Error-Resilient Coding for Underwater Video Transmission

Yang Zhang
Dept. of Ocean Eng.
Ocean University of China
Qingdao, China 266100
Email: yangzhang003@gmail.com

Shahriar Negahdaripour
ECE Department
University of Miami
Coral Gables, FL 33146
Email: shahriar@miami.edu

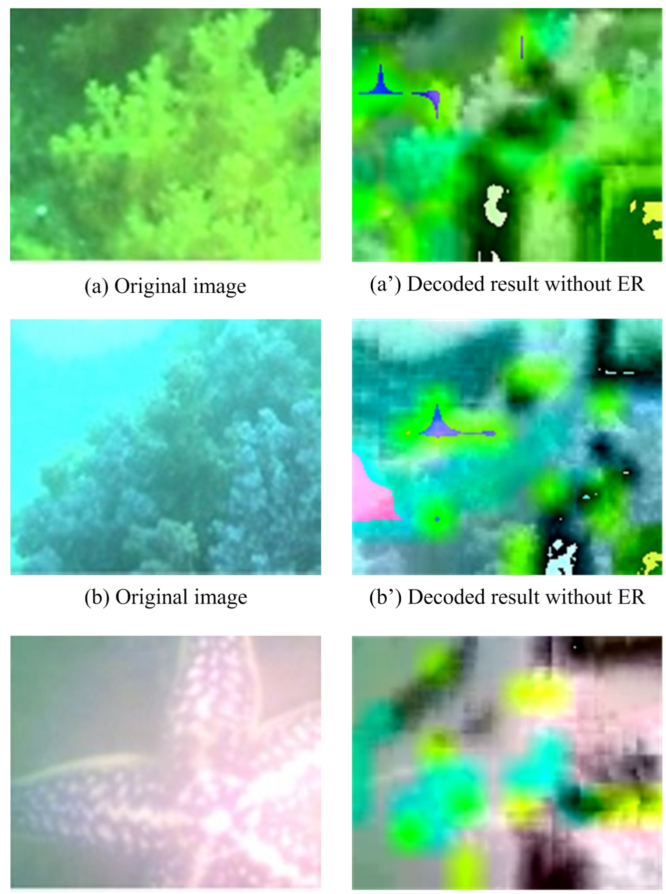
Qingzhong Li
Dept. of Automation Eng.
Ocean University of China
Qingdao, China 266100
Email: liqingzhong@ouc.edu.cn

***Abstract*—Addressing transmission errors in underwater acoustic channels is a key challenge for the real-time video communication between an autonomous underwater vehicle and a surface station. In this paper, we propose an error-resilient video compression technique based on hybrid multiple descriptions and redundant pictures to overcome impact of packet loss in underwater acoustic transmission. Video sequences are split into two descriptions with selective redundant pictures in order to achieve a balance between coding efficiency and error resiliency. Experiments with underwater video sequences are presented to assess the performance of the proposed approach under various packet loss rates, in comparison to state-of-the-art single and multiple description coding techniques.**

***Index Terms*—Underwater imagery, video compression, error resiliency, multiple description coding, redundant picture, packet loss.**

I. Introduction

Video contents captured by autonomous underwater vehicles (AUVs) are usually compressed at very low bit rates for transmission over acoustic channels with very limited bandwidth [1]–[6]. Compressed bitstreams typically become vulnerable to transmission errors, e.g. pocket loss and burst of bit errors, due to relatively high noise levels of underwater acoustic channels [7]. Moreover, high compression rates lead to an increased sensitivity to transmission errors, where error of one of a few bits can often degrade the video quality severely. Without error-resiliency (ER) capability, the reconstruction at 20% packet loss rate leads to serious degradation, as depicted in Fig. 1. For real-time applications, such as man-in-the-loop mission planning and operation, retransmission of lost packets is not readily feasible. It is clear that an error-resilient compression technique is necessary for underwater wireless video communication.



(a) Original image (a') Decoded result without ER
(b) Original image (b') Decoded result without ER
(c) Original image (c') Decoded result without ER

Fig. 1. Reconstruction results by video coding scheme without error-resiliency ability at 20% packet loss rate.

Most earlier studies on underwater video compression have been aimed at achieving better compression efficiency, based on either well-established compression standards or custom-designed methods; only a few address error resiliency and concealment. Collins et al. [7] improved the algorithm of set partitioning in hierarchical trees (SPIHT) [8] by means of error resilient organization and variable rate protection coding, taking three seconds to transmit each 256[pix]×192[pix] underwater image. Ribas et al. [9] designed an underwater video transmission system based on the MPEG-4 compression standard and the Orthogonal Frequency-Division Multiplexing (OFDM) modulator. This system offers error-resiliency through channel coding and interleaving in the modulator, but the target bit error rate (BER) is only 10^{-3}. Vall et al. [10] improved the error robustness performance using the same basic structure of the video transmission system in [9]. Here, the bit errors caused by the Doppler effect are suppressed by the proposed Doppler compensation algorithm incorporated in

978-1-5090-1537-5/16/$31.00 ©2016 IEEE

the OFDM modulation, and the reported improved target BER (1.3×10^{-2}) is mainly due to the error resiliency offered by MPEG-4 standard. The implementation of error resiliency at the video encoder offers the potential to reduce the amount of channel coding in the modulator, thus increasing the available data rate while maintaining a good visual quality.

Multiple Description Coding (MDC) has been an effective error-resilient technique, especially for real-time applications where the channel condition cannot be accurately estimated [11]–[15]. The MDC usually generates two or more equally important descriptions. If one description is missing after transmission, the other one can be decoded independently at an acceptable reconstruction quality. Redundant Picture (RP) is another widely used error-resilient solution [14], [16], [17], having properties similar to MDC. A redundant picture is produced by simply encoding the primary picture at a coarser quantization or lower bit rate. The corresponding redundant picture can be reconstructed at the decoder when a primary picture is corrupted or lost.

Both MDC and RP combat error-prone channels effectively at the expense of output bit-rate increase. Such error-resilient video encoders have to generate more transmission packets to achieve the same reconstruction quality. However, the bit-rate of underwater video communications is severely restricted due to the very limited bandwidth of acoustic channels. If the total input bit rate entering a link is more than the channel bandwidth, the upsurge of packet loss rate (PLR) will happen, resulting in serious video quality degradation. In order to meet the target bit rate of underwater acoustic communication, Avrashi et al. [18] proposed a low data-rate compression scheme by adaptively dropping frames, but the strategy is not designed for error correction coding. A desirable underwater video coding scheme should provide the robust error-resilient scalability with negligible bit rate increase.

In this paper, we propose an underwater video compression approach to strike a tradeoff between coding efficiency and error resiliency. We build on our previous work [5], [6] and design an error-resilient coding scheme based on hybrid multiple descriptions and redundant pictures at the encoder side. Video streams are split into the base and enhancement descriptions coded at different bit rates (as depicted in Fig. 2(b), and explained in detailed in Section II). According to the total target bit rate, redundant pictures are adaptively inserted to protect the basic descriptions in case of packet loss. Moreover, we selectively drop frames from the enhancement descriptions to keep the equivalent bit-rate budget. Extensive simulations demonstrate that our error-resilient scheme outperforms a state-of-the-art single video coding (SDC) method [6] in terms of average PSNR at various packet loss ratios.

The main contribution of the proposed error-resilient video compression technique is to achieve a more effective balance between bit rate increase and error resilient performance, which is of vital importance for low-delay applications in the presence of lossy underwater acoustic communication. Moreover, the hybrid MDC and RP scheme offers a more effective approach in the face of losses of both descriptions, thus significantly enhancing the robust performance against packet loss.

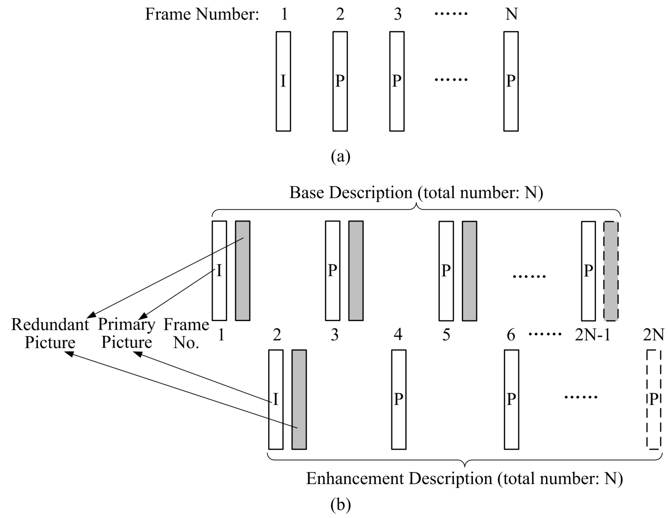


Fig. 2. (a) IPPP-frame structure; (b) Schematic diagram of proposed error-resilient video coding scheme.

In the remainder, we describe the implementation details of the proposed error-resilient video coding algorithm in Section II, and compare its performance with a state-of-the-art technique in Section III. We present a summary and conclusions from this investigation in Section IV.

II. Proposed Error-Resilient Video Coding Scheme

Traditional MDC methods usually generate two equally important descriptions to protect one when the other is lost. However, the video quality becomes seriously inferior when both descriptions are lost. Redundant pictures are efficient to address coded data loss, by simply using the same code modes as the corresponding primary pictures. We describe a hybrid error-resilient video coding scheme based on MDC and redundant pictures.

Fig. 2(a) depicts a simple N-frame IPPP-frame structure, comprising of an intra-coded ("I") frame and $N - 1$ forward predicted ("P") frames. The same frame structure is applied in devising our error-resilient video coding scheme, as illustrated in Fig. 2(b). It becomes apparent that the proposed method can be readily extended to a coding structure with B-frames.

The input video sequence is split into two subsequences composed of odd and even original frames, respectively. We treat the odd frames as the base and the even frames as the enhancement descriptions. In order to reduce the total output bit rate, we code enhancement descriptions at a lower bit rate in comparison to base descriptions:

$$R^e = \frac{2}{3} R^b \tag{1}$$

where R^e is the bit rate of each frame in the enhancement description and R^b is the bit rate of each frame in the base description. (Clearly, the base description bitstream is of more importance than the enhancement description bitstream due to more bit consumption.) The video sequence can be partly reconstructed at the price of reducing output frame rate when either base description or enhancement description is lost. If both descriptions are received at the decoder, the video quality can be improved.

We include redundant pictures in the descriptions to address the loss of both base and enhancement descriptions. The original frames in both the base and enhancement descriptions are the primary pictures. We code each redundant picture at the same bit rate R^r as the corresponding primary picture in enhancement descriptions:

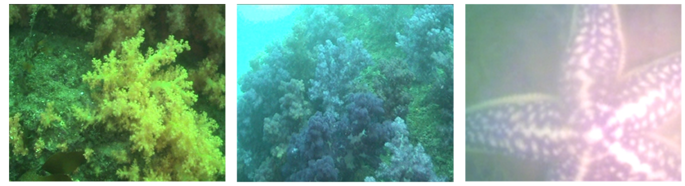
$$R^r = R^e = \frac{2}{3}R^b \tag{2}$$

Let m_n^b and m_n^e be the total number of primary plus redundant pictures in a base and enhancement description $n\,(n = 1, 2, \cdots, N)$, respectively. For example, $m_n^b = 0$ means that both primary and redundant pictures in the base description n are removed. As shown in Fig. 2(b), primary pictures in both the first base and enhancement descriptions are set to reference frames. Other primary pictures are coded as P-frames using motion compensation [2], [5], [6]. Therefore, we have one I-frame and $N-1$ P-frames in each of base and enhancement description. Since reference frames are extremely important for motion-compensated coding, we allocate a redundant picture to each I-frame in description 1 (i.e. $m_1^b = m_1^e = 2$). For P-frames, we only give the base descriptions enough redundant pictures to enhance the robustness of the output bitstream, as the base descriptions produce more bits than the enhancement descriptions do. The number of redundant picture for each P-frame in base descriptions is either zero or one, so $m_n^b = 1, 2\,(n = 2, 3, \cdots, N)$. In order to maintain the output bit rate equivalent to the bit rate of SDC, we do not allocate any redundant picture to the P-frames in enhancement descriptions, so $m_n^e = 1\,(n = 2, 3, \cdots, N)$.

We give a rate constraint to ensure that the output bitstream with error resiliency has approximately the same bit rate as the bitstream without redundant pictures within N-description ($2N$-frame) period:

$$\sum_{n=1}^{N} R_n(m_n^b, m_n^e) \approx \sum_{n=1}^{N} R_n(1,1) \tag{3}$$

where R_n is the total amount of bits used for coding both base and enhancement description n, and (m_n^b, m_n^e) denotes the redundancy mode of current description n, i.e. the number of primary plus redundant pictures in each description. To further reduce total bit rate, we exclude some primary pictures from enhancement descriptions. Typically, the error accumulation in the motion-compensated video coding becomes serious as the frame number increases. If errors occur at the beginning of motion-prediction period rather than at the end, then the video quality will deteriorate more severely. Therefore, considering a balance between reconstruction quality and coding complexity under constraint (3), we modify m_n^e as:



(a) Coral 1 (b) Coral 2 (c) Starfish

Fig. 3. Sample images from three underwater video sequences.

TABLE I
RECONSTRUCTION PERFORMANCE COMPARISON FOR THREE SEQUENCES WITHOUT PACKET LOSS

		SDC	Proposed	
			One Description	Two Descriptions
Sequence *Coral 1*	CR	533:1	807:1	323:1
	PSNR	34.41	31.37	35.82
Sequence *Coral 2*	CR	260:1	435:1	174:1
	PSNR	34.34	32.72	36.55
Sequence *Starfish*	CR	481:1	857:1	343:1
	PSNR	37.77	30.79	39.78

$$m_n^e = \begin{cases} 2; & n = 1 \\ 1; & 1 < n \leq 2N/3 \text{ and } \delta_n \leq T \\ 0; & \text{other} \end{cases} \tag{4}$$

where δ_n is the exceeded bit rate in percentage for current description n, as defined in [17]:

$$\delta_n = \frac{R_n(1,1) + \sum_{k=1}^{n-1} R_k(m_k^b, m_k^e)}{\sum_{k=1}^{n} R_k(1,1)} \tag{5}$$

We also use the same threshold $T = 0.05$ as in [17]. Such a modification yields the final output bitstream of at most $1+T$ times the bit rate of original input bitstream, which is only negligibly higher than the typical compressed video bit rate.

It should be noted that an error concealment algorithm is adopted, where both primary and redundant pictures are lost. A lost frame is concealed and reconstructed by motion interpolation between closest available descriptions. Moreover, the proposed error-resilient video compression scheme employs the same intra- and inter-frame coding algorithms as in our previous work [6].

III. EXPERIMENTAL RESULTS

To verify the effectiveness of the proposed error-resilient video coding scheme, we present the experimental results using three underwater video sequences, each with 96 frames. The sample images in Fig. 3 have a resolution of 320[pix]×240[pix], and depict the characteristics and texture contents of the sequences. We compare the proposed scheme with our previous video coding method [6] based on hybrid wavelets and directional filter banks (AHWD), which has been proved to outperform the traditional wavelet-based methods at very low bit rates. Since the AHWD-based method [6] does

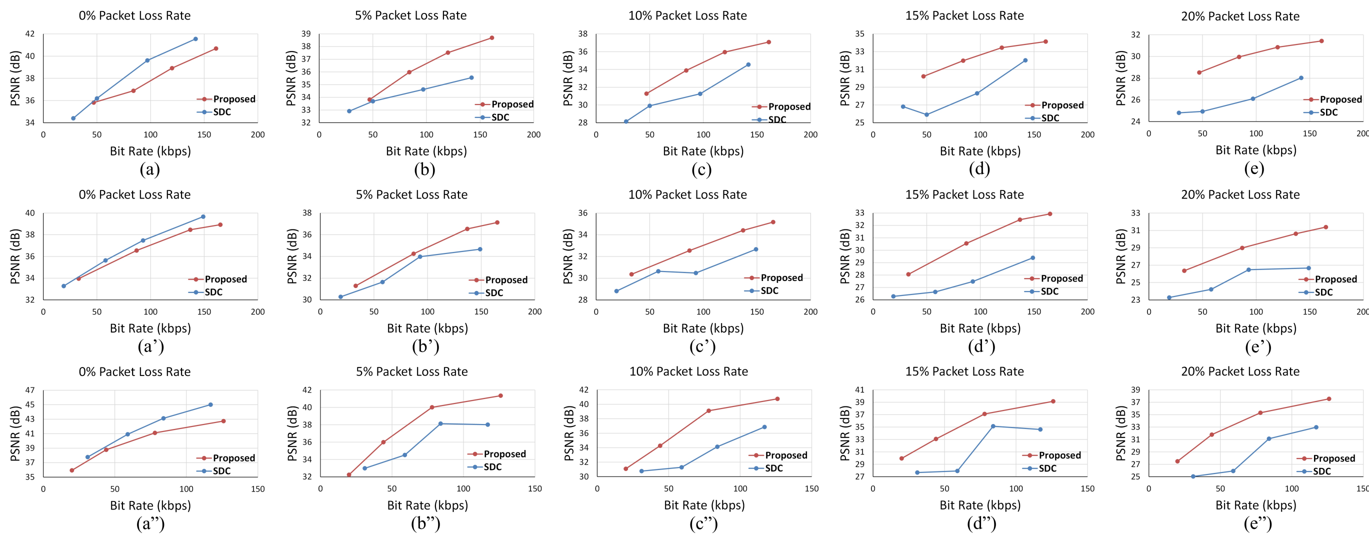


Fig. 4. RD performance comparison of the "Coral 1" (top row), "Coral 2" (middle row) and "Starfish" (bottom row) sequences at packet loss rates of (a-a") 0%, (b-b") 5%, (c-c") 10%, (d-d") 15%, and (e-e") 20%.

not have multiple descriptions for error-resilient coding, it is known as a single description coding (SDC) scheme. For both the proposed and SDC algorithms, each test sequence is coded using the IPPP-frame structure with frame-period $2N = 32$; see Fig. 2.

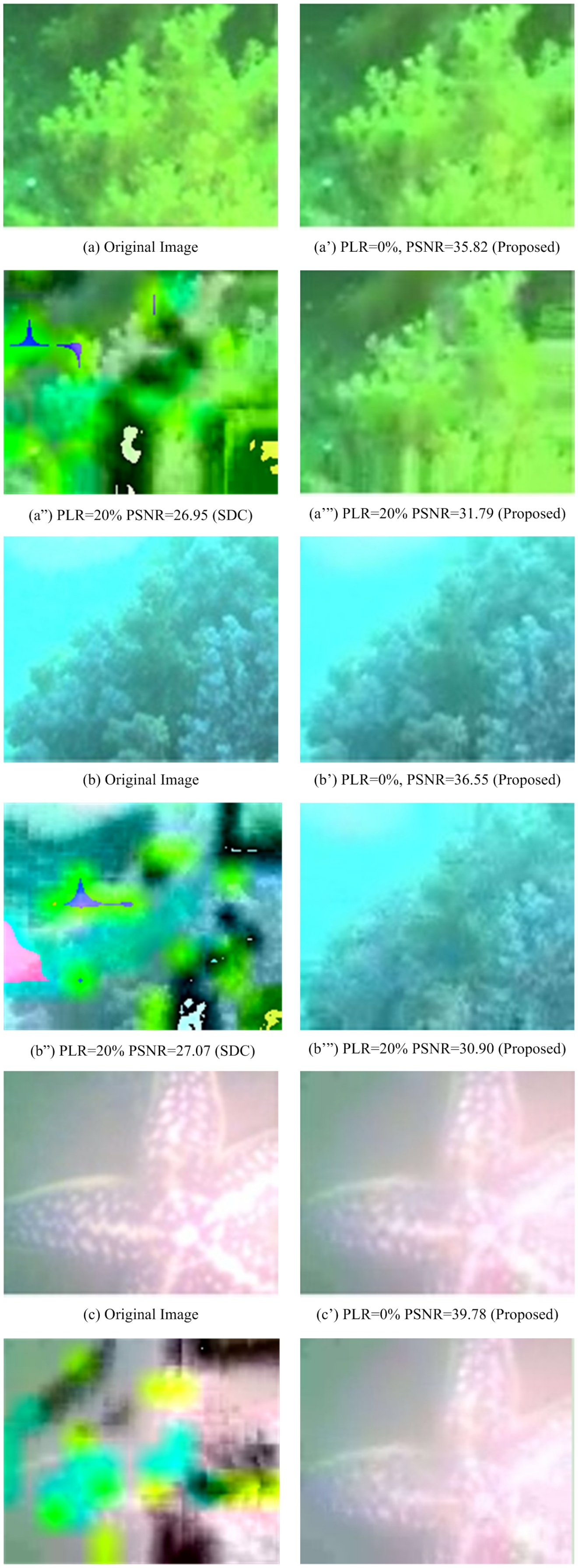
We first assess the coding performance without packet loss in terms of compression ratio and average PSNR. As shown in Table I, the proposed algorithm is tested in two cases: one description received and both descriptions received. Compared with two descriptions received, the proposed algorithm with one description received clearly provides much higher compression ratios at the expense of decreasing PSNRs. However, the method still produces acceptable reconstruction quality using only one description and motion-interpolation concealment. When the decoder receives both descriptions, the coding performance is comparable with the SDC method. Therefore, it is feasible for the proposed error-resilient coding scheme to reconstruct video contents based on multiple descriptions.

In order to verify the error-resilient performance of the proposed coding algorithm, the rate-distortion (RD) curves under different packet loss rates are given in Fig. 4. These RD curves represent how the average PSNRs change at various bit rates. A packet loss simulation tool [17], [19] is used to generate loss rates of 5%, 10%, 15% and 20% in the experiments. The final results are averaged over 200 repeated experiments for each loss rate due to the random loss pattern.

As shown in Figs. 4(a-a"), the SDC outperforms the proposed algorithm when there is no packet loss. This is because the proposed algorithm loses little coding efficiency using double bits for intra-frame coding in the two descriptions. In contrast, with packet loss, our proposed method achieves a better error-resilient coding performance than the SDC under each packet loss rate. The proposed scheme obtains significant PSNR gains, especially at the packet loss rate of 20%; see Figs. 4(e-e"). The reconstruction quality of the SDC method, where no error-resilient coding is utilized, degrades quickly as the bit rate decreases. The non-smooth RD curves generated by the SDC algorithm indicates that the increase in the data rate does not ensure a better decoding performance in terms of average PSNR; see Figs. 4(d), (c') and (d"). In contrast, the proposed algorithm produces smooth RD curves for different packet loss rates. It also maintains the similar output bit-rate range in comparison to the SDC method without error resiliency. Therefore, our video coding scheme strikes a good balance between data-rate increasing and error-resilient performance.

Fig. 5 shows sample frames from the three test sequences decoded by the SDC and proposed method. Without error-resiliency ability, the reconstruction by the SDC leads to severe degradation of color and details, as depicted in Figs. 5(a"-c"). However, incorporating the proposed redundancy for error resiliency, we obtain less distortions and higher PSNRs under 20% packet loss rate; see Figs. 5(a'"-c'"). As expected, the proposed error-resilient coding scheme outperforms the SDC in terms of both subjective quality and objective metric.

A more comprehensive assessment is presented in Fig. 6. Here, we compare the proposed coding scheme with the SDC and a state-of-the-art MDC scheme [13] where the input video is encoded into two independent threads without redundant pictures. For a fair comparison, the total bit rates for all the schemes are fixed to be 30 kbps, near the high end of the 10-40 kbps bandwidth, reported for underwater acoustic transmission. It can be seen in Fig. 6 that the proposed scheme performs generally the best at all packet loss rates, except at the 0% rate. At 20% loss probability, our scheme outperforms the MDC and SDC schemes by approximately 2 and 4 dB for the "Coral1" sequence. The PSNR gains between our scheme and the MDC gradually increases with the loss rate, as error resiliency achieved by the combination of coding descriptions and redundant pictures becomes notably beneficial, compared to joint coding with only two threads in



(a) Original Image (a') PLR=0%, PSNR=35.82 (Proposed)

(a") PLR=20% PSNR=26.95 (SDC) (a"') PLR=20% PSNR=31.79 (Proposed)

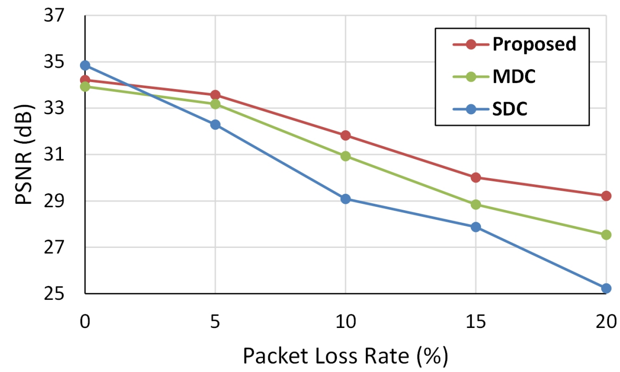
(b) Original Image (b') PLR=0%, PSNR=36.55 (Proposed)

(b") PLR=20% PSNR=27.07 (SDC) (b"') PLR=20% PSNR=30.90 (Proposed)

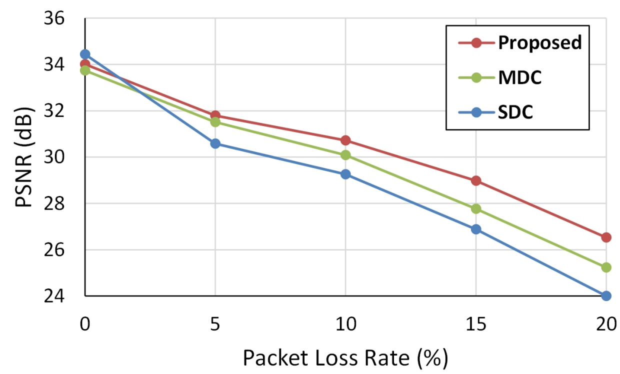
(c) Original Image (c') PLR=0% PSNR=39.78 (Proposed)

(c") PLR=20% PSNR=30.73 (SDC) (c"') PLR=20% PSNR=36.08 (Proposed)

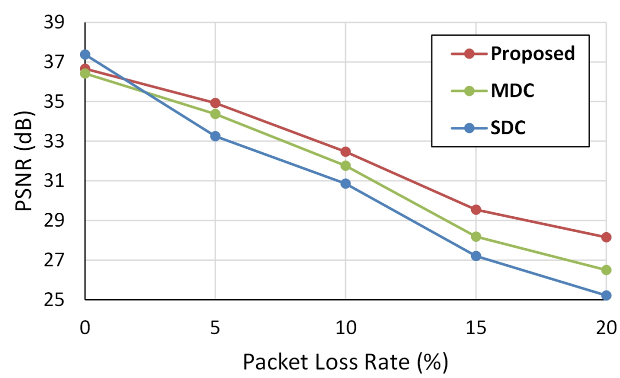
Fig. 5. Comparison among subjective qualities of decoded sample frames.



(a) Coral1 Sequence



(b) Coral2 Sequence

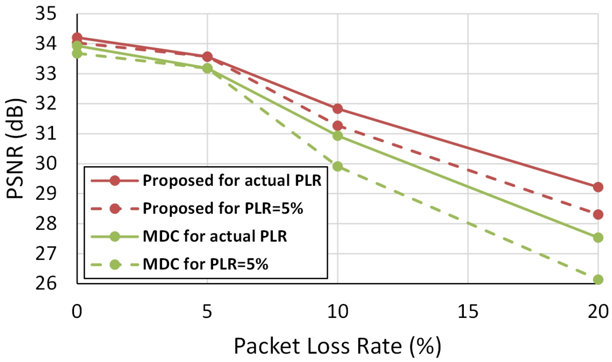


(c) Starfish Sequence

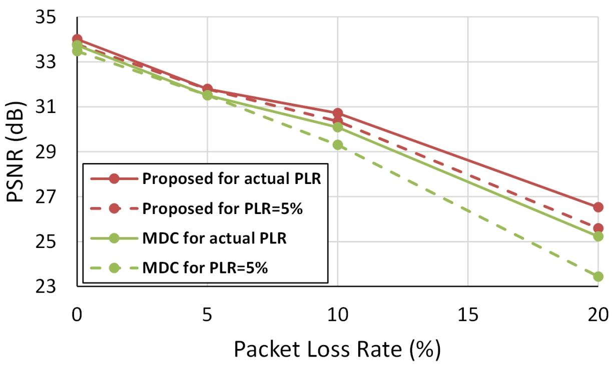
Fig. 6. Average PSNR versus packet loss rate for (a) "Coral1", (b) "Coral2" and (c) "Starfish" sequences at 30 kbps bit rate.

the MDC. We conclude that the proposed error-resilient coding method offers an efficient solution at the full range of low to high packet loss rate.

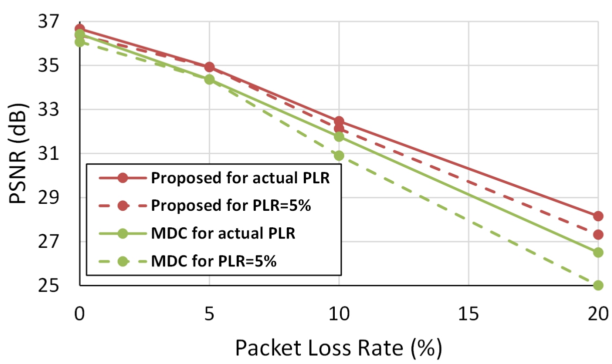
On practical scenarios, the packet loss rate is usually unpredictable. Accordingly, it is informative to evaluate performance and robustness, when assuming an inexact loss rate; this is usually the case under imprecise underwater acoustic channel conditions. Thus, we compare the proposed method with the MDC scheme using optimal and inexact loss-rate estimates. Here, we vary the actual loss rate from 0 to 20%, while setting



(a) Coral1 Sequence



(b) Coral2 Sequence



(c) Starfish Sequence

Fig. 7. Average PSNR versus actual packet loss rate using optimal and inexact loss-rate estimation for (a) "Coral1", (b) "Coral2" and (c) "Starfish" sequences at 30 kbps bit rate.

the estimated/assumed loss rate to 5%, in each experiment. Fig. 7 illustrates the end-to-end decoding performance for the three sequences. The optimal results by exact loss-rate estimation are represented by the solid lines while the PSNR curves by inexact estimation (5%) are plotted by the dash lines. As an example, Fig. 7(b) for the "Coral2" sequence depicts a PSNR degradation of 0.14 and 0.27 dB at 0% loss rate (assumed to be 5%) for our method and the MDC method, respectively. The gap expands but more gracefully for our method, as the assumed loss rate exceeds the actual 5% rate. At 20% loss rate, these PSNR gaps for both schemes are up to 0.94 dB and 1.79 dB, respectively. It is apparent that the proposed algorithm is more robust to unknown network conditions, which is desirable for video transmission through underwater acoustic channels.

IV. Conclusion

An error-resilient compression technique has been designed to combat packet loss in underwater acoustic transmission. Two descriptions are generated by splitting original sequences into odd and even frames. Redundant pictures are selectively inserted into descriptions to improve the resiliency to loss and to suppress the error propagation. In order to strike a trade off between bit-rate increase and error-resilient performance, some descriptions are removed at the end of inter-frame coding period. The experimental results with underwater video sequences have demonstrated that the proposed video compression technique based on hybrid multiple descriptions and redundant pictures can provide both desired coding efficiency and enhanced error resiliency against different packet loss rates, compared to state-of-the-art SDC and MDC algorithms.

Acknowledgment

This work was performed at the Underwater Vision and Imaging Lab (UVIL) of the ECE Department at the University of Miami. As a Visiting Scholar in UVIL, Yang Zhang's research work is being supported by the China Scholarship Council (CSC).

References

[1] D. F. Hoag, V. K. Ingle, and R. J. Gaudette, "Low-bit-rate coding of underwater video using wavelet-based compression algorithms," *IEEE Journal of Oceanic Engineering*, vol. 22, no. 2, pp. 393-400, Apr. 1997.

[2] S. Negahdaripour and A. Khamene, "Motion-based compression of underwater video imagery for operations of unmanned submersible vehicle," *Computer Vision and Image Understanding*, vol. 79, no. 1, pp. 162-183, Jul. 2000.

[3] Q.-Z. Li and W.-J. Wang, "Low-bit-rate coding of underwater color image using improved wavelet difference reduction," *Visual Communication and Image Representation*, vol. 21, no.7, pp. 762-769, May 2010.

[4] Y. Zhang, Q.-Z. Li, and S. Negahdaripour, "Seafloor image compression based on hybrid wavelets and directional filter banks," *in Proc. Oceans'15*, Genova, May 2015.

[5] Y. Zhang, Q.-Z. Li, and S. Negahdaripour, "Seafloor video compression using adaptive hybrid wavelets and directional filter banks," *in Proc. Oceans'15*, Washington D.C., Oct. 2015.

[6] Y. Zhang, S. Negahdaripour, and Q.-Z. Li, "Low bit-rate compression of underwater imagery based on adaptive hybrid wavelets and directional filter banks," *Signal Processing: Image Communication*, vol. 47, pp. 96-114, Sep. 2016.

[7] T. Collins, P. Atkins, J. Davies, and S. Dunn, "Robust video coding for underwater transmission," *in Proc. Vis. Commun. Image Process.*, San Jose, CA, Jan. 2001.

[8] A. Said and W. A. Pearlman, "A new fast and efficient image codec based on set partitioning in hierarchical trees," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, no. 3, pp. 243-250, Jun. 1996.

[9] J. Ribas, D. Sura, and M. Stojanovic, "Underwater wireless video transmission for supervisory control and inspection using acoustic OFDM," *in Proc. MTS/IEEE Oceans'10*, Seattle, WA, Sept. 2010.

[10] L. D. Vall, D. Sura, and M. Stojanovic, "Towards underwater video transmission," *in Proc. the Sixth ACM International Workshop on Underwater Networks*, Seattle, WA, Dec. 2011.

[11] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 74-93, Sep. 2001.

[12] Y. Wang, A. R. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proc. IEEE*, vol. 93, no. 1, pp. 57-70, Jan. 2005.

[13] J. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *in Proc. Vis. Commun. Image Process.*, San Jose, CA, Jan. 2001.

[14] I. Radulovic, P. Frossard, Y.-K. Wang, M. Hannuksela, and A. Hallapuro, "Multiple description video coding with H.264/AVC redundant pictures," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 1, pp. 144-148, Jan. 2010.

[15] N. Gadgil, H. Li, and E. J. Delp, "Spatial subsampling-based multiple description video coding with adaptive temporal-spatial error concealment," *in Proc. Picture Coding Symposium (PCS)*, Cairns, Australia, May 2015.

[16] Y.-K. Wang, M. M. Hannuksela, and M. Gabbouj, "Error resilient video coding using unequally protected key pictures," *in Proc. Int. Workshop Very Low Bitrate Video*, Madrid, Spain, Sep. 2003.

[17] D. Zhang, H.-Q. Li, and C.-W. Chen, "Robust transmission of scalable video coding bitstream over heterogeneous networks," *IEEE Trans. Circuits Syst. Video Technol.* vol. 25, no. 2, pp. 300-313, Feb. 2015.

[18] G. Avrashi, S. Museri, Y. Bucris, A. Sinai, and A. Amar, "Underwater video streaming using adaptive frame decimation," *in Proc. MTS/IEEE Oceans'14*, Taipei, Apr. 2014.

[19] Y. Guo, H. Li, and Y.-K. Wang, *SVC/AVC Loss Simulator Donation*, document JVTQ069, Joint Video Team, Nice, France, 2005.