Routing-aware Multiple Description Coding with Multipath Transport for Video Delivered over Mobile Ad-hoc Networks

Yiting Liao and Jerry D. Gibson
Department of Electrical and Computer Engineering
University of California, Santa Barbara, CA, 93106
Email: {yiting, gibson}@ece.ucsb.edu

***Abstract*—This paper proposes a cross-layer approach called routing-aware multiple description coding with multipath transport to support video communications over mobile ad-hoc networks. This approach establishes a packet loss model based on the MAC access mechanism and network parameters, and utilizes it along with the routing messages from multipath routing to estimate the packet loss probability of transmitted video packets. Then the estimated results are passed to the application layer to assist reference frame selection for multiple description coding in order to mitigate error propagation introduced in the motion-compensated loop. Results show that this is an effective approach to improve error resilience of video transmission over mobile ad-hoc networks and enhance the video experience for multiple users.**

I. Introduction

There has been a growing interest in video communications over mobile ad-hoc networks due to the emerging applications of ad hoc networks in military, homeland defense, and disaster recovery scenarios. However, mobile ad-hoc networks impose significant challenges to video transmissions since node mobility and the lack of infrastructure in the network can lead to frequent link failures and route changes. Furthermore, the link availability is affected by fading and interference in the wireless channels. As a result, video transmission over such networks experiences both random and burst losses, which can severely degrade the delivered video quality.

Among a number of solutions to address this problem, multiple description coding (MDC) with multipath transport (MPT) has been shown to be a very promising technique [1]–[4]. With MDC, a video sequence is encoded into two descriptions such that each description can be used to reconstruct the video with low but acceptable video quality while both descriptions together provide higher video quality. In addition, combining MDC with MPT can reduce the likelihood of simultaneous loss of both descriptions and enables load balancing in the networks.

When video packets transmitted over mobile ad-hoc networks suffer bursty packet losses that do not affect both descriptions simultaneously, MDC can effectively provide adequate quality by decoding the correctly received description. However, in a practical mobile ad-hoc network, both burst and random losses may appear in two descriptions at distinct times or simultaneously. Due to the motion compensated prediction loop employed in most of the MDC coders, the transmission errors can propagate to subsequent frames and cause great distortion in delivered video. Many approaches have been proposed to mitigate error propagation in MDC [5]–[7], yet it comes at the cost of coding redundancy. Other solutions to address the problem are traffic allocation and path selection for MDC with MPT [8]–[11], in which video packets are spread over different paths based on the error characteristics of paths to minimize end-to-end distortion.

In this paper, we propose a routing-aware MDC approach with MPT to alleviate error propagation caused by packet losses. This approach explores the relationship between packet losses and standardly available ad-hoc routing messages to estimate the packet loss probability of each transmitted video packet, and then apply a threshold-based algorithm to adaptively select reference frames in MDC. By avoiding using possible corrupted frames as a reference, this routing-aware approach can effectively mitigate error propagation in the motion-compensated prediction loop. Unlike common reference picture selection (RPS) work [12], [13], our approach does not require any extra channel feedback but retrieves information from normal routing messages. Thus it neither introduces extra cost nor additional delay to the transmission system and is suitable for real-time video applications over mobile ad-hoc networks. It is also different from a recent work [14] that jointly optimizes multipath routing and coding rate selection, because our approach does not require any change to routing protocols but just relies on standard routing messages.

In Section II, we discuss our routing-aware MDC approach with MPT, including the packet loss model based on the MAC access mechanism and network parameters, the multipath routing protocol with routing message feedback, and the reference frame selection scheme applied to the MD encoder. Then we use the QualNet simulator to simulate an ad-hoc network with 50 nodes and investigate the end-to-end video performance in Section III. We show that our proposed method not only provides consistent PSNR gains but also guarantees good perceptual video quality for multiple users.

This research has been supported by the UC Discovery Grant Program and Applied Signal Technology, Inc.

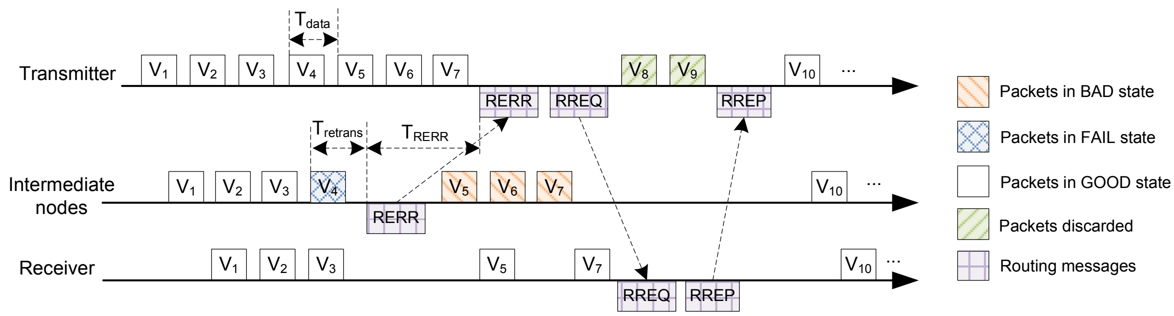


Fig. 2. An example to illustrate the packet losses in the network and the corresponding routing messages

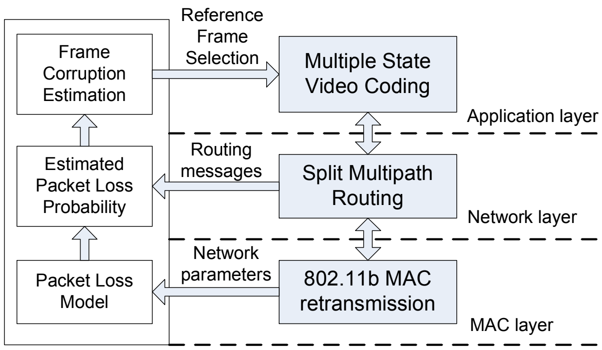


Fig. 1. Cross layer design: Routing-aware Multiple Description Coding with Multipath Transport

II. Routing-aware Multiple Description Coding with Multipath Transport

We present the system framework of our proposed approach in Fig. 1. As shown in Fig. 1, our approach includes an interplay of the MAC, network, and application layers. At the MAC layer, CSMA/CA scheme in 802.11b is used as the basic access mechanism. By considering the MAC layer protocol and network parameters, we establish a packet loss model for the preceding packets transmitted from the source node before a route error (RERR) message is received. At the network layer, a multipath routing protocol is implemented to support the transmission of two descriptions. Meanwhile, the routing messages along with the packet loss model are used to estimate the packet loss probability of transmitted video packets. The estimated results are then fed back to the application layer to determine the frame corruption probability and help the reference frame selection at the multiple description (MD) encoder. In the following, we discuss the proposed solutions in the three layers respectively.

A. MAC layer: Packet Loss Estimation Model

In the MAC layer, we build a packet loss estimation model based on the 802.11b MAC access mechanism and network parameters. This model models the packet loss probability of the packets previously sent from the source node before a RERR message is received.

As shown in Fig. 2, a RERR is initiated at the intermediate node when video packet v_4 exhausts all retransmission attempts and still fails to transmit to the next hop destination. We define the retransmission delay of this packet as T_{retrans}. After time T_{RERR}, the source node receives the RERR and stops transmitting video packets through the unreliable link. We see that packets v_5, v_6, v_7 sent during time period $T_{\text{retrans}}+T_{\text{RERR}}$ are still transmitted through the unreliable link and are very susceptible to packet loss. We derive a model to estimate the packet loss probability of the preceding packets sent through this unreliable link.

We denote $\text{Pr}(n)$ as the packet loss probability of the n^{th} preceding packet sent from the source node before the source node receives a RERR. Due to the random delay between link failure and RERR reception at the source, the n^{th} preceding packet before RERR can be sent at a time before, right at or after the link failure happens. We use three states to represent these three cases: GOOD means the packet is sent before the link failure, FAIL means the packet fails to transmit and triggers RERR, and BAD means the packet is sent after the link failure. We define $\text{Pr}(n)$ as

$$\text{Pr}(n) = \lambda_g \cdot p_g(n) + \lambda_f \cdot p_f(n) + \lambda_b \cdot p_b(n) \qquad (1)$$

where λ_g, λ_f, and λ_b represents the packet loss probability in GOOD, FAIL, or BAD state respectively, and $p_g(n)$, $p_f(n)$, and $p_b(n)$ denotes the probability of the n^{th} preceding packet in these three states, respectively. In our prior work [15], we discuss how to estimate the state probability distribution and the packet loss probability of λ_g, λ_f, and λ_b in detail. Based on the work in [15], we establish a packet loss model that estimates the values of $\text{Pr}(n)$. Next, we incorporate the packet loss model with the routing messages to estimate the packet loss probability of each transmitted packet.

B. Network layer: Multipath Routing

Combining MDC with MPT is an appealing approach because it provides error resilience as well as load balancing for video transmission over ad-hoc networks. To support MDC with path diversity, a multipath routing protocol is required to build multiple paths between the source and destination nodes in the ad hoc network. In our design, we implement split multipath routing (SMR) as our multipath routing protocol due to its popularity and simplicity [16]. SMR establishes multiple paths of maximally disjoint paths to avoid congestion in certain links and to efficiently utilize the network resources [17].

During the route discovery process in the SMR protocol, the destination sends a route reply (RREP) after receiving the first route request (RREQ) to construct the first route and waits a period of time to receive more RREQs. It then selects the route that is maximally disjoint to the first replied route and sends another RREP to create the second route. Once the source node receives the two RREPs, two routes are created and two video descriptions are sent through the two routes separately.

Whenever an intermediate node fails all retransmission attempts to send a packet to the next hop destination, it sends a RERR message to the source node to indicate a broken link. In Section II-A, we model the packet loss probability of preceding packets transmitted from the source node when the source node receives a RERR. At the network layer, we feed back every RERR and utilize the model to estimate the packet loss probability of each transmitted packet. Furthermore, before the source node receives a RREP to reconstruct the broken route, we discard the packets scheduled to be transmitted through that broken route and mark their packet loss probability as 1. Next, we use the estimated packet loss probability to assist the video coding at the application layer.

C. Application layer: Routing-aware Multiple Description Coding

MDC is an effective approach to enhance the error resilience of video transmission over lossy networks. The general idea is to encode the video sequence into several descriptions with equal importance. Each description can be decoded independently or combined with other descriptions for reconstruction. In general, the reconstructed video achieves better video quality when more descriptions are received.

Among the many proposed MDC algorithms [18], multiple state video coding (MSVC) proposed by Apostolopoulos in [1] is a very popular method since it is easy to implement and compatible with different video standards. Thus, we apply MSVC to our MD video encoder. At the encoder, the video sequence is temporally downsampled into two sub-sequences with odd and even frames, and the odd and even frames are encoded as two descriptions using an H.264 encoder.

During the encoding process, we utilize the packet loss probability of each video packet retrieved from network layer to determine the frame corruption probability by

$$p(f_k) = 1 - \prod_{\{v_i | v_i \in f_k\}} (1 - p(v_i)) \qquad (2)$$

where $p(v_i)$ is the packet loss probability of packet v_i, $p(f_k)$ is the frame corruption probability of frame f_k, and $\{v_i | v_i \in f_k\}$ is the set of packets that contain information of frame f_k.

We define a threshold p_{thres} to determine whether a frame is corrupted or not, i.e. if $p(f_k) \geq p_{\text{thres}}$, we consider frame f_k as corrupted. Then we utilize the frame corruption estimation results to assist the reference frame selection for MDC. We initialize the reference list that consists of previously encoded frames in the same description. Next, we remove the estimated corrupted frames from that list. If all frames are removed from the reference list, we check the previously encoded frames in the other description and add frames that are not estimated as corrupted to the list. The current frame is encoded using the reference frames in the list and transmitted over the network. By not using the possible damaged frame as reference, we expect to reduce error propagation due to packet losses. Moreover, our proposed approach only relies on the standard ad-hoc routing messages and it does not incur any extra overhead. We refer this method as routing-aware multiple description coding (RA-MDC) with MPT.

At the decoder, we utilize the MSVC decoder with the refined error concealment method as proposed in [19]. When the decoder receives the corrupted descriptions, it decodes the correctly received MBs and conceals the lost MBs with the refined MB concealment method that considers the information from both descriptions for better recovery. The refined intra MB concealment reconstructs the lost MBs in the intra frames by using the temporal correlation between adjacent intra frames in two descriptions, while the refined inter MB concealment uses an additional reference list to perform the motion-compensated concealment. Finally, the concealed descriptions are interleaved to achieve the final reconstruction.

III. Experimental Results

A. Network Setup

We simulated our proposed RA-MDC method with MPT using the modified JM codec and the Qualnet simulator and compared the end-to-end performance with single description coding (SDC) and MDC with MPT. For these three methods, we use the same MPT strategy such that even and odd frames are transported through two separate routes.

In the network, 50 nodes are uniformly placed in a $500m \times 500m$ region. The movement of each node is characterized by a random waypoint model [20] with the node speed in the range of $0 \sim 10$ m/s and a pause time of 120 s. The transmission power of each node is 15 dBm. We use IEEE 802.11b with 5.5 Mbps PHY transmission rate and CSMA/CA basic access protocol. The values of IEEE 802.11b parameters used for the packet loss probability model can be found in [21]. A pair of source and destination nodes is randomly chosen to transmit video packets and packets are dropped if they do not reach the destination by the playout deadline of 350 ms. Unless otherwise specified, the above settings are chosen in the simulations.

We consider five video sequences “Foreman”, “Coastguard”, “Mother-daughter”, “News”, and “Silent”, which are all at CIF format with 150 frames at a frame rate of 15 fps. The video sequences are encoded into RTP packets with a packet size of 500 bytes. We generate two descriptions for each video sequence and the bitrate of each video sequence is 400 kbps, which corresponds to a bitrate of 200 kbps for each description. The two descriptions are transmitted through two paths over the network. For each network scenario, each video sequence is sent repeatedly 500 times to generate a statistically meaningful quality measure.

TABLE I

AVERAGE PSNR FOR CODED FOREMAN SEQUENCE AT 400 KBPS WITHOUT AND WITH TRANSMISSION LOSSES

PSNR (dB)	SDC	MDC	RA-MDC
Without losses	35.77	34.56	34.50
With losses (p=4.5%)	32.20	32.45	33.51

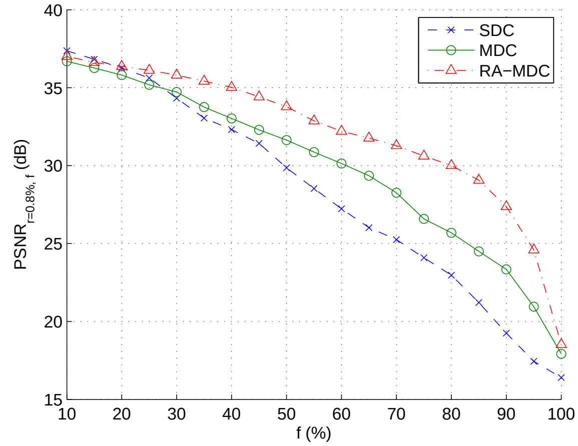
We use average PSNR of all frames over all realizations to evaluate the objective video quality of the decoded video sequences. In addition, we introduce PSNR$_{r,f}$ proposed in [22], [23] as a multiuser perceptual video quality indicator. PSNR$_{r,f}$ is defined as the PSNR surpassed by $f\%$ of the video frames for $r\%$ of the realizations, which shows the video quality guaranteed for $r\%$ of realizations among $f\%$ of the frames.

B. Performance Evaluation

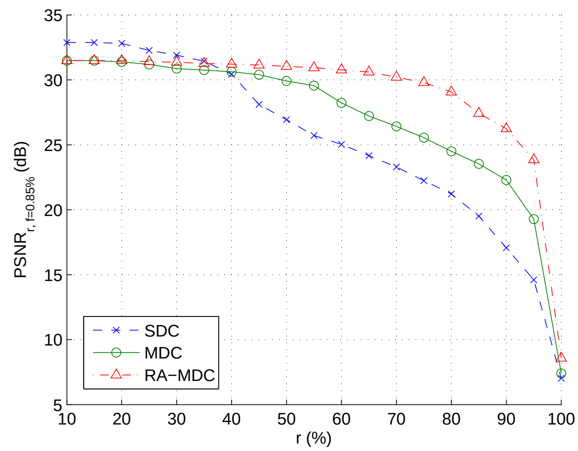
First, we examine the case that the transmission power of each node is 15 dBm, which leads to an overall packet loss rate around 4.5%. We look at the PSNR performance of the three methods for the Foreman sequence without and with transmission losses in Table I. We see that for the coded Foreman sequence without transmission losses, SDC achieves highest PSNR for the same bitrate, while MDC has a PSNR slightly higher than RA-MDC. On the other hand, RA-MDC achieves the highest average PSNR in the presence of moderate transmission losses. The results show that both MDC and RA-MDC trade coding efficiency for the reconstructed video quality under transmission losses, while RA-MDC provides a better tradeoff between coding efficiency and error resilience. Based on the frame loss estimation results in RA-MDC, fewer frames are used as reference for RA-MDC, which leads to a 0.06 dB lower PSNR than MDC when there is no transmission loss. However, the RA-MDC achieves 1 dB gain in PSNR under transmission losses, since it can effectively stop error propagation by not using corrupted frames as a reference frame.

Next, we compare the PSNR$_{r,f}$ results with fixed values of r and f. Figure 3(a) shows the PSNR$_{r,f}$ values for the three coding methods with fixed $r = 80\%$, which indicates the delivered video quality guaranteed for 80% of the realizations (users) for f percentage of the frames. In Fig. 3(a), we see that about 28%, 16%, and 5% of the frames in 80% of the realizations have a PSNR lower than 25 dB for SDC, MDC, and RA-MDC respectively. This suggests that RA-MDC has the fewest bad-quality frames that may attract the viewer's attention and make the video visually annoying.

Figure 3(b) presents PSNR$_{r,f}$ for SDC, MDC, and RA-MDC with fixed $f = 85\%$. This figure shows the video quality achieved for $r\%$ of the users in 85% of the frames. We see that RA-MDC guarantees a better video quality for most of the realizations compared to the other two methods. For example, RA-MDC guarantees a PSNR of 29.07 dB for 85% of the frames in 80% of the realizations, while SDC and MDC can only guarantee a PSNR of 24.50 dB and 21.22 dB for the same values of r and f. We also see that RA-MDC has a flatter curve



(a) PSNR$_{r=80\%,f}$



(b) PSNR$_{r,f=85\%}$

Fig. 3. Comparing PSNR$_{r,f}$ of SDC, MDC, and RA-MDC, Foreman sequence (CIF, 15fps) at 400 kbps, packet loss rate 4.5%. Average PSNRs of SDC, MDC, and RA-MDC are 32.20 dB, 32.45 dB, and 33.51 dB respectively

than SDC and MDC, which indicates that for RA-MDC, the video experience of multiple users has a smaller variance such that PSNR$_{f=85\%}$ for over 70% of the users is greater than 30 dB. On the other hand, PSNR$_{f=85\%}$ for 70% of the users for SDC varies in the range of 23.29 to 32.87 dB, which means that some of the users experience good video quality while the others have a bad video experience. Figure 3 shows that our proposed RA-MDC not only achieves the highest objective video quality but also provides good performance for multiple users.

Finally, we present the performance of the three methods for different video sequences under two typical packet loss rates in Table II. We see that RA-MDC consistently achieves the best PSNR and PSNR$_{r=80\%,f=85\%}$ for all video sequences under different packet loss rates. RA-MDC achieves gains in PSNR in the range of 0.7-2.3 dB compared to SDC, and gains in PSNR in the range of 0.7-1.4 dB compared to MDC. The performance improvement of RA-MDC in PSNR increases as packet loss rate increases since both SDC and

TABLE II
PERFORMANCE FOR DIFFERENT VIDEO SEQUENCES UNDER DIFFERENT PACKET LOSS RATES

	packet loss rate 4.5%						packet loss rate 8.8%					
	PSNR			$PSNR_{r=80\%,f=85\%}$			PSNR			$PSNR_{r=80\%,f=85\%}$		
	SDC	MDC	RA-MDC	SDC	MDC	RA-MDC	SDC	MDC	RA-MDC	SDC	MDC	RA-MDC
Coastguard	28.30	28.38	29.03	21.42	23.77	27.76	27.16	27.62	28.52	18.79	22.30	25.16
Foreman	32.20	32.45	33.51	21.22	24.49	29.07	30.51	31.23	32.67	18.67	22.49	26.84
Mother-daughter	39.40	39.86	40.63	27.86	34.20	38.66	37.91	38.91	39.99	24.29	32.00	34.28
News	37.24	37.50	38.53	22.47	29.63	36.06	35.43	36.38	37.72	20.42	27.89	31.41
Silent	35.13	35.20	35.89	24.93	29.66	33.85	33.61	34.32	35.24	20.32	26.59	29.32

MDC suffer more error propagation under more transmission errors. Furthermore, RA-MDC increases $PSNR_{r=80\%,f=85\%}$ by up to 13.6 dB as compared to SDC and by up to 6.4 dB as compared to MDC, which indicates substantial improvement for multiple users over the network.

IV. CONCLUSION

In this paper, we propose a cross-layer approach to support video transmission over mobile ad-hoc networks. The approach incorporates a routing-aware reference frame selection method with multiple description coding and multipath routing to enhance the error resilience of the delivered video. We discuss the transmission and coding mechanism at the MAC, network, and application layers and how the interplay among these layers can be beneficial to video transmission. The main advantage of our proposed approach is that it neither requires extra feedback channel nor incurs any extra overhead, but just utilizes the known network parameters and routing information to retrieve packet loss information embedded at the network layer. In addition, it only introduces negligible reduction in coding efficiency and does not cause extra delay. Therefore, our proposed approach is suitable for real-time video applications over mobile ad-hoc networks.

REFERENCES

[1] J. G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," in *SPIE Proceedings on Visual Communications and Image Processing*, vol. 4310, no. 1, 2001, pp. 392–409.

[2] N. Gogate, D. M. Chung, S. S. Panwar, Y. Wang, F. N. Commun, and P. River, "Supporting image and video applications in a multihop radio environment using path diversity and multiple description coding," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 9, pp. 777–792, 2002.

[3] S. Mao, S. Lin, S. S. Panwar, Y. Wang, and E. Celebi, "Video transport over ad hoc networks: multistream coding with multipath transport," *IEEE Journal on Selected Areas in Communications*, vol. 21, no. 10, pp. 1721–1737, 2003.

[4] B. A. Heng, J. G. Apostolopoulos, and J. S. Lim, "End-to-end rate-distortion optimized MD mode selection for multiple description video coding," *EURASIP Journal on Applied Signal Processing*, 2006.

[5] V. A. Vaishampayan and S. John, "Balanced interframe multiple description video compression," in *Proceedings of International Conference on Image Processing*, vol. 3, 1999, pp. 812 –816.

[6] T. Nguyen and A. Zakhor, "Matching pursuits based multiple description video coding for lossy environments," in *Proceedings of International Conference on Image Processing*, vol. 1, Sept. 2003, pp. I–57–60.

[7] Y.-C. Lee, Y. Altunbasak, and R. M. Mersereau, "A drift-free motion-compensated predictive encoding technique for multiple description coding," in *Proceedings of International Conference on Multimedia and Expo*, vol. 3, July 2003, pp. III–581–584.

[8] A. C. Begen, Y. Altunbasak, and O. Ergun, "Multi-path selection for multiple description encoded video streaming," in *IEEE International Conference on Communications*, vol. 3, May 2003, pp. 1583–1589.

[9] S. Mao, Y. T. Hou, X. Cheng, H. D. Sherali, S. F. Midkiff, and Y. Q. Zhang, "On routing for multiple description video over wireless ad hoc networks," *IEEE Transactions on Multimedia*, vol. 8, no. 5, pp. 1063–1074, 2006.

[10] S. Kompella, S. Mao, Y. T. Hou, and H. D. Sherali, "Path selection and rate allocation for video streaming in multihop wireless networks," in *Proceedings of Military Communication Conference*, Oct. 2006, pp. 1–7.

[11] S. Murthy, P. Hegde, V. Parameswaran, B. Li, and A. Sen, "Improved path selection algorithms for multipath video streaming in wireless ad-hoc networks," in *Proceedings of Military Communications Conference*, 2007, pp. 1–7.

[12] S. Lin, S. Mao, Y. Wang, and S. Panwar, "A reference picture selection scheme for video transmission over ad-hoc networks using multiple paths," in *Proceedings of IEEE International Conference on Multimedia and Expo*, Aug. 2001, pp. 96–99.

[13] Y. J. Liang, E. Setton, and B. Girod, "Channel-adaptive video streaming using packet path diversity and rate-distortion optimized reference picture selection," in *IEEE Workshop on Multimedia Signal Processing*, Dec. 2002, pp. 420–423.

[14] L. Zhou, B. Geller, B. Zheng, A. Wei, and J. Cui, "System scheduling for multi-description video streaming over wireless multi-hop networks," *IEEE Transactions on Broadcasting*, vol. 55, no. 4, pp. 731–741, Dec. 2009.

[15] Y. Liao and J. D. Gibson, "Frame corruption estimation from route messages for video coding over mobile ad hoc networks," accepted by the 44th Annual Asilomar Conference on Signals, Systems, and Computers, 2010.

[16] S. Mueller, R. P. Tsang, and D. Ghosal, "Multipath routing in mobile ad hoc networks: Issues and challenges," in *Performance Tools and Applications to Networked Systems*. Springer, 2004, vol. 2965, pp. 209–234.

[17] S. J. Lee and M. Gerla, "Split multipath routing with maximally disjoint paths in ad hocnetworks," in *IEEE International Conference on Communications*, vol. 10, 2001, pp. 3201–3205.

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

[19] Y. Liao and J. D. Gibson, "Refined error concealment for multiple state video coding over ad hoc networks," in *Proceedings of the 42nd Asilomar Conference on Signals, Systems and Computers*, Oct. 2008, pp. 2243 –2247.

[20] J. Broch, D. A. Maltz, D. B. Johnson, Y. C. Hu, and J. Jetcheva, "A performance comparison of multi-hop wireless ad hoc network routing protocols," in *Proceedings of the 4th annual ACM/IEEE international conference on Mobile computing and networking*, 1998, pp. 85–97.

[21] "IEEE Standard Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications," *IEEE Std 802.11-2007 (Revision of IEEE Std 802.11-1999)*, pp. C1–1184, 2007.

[22] J. Hu, S. Choudhury, and J. D. Gibson, "$PSNR_{r,f}$: Assessment of delivered AVC/H. 264 video quality over 802.11a WLANs with multipath fading," in *Proceedings of the First Multimedia Communications Workshop*, 2006.

[23] J. Hu, S. Choudhury, and J. Gibson, "Video capacity of WLANs with a multiuser perceptual quality constraint," *IEEE Transactions on Multimedia*, vol. 10, no. 8, pp. 1465–1478, Dec. 2008.