Adaptive Underwater Video Transmission via Software-Defined MIMO Acoustic Modems

Mehdi Rahmati, Adam Gurney, and Dario Pompili
Department of Electrical and Computer Engineering, Rutgers University–New Brunswick, NJ, USA
Emails: {mehdi_rahmati, pompili}@cac.rutgers.edu; adam.gurney@rutgers.edu

***Abstract*—Video signal transmission enables a wide range of applications in the underwater environment; such as coastal and tactical multimedia surveillance, undersea/offshore exploration, oil pipe/bridge inspection, video monitoring of geological/biological processes from the seafloor to the air-sea interface—that all require real-time multimedia acquisition and classification. Yet, it is a challenge to achieve an efficient and reliable video transmission, due to the spectrum limitations underwater and also the error prone nature of the acoustic channel. In this paper, we propose a pairwise scheme to manage the video distortion-rate tradeoff for underwater video transmission. The proposed Multi-input Multi-output (MIMO)-based Software-Defined Acoustic Radio (SDAR) system adapts itself to meet the needs of both video compression and underwater channel in a timely manner from one hand, and keeps the overall video distortion—caused by the coder/decoder and channel—under an acceptable threshold from the other hand. The scalability of Universal Software Radio Peripheral (USRP) with high processing capabilities is exploited in the proposed structure along with the temporal, spatial and quality scalability of Scalable Video Coding (SVC) H.264/MPEG-4 AVC compression standard. Experimental results at Sonny Werblin Recreation Center, Rutgers University, as well as simulations are presented, while more experiments are in-progress to evaluate the performance of our testbed in more challenging environments such as in the Raritan River, New Jersey.**

***Index Terms*—Scalable video coding; software-defined modem; underwater acoustic communications; video transmission.**

I. Introduction

Overview: Underwater Acoustic Networks (UAN), including static sensors/actuators and mobile vehicles, underpin the underwater world and are instrumental to support next-generation ocean-interconnected networks [1], [2]. To enable these applications, and to pave the way towards the Internet of Underwater Things (IoUTs) paradigm [3], multimedia transmission is a necessity, especially from places where humans cannot easily/safely go to store/process/compress it, and to provide support and service to these applications with different Quality of Service (QoS) requirements ranging from delay sensitive to delay tolerant and from loss sensitive to loss tolerant [4].

Motivation: In most cases, Remotely Operated Vehicles (ROVs) are often tethered to the supporting ship by a fiber cable or have to rise periodically to the surface to communicate with a remote station via Radio-Frequency (RF) waves. Tethering is a serious limitation in the development of underwater systems for future applications involving multiple underwater vehicles as it constrains the maneuverability and range of the vehicles engaged in the mission, which run the risk to get tangled and compromise the mission itself. Resurfacing periodically, on the other hand, does not guarantee interactivity and leads to energy/time inefficiencies. Having a reliable and fast wireless transmission is a necessity in such an environment, in which RF waves are absorbed for distances above a few tens of meters, optical waves require narrow laser beams and suffer from scattering and ocean wave motion, and acoustic waves—while being able to propagate up to several tens of kilometers—lead to a communication channel that is very dynamic, prone to fading, spectrum limited with passband bandwidths of only a few tens of kHz due to high transmission loss at frequencies above 50 kHz, and affected by non-Gaussian and non-white noise [1], [5].

State of the Art: While traditional commercial acoustic modems with their fixed-hardware designs hardly meet the required data-rate and flexibility to support these futuristic applications, recently other solutions based on adaptive, open source, and reconfigurable architectures employing Software-Defined Acoustic Radios (SDAR) have been proposed by many research groups. The Centre for Maritime Research and Experimentation (CMRE) proposed a high-level architectural structure [6] towards designing a Software-Defined Open-Architecture Modems (SDOAM) that is compatible with the JANUS standard [7]. Authors in [8] proposed a new networking platform for short-range acoustic SDARs called SEANet. A thorough survey on the past and current SDARs was presented in [9], where the joint project between The Netherlands Organization for Applied Scientific Research (TNO) and the Norwegian Defense Research Establishment (FFI) focusing on building a flexible and programmable modem, called NILUS softmodem, was discussed. The other platform, presented in [10], tries to adapt some of the terrestrial radio and network development with the underwater acoustic environment.

Related Work: Conventional video coding might not meet the underwater video streaming and the futuristic applications requirements, since they require higher and more flexible data-rates. This is hard to achieve in distances above hundred meters through the acoustic channel, as acoustic waves suffer from attenuation, limited bandwidth, Doppler spreading, high propagation delay, and time-varying propagation characteristics [1], [4]. To achieve higher data-rates in the bandwidth-limited underwater acoustic channel, several techniques should be combined together. In [11], a method based on Hybrid Automatic Repeat Request (HARQ) is proposed that exploits

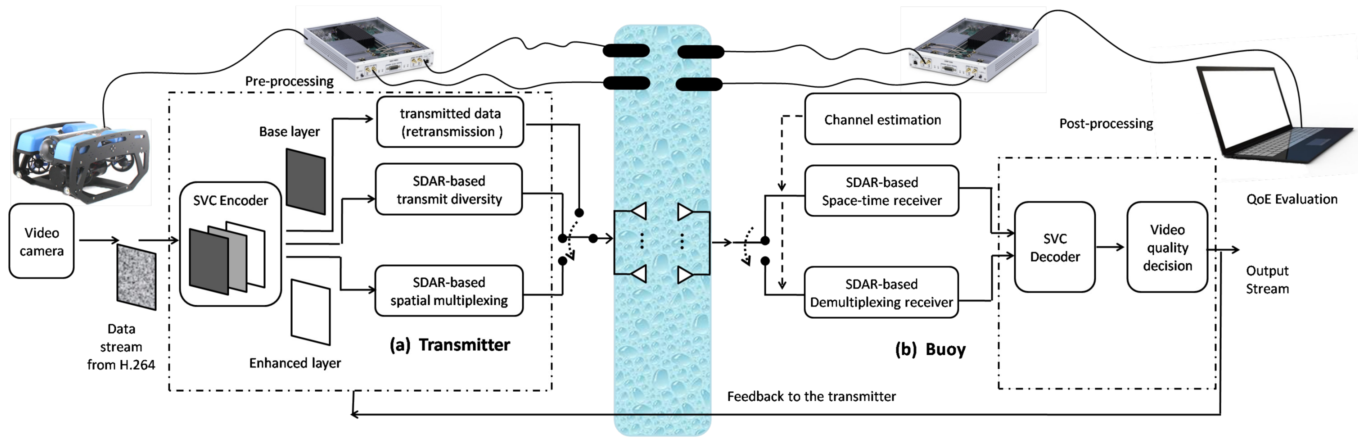


Fig. 1: Proposed system structure, which uses Scalable Video Coding (SVC) and a flexible scheme on MIMO Software-Defined Acoustic Modem. Transmission techniques that utilize diversity, spatial multiplexing, and retransmission are the modalities.

the diversity gain offered by independent links in an underwater acoustic Multiple Input Multiple Output (MIMO) system. Authors in [12] discuss the relationship between underwater acoustics and optics for long-range and short-range distances, respectively to find out the correlation between the properties and the reliability of the acoustic and optical links. In [13], we propose a signaling method for video transmission that makes use of multiple domains to leverage the benefits of Acoustic Vector Sensors (AVS). Scalable Video Coding (SVC), as an extension of H.264/MPEG-4 AVC, can be the solution for video delivery in harsh environments, such as in underwater, by offering more flexibility via different modalities—temporal (frame rate), spatial (frame size), and quality (fidelity or SNR)—to match the lossy video compression and erroneous transmission environments. It also can support the scalability in the complexity and in the Region Of Interest (ROI) [14].

Our Contributions: We introduce a reconfigurable SDAR testbed based on the latest Universal Software Radio Peripheral (USRP) family product designed by National Instrument (NI). We propose a pairwise solution for underwater video transmission that leverages the characteristics of the MIMO structure to balance the data-rate and reliability, i.e., the *multiplexing-diversity tradeoff*, in order to transmit a scalable video with an acceptable level of distortion. Scalable Video Coding (SVC) provides the scalability in the video bit stream processing to adapt to the preference of end users as well as to the varying characteristics of the network. The common types of scalability—temporal (frame rate), spatial (frame size), and quality (fidelity or SNR)—should be adaptively assigned according to the software-defined modem parameters for signal transmission in the proposed MIMO structure. Therefore, the parts of video with more importance are transmitted with a higher reliability than the enhancement parts, so that the possible error in the transmission does not propagate to the upper layers and the total distortion is kept under a desirable threshold. Several experiments and simulations have been conducted on a camera-equipped, SDR-based underwater vehicle, and the results are presented in this paper.

Paper Organization: The remainder of the paper is organized as follows. In Sect. II, we propose our solution towards underwater video transmission. In Sect. III, we present the performance evaluations and results based on the conducted field experiments and computer simulations; finally, in Sect. IV, we conclude the paper and discuss the future path of this research.

II. Proposed Solution for Video Transmission

In this section, first we introduce the structure to support the proposed system. Then, we describe the proposed adaptive pairwise application-physical-layer optimized algorithm to mitigate the overall distortion effects, caused by the coder/decoder and by the error prone underwater acoustic channel, while we are transmitting a scalable video with the highest possible data rate.

System Model: We present here how our proposed scheme leverages the MIMO structure and the scalability characteristic of the compressed video. The required blocks of the system are depicted in Fig. 1. Video is initially recorded by the underwater robot, then it is encoded and prepared for transmission in the pre-processing block. The SVC encoder produces the required base and enhancement layers. The Data should be protected against noisy channels by an appropriate channel coding scheme [11]. Finally, one of the schemes, i.e. transmit diversity or spatial multiplexing, is chosen for transmission, based on the output of the optimization process. At the receiver, the post-processing and decoding operations will be performed in space-time/demultiplexing. Humans will participate dynamically in a closed-loop fashion to finely tune the system based on the required Quality of Experience (QoE) of the received video stream. The appropriate decision is optimized and the transmitter is notified accordingly.

We assume that the video bitstream is divided into chunks, and then in SVC, each chunk includes a *base layer* plus L *enhancement layers* with different modalities to enhance the quality of the video. This modality is being decided at the pre- and post-processing, regarding the Rate-Distortion (RD) requirements of the system. For a compressed video, we can write [15],

$$D_e(R_e) = \frac{\theta}{R_e - R_0} + D_0, \tag{1}$$

where D_e represents the distortion of the encoded video and R_e is the output rate of the encoder; the other remaining variables θ, R_0, and D_0 depend on the encoded video and on the model, and are estimated empirically. To quantify and measure the video streaming distortion over lossy channels, the Peak Signal-to-Noise Ratio (PSNR) is used as a generic metric for measuring the video distortion D_e based on the overall Mean Square Error (MSE). To reduce the amount of distortion, SVC provides hierarchical prediction structures for temporal scalability, inter-layer prediction of motion for spatial and quality scalability, and key pictures definition for drift control in packet-based quality scalable coding with hierarchical prediction structures [14]. Note that the total amount of distortion at the decoder is caused by the errors at the encoder compression D_e and the errors caused by the channel.

Diversity-Multiplexing Modalities: We utilize two extreme coding modalities of a MIMO system, i.e., Space-Time Coding (STC) and Spatial Multiplexing (SM) in the proposed software-defined modem. This provides spatial diversity gain or multiplexing gain, respectively, in order to adapt to the acoustic channel conditions. Because of this fundamental trade-off between the error-rate and data-rate of a MIMO channel, it is not possible to reach the maximum diversity and multiplexing gains at the same time. Authors in [16] calculated the achievable diversity-multiplexing tradeoff in a MIMO system with m transmit antennae and n receive antennae as $d^*(r) = (m-r)(n-r)$, where $d^*(r)$ shows the diversity gain and integer $r \in \mathbb{Z}$ represents the multiplexing gain, which can be defined as, $r = 0, 1, ..., \min(m, n)$. As two special cases, $d^*_{max} = mn$ and $r^*_{max} = \min\{m, n\}$. Besides, approaching one of these two extremes means decreasing the other one. In other words, the tradeoff curve confirms that while the rate increases by r bps/Hz over an increase of 3 dB in SNR, the error rate is reduced by an order of $2^{-d^*(r)}$. However, this tradeoff is achieved under *ideal conditions* in which SNR approaches infinity for i.i.d. Rayleigh-fading channels. This asymptotic definition usually breaks if the SNR is limited as it is the case in real scenarios. As an example, in an AWGN channel, effective diversity and multiplexing gains are calculated for low SNRs, γ, with array gain, g, as follows [17],

$$r = \frac{R}{\log_2(1 + g\gamma)}, \quad d(r, \gamma) = -\frac{\partial \ln P_{out}(r, \gamma)}{\partial \ln \gamma}. \tag{2}$$

The spectral efficiency and the outage probability are represented by r and $P_{out}(r, \gamma)$, respectively, and are related as,

$$P_{out}(r, \gamma) = \Pr\left[\log_2 \det(\mathbf{I}_n + \frac{\gamma}{m}\mathbf{H}^*\mathbf{H}) < R\right], \tag{3}$$

where $\mathbf{I}_n$ represents the $n \times n$ identity matrix and superscript * stands for the conjugate transpose of the $n \times m$ channel matrix $\mathbf{H}$. For instance, when STC is exploited to achieve maximum diversity, the outage probability can be approximated given the fading distribution of the channel. It was shown in [18] that for uncorrelated MIMO channels, $\mathbf{H}$ can be represented by the variances of the power gains of channel as $\mathrm{var}\, \|\mathbf{H}\|_\mathrm{F}^2 = \sum_{\mathrm{i}=1}^{\mathrm{n}} \sum_{\mathrm{j}=1}^{\mathrm{m}} \mathrm{var}\, |\mathrm{h_{ij}}|^2$. These values can be obtained by estimating the mean powers of the channel matrix in the experiment.

Pre-processing and Optimization: Let $R_e(c, l)$, with SVC layers $\{l = 1, ..., L+1\}$, denote the rate for layer l of video chunk c. An appropriate data-rate $R_i \geq R_e(c, l)$ for reliable communication should be assigned to layer l in order to maximize the total transmission rate, i.e., transmitting as many video layers as possible without getting an outage or erroneous reception, given the bandwidth limitations and the quality of the underwater channel as well as the distortion.

The following optimization problem justifies the aforementioned discussion. The first problem is a knapsack program, which defines the enhancement layers of rate $R_e(c, l)$ that could be transmitted over the underwater channel with maximum achievable communication data-rate R_{max}.

$$\max_{\alpha_l} \ \mathcal{F}_\mathcal{R} = R_e(c, 1) + \sum_{l=2}^{L+1} \alpha_l \alpha_{l-1} R_e(c, l) \tag{4a}$$

$$\text{s.t.} \quad \sum_{i=1}^{L+1} \alpha_i \alpha_{i-1} R_i \leq R_{max}, \tag{4b}$$

$$\alpha_1 = 1, \ \alpha_i \in \{0, 1\}, \forall i \in \{2, ..., L+1\}. \tag{4c}$$

Coefficients $\{\alpha_i\}$ determine the set of enhancement layers that can be passed through the channel given the mentioned constraints. Selecting each layer depends on the presence of its preceding layer. This optimization guarantees that the base layer of each chunk is transmitted first, then the enhancement layers are transmitted based on the capacity of the underwater communication system.

Video Quality Decision: At the receiver side, to ensure that the threshold of Quality of Experience (QoE) is fulfilled by the algorithm, an optimization problem is cast to find the required transmission rate for every layer that leads to minimum possible value of distortion. The total distortion is modeled as $D = D_e + D_c$, where D_e is the distortion imposed by the codec as presented in (1) and D_c is the acoustic channel distortion, which is related to channel effective loss rate (λ). The following optimization problem makes decision for the next round of transmission.

$$\min_{R_i} \ \mathcal{F}_\mathcal{D} = D_e(\sum_{l=1}^{L+1} \alpha_l R_e(c, l)) + \delta \frac{\sum_{i=1}^{L+1} \lambda_i \alpha_i R_i}{\sum_{i=1}^{L+1} \alpha_i R_i} \tag{5a}$$

$$\text{s.t.} \quad \sum_{i=1}^{L+1} \alpha_i R_i \geq R_{min}, \tag{5b}$$

$$R_i \geq R_e(c, l), \tag{5c}$$

$$D \leq D_T. \tag{5d}$$

The second term in (5a) represents the distortion imposed by the channel, in which δ is a constant that is defined

Algorithm 1 Adaptive video transmission procedure.

```
Layers = scalableVideoCoder(U)
Decide on video layers and data-rate based on the optimization
Transmit(baseLayer); s ← 1          % s is the number of trials
while t < Chunk Time do
    Video quality decision process
    Get feedback from receiver
    if feedback is resend request then
        transmit(feedback.requestedLayer)
        channelState.update(BAD)
        s ← s + 1
    else
        transmit(Layers.nextLayer)
        s ← s + 1
        channelState.update(GOOD)
        if s = ∑ iα_i % α_i stands for the layer coefficient then
            Goto 1          % Done Transmitting this chunk
        end if
    end if
    if channelState.rollingAverage > threshold then
        transmitter.switchTo('Multiplexing')
    else
        transmitter.switchTo('Space-Time Coding')
    end if
end while
```

by the video codec [19], and λ_i is the effective loss rate of the i^{th} communication channel. In the constraints, R_{min} stands for the minimum required rate to avoid P_{out}, and D_T represents the acceptable distortion threshold perceived by end user. This problem can be optimized through piecewise linear approximation method which leads to a convex approximation function for (5a).

System Algorithm: Algorithm 1 describes the general procedure in order to compress and transmit the video. The transmitter decides on the communications method and compression scalability based on the feedback received from measuring the QoE by the user. Our SDAR adaptively reconfigures itself by solving the optimization problem to be able to switch between the two extreme transmission modes and decide on the required number of video layers to achieve the required video scalability. The base-layer stream, which contains the highest priority information of the video, requires the most reliability, while the enhancement layers requires a higher data-rate R_{max}. This goal can be achieved by switching between the diversity-based and multiplexing-based modes. This, on the other hand, might result in more communications errors the receiver side, if the channel condition is not good.

III. Experiments and Simulations

Testbed: We evaluated the proposed approach by conducting preliminary field experiments. We modified an existing tethered Remotely Operated Vehicle (ROV), called BlueRov2 [20], to make it wireless and autonomous. This underwater robot is equipped with a 1080p camera. A video feed is passed to the modem and the acoustic transducer to perform the transmission to the buoy in order to provide the required feedback to the vehicle. A high-performance and scalable platform with a programmable Kintex-7 FPGA, called X-300 designed by Ettus Research Group [21], is exploited in this research. It contains a main-board to provide basic functionalities of the modem, while the daughter-boards take care of up/down conversions and of the other required band-pass signal processing procedures. Teledyne Marine TC4013 transducers [22] with a frequency range of 170 kHz are used in the proposed testbed. The specifications of the system are summarized in Table I.

TABLE I: Hardware Specifications.

Part	**Parameter**	**Value**
Transducer	Frequency range	1–180 kHz
	Directionality	Omnidirectional
	Receiving sensitivity	-211 dB $\pm$ 3 dB re 1 V/μPa
	Transmit sensitivity	130 dB $\pm$ 3 dB re 1 V/μPa
Preamplifier	Frequency range	0.5–500 kHz
	Output gain	0–50 dB
	High pass filters	1 Hz–250 kHz
	Low pass filters	1 kHz–1 MHz
Modem	Mainboard	Kintex-7 FPGA
	Frequency range	0–30 MHz
	Clocking	10 MHz/1 PPS
	ADC sample rate	2 yuchannels, 200 MS/s
	ADC resolution	14 bits
	DAC sample rate	2 channels, 800 MS/s
	ADC resolution	16 bits
MIMO	Uplink Structure	2x2 MIMO
	Feedback Structure	1x1 SISO
	Duplexing	FDD
Camera	Standard	H.264 1080p
	Lens	1×1.7 mm
	Tilt range	$\pm 90°$
	Horizontal FoV	110°

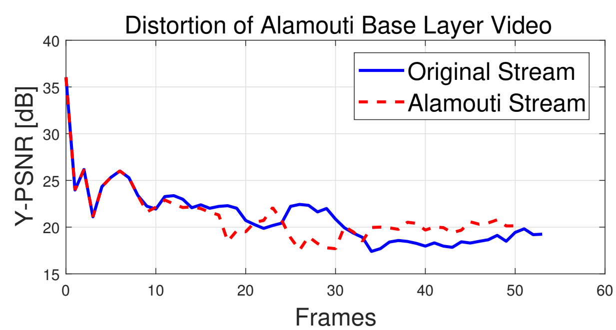
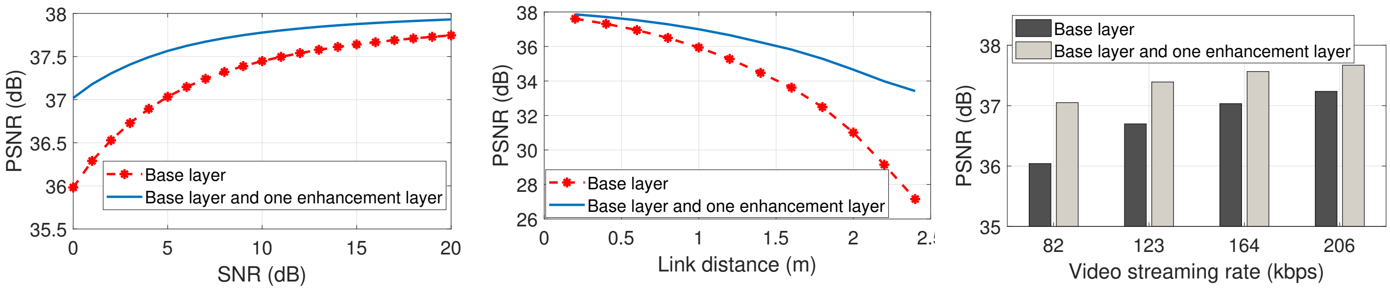


Fig. 2: PSNR metric in base layer video stream vs. distorted stream encoded with Alamouti and BCH(255,163).

We use the Joint Scalable Video Mode (JSVM) software as the reference package for implementing SVC. Using the FixedQPEncoder program, test videos were down-sampled and then encoded into 3 layers of different qualities. Each layer has a target fixed bit rate, and the Quantization Parameter (QP) is varied in order to optimize the Peak Signal-to-Noise Ratio (PSNR) metric while staying under the target bitrate. Additionally, 2 different temporal layers can be extracted for each quality layer, giving 6 total layers. Open SVC Decoder is used for decoding due to its implementation of error concealment and its integration with Mplayer for video streaming.

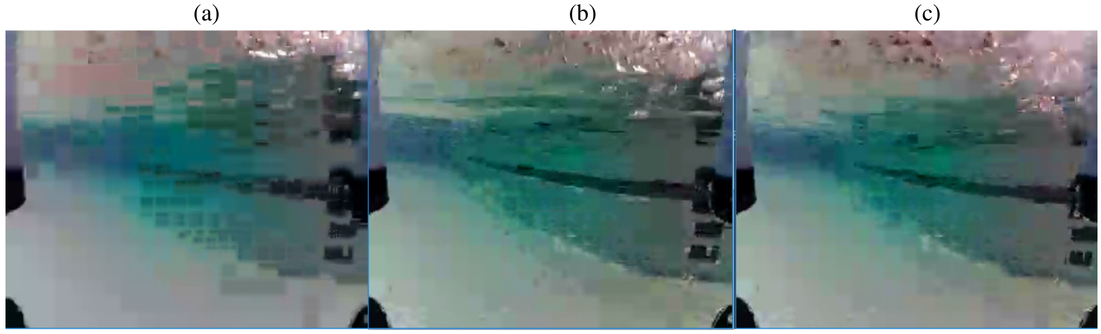
Simulation: Video streams encoded with Alamouti and Spatial Multiplexing were each put through a random



(a) (b) (c)

Fig. 3: Simulated distortion (PSNR) for the received video (a) versus SNR; (b) versus link distance; (c) for different video streaming rate when $SNR = 5\ dB$.

(a) (b) (c)



(d) (e) (f)

Fig. 4: (a) Frame from original video; (b) Frame of video reconstructed from only the base layer; (c) Frame of video reconstructed from the base layer and enhancement layers (all layers); (d),(e), and (f) show the reconstructed frames of the base and enhancement layers after experiencing a harsh error in the received signal.

frequency-flat 2x2 channel and introduced to AWGN noise. Fig. 2 shows the change in PSNR of the Alamouti video due to distortion. Blind channel estimation techniques [23] were used to extract the channel gains for Zero Forcing equalization.

Fig. 3 shows the result of solving optimization problems (4a) and (5a) in the pre- and post-processing parts of the system. In these figures, PSNR is plotted as the received distortion measure given the decision of the transmitter on transmitting base layer or base and one enhancement layers. In Fig. 3(a), the effect of SNR–as a metric of communication channel quality– on distortion is investigated. Fig. 3(b) shows the decreasing trend of PSNR when the link distance increases. Fig. 3(c) shows the distortion metric for different video streaming rates.

Pool Experiment: For our preliminary experiment, hydrophones were placed in a large pool at a distance varying from 15 to 100 cm apart and in the range of $20 - 50$ cm depth. For multiple distances and depths in this range, one-way transmission was then done with maximum data-rate of 100 kBd and with H264/AVC codec JSVM for some video test signals (Fig. 4).

Fig. 5 shows the PSNR metric of the video reconstructed from a certain error rate, using all layers up to 0, 1, and 2.

IV. Conclusion and Future Work

We proposed a novel scheme to layerize and transmit the video underwater using a Multi-input Multi-output (MIMO)-based Software-Defined Acoustic Radio (SDAR). The balance between the data-rate and reliability, i.e., the *multiplexing-diversity tradeoff* is required. In order to transmit a scalable video with a pre-defined level of distortion, Scalable Video Coding (SVC) is exploited which provides the scalability in the video bit stream processing to adapt to the preference of end users as well as to the varying characteristics of

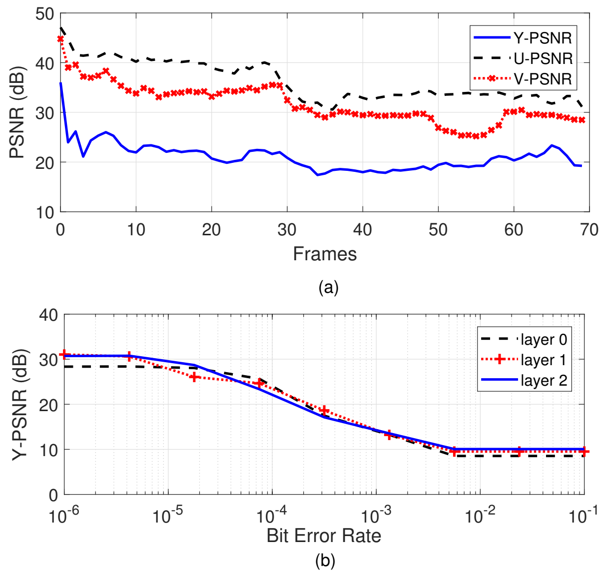


Fig. 5: Distortion metric for video; (a) over time, Y-PSNR is associated with luminance and U- and V- PSNR with chrominance; (b) Y-PSNR of video distorted by random bit errors, reconstructed from layers up to 0, 1, and 2.

the network. Experimental results at Sonny Werblin Recreation Center, Rutgers University were presented, while more experiments are in progress to evaluate the performance of our testbed in more challenging environments such as in the Raritan River, New Jersey. In the future, we hope to extend to novel encoding schemes such as High Efficiency Video Coding (HEVC) to maximize the quality we can achieve with limited bandwidth.

ACKNOWLEDGMENT

This work was supported by the US National Science Foundation (NSF) under the CAREER Award No. OCI-1054234.

The authors thank Jeffrey Zeszotarski, aquatics coordinator, at Rutgers University for his help and support during field experiments.

REFERENCES

[1] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater acoustic sensor networks: Research challenges," *Ad Hoc Networks*, vol. 3, no. 3, pp. 257–279, Mar. 2005.

[2] M. Rahmati and D. Pompili, "Unisec: Inspection, separation, and classification of underwater acoustic noise point sources," *IEEE Journal of Oceanic Engineering*, vol. 43, pp. 777–791, 2018.

[3] M. Rahmati, D. Pompili, and R. Petroccia, "Collaborative hybrid ARQ for CDMA-based reliable underwater acoustic communications," in *Proceedings of Underwater Communications and Networking Conference (UComms)*. IEEE, 2018, pp. 1–5.

[4] D. Pompili and I. F. Akyildiz, "A multimedia cross-layer protocol for underwater acoustic sensor networks," *IEEE Transactions on Wireless Communications*, vol. 9, no. 9, pp. 2924–2933, 2010.

[5] M. Rahmati and D. Pompili, "Interference cancellation in multiuser acoustic underwater networks using probabilistic SDMA," in *Proceedings of 12th International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE, 2015, pp. 307–315.

[6] J. Potter, J. Alves, T. Furfaro, A. Vermeij, N. Jourden, D. Merani, G. Zappa, and A. Berni, "Software defined open architecture modem development at cmre," in *Proceedings of Underwater Communications and Networking Conference (UComms)*. IEEE, 2014, pp. 1–4.

[7] J. Potter, J. Alves, D. Green, G. Zappa, I. Nissen, and K. McCoy, "The JANUS underwater communications standard," in *Proceedings of Underwater Communications and Networking Conference (UComms)*. IEEE, 2014, pp. 1–4.

[8] E. Demirors, J. Shi, R. Guida, and T. Melodia, "SEANet G2: Toward a high-data-rate software-defined underwater acoustic networking platform," in *Proceedings of the International Conference on Underwater Networks & Systems*. ACM, 2016, p. 12.

[9] H. S. Dol, P. Casari, T. Van Der Zwan, and R. Otnes, "Software-defined underwater acoustic modems: Historical review and the nilus approach," *IEEE Journal of Oceanic Engineering*, vol. 42, no. 3, pp. 722–737, 2017.

[10] D. Torres, J. Friedman, T. Schmid, M. B. Srivastava, Y. Noh, and M. Gerla, "Software-defined underwater acoustic networking platform and its applications," *Ad Hoc Networks*, vol. 34, pp. 252–264, 2015.

[11] M. Rahmati and D. Pompili, "uwMIMO-HARQ: Hybrid ARQ for reliable underwater acoustic mimo communications," in *Proceedings of the 10th International Conference on Underwater Networks and Systems (WUWNET)*. ACM, 2015, pp. 1–8.

[12] R. Diamant, F. Campagnaro, M. D. F. De Grazia, P. Casari, A. Testolin, V. S. Calzado, and M. Zorzi, "On the relationship between the underwater acoustic and optical channels," *IEEE Transactions on Wireless Communications*, vol. 16, pp. 8037–8051, 2017.

[13] M. Rahmati and D. Pompili, "SSFB: Signal-space-frequency beamforming for underwater acoustic video transmission," in *Proceedings of International Conference on Mobile Ad Hoc and Sensor Systems (MASS)*. IEEE, 2017, pp. 180–188.

[14] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H. 264/AVC standard," *IEEE Transactions on circuits and systems for video technology*, vol. 17, no. 9, pp. 1103–1120, 2007.

[15] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 1012–1032, 2000.

[16] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Transactions on information theory*, vol. 49, no. 5, pp. 1073–1096, 2003.

[17] R. Narasimhan, "Finite-SNR diversity–multiplexing tradeoff for correlated rayleigh and rician mimo channels," *IEEE Transactions on Information Theory*, vol. 52, no. 9, pp. 3965–3979, 2006.

[18] J. Pérez, J. Ibánez, L. Vielva, and I. Santamaria, "Closed-form approximation for the outage capacity of orthogonal stbc," *IEEE Communications Letters*, vol. 9, no. 11, pp. 961–963, 2005.

[19] J. Wu, B. Cheng, C. Yuen, Y. Shang, and J. Chen, "Distortion-aware concurrent multipath transfer for mobile video streaming in heterogeneous wireless networks," *IEEE Transactions on Mobile Computing*, vol. 14, no. 4, pp. 688–701, 2015.

[20] Blue Robotics. [Online]. Available: http://www.bluerobotics.com.

[21] USRP X Series. [Online]. Available: https://www.ettus.com

[22] RESON TC4013 Hydrophone Product Information . [Online]. Available: http://www.teledynemarine.com/reson-tc4013

[23] V. Choqueuse, A. Mansour, G. Burel, L. Collin, and K. Yao, "Blind channel estimation for stbc systems using higher-order statistics," *IEEE Transactions on Wireless Communications*, vol. 10, no. 2, pp. 495–505, February 2011.