Channel-Aware Robust Video Streaming over Wireless LANs Using Multiple-Description FGS Coding

Chih-Ming Chen[1,3], Chien-Min Chen[1], Chia-Wen Lin[2], and Yung-Chang Chen[1]

[1]*Department of Electrical Engineering, National Tsing Hua University, Hsinchu, Taiwan 30013, R.O.C., ycchen@ee.nthu.edu.tw*
[2]*Department of Computer Science & Information Engineering, National Chung Cheng University, Chiayi, Taiwan 621, R.O.C., cwlin@cs.ccu.edu.tw*
[3]*Chunghwa Telecom Co., Ltd.*

Abstract

In this paper, a novel error-resilient coding technique, named multiple description scalar quantization for fine granularity scalability (MDSQ-FGS), is presented to enhance error resilience capability and to improve the temporal prediction efficiency for video coding over lossy packet networks and wireless channels. MDSQ-FGS video coding scheme based on multiple descriptions coding (MDC) with modified multiple description scalar quantization (MDSQ) and partial predictions are proposed to control drifting error without too much reduction in efficiency. We also combine the benefits of multiple stream coding and multiple path transmission to maintain best constant video quality despite packet loss when roaming in wireless LANs (WLANs). Simulation results indicate that the proposed coding system outperforms the normal MPEG-4 FGS coder for coding efficiency, and can significantly improve the quality of streamed video when mobile client roams in WLAN.

1. Introduction

The benefits for wireless application are not only without wire connections but also mobile ability, especially for mobile computing devices. However, the heterogeneity of client networks and devices makes it very difficult to adapt the video contents to a wide degree of different client channel conditions, especially for mobile users who usually roam in wireless networks. In order to achieve error robustness for transmitting video over wireless networks, a video transcoder [1] can be placed in an intermediate network node (e.g., mobile switch/base-station, proxy server, and media gateway) connected to a high-loss network (e.g., wireless network or highly congested network) to adapt the non-error-resilient compressed video streams into error-resilience-capable streams. Figure 1 illustrates an exemplary three-tier streaming system using a wireless LAN as an extension to the existing wired infrastructure, offering local end-point devices the convenience of wireless connections. This three-tier streaming system involves a streaming server, a media gateway with a transcoder inside, and mobile client terminals. It can be divided into two parts: the first including the path from the streaming server to the media gateway and the second containing the path from the media gateway to wireless terminals via the wireless access points (APs). The media gateway serves as a proxy/adaptation engine to adapt multimedia contents to heterogeneous client devices. To achieve a high level of accessibility, it needs to provide a reliable and efficient transmission.

Video transport over wireless networks may suffer from signal fading, noise interference, network congestion, and handoffs which will usually lead to data error or packet loss. Even a single bit error in a

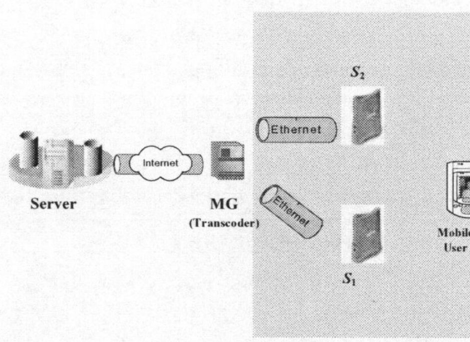


Figure 1. Illustration of a three-tier video streaming with two diverse wireless LAN links between the media gateway and the mobile terminals.

packet can cause the loss of a whole packet, if the number of corrupted bits goes beyond the error correction capacity of error correction codes. This packet-loss problem may lead to serious video quality degradation, which not only affects the quality of current frame, but also leads to error propagation to subsequent frames because of the motion compensated prediction used [2].

For video streaming to handle the problem of packet loss, the most popular approach is to add redundancy among packets. MDC is a kind of joint source and channel coder. The objective of MDC [1][3] is to increase the reliability (or error resilience) of data transmission under channel failures, by employing the diversity of channels. MDC inserts an additional rate to make the bitstream more resilient to transmission errors, and it always incorporates motion-compensated prediction to enhance coding efficiency in a video coder. If data are sent over multiple independent channels, the possibility of failures of all channels is greatly reduced, leading to a high possibility of receiving correct data from at least one channel, thereby making it easier to restore the original data with acceptable quality. Several MDC schemes have been proposed recently to address the error resilience coding problem. In [4], a video coder based on MD scalar quantization (MDSQ) with two independent prediction loops was proposed. MDC coders, however, are usually associated with drift when the signal used for prediction in an encoder is unavailable to the decoder due to loss of descriptions, which can lead to serious error propagation. The drift can be effectively eliminated using multiple predictors as proposed in [5] A drift-free MDC coder based on wavelet transform was proposed in [6]. In addition to error resilience, multiple path transport (MPT) is always applied to MDC, which can mitigate network congestion as well as can increase overall network utilization. Several research works integrating MPT with MDC have been proposed for video communication over the Internet [5] [7] or wireless networks [8] with diverse transmission paths. In this paper, we propose a novel error-resilience coding technique, named multiple description scalar quantization for fine granularity scalability (MDSQ-FGS), to enhance error resilience capability and improve the temporal prediction efficiency for video coding. We also combine MDSQ-FGS and multiple path transmission (MPT) to maintain best constant video quality despite packet loss when roaming in wireless LANs (WLANs). The rest of this paper is arranged as follows. The basic idea of Base-MDSQ coding will be described in Section 2. In Section 3, a novel MDSQ-FGS coder is proposed to reduce drift. Section 4 gives an overview of the proposed system. In Section 5, experimental results are presented. Finally, conclusion is drawn in Section 6.

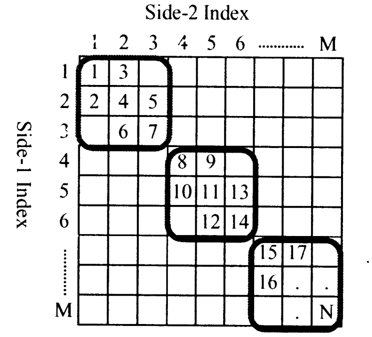


Figure 2. Example of the modified index assignment for Base MDSQ.

2. Base-MDSQ with partial prediction

In this paper, a Base-MDSQ video coder is used to mitigate error propagation in video coding, which was proposed in our previous work [9]. There we introduced a concept of the base part of MDSQ that is referred to as "Base-MDSQ". Encoder extracts the common part of all descriptions in concept, which will be referred to as the "base part" of multiple descriptions. We use the terms "base part" instead of "base layer" to avoid confusion with layer coding. The concept of base part is very useful in video coding, since encoder/decoder may use it as the prediction reference to reduce drift if the base part is always received in the decoder. Figure 2 shows an example of a modified index assignment of Base-MDSQ. In the index assignment matrix, we group all the central indexes to form a number of "base part groups". Each group corresponds to different reconstruction value and is bounded by a rectangle, which is a square generally. We refer the width or height of the square rectangle as "base part size".

Assume that the scalar quantizer at the encoder partitions the support of the source value into N central indexes, and the partitioning thresholds $\bar{t} = (t_0, t_1, \cdots, t_N)$, where $t_0 \le t_1 \le \cdots \le t_N$. Let the probability distribution function of the source be $p(x)$. The j-th base part group B_j contains a set of central indexes $\{i_o^{j,1}, i_o^{j,2}, \cdots, i_o^{j,bn_0}\}$ and corresponds to a set of side-1 indexes $\{i_1^{j,1}, i_1^{j,2}, \cdots, i_1^{j,bn_1}\}$ and a set of side-2 indexes $\{i_2^{j,1}, i_2^{j,2}, \cdots, i_2^{j,bn_2}\}$, where bn_0 is smaller than or equal to the square of the base part size, bn_1 and $bn_1$are equal to the base part size. The reconstructed value of the base part $\tilde{x}_{B_j}$ is calculated as the mean value of cell intervals of all the central indexes $\{i_o^{j,1}, i_o^{j,2}, \cdots, i_o^{j,bn_0}\}$ as follows:

$$\widetilde{X}_{B_j} = \sum_{i_0 \in \{i_0^{j,1}, \cdots, i_0^{j,bn_0}\}} \int_{i_0 - 1}^{i_0} x \cdot p(x) dx \qquad (1)$$

where a possible pair of side indexes (i_1, i_2) is mapped from a central index i_0. We may describe the reconstruction procedure of base part from central and side indexes by three mappings. The first mapping $b_0 : \{1,2,\cdots,N\} \rightarrow \Re$ maps a central index to the base part, while the second $b_1 : \{1,2,\cdots,M\} \rightarrow \Re$ and the last $b_2 : \{1,2,\cdots,M\} \rightarrow \Re$ map the side-1 and side-2 indexes to the base part respectively. Now we can formally define the three mappings as follows:

$$b_o(i) = \widetilde{X}_{B_j}, \quad \forall\, i \in \{i_o^{j,1}, i_o^{j,2}, \cdots, i_o^{j,bn_0}\} \qquad (2)$$

$$b_1(i) = \widetilde{X}_{B_j}, \quad \forall\, i \in \{i_1^{j,1}, i_1^{j,2}, \cdots, i_1^{j,bn_1}\} \qquad (3)$$

$$b_2(i) = \widetilde{X}_{B_j}, \quad \forall\, i \in \{i_2^{j,1}, i_2^{j,2}, \cdots, i_2^{j,bn_2}\} \qquad (4)$$

As long as the decoder receives at least one description, we can obtain the base part by the mapping $b_o(\cdot)$ or $b_1(\cdot)$ or $b_2(\cdot)$ depending on which descriptions are received. The base part of the source is available with high probability at the decoder, so we may use it as the prediction reference for video coding. By this method, there is almost no drift in the video decoding over lossy networks.

Inspired by the concept of RFGS [10], we propose the idea of partial prediction over Base-MDSQ. Partial prediction is a very simple but effective technique for layer coding to combat loss in the predictive coding, at the expense of a little increased bit rate. The prediction reference $\hat{x}[n]$ of the source sequence $x[n]$ is given by

$$\hat{x}[n] = \hat{x}[n-1] + \alpha_B \cdot \Delta x_B[n-1] + \alpha_E \cdot \Delta x_{EB}[n-1]. \qquad (5)$$

where $\Delta x_B[i]$ means the base layer of residue, $\Delta x_{EB}[i] \equiv \Delta x_E[i] - \Delta x_B[i]$ while $\Delta x_E[i]$ means the enhancement layer of residue. $\Delta x_{EB}[n]$ and $\Delta x_E[n]$ are multiplied by the factors B and E respectively, where $0 \le \alpha_E \le \alpha_B \le 1$, which are referred to as partial factors.

3. MDSQ-FGS with partial prediction

The MPEG-4 FGS video coding scheme is a scalable coding technique with bit-plane coding in the enhancement layer to adapt to channel-bandwidth fluctuations. Its coding efficiency is usually significantly lower than that of a non-scalable coder because the motion prediction of FGS is from the low-quality base-layer video only. Using higher quality reference frames in the enhancement layer can improve the coding efficiency, but will cause "drift" error propagate, leading to a trade-off between coding efficiency and drift error. To handle the problem of packet loss, the most popular approach is to add redundancy among packets. MD coding is suitable for multimedia transmission, since its decoder can restore the original data with tolerable distortion when only partial descriptions are received. We propose an MDSQ-FGS scheme with partial predictions coding to achieve higher coding efficiency while keeping low drift-error. The MDSQ-FGS scheme is based on the MD coding scheme to enhance error resilience capability. The proposed MDSQ-FGS scheme uses a

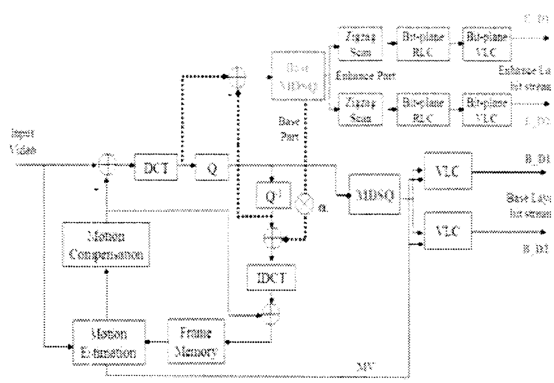


Figure 3. Encoder architecture of MDSQ-FGS video coder with partial prediction.

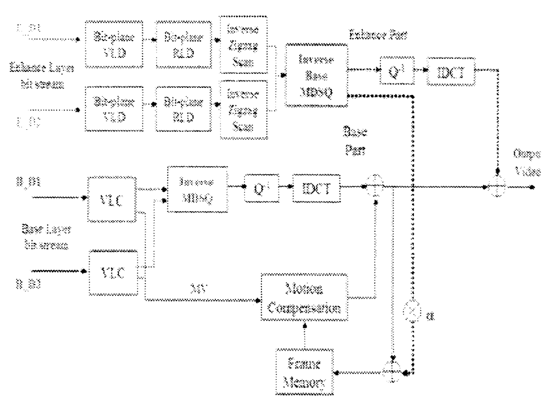


Figure 4. Decoder architecture of MDSQ-FGS video coder with partial prediction.

base part of the enhancement-layer for motion prediction of the base-layer to increase coding efficiency.

The encoder and decoder architectures of MDSQ-FGS with partial prediction is shown in Figure 3 and Figure 4, respectively. The quantized DCT coefficients of the base layer are split into two descriptions by an MDSQ coder, and further encoded with run-length coding and variable-length coding prior to be transmitted. For the base layer, the intra DC coefficients and motion vectors are duplicated in both descriptions. The enhancement layer bitstreams are produced first by calculating the difference between the original DCT coefficients and the reconstructed base-layer video, and then encoded by the Base-MDSQ coder to split the residues into two descriptions (i.e. side indexes), followed by the run-length coding and variable-length coding on the side indices prior to being transmitted. The resulting bit streams of the base layer and the enhancement layer are subsequently both packetized. The base part of enhancement layer also is extracted from the Base-MDSQ coder and added into the feedback loop of the base layer to get higher quality reference for motion prediction. The leaky factor α is used to attenuate the drift error when both descriptions are not received correctly, but it still can be used to reconstruct partial enhancement reference. However, the smaller α will lead to the lower performance when the base part of Base-MDSQ coder can be reconstructed correctly.

4. The proposed system

Figure 5 illustrates our streaming testbed which is composed of four personal computers to emulate a media gateway which implements an MDC video transcoder, two base-stations (S_1 and S_2), and one client, to emulate possible roaming situations among the base-stations and the wireless clients. In the testbed, stations S_1 and S_2 are responsible for detecting the clients located in its communication zone, and then reporting to the media gateway the channel statistics about the clients they are serving. According to the information, the media gateway maintains a path-list table that records which clients can be reached by the station. These paths are sorted according to their channel conditions. The status of a path to a client is updated when any new channel information about the client is received. If the server does not receive any information about the client from any station within a pre-specified timeout interval (namely, the lifetime of the client), the client will be marked as temporarily non-connectable. Based on the path-list table, the

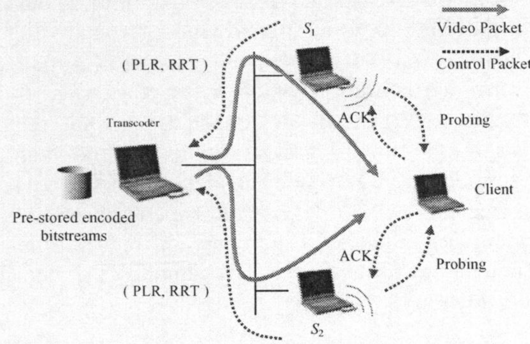


Figure 5. Proposed testbed for realtime multipath streaming with one media gateway (video transcoder), two base-stations, and one mobile client.

media gateway can chooses to deliver the video data to the client via either only one channel with good quality, or two channels with acceptable qualities but may not be very reliable. In this work, the stations use the User Datagram Protocol (UDP) to send the probe frames and receive the ACK frames from the client. Only a departure timestamp and a sequence number are carried in the probe frames and also in the ACK frames. For the probing, it is difficult to measure the delay time accurately unless the clock times of the station and the client are well synchronized. Using the round-trip time (RTT) to replace the one-way delay to estimate the channel condition is more reliable, since the receiver sends back the ACK frame to the station immediately via the same channel. In this work, the base-stations calculate the average RTT $\overline{RTT_k^i}$ and the average packet loss rate $\overline{PLR_k^i}$, respectively, of each channel between the stations and the client in a sliding time interval C of a fixed number of packets and report these channel statistics to the media gateway to guide the selection of channels, as follows.

$$\overline{RTT^i_k} = \frac{1}{C} \sum_{j=k-C+1}^{k} RTT_j^i \tag{6}$$

$$\overline{PLR^i_k} = \frac{1}{C} \sum_{j=k-C+1}^{k} PL_j^i \tag{7}$$

$$\max RTT^i = \max(RTT_k^i, \max RTT^i) \tag{8}$$

where RTT_j^i stands for the round-trip time of the jth packet through the ith channel, $PL_j^i = 1$ indicates the jth packet through the ith channel gets lost or corrupted, and $RTT_j^i = \max RTT^i$ if $PL_j^i = 1$.

The proposed mode decision algorithm chooses to use the single-description (SD) mode to send video via a single channel if at least one channel is estimated to be in good condition which will not sacrifice coding efficiency. When all the channels are not reliable, instead, the media gateway will use the multiple-description (MD) mode to deliver video data via diverse channels to provide strong error protection by adding redundancy at the cost of reducing coding efficiency to some degree. We summary the algorithm as follows:

Algorithm: Channel-Aware Mode Decision for Error Resilience Transcoding

---

At the SD mode:

if $(\overline{PLR_k^i} - \overline{PLR_k^j} > PLR_{LowerBound})$

Switch to the MD mode: transcode the video data into two descriptions, and send the two descriptions via both channels

else

{

if $(\overline{RTT_k^i} - \overline{RTT_k^j} > RTT_{LowerBound} \times \overline{RTT_k^i})$

Switch to the MD mode: transcode the video data into two descriptions, and send the two descriptions via both channels

else

Stay at the SD mode: forward the video data through channel $\boldsymbol{i}$ (via S_i) without transcoding

}

At the MD mode:

if $(\overline{PLR_k^i} - \overline{PLR_k^j} > PLR_{UpperBound})$

Switch to the SD mode: forward the video data through channel 2 (via S2) without transcoding

else

{

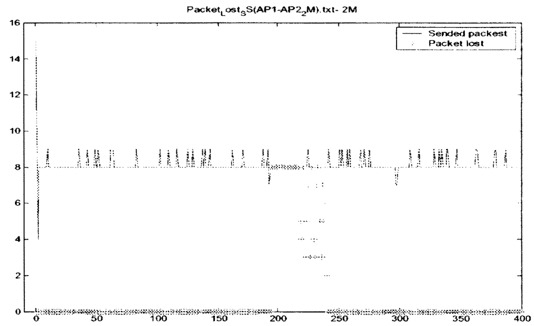
if $(\overline{RTT_k^i} - \overline{RTT_k^j} > RTT_{UpperBound} \times \overline{RTT_k^i})$

Switch to the SD mode: forward the video data through channel $\boldsymbol{j}$ (via S_j) without transcoding

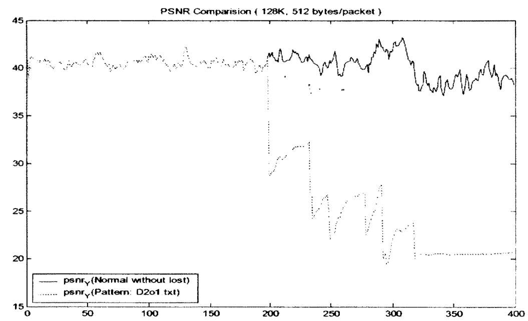
else

Stay at the MD mode: transcode the video data into two descriptions, and send the two descriptions via both channels

}



(a)



(b)

Figure 6. Effect of handoff on video streaming across two access points in wireless LANs: (a) packet loss due to handoff, and (b) PSNR distortion during handoff.

5. Experimental results

In the wireless LAN environment, a handoff to find another AP for a mobile device will occur when a channel cannot provide satisfactory connection quality. In general, a sequence of events must transpire for roaming. After the long roaming duration, it performs a nomadic roaming and cause very heavy transmission data loss, leading to severe transient video degradation, as shown in Figure 6. In order to construct a roaming environment in a lab environment, the stations are physically separated by about 30 meters and their transmitting power is reduced to 25%. A factory default transmission retry limit of 7 is configured for the stations. The parameters of the algorithm for channel-aware mode decision are set as $PLR_{LowerBound}$ =0.2, $PLR_{UpperBound}$ =0.4, $RTT_{LowerBound}$ =0.5 * T_{probe} and $RTT_{UpperBound}$ =1.25 * T_{probe}. Where, the probing period T_{probe} for stations to send the probe frames is set as 100ms. The first frame of each test sequence is coded as an I-frame, and all subsequent frames are coded as P-frames. The test bistreams are then transcoded into two descriptions by using the proposed MDSQ-FGS coding scheme when the transcoder decides to use the MD mode to deliver data to the client through both

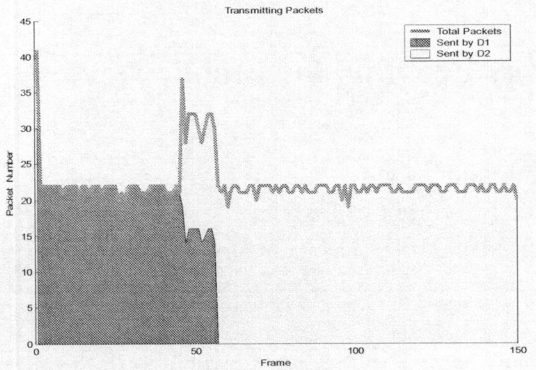


Figure 7. Frame-by-frame channel utilization.

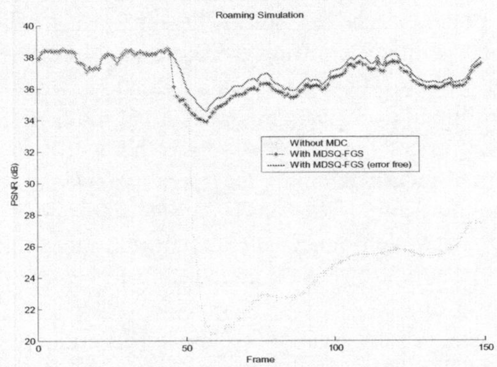


Figure 8. PSNR Performance comparison between with and without the proposed MDSQ-FGS scheme.

paths. The video packet length is 480 bytes. We simulate a roam between stations S_1 and S_2 with a scenario that a client moves from S_1 toward S_2 at a pedstrain speed. Error concealment by zero-motion replacement is adopted for restoring lost video data. In order to provide different protection, we simply duplicate the packets of base layer bit stream and assign the duplicated packet behind enhancement layer bitstream packets to deal with burst error. Figure 7 shows that, between frame 46 and 56, the MD mode will consumes about 45% more data packets more than the SD mode because two-description bitstreams with redundancy are sent. Nevertheless, the proposed roaming architecture with multipath transmission can greatly reduce the distortion caused by roaming with transition, even there still are some unavoidable packet losses at the MD mode. The quality comparison is shown in Figure 8.

6. Conclusion

In this paper, we combined the proposed MDSQ-FGS video coder and multipath routing to tackle packet loss problem caused by a roam of a mobile client in WLAN to improve the visual quality. The experimental results obtained from our multipath streaming test-bed show that the use of path diversity as a temporary step for roaming can provide significant benefits compared to roaming with single-description streaming. The proposed methods can also be extended to the cellular applications, such as the 3G systems, to achieve error robust handoffs among base-stations for mobile users in wireless Wide Area Networks (WANs).

7. References

[1] J. Xin, C. W. Lin, and M. T. Sun, "Digital video transcoding," Proceedings of the IEEE, vol. 93, no. 1, pp. 84-97, Jan. 2005.

[2] Y. Wang and Q.-F. Zhu, "Error control and concealment for video communication: A review," Proceedings of the IEEE, vol. 86, no. 5, pp. 974–997, May 1998

[3] Y. Wang, A. Reibman and S. Lin, "Multiple description coding for video delivery," Proceedings of the IEEE, vol. 93, no. 1, pp. 57-70, Jan. 2005.

[4] V. A. Vaishampayan, "Design of multiple description scalar quantizers," IEEE Trans. Information Theory, vol. 39, no. 3, pp. 821-834, May 1993.

[5] A. R. Reibman, H. Jafarkhani, Y. Wang, M. T. Orchard, and R. Puri, "Multiple description coding for video using motion compensated prediction," IEEE Trans. Circuits and Systems for Video Technology, vol. 12, no. 3, pp. 193–204, Mar. 2002.

[6] J. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in Proc. SPIE Conference on Visual Communications and Image Processing, pp. 392–409, 2001.

[7] N. V. Boulgouris, K. E. Zachariadis, A. N. Leontaris, and M.G. Strintzis, "Drift-free multiple description coding of video," in Proc. IEEE Workshop on Multimedia Signal Processing, pp.105-110, Oct. 2001.

[8] N. Gogate, D. M. Chung, S. S. Panwar, and Y. Wang, "Supporting video/image applications in a mobile multihop radio environment using route diversity and multiple description coding," IEEE Trans. Circuits and Systems for Video Technolog, vol. 12, no. 9, pp. 777–792, Sept. 2002.

[9] C. M. Chen, C. M. Chen, and Y. C. Chen, "Multiple description motion compensation video coding with leaky prediction," Proc. IEEE International Symposium on Circuits and Systems, vol. 3, pp. 825-828, May 2004

[10] H. C. Huang, C. N. Wang, T. Chiang, "A Robust Fine Granularity Scalability Using Trellis-Based Predictive Leak," IEEE Trans. CSVT, vol.12, No.6, pp.372-385, June 2002.