A REFERENCE PICTURE SELECTION SCHEME FOR VIDEO TRANSMISSION OVER AD-HOC NETWORKS USING MULTIPLE PATHS*

Shunan Lin, Shiwen Mao, Yao Wang, Shivendra Panwar

Department of Electrical Engineering, Polytechnic University
6 Metrotech, Brooklyn, NY, 11201, USA

ABSTRACT

Enabling video transmission over ad-hoc networks is more challenging than over conventional mobile networks because a connection path in an ad-hoc network is highly error-prone and the path can go down frequently. On the other hand, it is possible to establish multiple paths between a source and a destination, which provides an extra degree of freedom in coding algorithm design. This paper presents a feedback-based reference picture selection scheme for video transmission over ad-hoc networks. Encoded video streams are transmitted over multiple paths and the reference frames for motion compensated prediction are selected according to the feedback information about the paths' condition. Simulations under the two paths scenario have shown significant improvement over two standard techniques, layered coding and video redundancy coding, which do not use feedback. A novel statistical model for the ad-hoc multi-path environment is also proposed and used in our simulation of transmission loss.

1. INTRODUCTION

An ad-hoc network is a collection of mobile nodes that will create the network "on demand". The main differences between ad-hoc networks and conventional cellular technology are the lack of a centralized entity within ad-hoc networks and the independence from pre-existing infrastructure. Consequently, ad-hoc networks are an appealing option in applications where the desired infrastructure does not exist whether due to sparse population, economic condition, or after a disaster such as an earthquake.

Video transmission over ad-hoc networks is more challenging than over conventional mobile networks. Besides having quite high transmission bit error rates during fading periods, an ad-hoc network may undergo frequent and unpredictable changes in the network topology, which results in a relatively short lifetime of the network paths. Thus, paths may become frequently invalid during connections, which may cause severe degradation in decoded video quality. On the other hand, one characteristic of an ad-hoc network is that it is usually possible to establish more than one path between a source and a destination given their mesh topology. In this paper, we propose a video coding and transmission scheme that takes advantage of the availability of multiple paths for combating transmission errors.

Many error resilient video coding techniques have been proposed in recent years [1,2]. One approach is to encode the video signal into two layers and transmit them with unequal error protection. In ad-hoc scenario, these two layers can be fed into two different paths, with the base layer assigned to the better path. This technique, however, cannot stop error propagation caused by the predictive coding scheme used in most coding standards, if an error occurs on the path carrying the base layer. Another approach is video redundancy coding (VRC) [3], in which the image sequence is divided into two or more threads in a round-robin fashion and each sequence is encoded independently. It uses a so-called sync frame, which is encoded by all threads at regular intervals to limit error propagation within individual intervals. The penalty of this scheme is the high redundancy (i.e. loss in coding efficiency) due to the longer prediction distance. Furthermore, error propagation can still extend beyond an interval if all the threads encounter errors in the same sync interval. These schemes are designed to accommodate a range of operating conditions and cannot be very efficient for a specific environment. To handle the fast varying path conditions in ad-hoc networks, we propose to use feedback-based error resilient video coding techniques, which can offer higher coding efficiency and minimize error propagation and distinguishes this scheme from our previous work in the area of radio transmission over ad-hoc networks [4], which used multiple description coding with no feedback.

Feedback-based video coding techniques include error tracking, error confinement and reference picture selection (RPS) [5]. The RPS technique encodes the current video frame with reference to a previous frame selected based on the feedback information instead of always using the last transmitted frame. It can stop error propagation without incurring excess loss of coding efficiency, and has been included in the H.263+ standard as an optional tool [6]. A drawback of feedback-based approaches is that they can only be used in applications where a feedback channel is available. In ad-hoc networks, feedback can be either sent back through a different path, or it can be piggy-backed in the packets of the reverse video stream. The bandwidth it takes is minimal in either case since even in the first case the feedback packet size is much shorter than data packets.

The efficiency of feedback-based schemes depends on the delay involved in receiving the feedback information. The shorter the delay is, the more quickly the encoder can stop the error propagation, and the less distant frames are used as reference pictures. Note that the delay in receiving the feedback

---

* This work was supported in part by NSF Grant ANI 0081357 and by the New York State Center for Advanced Technology in Telecommunications (CATT).

2001 IEEE International Conference on Multimedia and Expo
ISBN 0-7695-1198-8/01 $17.00 © 2001 IEEE

0-7695-1198-8/01/$10.00 (C) 2001 IEEE

information does not cause extra encoding delay or the overall end-to-end delay. Therefore, feedback-based approaches are viable options even for interactive applications.

In an ad-hoc network, the propagation delay is usually negligible since the nodes are relatively near to each other, and the transmission delay for a packet essentially depends on the product of the number of links in a path and the transmission time of a packet plus any queuing delays. Leaving aside queuing delays, which are controllable by appropriate traffic control mechanisms, the end-to-end delay of a path is determined by the number of links, the bit rates of the links, and the size of the packets. Assuming that each packet contains data from a group of blocks (GOB) and that the number of links is lower than the number of GOBs in a video frame (this is usually true for ad-hoc networks because ad-hoc links only cover a small region), then the transmission delay would be shorter than a frame interval. If we further assume that the encoder and decoder each has a smoothing buffer equivalent to a frame interval, then the maximum delay between the time when a frame is encoded at the source and the time when the decoder can detect any loss in the frame will be less than 3 frames. Assuming further that the feedback information requires negligible time to send because of its small size, and assuming that there is no feedback error, we can legitimately assume that the maximum delay for receiving the feedback about a coded frame is 3 frames. If the actual encoding buffer delay and the transmission delay are shorter, than the feedback delay may be within 2 frames. As will be shown by simulation results, the feedback delay critically affects the encoding efficiency of the proposed scheme.

The paper is organized as follows. In section 2, the proposed coding scheme using reference picture selection is introduced. In section 3, the proposed model for the ad-hoc network is described. In section 4, we present simulation results and compare them with layered coding and VRC. Section 5 concludes the paper.

2. REFERENCE PICTURE SELECTION BASED ON FEEDBACK AND PATH STATUS PREDICTION

One of the main challenges in video coding for ad-hoc networks is how to limit the extent of error propagation caused by one bad path. Obviously, one way is to send independent bit streams among those paths. This can be accomplished by, say, the VRC scheme. The problem is, as mentioned above, the coding efficiency is low because a frame is always predicted from a frame that is more than one frame away. We note that there is no reason to forbid one path from using another path's frames as reference if all the paths are good. Motivated by this observation, we propose to choose the reference frames as follows: *always choose the last frame that is believed to be transmitted reliably (based on the feedback message and path status prediction) as the reference frame.*

Specifically, we sent the coded frames on separate paths. The mapping of frames to paths depends on the available bandwidth on each path. For example, in the two paths case, if the two paths have the same bandwidth, then even frames are sent on path one, and odd frames on path 2. We assume that a feedback message is sent back for each received frame by the decoder. If any packet (containing one GOB) in a frame is detected as lost, the decoder sends a negative feedback (NACK) for that frame. Otherwise, it sends a positive feedback (ACK). The feedback information for a frame may be sent on the same path in which the frame data are sent, or on a different path. We assume that the feedback message is always delivered intact by using appropriate transport control. This, for example, can be accomplished through the Receiver Report in the RTP protocol. An encoder receives the feedback message for frame *n-d* when coding frame *n*, where *d* is the expected maximum delay in terms of frame intervals. Furthermore, once a NACK is received for a frame delivered on one path, we assume that path remains "bad" until an ACK is received. Similarly, we assume the path stays in the "good" status until a NACK is received. When encoding a new frame (which can be sent on either path), the encoder deduces the last possibly correctly decoded frame, based on the feedback messages received up to this time and the above method of path status prediction, and uses that frame (which may have been sent on either path) as the reference frame.

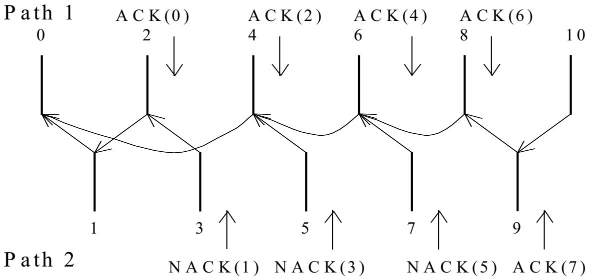

Figure1 Illustration of the RPS scheme for the two-path scenario. The arrow associated with each frame indicates the reference used to code that frame.

Figure 1 is an example of the proposed reference selection scheme. Here we assume *d=3.* When NACK (1) is received at the time for coding frame 4, the encoder knows that frames 2 and 3 cannot be decoded correctly due to error propagation. Therefore, frame 0 is chosen as reference for frame 4. Furthermore, path 2 is set to bad status. When encoding frame 6, because path 2 is still in the "bad" status the encoder uses frame 4 instead of frame 5 as reference frame. On the other hand, when ACK (7) is received, path 2 is changed to "good" status, so frame 9 is chosen as the reference of frame 10.

The proposed coding scheme offers a good trade-off between coding efficiency and error resilience. When both paths are good, it uses the immediate neighboring frame as the reference, thereby achieving the highest possible prediction gain and consequently coding efficiency[*]. On the other hand, when one path is bad, the encoder avoids using any frames that are affected by path errors, thereby minimizing the error propagation period.

---

[*] For certain sequences such as those with periodically repeated backgrounds, it is possible that using a frame further away can achieve higher prediction accuracy.

3. MODELING OF THE AD-HOC NETWORK

To verify the performance of our RPS scheme, we simulate the transmission over two lossy paths. To simulate the path errors, we propose a new model for the ad-hoc network.

Link Model Although it is well-known that wireless channels are time varying because of user mobility and multi-path propagation, experiments for various types of channels show that the basic channel parameters can be stable for short time intervals. Therefore a non-stationary wireless channel can be adequately represented by a set of stationary channel models[8,9]. There has been extensive literature on wireless channel modeling [7,8,9]. The classical two-state Gilbert-Elliott model for burst noise channels has been widely used [7]. In [8,9], K-state Markov models are used in which the received instantaneous signal-to-noise ratio (SNR) is partitioned into K ranges where each range of SNR corresponds to one state. Various techniques are proposed to derive the transition probabilities of the Markov process given the partitioning.

In this work, we use an N+1 state discrete time Markov process to model a wireless link in the ad-hoc network. Let $S=\{s_0, s_1, ..., s_N\}$ denotes the finite state space of the Markov process. As shown in Fig 2, the transition probabilities are $\{a_i, i=1,2,...,N\}$ and $\{b_i, , i=1,2,...,N\}$, respectively. Each state i is associated with a Packet Loss probability (PLP) P_i. Without loss of generality, we assume Pi's are in decreasing order, with $P_0=1$ (the link is down) and $P_N=0$ (no loss at all). Equilibrium probabilities of the states can be easily derived from the one-step transition matrix. The parameters can be derived from measurement data available.

We assume that there is no abrupt change of the quality of a link, i.e., transitions only happen between adjacent states and only occur at the beginning of the transmission of each packet. So each packet will experience a fixed loss probability when it is transmitted over a link. This is reasonable because the movement of the mobiles, fading, the power drain on batteries, and interference of the channels are all happening in a relatively slow time scale. Abrupt changes are not likely if we don't consider events such as the sudden failure of a mobile node.

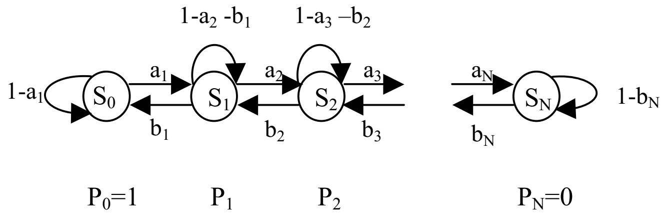

Figure 2. State Diagram of the Link Model

Path Model Most previous work on wireless channel modeling focuses on modeling a fading channel with a single hop [7,8,9]. In Ad-hoc networks, a path consists of a number of wireless links, which can be dependent or independent of each other. Furthermore, in the multi-path case, the availability of a path, and the correlation of the paths are also important factors in addition to the loss rates and the available bandwidth [7]. Therefore the single hop model may be inadequate if we wish to take into account such issues. We model an ad-hoc path as follows: Let $\{L_i, i=1,2,...,N\}$ represents a pool of links that can be used to form a set of paths from a source node to a destination node, e.g., the link pool in [10]. Each link in the pool is modeled as Fig 2. A path is constructed by randomly choosing a subset of links from the pool. A packet goes through each link along the path and experiences a different loss probability on each of them. The number of common links shared by the two paths determines the degree of their correlation. Common links are perfectly correlated while different links are assumed independent. More common links therefore implies higher correlation. When one of the links of a path is in state S_0, the path is down and not available. We can either choose a new set of links after a rerouting delay, i.e., a routing process is activated and a new path is found; or we can wait until the link transits from state S_0 to S_1, i.e., the link is up again. In our simulation, when the link is down, all packets are dropped and we wait for the link to go up again without switching to another path.

4. SIMULATION RESULTS

The proposed scheme was tested on the sequence "Foreman". The format is QCIF, the original frame rate is 30 fps, and the target coded frame rate is 10 fps. We implemented the proposed RPS scheme by modifying the public domain H.263+ codec [11,6]. The TMN 8.0 rate control algorithm was used to produce constant rate bit streams. In the encoder, we limited the reference frame buffer to L=12 frames. A frame was coded as I frame if the chosen reference frame is beyond 12 frames. We only simulated the case when two paths with the same bandwidth are available. The number of links in each path and the model parameters for the links are varied to simulate different path environments. The coded frames are sent alternately on the two paths. The first two frames are coded in the intra-mode and transmitted on separate paths. The feedback time was fixed at 300ms (d=3 frames time). For comparison, VRC and the two-layer SNR scalable mode in H.263+ were also tested for the same network parameters. For layered coding, we use intra macroblock refreshment technique with 5% refresh rate in both layers to mitigate the error propagation effect. For VRC, we used the 2-5 mode [3], which means there are two threads and one sync frame in every five frames in each thread.

One GOB was put into one packet but the packet header overhead was not counted in the bit rate. In the decoder, different operations were performed for different coding scheme. For the proposed RPS scheme, if any one GOB was lost, a NACK of that frame was sent to the encoder and the GOB was decoded by copying the corresponding GOB in the last decoded frame. For VRC, a lost GOB was decoded by copying the GOB from the same position of last frame if current frame is not the sync frame, otherwise, the decoder would check another sync frame and chose the better GOB. For the layered coding, the base layer's error concealment strategy was the same as RPS whereas the enhancement layer just copies the corresponding GOB from the base layer. To get the best quality from layered coding, we transmitted the base layer over the better path and the enhancement layer over the other path.

A three-state Markov model was used for the links on a path. The loss probability for the states are: P_0=1, P_1, and P_2=0, respectively, which implies that each link can be down, error free, or up but with some loss probability P_1. Each path has less than 5 links, so the feedback delay will be less than 3-frame times as discussed in the introduction. In experiments 1 and 2, we use two symmetric paths, while in experiments 3 and 4 two asymmetric paths are used. For each set of path model parameters, 30 different realizations of the packet loss patterns are run. The average packet loss rates on the two paths for the first two experiments are (7%, 7%) and (15%,15%), respectively, and are (7%,15), (6%,25%) for the last two experiments. Table 1 gives the average (over different realizations) PSNR of the reconstructed frames at the decoder for different path parameter sets. The bit rate of RPS is 70kbps on each path and bit rates of VRC and layered coding are 71.6kbps and 71.5kbps for each path respectively. It can be seen that the proposed RPS scheme outperforms the other two schemes in all the scenarios. The layered coder outperforms VRC when the two paths have asymmetric loss characteristics, but is still worse than the proposed scheme.

Figure 3 shows the PSNR vs. frame number for a particular realization of a path model in which the two paths have identical model parameters and consequently the same average packet loss rate (7%). We can see that all the schemes suffer from severe quality degradation when both paths fail (after frame 110). But the RPS scheme recovers from this degradation and stops the error propagation more quickly. When both paths are good (between frames 60 and 110), our scheme also outperforms the other two because in this case the coder reduces to a conventional coder that is optimized for coding efficiency.

To evaluate the effect of feedback delay, we also performed another set of simulation of the proposed RPS scheme assuming d=2. The average PSNRs are also included in Table 1. On average, the improvement in PSNR is about 0.5 dB when the feedback delay changes from d=3 to d=2.

Table 1: Average PSNR of the Y component of the decoded frames

Packet loss rate(%)	7,7	15,15	7,15	6,25
RPS (d = 2)	29.3	27.0	27.8	26.5
RPS (d = 3)	28.8	26.6	27.6	26.2
VRC	26.6	21.7	24.0	22.6
Two layered	25.0	22.0	24.5	23.7
h.263 w/o option	22.5	19.4	20.4	19.4

5. CONCLUSION

In this paper, we proposed a feedback-based reference picture selection scheme in video coding for transmission over ad-hoc networks. By selecting reference pictures according to the predicted status of the path condition, which in turn depends on the feedback message, we can achieve high resilience to path errors (packet loss as well as path failures) at a slight cost of coding efficiency. A practical advantage of this approach is that it is compatible with the H.263+ standard. A novel model for an ad-hoc network is developed and used in our simulation of packet loss patterns.

REFERENCES

[1] Y. Wang and Q. Zhu, " Error control and concealment for video communication: a review," Proc. IEEE, vol. 86, pp. 974-997, May, 1998.

[2] Y. Wang, S. Wenger, J. Wen and A. K. Katsaggelos, "Error resilient video coding techniques," IEEE Signal Magazine, vol 17, pp. 61-82, July, 2000.

[3] S. Wenger, " Video redundancy coding in H.263+," in Proc. AVSPN, Aberdeen, UK, Sept.1997

[4] N.Gogate,S.Panwar, "Supporting video/image applications in a mobile multihop radio environment using route diversity," IEEE ICC, Vancouver CA, June. 1999.

[5] B. Girod and N. Färber, " Feedback-based error control for mobile video transmission," Proc. IEEE, vol 87, pp. 1707-1723, Oct. 1999.

[6] ITU-T Recommendation H.263, "Video coding for low bit rate communication," 1998.

[7] E. N. Gilbert, "Capacity of a Burst-Noise Channel", Bell Syst. Tech. J., Vol.39, No.9, pp.1253~1265, Sept. 1960.

[8] H. S. Wang, and N. Moayeri, "Finite-State Markov Channel-A Useful Model for Radio Communication Channels", IEEE Transaction on Vehicular Technology, Vol.44, No.1, pp. 163-171, Feb. 1995.

[9] Q. Zhang et al, "Finite-State Markov Model for Rayleigh Fading Channels", IEEE Transactions on Communications, Vol.47, No.11, pp.1688-1692, Nov. 1999.

[10] M. R. Pearlman et al, "On the Impact of Alternate Path Routing for Load Balancing in Mobile Ad-Hoc Networks", Proceedings of Mobile and Ad Hoc Networking and Computing 2000, pp. 3-10, Aug. 2000.

[11] University of British Columbia, "H.263+ codec", ftp://dspftp.ece.ubc.ca

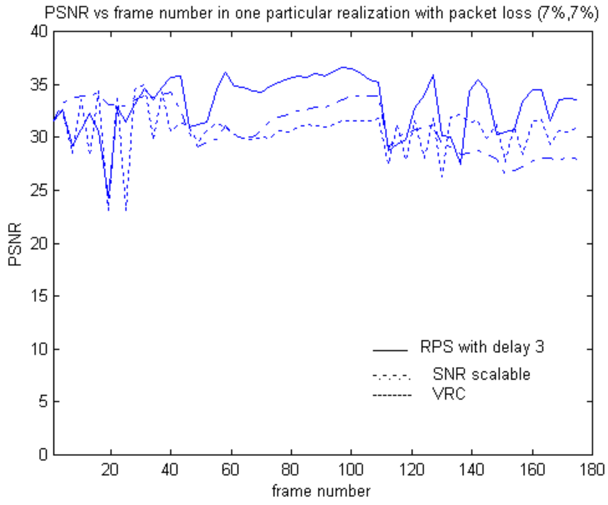

Figure 3 PSNR vs frame number in one realization of packet loss for the network model (7%,7%)