding than other bits in the motion part, MUEP can provide better decoded video quality than TUEP can.

This paper is organized as follows. In Section 2, the basic ideas of the TUEP technique are introduced. Section 3 describes mutual information between consecutive video frames. The proposed scheme is addressed in Section 4. Section 5 gives details about the simulations followed by the conclusion in Section 6.

2 Error Resilient with the TUEP Technique

The output of a simple MPEG-4 video encoder that uses all error resilience tools is a bitstream that contains video packets that begin with a header, followed by the motion information and the texture information (DCT coefficients). The header bits are the most important information of the video packet, since the whole video packet will be discarded if the header is received with error. The motion information has the next level of importance, as motion compensation cannot be performed without it. The texture information is the least important of the three segments of the video packet since, motion compensated concealment can be performed without too much degradation of the reconstructed image if the texture information is lost. Since the video packets can be broken into sections with different levels of importance, the number of bits for the channel coding of each section should be proportional to its relative importance, creating an unequal error protection channel coder.

Using unequal error protection implies that different rate coders are applied to different sections of the video packet. When using unequal error protection, the header bits would get the highest amount of protection. The motion bits would get the next highest level of

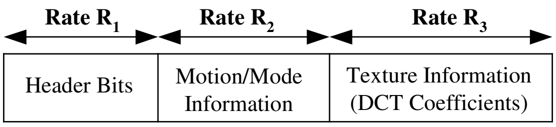

Fig. 1 Unequal protection of an MPEG-4 video packet in the P-frame.

protection, and the texture bits would receive the lowest level of protection.

Fig. 1 illustrates the protection scheme described above, in which R_1, R_2 and R_3 represent the code rates of the three partitions of the video packet in the P-frame. To match the coders to the level of importance of the different partitions of the video packet, the coder rates are chosen in such a way that $R_1 < R_2 < R_3$. Thus, by using this technique, the errors are less likely to occur in the important sections of the video packet, thereby improving application perceived quality.

3 Mutual Information

The video is a series of still images (referred to as frames). It makes sense to simply display each full image consecutively, one after the other. Mutual information between consecutive video frames (in the position of random variables X and Y) is given by [23]:

$$I(X;Y) = \sum_{x}\sum_{y} p(x,y)\log[p(x,y)/[p(x)p(y)]] \qquad (1)$$

where p(x,y) is the joint probability mass function (pmf) of random variables X and Y, p(x) and p(y) represent the marginal pmf of X and Y respectively.

Let p(x,y) be defined by coupling the values of adjacent pixels in any two consecutive frames. The allowed range of values for p(x,y) is (0, 0), (0, 1), (0, 2), …, (255, 255). Hence, the total number of outcomes is $(255)^2$. In addition, let p(x) and p(y) be the normalized histograms of frame-X and frame-Y respectively. It is obvious that if the two consecutive video frames represent a scene change or high motion, then the mutual information between these is reduced. Therefore, the mutual information between video frames determines the scene changes and high motion. This is clearly shown in Fig. 2. In this figure, the mutual information frame-1 and frame-2, and also frame-2 and frame-3 are 3.4386 and 1.0628, respectively. The first two frames exhibit high temporal correlation. Hence, the mutual information between these frames is a big value. On the contrary, frame-2 and frame-3 represent a change in video scene. Therefore the information overlap between these frames is relatively smaller.

4 The Proposed MUEP Technique

As described in Section 2, the technique of TUEP applies different but fixed rate coders to different sections of the video packet. But, the MPEG-4 video packets are not exactly the same length and partitions have different lengths in different packets, due to the variable length coding used and to the requirement of having an integer number of macroblock in each packet. This implies that a fixed UEP scheme cannot be suited. Hence, the proposed MUEP technique uses dynamic coder rates for motion section in the P-frames based on the content of video scene that gets mutual information between video frames.

On the one hand, abrupt scene changes and irregular motion require a higher number of bits compared to stationary/gradually changing scenes. On the other hand, the areas with the higher number of bits are more likely hit by errors. Moreover, temporal predictive coding technique aggravates error propagation for high motion areas. Hence, to provide more robust MPEG-4 video transmission, the high motion areas or scene changes must be sent to a lower code rate. Thus, in MUEP technique mutual information between video frames is used to determine the channel coder rates for the motion section in the P-frames, since P-frames are much more important than B-frames.

The block diagram of the proposed technique is shown in Fig. 3. Initially with above information two-threshold levels are defined, one high-threshold

Fig. 2 Frames 1, 2 and 190 in test video sequence “Foreman”.

level (T_h) and the other low-threshold level (T_l). The value of these threshold levels, which have been obtained in this paper for the test video sequences with simulations are related to the nature of video and total transmission channel rate. Then for motion partition in the P-frames considers three different code rates (instead of Rate-R_2 in TUEP technique). If the mutual information between two consecutive video frames is higher than T_h, it will use the high code rate (R_h). If it is lower than T_l, it will use the low code rate (R_l). If it is between the T_h and T_l, it will use the mean code rate (R_m). To match the coders to the level of importance of the motion partition of the video packet in the P-frames, the coder rates are chosen in such a way that $R_l<R_m<R_h$ so that irregular and high motion areas may be protected more against transmission channel errors.

In the proposed method, if T_h is set to small value, most frames will be encoded with the high rate R_h. This will increase the compression efficiency of the proposed method (with the worse error resilient results), and vice versa. On the other hand, if T_l is set to large value, most frames will be encoded with the low rate R_l. This will decrease the compression efficiency of the proposed method (with the better error resilient results), and vice versa. Hence, the threshold is empirically selected, considering the trade-off between error resilient performance and compression efficiency. There is no difference between MUEP and TUEP techniques for utilizing code rates in I and B frames.

5 Simulations Results

In order to simulate the proposed MUEP technique we used an MPEG-4 codec [24] and Rician wireless channel model. The Rician channel characteristics were three fading paths with the delays [0 0.5 1] μs, the gains [0 -5 -10] dB, Doppler maximal frequency f_d = 20 H_z and Rician factor k = 1. In order to test the performance of the MUEP technique, the “Suzie”, “Carphone” and “Foreman” video sequences at QCIF resolution, and “Stefan” video sequence at CIF resolution have been considered. The compressed video stream is composed of one I frame followed by one P-frame and two B-frames (IPBBPBBP…). The video sequences coded at 10 fps with the quantization parameter equal to 10.

The sequences were coded using all the MPEG-4 error resilience tools. The compressed bitstreams were then channel coded using convolutional encoding of the data with either TUEP technique using a fixed rate ¾ for the motion partition in the P-frames or MUEP technique using the dynamic different rates for the motion partition in the P-frames. Table 1 reports the video packet size (VP), number frames (NF) for both TUEP and MUEP techniques and also the rest of parameters utilizing in the MUEP technique.

These TUEP and MUEP rates were chosen because they both give approximately the same amount of the total output rates, and they were obtained by simulations based on trial-and-error method. It should be noted that the rates of R_1 and R_3 for both techniques are the same.

Coded MPEG-4 bitstream is transmitted on a Rician channel with above characteristics and assuming DPSK (Differential Phase Shift Keying) modulation. The transmission channel bit rate in reference to the available bandwidth of the wireless channel (channel capacity) in our simulation varies from 9 kbps to 10 kbps. Then the bitstream is decoded by a convolutional decoder to generate the bitstream fed to the MPEG-4 decoder.

Figs. 4, 5, 6 and 7 show the values of PSNR for the reconstructed video sequences “Suzie”, “Carphone”, “Foreman” and “Stefan” with respect to the number of frames, respectively. These plots show that the performance of MUEP technique is much better than TUEP technique for both CIF and QCIF resolution. This is because in MUEP technique high motion and scene change areas, which are more sensitive to error, are more protected. On the contrary, in TUEP technique the errors occur more on these areas and aggravates error propagation for it. Hence, it produces lower quality reconstructed video.

In Table 2, the rest of parameters results have been summarized. In this Table, T-APSNR, M-APSNR, T-R_t, M-R_t, T-BER and M-BER parameters are average PSNR, total output rate in P-frame and channel bit error rate for TUEP and MUEP techniques respectively. The simulations for every method are performed 20 times, and the average value will be used as the final results. Several important points exist about the reported results in Table 2. In the first place, the more video scene changes are, the more effective MUEP technique will be. In the second place, total output rate for MUEP

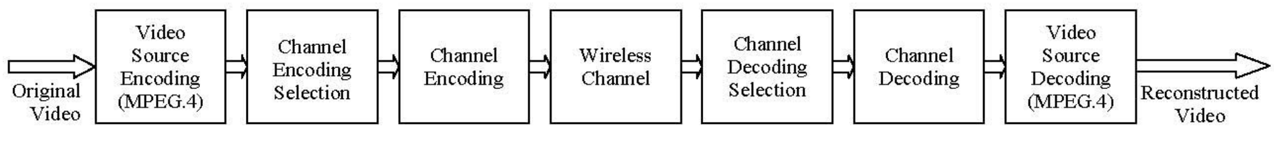

Fig. 3 Block diagram the proposed technique.

technique is always higher than TUEP technique. Hence MUEP technique is more efficient. Thirdly, some of the video sequences in TUEP technique actually produce fewer errors in the channel decoded bitstream than the MUEP technique (as shown in Table 2), yet it still produces lower quality reconstructed video. This is because the errors in TUEP technique occur more on high motion and scene change areas that are more sensitive to error since in these areas there is less protection. Hence, in the TUEP technique is aggravated error propagation so that reduces the quality of reconstructed video. In Fig. 8, an example is given to show the improvement for reconstructed video quality after transmission over wireless channels.

Table 1 The MUEP and TUEP techniques parameters.

Parameter / video	VP	NF	T_h	T_l	R_l	R_m	R_h
Suzie	400	140	4.15	2.7	$^{8}/_{11}$	$^{16}/_{21}$	$^{8}/_{10}$
Carphone	500	120	4.25	3.7	$^{8}/_{11}$	$^{16}/_{21}$	$^{8}/_{10}$
Foreman	500	120	4.5	3.27	$^{8}/_{11}$	$^{16}/_{21}$	$^{8}/_{10}$
Stefan	750	32	3.5	1.65	$^{8}/_{11}$	$^{16}/_{21}$	$^{8}/_{10}$

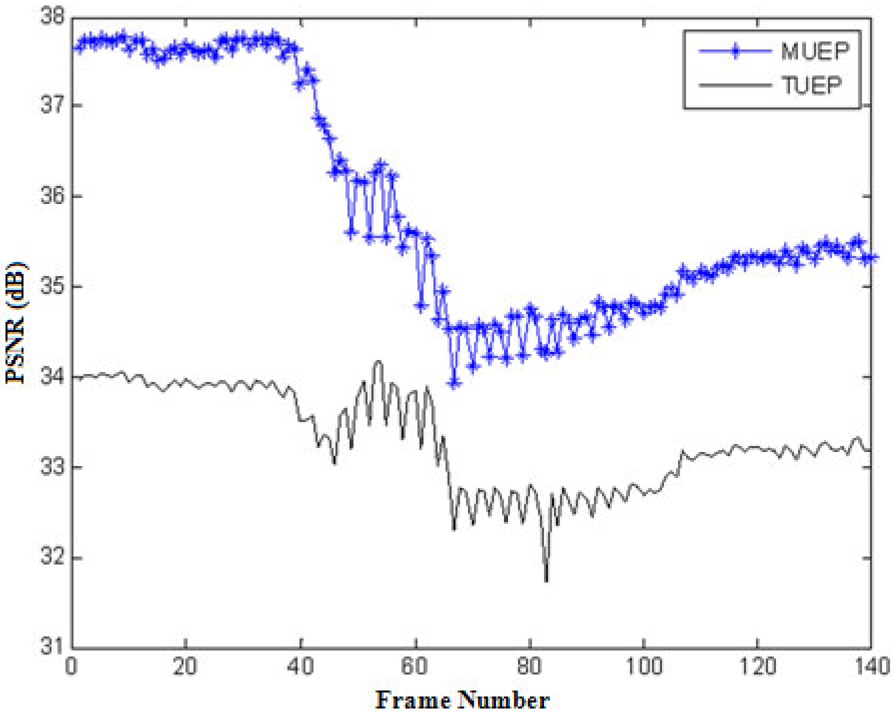

Fig. 4 Average PSNR versus frame number for video sequence “Suzie”.

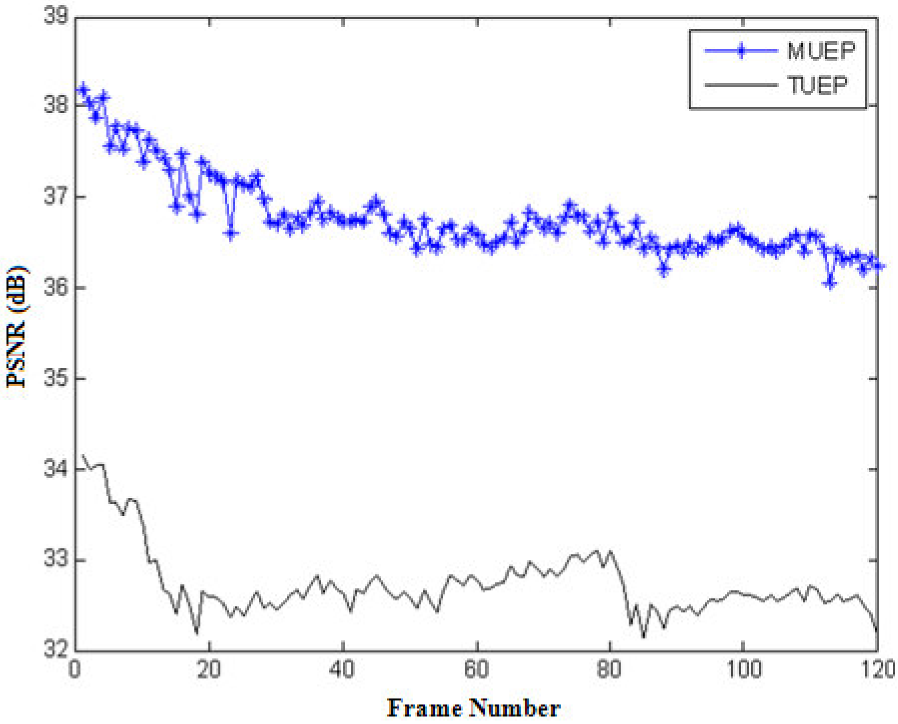

Fig. 5 Average PSNR versus frame number for video sequence “Carphone”.

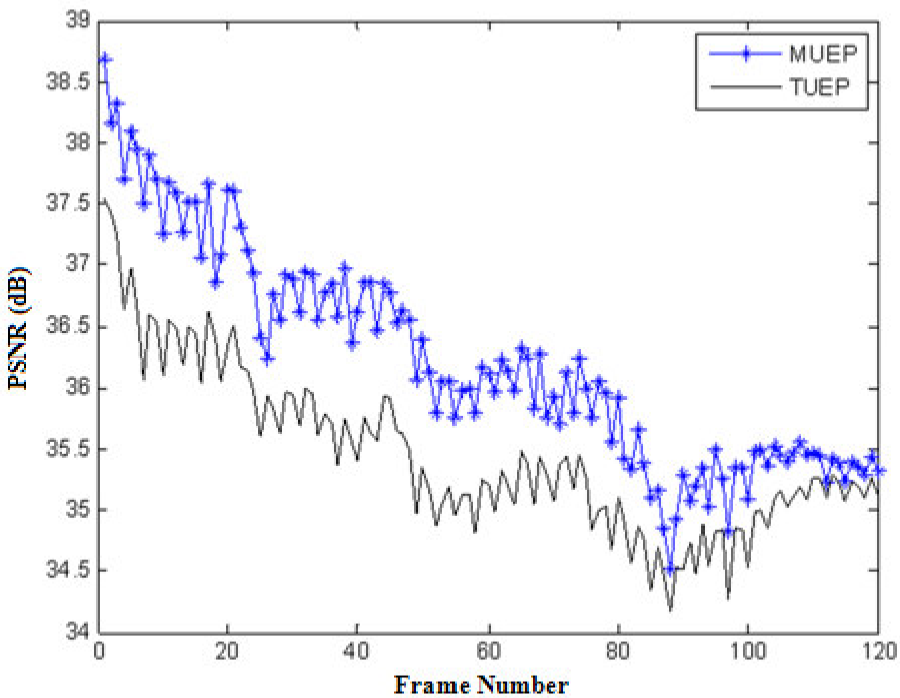

Fig. 6 Average PSNR versus frame number for video sequence “Foreman”.

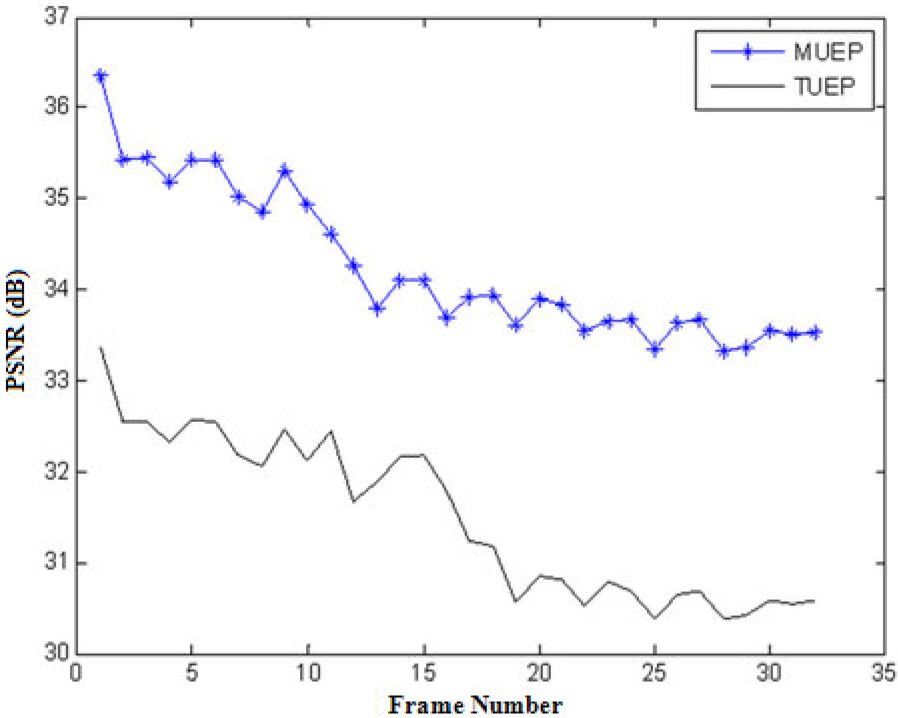

Fig. 7 Average PSNR versus frame number for video sequence "Stefan".

Table 2 The simulation results for video sequences.

Parameter / Video	T_APSNR (dB)	M_APSNR (dB)	T-R_t	M-R_t	T-BER	M-BER
Suzie	33.3065	35.8955	$^8/_{10.2897}$	$^8/_{10.2877}$	4.1828e-004	4.1785e-004
Carphone	32.7292	36.7848	$^8/_{10.2456}$	$^8/_{10.2419}$	3.8878e-004	3.8806e-004
Foreman	35.4475	36.2208	$^8/_{10.2510}$	$^8/_{10.2452}$	2.2317e-004	2.2484e-004
Stefan	31.4907	34.2508	$^8/_{10.1204}$	$^8/_{10.1177}$	2.5518e-005	2.5593e-005

Fig. 8 Foreman sequence simulation result for 20th frame, (a) original image, (b) reconstructed image by MUEP and (c) reconstructed image by TUEP.

6 Conclusion

In this paper, we have proposed a new modified unequal error protection (MUEP) technique based on the mutual information of two consecutive video frames to strengthen the robustness of the transport of MPEG-4 video over wireless channels. In MUEP technique, instead of using one channel coder for motion partition in the P-frame, three channel coders with three different rates are used. For this purpose, two-threshold levels T_h and T_l are defined**.** If the mutual information between two consecutive video frames is higher than T_h, channel coder will use the high rate R_h. If it is lower than T_l, it will use the low rate R_l. If it is between the T_h and T_l, it will use the mean rate R_m. Thus, more sensitive to errors areas are protected more against channel errors. Simulation results show that the proposed technique enhances both subjective visual quality and average PSNR about 2.5 dB and 1 dB, comparing to the TUEP method and our previous work, respectively.

The significant point is that some of the video sequences in TUEP technique actually produce fewer errors in the channel decoded bitstream than the MUEP technique, yet it still produces lower quality reconstructed video. Thus, it confirms the accuracy of the proposed idea. The other point is that the MUEP technique distinguishes scene changes accurately, comparing to the previous work. Therefore, it increases quality of the reconstructed video even though there is the scene change. However, a possible disadvantage may lie in the fact that the MUEP technique requires an additional time for the calculation of mutual information between video frames, which may limit applications in the real time. Of course, there are techniques that decrease this additional time highly, which will be surveyed by the authors in future.

References

[1] Koenen R., "*Overview of the MPEG-4 standard*," ISO/IEC JTC1/SC29/WG11 M4030, 2001.

[2] Yan B. and Ng K. M., "A survey on the techniques for the transport of MPEG-4 video over wireless networks," *IEEE Trans*. on Consumer Electronics, Vol. 48, No. 4, pp. 863-873, Nov. 2002.

[3] Soares L. D. and Pereira F., "MPEG-4: A flexible coding standard for the emerging mobile multimedia applications," *The Ninth IEEE Int. Sys. on*, Vol. 3, pp. 1335-1339, 1998.

[4] Zhu J., Matrawy A. and Lambadaris I., "Models and tools for simulation of video transmission on wireless networks," *CCECE* 2004 - *CCGEL* 2004, Niagara Falls, pp. 781-784, May 2004.

[5] ISO/IEC 14496-2, "Information technology-generic coding of audio visual objects- Part 2: Visual", *Proposed Draft Amendment1*, Seul, Mar. 1999.

[6] Requirements Group, "*MPEG-4 applications*," ISO/IEC JTC1/SC29/WG11 MPEG 99/N272, Mar. 1999.

[7] Sikora T., "Trends and perspectives in image and video coding," *Proceeding of the IEEE*, Vol. 93, No. 1, pp. 6-17, Jan. 2005.

[8] Yan B. and Ng K. W., "Mode-base error-resilient techniques for the robust communication of MPEG-4 video," *IEEE. Trans. on Circuits and Systems for Video Technology*, Vol. 14, No. 6, pp. 874-879, June 2004.

[9] Du B., "Multimedia communication in wireless environment," *Multi-Media Modeling Conference Proceedings*, 2006 12th International, pp. 460-464, 4-6 Jan. 2006.

[10] Heinzelman W. R., Budagavi M. and Talluri R., "Unequal error protection of MPEG-4 compressed video," *International Conference on Image Processing, ICIP 1999*, Vol. 2, pp. 530-534, 24-28 Oct. 1999.

[11] Hagenauer J., "Rate-Compatible Punctured Convolutional codes (RCPC codes) and their applications," *IEEE Trans. on Comm.*, Vol. 36, No. 4, pp. 389-399, Apr. 1988.

[12] Li C., Lu J. and Letaief K. B., "Adaptive error resilient video transmission over wide-band CDMA networks," *IEEE VTS 53rd Vehicular Technology Symp.*, Vol. 13, pp. 2086-2090, 2001.

[13] Adsumilli C. and Hu Y. H., "A dynamically adaptive constrained unequal error protection scheme for video transmission over wireless channels," *IEEE Workshop on Multimedia Signal Processing*, *2002*, pp. 41-44, 9-11 Dec. 2002.

[14] Mohr A. E., Riskin E. A. and Lander R. E., "Unequal loss protection: graceful degradation over packet erasure channels through forward error correction*,*" *IEEE Journal on Selected Areas in Comm.*, Vol. 18, No. 7, pp. 819-828, June 2000.

[15] Grangetto M., Magli E., Marzo M. and Olmo G., "Guaranteeing quality of service for image transmission by means of hybrid loss protection," *Proc. Intl., Conf. on Multimedia and Expo., In Press*, Vol. 2, pp. 469-472, 2002.

[16] Regunathan S. L. and Rose K., "Robust video compression for time-varying wireless channels," *Proc. Visual Comm. and Image Proc.*, Jan. 1999.

[17] Cheung G. and Zakhor A., "Bit allocation for joint source/channel coding of scalable video," *IEEE Trans. on Image Proc.*, Mar. 2000.

[18] Martini M. G. and Chiani M., "Proportional unequal error protection for MPEG-4 video transmission", *IEEE International Conference on*

Communications, *ICC* 2001*,* Vol. 4, pp. 1033-1037, 11-14 June, 2001.
[19] Yan B. and Ng K. W., "An improved unequal error protection technique for the wireless transmission of MPEG-4 video," *ICICS-PCM*, pp. 513-517, 15-18 Dec. 2003.
[20] Yin T. C., Huang Y.-C., Lin M.-H. and Chen W.-C., "Error-resilient MPEG-4 video communication over error-prone wireless networks," *International Conference on Consumer Electronics, ICCE 2005, Digest of Technical Papers,* pp. 313-314, 8-12 Jan. 2005.
[21] Qu Q., Pei Y. and Modestino J. W., "An adaptive motion-based unequal error protection approach for real-time video transport over wireless IP networks," *IEEE Trans. on Multimedia*, Vol. 8, No. 5, pp. 1033-1044, Oct. 2006.
[22] Lie C., "On the unequal error protection for progressive image transmission," *IEEE Trans. on Image Processing*, Vol. 16, No. 9, pp. 2384-2388, Sept. 2007.
[23] Cover T. M. and Thomas J. A., "*Elements of information theory*", John Wiley & Sons, New York, 1991.
[24] Winder S. A. J., "*ISO/IEC 14496 (MPEG-4) video reference software*", version: Microsoft-FPDAM1-1.0-000403, last update: Feb. 5, 2007.

Hossein Ghanei Yakhdan was born in Yazd, Iran in 1966. He received the B.Sc. degree in electrical engineering from Isfahan University of Technology in 1989, the M.Sc. degree in 1993 from K. N. Toosi University of Technology and the Ph.D degree in 2009 from Ferdowsi University of Mashhad. Since 1994, he has been with the Department of Electrical Engineering, Yazd University. His research interests are in digital video and image processing and error resilient and concealment for video communication.

Morteza Khademi was born in Iran, in 1958. He received the B.Sc. and M.S. degrees from Isfahan University of Technology, Isfahan, Iran, in 1985 and 1987, respectively and the Ph.D. degree from the Wollongong University, Australia, on video communications in 1995, all in Electrical Engineering. He joined Ferdowsi University of Mashhad, Iran in 1987. He is currently Associate Professor at the Department of Electrical Engineering, Ferdowsi University of Mashhad. Since then, he has co-chaired two conferences on "Electrical Engineering (ICEE2004)" and "Machine Vision and Image Processing (MVIP2006)" in Iran. His current research interests are in the areas of video communications, biomedical signal processing and soft computing. He has published over 70 articles in these research fields.

Jalil chiti-zadeh received the B.S degree in communication engineering from the Napier University, Scotland in 1980, and the M.S degree in communication engineering from the Essex University of Technology, UK in 1985. He received the PH.D degree in communication engineering from the Imperial College of London University in 1989. He was a member of computer network engineers team with several networking projects in Iran Telecommunication Research Center (ITRC) since 1981 to 1983. He is currently an associate professor and a project leader at Computer & Communication Center at Electrical and Engineering Department of Ferdowsi University of Mashhad. His research interests are in cellular wireless networks, sensor networks, mobile ad-hoc networks, high-speed networking and QoS.