Adaptive Video Transmission Over A Single Wireless Link+

Hao Wang, S. Venkatesan

Telecommunication Engineering Department
The University of Texas at Dallas, P.O. Box 830688
Richardson, TX 75083-0688
{haowz,venky}@utdallas.edu

Abstract

In this paper, we propose an adaptive priority-based selective repeat ARQ (PSR-ARQ) transmission scheme for transmitting video over a single-hop lossy wireless link. If transmission errors occur in a lower layer of the video, the errors propagate to all the higher layers of related frames resulting in significant degradation of the video quality at the receiver. Due to the limited channel bandwidth, retransmission of every corrupted packet may be unacceptable for video streaming applications. Our proposed on-line priority-based scheme prioritizes all layers of various frames of a group of pictures and transmits them over a lossy wireless link. The on-line transmission policy achieves good video quality (of small distortion). Simulation results show that the proposed scheme performs better than the traditional scheme for video transmission over wireless channels.

I. Introduction

The trend to streaming stored media over wireless networks is emerging. Streaming video stored at the video server is the technique to achieve smooth playback of video directly at wireless terminals without downloading the entire file before the playback begins. While the demand for such video communication over wireless links has increased considerably, the transmission rate of the wireless channels remains low. In addition, wireless links are error prone. The limited bandwidth and high bit error rate (BER) can be problematic for video streaming over wireless networks. In this paper, we consider the problem of transporting video over a single-hop wireless channel. The single-hop wireless channel can be a dedicated channel as in mobile cellular scenarios or a shared channel as in Wi-Fi (802.11).

Video signals transmitted through the wireless channels are corrupted by two types of errors: (1) Stationary random errors that depend on the mean strength of the received wave or on the distance between the station and the wireless terminal and, (2) Variable error that is caused by the motion of the portable terminal or Rayleigh fading that is caused by the wave strength dips resulting in burst errors [8]. If the wave strength is 10db smaller than the mean wave strength, BER can be 20 times worse. Faster the portable terminal moves, higher the frequency of distortion caused by burst errors and more the deterioration in video quality.

While forward error-correcting codes (FEC) can be used to reduce the effects of transmission errors at the decoder, the associated increase in bit payload is often unacceptable, particularly in a wireless environment, where bandwidth resources are severely limited. Error concealment techniques have been applied to decoders such that they can extrapolate information from received bits to reconstruct the lost information. However, when large amounts of data are lost due to burst errors, the reconstruction process has insufficient information to work on.

Closed-loop error control techniques like automatic repeat request (ARQ) have been shown to be very effective and have been successfully applied to wireless video transmission [2,4,5]. However, retransmission of corrupted data frames introduces additional delay, which might be unacceptable for real-time conventional services. A combination of ARQ and FEC (hybrid ARQ) is recognized as a good choice for video over wireless.

Liu and Zaiki [7] classify the hybrid ARQ scheme into two categories: type-I and type-II schemes. A general type-I ARQ scheme uses error detection and correction codes for each transmission and retransmission. The receiver attempts to correct the errors by decoding the current received code words. Previously received code words that contain uncorrectable errors are discarded. A general type-II ARQ scheme uses a low rate error correction code. An information packet is first transmitted with some correction bits for error correction. Incremental redundancy bits are transmitted upon retransmission request. The receiver combines the transmitted and retransmitted packets together to form a more powerful error correction code to recover the information. For time-varying wireless channels with burst errors, type-II ARQ scheme may suffer from poor performance. Refinements of ARQ schemes for video have been proposed to alleviate this

+ This research is partially supported by a grant from the Department of Navy.

problem. Pyun.et.al [2] proposed a hybrid ARQ with interleaving scheme, which reduces the number of retransmission to minimize the delay. Zhang and Kasam [10] proposed a hybrid ARQ scheme, which takes advantages of both type-I and type-II hybrid ARQ, and this partially solves the problem of inadequacy of error correction code to recover the information. For time-varying wireless channels with burst errors, type-II ARQ scheme may suffer from poor performance. Wang, Zheng and Copeland [9] proposed QOS selective repeat ARQ (QSR-ARQ) scheme to enhance the existing data link layer protocols in wireless mobile environments. This scheme delivers I and P frame packets with limited numbers of retransmission to guarantee the delay bound while it transmits B frame packets only once to avoid long delays caused by retransmission [1, 14].

In this paper, we propose a transmission scheme that is priority-based selective repeat (PSR) ARQ. Layered video packets are first sequenced at the sender based on the contribution to the entire video quality and inter-dependency. The base layer of the I-frame is the most significant one and is at the head of the sequence. The highest layer of P/B frame is of least significance and is placed at the tail of the sequence. Our PSR-ARQ scheme increases the chance of successful transmission of important packets by maximizing the number of retransmissions within the delay bound. The PSR-ARQ scheme works at the application layer on tops of the UDP/IP stack. The link layer provides the acknowledgement scheme. For each transmission of data frame at the link layer, the link layer will send the ACK/NACK message to the PSR-ARQ. Based on the link layer acknowledgement, PSR-ARQ controls the video packet retransmission. This requires control traffic between the link layer and the application layer.

II. PSR-ARQ Transmission Scheme

The output of H.263 compression is a sequence of three types of frames: I frame (Intra-picture), P frame (forward Predicted) and B frames (Bidirectional predicted). I frames are reference frames and are self-contained. A P frames specifies the difference between the previous I frame; a B frame is an interpolation between the previous and subsequent frame of I or P type. If an I frame is lost, all the subsequent P and B frames (until the next I frame) are of no value. If a P frame is lost, all the previous and subsequent consecutive B frames (till the next I or P frame) are of no value. Clearly I-frames that start each scene are statistically larger (in size) than the P or B frames because I-frames are self contained.

Layered video coding is very useful in coping with the time-varying nature of wireless channel conditions [3]. A scalable encoder encodes video into several layers. The base layer guarantees a basic display quality and each enhancement layer (correctly received by the receiver) improves the video quality. Only base layer and few enhancement layers are received correctly if the channel exhibits transmission errors. If the channel condition is good, many enhancement layers (in addition to the base layer) are successfully transmitted and good video quality is achieved at the receiver.

We assume that layered coding encodes an I-frame into a hierarchy of X layers {I (BL), I (EL_1)… I (EL_{X-1})}, where I (BL) is the base layer while layers I (EL_1) to I (EL_{X-1}) are enhancement layers. Layered coding encodes P-frame and B-frame into a hierarchy of Y and Z layers. Accordingly, {P (BL), P (EL_1), P (EL_{Y-1})} are generated for each P frame and {B (BL), B (EL_1), B (EL_{Z-1})} are generated for each B frame. The layered video is shown in Figure 2. Note that we have X>Y>Z since I frames are larger (in size) than the related P frames and P frames are larger than B frames.

To reduce the amount of the computation performed at the sender when transmitting, we order all the packets (off-line). The rule is based on the computed distortion (to the entire GOP (Group of Picture) of each packet. A sequence of ordered video packets is available at the sender. The packets contribute to the entire video quality with decreasing significance from the left side to the right side in the ordered sequence. If the given video sequence is I B_1 B_2 P_1 B_3 B_4 P_2 B_5 B_6 P_3, the ordered packet sequence could be similar to the following:

I(BL), I(EL_1), P_1(BL), P_2(BL), P_3(BL), B_1(BL), B_2(BL), B_3(BL), B_4(BL), B_5(BL), B_6(BL), I(EL_2), I(EL_3), P_1(EL_1), P_2(EL_1), P_3(EL_1), B_1(BL), B_2(EL_1), B_3(BL), B_4(EL_1), B_5(BL), B_6(EL_1)…..

In the above list, I (BL) represents the base layer of I frame. P_2 (EL_1) represents the first enhancement layer of P_2 and B_2 (EL_1) represents the first enhancement of B_2. By changing X, Y and Z values properly, it is reasonable to assume that each video packet is of the same size statistically.

The algorithm presented in Figure 1 constructs the sequence. Note that if frame P_1 is ordered prior to P_2, its related B frames of P_1 should be ordered prior to the B frames related to P_2 as well.

In our scheme, packets are transmitted by on-line priority-based selective repeat ARQ with acknowledgement of ACK or NACK. The video streaming server keeps two transmission queues. Initially, video packets of first two GOPs put into each transmission queue separately according to the off-line ordering algorithm. Retransmission is required if no ACK is received before timeout or a NACK is received. Upon receiving an ACK, the server deletes the transmitted video packet from the transmission queue. The server selects the first video packet from the first queue for transmission. The reason we choose PSR-ARQ is because of the strong data dependency among video packets. Packets P_1(BL), P_2(BL) and P_3(BL) are useless unless I(BL) is transmitted successfully.

```
j = 1;   //notation: j = 1: BL, j =2: EL1, j=3: EL2 …
L = {all packets of all frames of current GOP}
Create an empty queue.

// X, Y and Z are the total number of layers of I, P or B
// frames respectively.

DO
{
Delete X/Z layers of I frame (from the bottom to the top)
from L and add them to the queue (again order it from
bottom to top)

Delete Y/Z layers of all P frames from L and add the deleted
layers to the queue (in the order of decreasing significance)

Delete layer j of all B frames and add the deleted layers to
the queue (in the same order as their related P frames).

j = j+1;

} WHILE (j<=Z)

// The queue now has the ordered the video packet sequence
```

Fig. 1 Video packet sequence off-line ordering algorithm

Each packet will be retransmitted in case of loss or packet corruption, but the maximum number of times a packet can be transmitted is limited. The number of retransmissions of a packet is bounded by the timestamp of current GOP. Packets belonging to the same GOP are buffered at the receiver before the entire GOP is scheduled to be decoded and displayed.

An important component of our on-line PSR-ARQ transmission scheme is the on-line transmission policy at the sender. The on-line policy dictates which packet at the source should be (re)transmitted. For every feasible set of unsent (or retransmission required) video packets, the decision made by the policy must increase the video quality by minimizing the expected distortion of the entire video. The expected distortion of a video packet is defined as the multiplication of absolute distortion of the video packet and the overall transmission failure probability of maximum number of retransmission.

The on-line transmission policy is simple when deciding which packet should be transmitted in current GOP. It selects the first packet of current GOP in the transmission queue. However, the decision is not clear when choosing between the first packet of the current GOP and the first packet of the next GOP. For instance, as shown in the Figure 3, packet I (BL) of GOP M+1 is of greater significance than the unfinished packet, say $B_k(EL_Z)$ in GOP M. But packet $B_k(EL_Z)$ has less number of chances to transmit. Our on-line transmission policy computes the expected distortion of the next packet of the current GOP and of the first packet of the next GOP. Then it chooses the packet with the smaller expected distortion value to transmit. The computation of expected distortion is given by equations (5) and (6) at the end of section II.

The on-line PSR-ARQ transmission algorithm is shown in Figure 4. Next we evaluate our scheme. A wireless channel can be described as a packet-based burst error channel [11, 12]. One commonly studied analytical model is the Gilbert model (two-state Markov model), which characterizes the bursty nature of wireless channel errors. The model has two states, a good state (S_0) and a bad state (S_1). P_{ij} is the transition probability from state i to state j. P_{00}, P_{01}, P_{10}, and P_{11}, are the state transition probabilities. P_{00} represents the probability of transition from good state to good state. P_{10} represents the

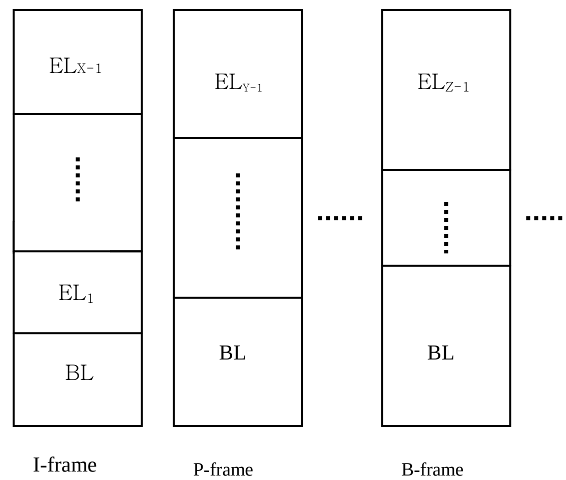

Fig. 2 The layered video frame sequence of a GOP

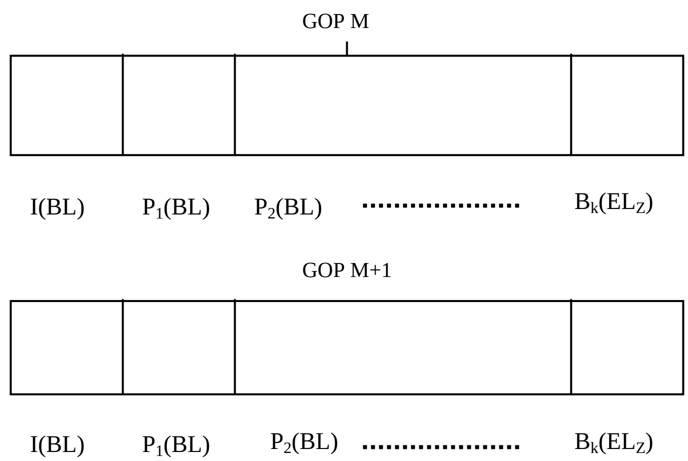

Fig. 3 The on-line transmission scheme

Sequence all GOPs in the video by number 1, 2...in their temporal order.

Pre-buffer M-1 GOPs at the receiver. The time-bound for following GOP transmission is given by the time-stamp of that GOP.

At sender, for each GOP, order all packets off-line using the algorithm of Figure 1. Initially, put packets of GOP M and M+1 into two transmission queues respectively.

```
K=M;

DO
{ IF (GOP K exceeds its time-stamp)   {
     Dump all packets of GOP k from the queue.
     Put packets of GOP K+2 into that empty queue.
     K++;
     }
  ELSE {
    Compute the expected distortion by sending the first
    active packet of the GOP K and denotes it as  D(1).

    Compute the expected distortion by sending the first
    active packet of the GOP K+1 and denotes it as D(2).

    IF( D(1)<D(2) )
     Transmit the first active packet of the GOP K.
    ELSE
      Transmit the first unmarked packet of the GOP K+1.
    }
}WHILE (the video transmission is not completed)
```

Fig. 4 On-line PSR-ARQ transmission algorithm

probability of transition from bad state to good state. State transitions occur at discrete time instants. In our work, we assume the time between two successive time instants is equal to the time taken for transmitting one packet at the MAC layer. A packet is transmitted correctly when the channel is in good state and errors occur when the channel is in bad state. The transitions between states occur at each packet instant. The channel state-transition probability matrix can be set up as

$$P = \begin{bmatrix} P_{00} & P_{01} \\ P_{10} & P_{11} \end{bmatrix} = \begin{bmatrix} 1-P_{01} & P_{01} \\ P_{10} & 1-P_{10} \end{bmatrix}; \quad P(S_0) = \pi_0; \quad P(S_1) = \pi_1 .$$

The packet error statistics vary according to the values of the transition probabilities. The transition probabilities can be calculated from the channel characteristics, i.e. the packet-error-rate (PER) and average-burst-length (ABL). Given the same PER, large ABL implies that errors are bursty. We have

$$ABL = \frac{1}{P_{10}}; \qquad (1)$$

$$PER = \frac{P_{01}}{P_{10} + P_{01}}. \qquad (2)$$

To simulate the burst error, we generate a uniformly distributed random number r in (0,1). The transition from S_0 to S_1 occurs when r is less than P_{01} and the transition from S_1 to S_0 occurs when r is less than P_{10}.

With the assumption that any bit error will cause the packet error. Therefore PER = 1- (1-BER)^{L}, where L is the packet length. When the packet length becomes longer, PER becomes higher.

The time-bound for transmitting current GOP is determined by the duration T_1 before the last frame of the current GOP is scheduled to decode and display. Accordingly, the maximum number of packets F (including retransmitted packets) transmitted for the current GOP can be computed as:

$$F = \frac{T_1 * B}{S}; \qquad (3)$$

Where B is the estimated bandwidth of the wireless link and S is the packet size. B can be estimated fairly accurately and the details appear in the next section.

The maximum number of packets G (including retransmitted packets) transmitted for the next GOP with duration T_2 can be computed as:

$$G = \frac{T_2 * B}{S}; \qquad (4)$$

Suppose in the process of current GOP transmission, the next scheduled layer of current GOP is X_i (Y_j), which has the absolute distortion D (X_i (Y_j)). The expected distortion of that layer can computed as:

$$D(1) = D_E(X_i(Y_j)) = PER^F * D(X_i(Y_j)); \quad (5)$$

The scheduled layer from next GOP is $X_a(Y_b)$ with absolute distortion D($X_a(Y_b)$). The expected distortion of that layer can computed as:

$$D(2) = D_E(X_a(Y_b)) = PER^G * D(X_a Y_b); \qquad (6)$$

III. Simulation Studies

We use three video sequences (foreman, waterski and wg_cs_9) to measure the simulated performance. In our experiments, only the DC value is included in the base layer. Each enhancement layer contains one or more AC values. The higher enhancement layers include AC values with larger frequency than AC values contained in lower enhancement layers. PSNR (peak signal-to-noise ratio) is used to evaluate the layered video quality. Higher the value of PSNR, better the video quality is. Also, we assume that all the motion vectors are decoded correctly.

First, we analyzed video quality with respect to different number of AC values. The PSNR values are shown in table 1 and Figure 5. Second, the gains of PSNR by adding more DCT values are computed by deducting the PSNR value from the previous one in Table 1. Note that Table 1 and Figure 5 show the PSNR values of a typical frame.

Based on the observation of Table 1, we use the video layering policy shown in Table 2.

The layered video packets are transmitted over a wireless link. In equation (5) and (6), estimated wireless bandwidth is required for computation. The available wireless bandwidth can be estimated by the equation (7).

$$B_{next} = k * B_{past} + (1-k) * B_{curr} \qquad (7)$$

Where B_{next} is the estimated wireless channel bandwidth for the next period. B_{past} is the past estimate of wireless link bandwidth and B_{curr} is the measured net bandwidth in the current estimation period. Note that B_{past} of next estimation period equals to B_{next} of current estimation period. The estimation coefficient is k ($0<k<1$).

The estimated and the actual wireless channel bandwidths are shown in Figure 6. Figure 6 shows a good estimation of wireless bandwidth given by equation (7).

We compare our proposed on-line PSR-ARQ transmission scheme with the traditional scheme (SR ARQ). In our simulation, the PSNR values are obtained by averaging the absolute PSNR values of video GOPs transmitted by our scheme and the traditional scheme. The PSNR values of the proposed and the traditional schemes for different PER wireless link are presented in Figure 7. Figure 7 shows that the proposed transmission scheme performs better than the traditional scheme. As the PER increases, PSNR gain of our scheme over the traditional scheme increases.

IV. Conclusion

In this paper, we have presented a scheme for layered video streaming over a wireless link. Our PSR-ARQ scheme minimizes the significant packets' failure ratio by maximizing the number of retransmission within the delay bound. Our transmission policy achieves good video quality (of low distortion). Simulation results show that our scheme performs better than traditional scheme for video transmission over a single wireless link with limited bandwidth and high burst errors. If the wireless link provides QoS, for example when using 802.11e, it is clear that the wireless link as viewed by our scheme exhibits good behavior (low latency, etc) and hence the video quality will be good. Note that since our scheme is adaptive, the video quality is automatically adjusted based on the quality of the wireless link. Many concurrent schemes can be supported on a single wireless channel of sufficient bandwidth due to the very small real-time processing overhead. Since significant preprocessing is done, supporting live video requires large buffering delay at the sender and hence the end-to-end latency can be large. If audio is encoded in a layered way, our approach can easily be adapted to audio streaming also. The details of the encoding method used in audio will be needed to accomplish this.

Our future work will be video transmission over multi-hop wireless links or wired/wireless links.

Table 1 PSNR with different DCT values

	wg_cs_9	foreman	waterski
DC	24.25	26.47	29.04
DC+AC1	24.66	26.35	29.09
DC+AC1~2	32.84	34.14	34.28
DC+AC1~3	36.02	36.4	36.7
DC+AC1~4	36.66	37.19	36.98
DC+AC1~5	36.89	37.1	37.04
DC+AC1~6	36.78	37.19	37.06
DC+AC1~7	36.87	37.18	37.11
DC+AC1~8	37.48	37.78	37.21
DC+AC1~9	40.65	39.68	38.81
DC+AC1~10	42.68	40.73	40.12
DC+AC1~11	44.14	41.57	40.46
DC+AC1~12	44.74	41.98	40.59
DC+AC1~13	44.61	41.73	40.55
DC+AC1~14	43.82	41.71	40.56

Table 2 Video layering policy

I-Frame	DCT values
Base layer	DC
Enhancement 1	AC 1 - AC 3
Enhancement 2	AC 4 - AC 8
Enhancement 3	AC 9 - AC 14
P/B Frame	DCT values
Base layer	DC + AC 1- AC 2
Enhancement 1	AC 3 - AC 14

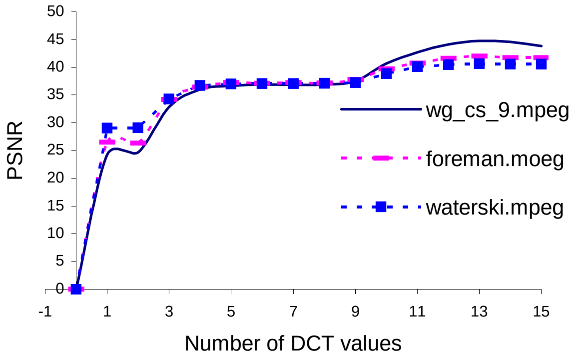

Fig. 5 PSNR with different DCT values

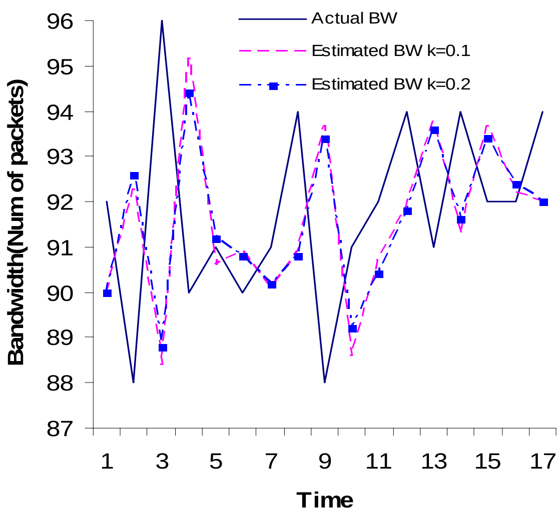

Fig. 6 Wireless link bandwidth estimation

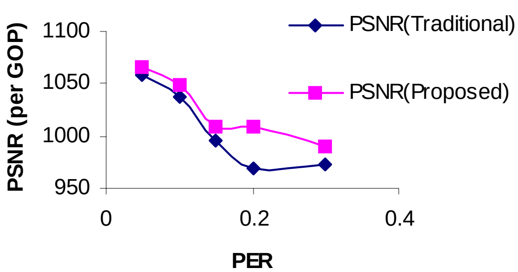

Fig. 7 Proposed vs. Traditional Transmission Scheme

REFERENCES

1. S.Varadarajan, H.Q.Ngo, J.Srivastava, "An adaptive, perception-driven error spreading scheme in continuous media streaming" *Distributed Computing Systems, pp.475-483, Apr. 2000.*
2. J.Y.Pyun, J.J.Shim, S.J.Ko, S.H.Park, "Packet loss resilience for video stream over the Internet" *IEEE Trans. Consumer Electronics Vol.48 No.3.pp.556-563 Aug. 2002.*
3. T.A.Lee, S.G.Chan, Q.Zhang,W.W. Zhu, Y.Q.Zhang, "Allocation of layer bandwidths and FECs for video multicast over wired and wireless networks" *IEEE Trans. Circuits and Syst. Video Technol Vol 12, No.12,pp1059-1070, Dec,2002.*
4. B.Girod, N.Farber, "Feedback-based error control for mobile video transmission" *PROCEEDINGS OF THE IEEE, VOL. 87, NO. 10, pp.1707-1723.Oct. 1999.*
5. M.Podolsky, S.McCanne, M.Vetterli "soft ARQ for layered streaming video" *Tech.rep. UCB/CSD-98-1024, University of California, Computer Science Department, Berkeley, CA, Nov.1998.*
6. N.Y.Yin,S.Q.Li,T.E.Stern "Congestion control for packet voice by selective packet discarding" *IEEE Trans. Communications, Vol.38, No.5, pp.674-683.May.1990.*
7. H.Liu, M.E.Zarki "Performance of H.263 video transmission over wireless channels using hybrid ARQ" *IEEE Journal of Selected Areas in Communications, Vol.15,No.9 pp.1775-1786 Dec.1997.*
8. C.H.Chou, C.W.Chen "A perceptually optimized 3-D subband codec for video communication over wireless channels" *IEEE Transactions on Circuits and Systems for Video Technology Vol.6, No.2pp.386-397 Apr.1996.*
9. *S. Wang, H. .Zheng, J.A.Copeland, "An error control design for multimedia wireless networks," IEEE Vehicular Technology Conference (VTC2000-Spring).Vol, 2 pp.795-799, 2000, Tokyo.*
10. *Q. Zhang, S.A. Kasam, "Hybrid ARQ with selective combining for video transmission over wireless channels," IEEE International Conference on Image Processing, Vol.2, pp. 692-695, 1997.*
11. *Deepak S.Turaga, Tsuhan Chen, "Model-Based Error Concealment for Wireless Video"IEEE Transactions on Circuits and Systems for Video Technology, Vo.l12,No.6, pp.483-495 June 2002.*
12. *Sunghyun Choi,Kang G.Shin, "A Class of Adaptive Hybrid ARQ Schemes for Wireless Links",IEEE Transactions on Vehicular Tech,Vol.50,No.3,pp.777-790May 2001.*
13. *Joan L. Mitchell, William B.Pennebaker,"MPEG video compression standard".*
14. S.Varadarajan, H.Q.Ngo, J.Srivastava, "Error spreading: a perception-driven approach to handling error in continuous media streaming",*IEEE*/ACM*Transaction on Networking Vol.10 Issue Feb 2002 Page 139-152.*