High Rate Acoustic Link for Underwater Video Transmission

Costas Pelekanakis,[1] Milica Stojanovic[1] and Lee Freitag[2]

[1] Massachusetts Institute of Technology, Cambridge, MA 02139
[2] Woods Hole Oceanographic Institution, Woods Hole, MA 02534

Abstract **A high bit rate acoustic link for video transmission over an underwater channel is investigated. The key to achieving this objective lies in two approaches: use of efficient data compression algorithms and use of high-level bandwidth-efficient modulation methods. Currently available video encoding standards allow video transmission at bit rates as low as 64 kbps. While this rate is still above the limit of commercially available acoustic modems, prototype acoustic modems based on phase coherent modulation / detection have demonstrated successful transmission up to 30 kbps over a deep-water vertical path. To bridge the final gap and provide acoustic transmission capability needed for near real-time video, we focus on the use of high-level bandwidth-efficient modulation methods. An experimental system, based on discrete cosine transform and Huffman entropy coding for video compression, and variable rate M-ary QAM was implemented. Phase-coherent detection is accomplished by decision-directed synchronization and adaptive equalization. System performance is demonstrated experimentally, using 25000 symbols/sec at a carrier frequency of 75 kHz over a short vertical path. Excellent results were obtained using modulation methods of 16, 32 and 64-QAM, thus achieving bit rates as high as 150 kbps, which are sufficient for real-time transmission of compressed video.**

Keywords **Acoustic communications, bandwidth efficient modulation, compressed video, autonomous underwater systems.**

I. INTRODUCTION

The problem of real-time wireless video transmission from an underwater vehicle to a surface platform represents one of the last milestones in the development of autonomous systems for ocean exploration and monitoring. Underwater exploration has emerged as an area of great interest to many scientists and engineers, as well as to general population. Crucial components in the development of ocean-monitoring systems are the autonomous underwater vehicles and the means by which they can communicate to the surface. While the majority of today's underwater imaging is performed by transmitting signals from submersibles to the surface via optical cables, advances in acoustic underwater communications make it possible to conceive of a scenario in which video signals are transmitted in a wireless manner.

A particular scenario of interest to this work is transmission over a vertical channel at depths between several tens and several hundreds of meters. This scenario is relevant to applications where an autonomous underwater vehicle (AUV) is overshadowed by an autonomous surface craft (ASC) so that the two remain in communication over a near-vertical link. Vertical transmission minimizes multipath propagation, and the achievable bit rate is limited only by the system bandwidth.

The objective of this work is to demonstrate the capability to acoustically transmit digital signals over an underwater channel at rates that are sufficiently high to accommodate real-time video information. The challenge encountered in achieving this goal is that the data rate required for video transmission is high as compared to the bandwidth of the acoustic channel. Currently available commercial acoustic modems provide transmission rates up to several kilobits per second (kbps). While these rates may be sufficient for navigation and control, data rates that are at least ten, if not a hundred times higher are required for reasonable quality video transmission.

The key to achieving video transmission over the band-limited underwater channels lies in two approaches: (1) efficient data compression and (2) use of highly bandwidth-efficient modulation methods. The goal in combining these two approaches is to close the gap between the bit rate needed for video transmission and that supported by the acoustic channel.

The focus of this paper is on the design and experimental demonstration of high-level bandwidth-efficient modulation methods. Namely, a system employing variable rate modulation, ranging from 8 to 64-QAM was implemented in software and tested over a 10 meter vertical path using transducers with 60 kHz – 90 kHz bandwidth. The system was used to transmit an underwater video clip, encoded as a sequence of still images using standard discrete cosine transform and efficient entropy coding. Receiver processing includes phase-coherent detection and equalization, using both linear and decision-feedback adaptive equalizers. Preliminary results were achieved over a very short link, with bandwidth-efficiency as high as 6 bits/sec/Hz.

The paper is organized as follows. In Sec.II, an overview is given of the existing work on image and video compression for transmission over underwater acoustic channels. Sec.III presents design principles used in the experimental system. Results of data processing are given in Sec.IV. Finally, Sec.V summarizes conclusions and directions for future work.

II. Background

A. Acoustic communication link

The rate at which information (compressed video) can be transmitted through the ocean is limited by the bandwidth of the acoustic channel and by the distortions caused by multipath propagation. Multipath distortion is particularly severe on horizontal transmission links, where it is the ultimate factor that limits the achievable transmission rate. However, on carefully positioned vertical links this problem can be eliminated, or at least minimized. It then becomes possible to fully utilize the system bandwidth, provided that efficient signaling method is used. Bandwidth-efficient signaling is based on high-level PSK and QAM modulation methods. Phase-coherent detection methods, necessary for these modulations, are based on adaptive equalization and synchronization techniques [1], which have been favored in the majority of high-rate acoustic communication links [2]. These techniques also form the basis of the DSP implementation in the WHOI Utility Acoustic Modem, which has been successfully tested in a great number of horizontal acoustic channels over the past several years. Recently, the WHOI modem operation was also demonstrated over a 3000 m deep vertical link, where successful operation was achieved at 15 kbps [3]. In these applications, modulation methods used were 4-PSK or 8-PSK. Thus, the bandwidth efficiency of 3 bps/Hz was achieved. Having little multipath distortion, vertical acoustic channel is capable of supporting modulation methods of higher bandwidth efficiency.

B. Video compression

To achieve an acceptable frame rate (the number of image frames transmitted per second) images must be represented with as few bits as possible. Video transmission over underwater acoustic channels requires extremely high compression ratios. The approach currently favored by most of the experimental underwater image transmission systems is that of transmitting a sequence of still images. In this approach, each image from a sequence is encoded independently. Encoding is performed in an efficient manner to provide a certain compression ratio. The standard method for image coding is the transform domain coding, using the discrete cosine transform (DCT). In this method, the image is first transformed into a set of DCT coefficients. By eliminating the (spatial) redundancy between pixels, this transformation provides energy compaction, i.e., the number of coefficients needed to represent the image is generally much smaller than the number of original pixel levels. An alternative to transform domain coding is subband coding. In this approach, a discrete wavelet transform (DWT) is taken, which effectively decomposes the signal (pixel levels) into subbands of unequal length, where each subband is represented by its transform coefficients. In this manner, subbands that contain more information can be represented more precisely, thus achieving energy compaction. The coefficients (DCT or DWT) are then quantized, using, in the simplest form, a scalar quantizer. Vector quantization can be applied instead to groups of coefficients, to ultimately provide better compression. Finally, the quantized levels are encoded, using an efficient method such as entropy coding (the levels that occur most frequently are represented by fewest bits). The resulting frame of bits is then transmitted. It is, of course, desirable to support as high a frame rate as possible. Commonly, it is required to have a frame rate on the order of 10 frames per second for an acceptable video quality.

Video compression is different from compression of still images in that it considers the incremental difference between images, rather than compressing each image individually. In video compression, the first image of a sequence is transmitted as usually, but from then on, only the difference between images is encoded and transmitted. Because there is (temporal) redundancy among adjacent images in a video, the differential information can be transmitted at a lower rate for the same video quality. Motion-compensated prediction is a method frequently used for low-bit-rate video coding. In this method, further gains in compression are achieved by encoding the difference between the image and its predicted value, rather than the difference between the current and previous image. Prediction is performed in an optimal manner based on the history of images. The coefficients of the prediction filter are then transmitted along with the encoded signal. The encoding of the prediction error signal is performed using any of the usual image coding methods. These principles are used in the standardized video compression algorithms.

During the past several years, there has been a proliferation of work in the domain of low-bit-rate image coding, driven largely by the demand for video conferencing over band-limited channels—both wireline and wireless. ITU standards H.263, and the efforts of MPEG-4 group are concerned with video transmission at bit rates below 64 kbps. Coding and decoding algorithms that use reduced size images and reduced frame rates to comply with bit-rate requirements as low as 9.6 kbps are commercially available, although the software is often proprietary. In addition, there is on-going research on video coding algorithms whose design targets a pre-specified bit-rate on the order of 10-20 kbps. For example, reference [4] describes a compression scheme that transmits 144 x 176 pixel images, with 8 bits per pixel and 10 frames per second using 16 kbps. Bit rates in this range can be well supported by a carefully designed acoustic link.

Despite the advances in low bit rate coding for video transmission over band-limited channels, all but the most recent experimental underwater systems rely on encoding of still images using JPEG principles.

C. Experimental systems

The first system to demonstrate image transmission over a vertical path was developed in Japan [5]. The JPEG standard DCT was used to encode 256 x 256 pixel still images with 2 bits per pixel. Transmission of about one frame per 10 seconds was achieved using 4-PSK at 16 kbps. The remarkable results obtained with this system included a video of a slowly-moving object, transmitted acoustically from a 6,500 m deep ocean trench. Another vertical-path image transmission system was developed in France, and successfully tested in 2,000 m deep water [6]. This system was also based on the JPEG standard, and used 2-PSK for transmission at 19 kbps.

More recently, an image transmission system has been developed in a Portuguese effort called ASIMOV [7]. In this project, a vertical transmission link is secured by a coordinated operation of an AUV and an ASC. Once the site is chosen and the vehicles are positioned, transmission of a sequence of still images at about 2 frames/sec is accomplished at 30 kbps using an 8-PSK modulation method.

While the approach of still image coding suffices for many underwater applications, improvements are available from dedicated algorithms which combine image coding with motion compensation and prediction. The most recent experimental underwater video transmission system, developed in Japan [8], employs 4-PSK, 8-PSK and 16-QAM signals in a 40 KHz bandwidth to achieve transmission at up to 128 kbps. The system used 100 kHz carrier frequency and was tested over a short vertical path of 30 m. The MPEG-4 standard was employed for video compression, and a frame rate of 10 frames/sec was supported.

D. Compression of underwater imagery

Efficient compression can be achieved if there is *a-priori* information available about the images to be taken. Algorithms that exploit the properties inherent specifically to underwater images are such an example. Because underwater images have low contrast, their information is concentrated at low frequencies. Thus, by decomposing the image information into low and high frequency subbands, and encoding the low bands with more precision, it is possible to achieve higher compression ratios. This is the basic motivation behind the work in [9] which used the DWT in place of the standard DCT. The DWT is combined with entropy-constrained vector quantization (ECVQ) and motion-compensated prediction to achieve an average of 0.08 bits/pixel. This algorithm was applied to a sequence of underwater images, taken at 30 frames per second, each having 256 x 256 8-bit pixels. The achieved compression ratio of 100:1 provided very good quality monochrome video. The resulting bit rate needed to support such high quality is on the order of 160 kbps, which surpasses the capabilities of the current acoustic modem technology. However, the algorithm is equally applicable to reduced-size images. For example, a 144 x 176 pixel image would require 60 kbps with 30 frames per second, or 20 kbps with 10 frames per second. These values are approaching the capabilities of an acoustic modem, provided that a bandwidth-efficient modulation / detection scheme is used.

Another system that exploits wavelet-based compression together with motion-compensation is proposed in [10]. Although it attains approximately the same compression ratio (100:1) as in [9], it has better visual intelligibility because it employs a generalized dynamic image model (GDIM) that decouples the geometric and photometric variations in an image sequence commonly encountered in deep sea imagery. This approach is in contrast with ordinary terrestrial motion-compensated algorithms, where steady and uniform illumination is the underlying assumption. Using 128 x 128 pixel frames and 30 frames/sec, the resulting bit rate needed to support real-time video transmission is on the order of 40 kbps.

The traditional methods described above fall into the category of hybrid methods, because they combine image compression with motion compensation. A different approach is emerging in the form of model-based video compression methods. These methods exploit the *a-priori* knowledge of shapes that appear in a particular video segment. A unique example is the algorithm proposed in [11], which is designed especially to capture the bubble emission process from hot vents. Because they rely on the assumptions about a model, these algorithms have limited applicability; however, they can yield high compression ratios. At the moment, custom-design of compression algorithms for real-time transmission of underwater video remains an open research area.

III. SYSTEM DESCRIPTION

The proposed system is based on the use of variable level PSK and QAM modulation whose goal is to achieve as high bandwidth-efficiency as possible over a given acoustic channel. The system block diagram is shown in Fig. 1. At the input to the system are video frames, which are first compressed using a selected method. In the current implementation, images in the video sequence are compressed individually, by applying DCT to 8 x 8 pixel subblocks of a frame, followed by scalar quantization of coefficients and constrained-length Huffman coding. The quantizer and the Huffman tables can be found in [12]. The resulting bit stream is mapped into the symbols of desired signal constellation. 8-PSK, 16-QAM, 32-QAM and 64-QAM modulation methods with rectangular constellation shapes have been implemented so far. After addition of training data and packetizing, transmitter filtering is performed using a square root raised cosine pulse with roll-off factor 0.25 and truncation length of ± 4 symbol intervals. The signal is then modulated onto the carrier and passed to the output stages of the transmitter.

The received signal, after amplification and A/D conversion is shifted to baseband, low-pass filtered and down-sampled to 2 samples/symbol. Packet synchronization is achieved by matched-filtering to a 28-bit Barker sequence. Adaptive filtering by a T/2 fractionally-spaced equalizer with integrated phase tracking is then employed. Both linear and decision-feedback equalizers were considered, operating under LMS and RLS algorithms, respectively. The detected data symbols are finally converted to bits and passed on to the video decoder.

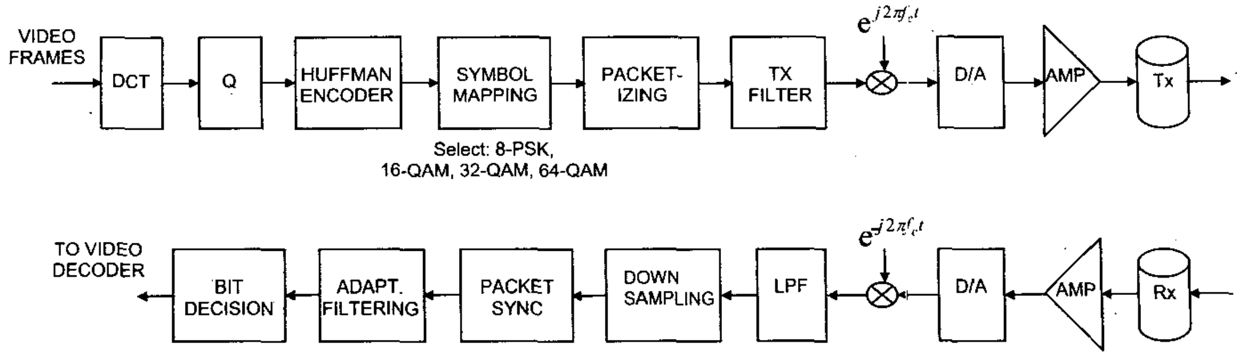

Fig. 1. Block diagram of the transmitter and receiver.

IV. Experimental Results

The experiment was conducted in Woods Hole, MA in November 2002. Below, we describe the experiment set-up, the signals used in the experiments, and the results of data processing.

A. Experiment set-up

Fig. 2 illustrates the two RD Instruments Long Ranger transducers used for the experiment. The transducers have a 6^0 degrees conical beampattern providing approximately 15 kHz of usable bandwidth at a carrier frequency of 75 kHz.

Fig. 3 shows a 10 m pole, on which the two transducers were mounted. The pole was vertically submerged, with the receiver 2 meters below the surface, and the transmitter at the lower end.

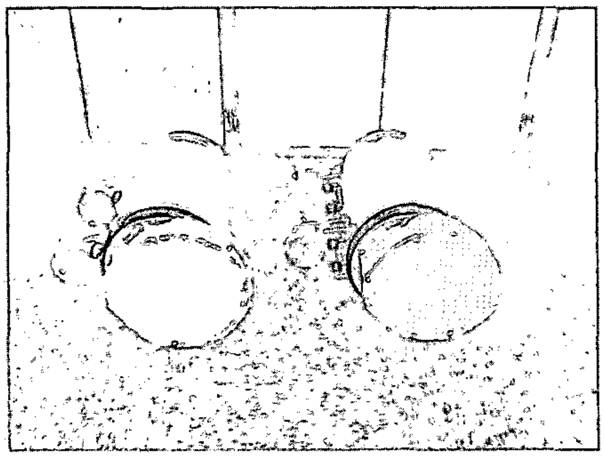

Fig. 2. Transducers used in the experiment.

All the signal processing described in Sec.III was implemented in Matlab, running on two lap-top computers, one acting as the transmitter and the other as the receiver. Each computer had a National Instruments DAQ Card-6062E which was used for digital-to-analog and analog-to-digital conversion of the signals used during the experiment. The passband sampling rate was 250 kHz.

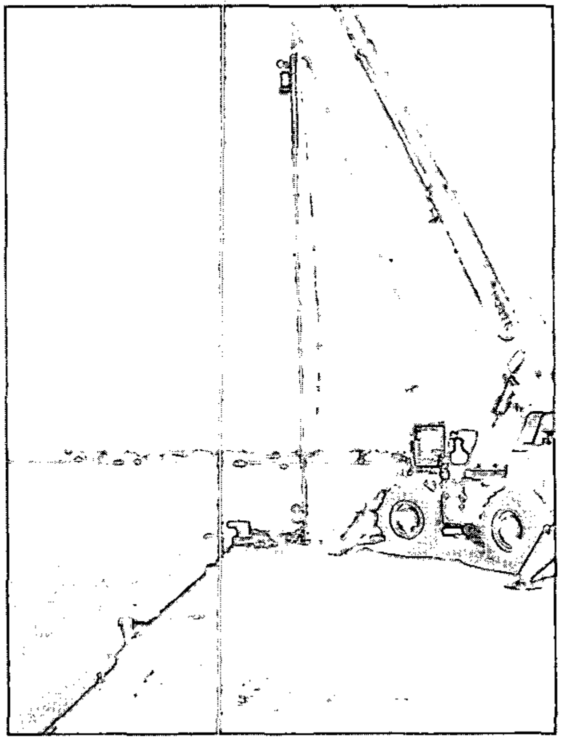

Fig. 3. Both transducers are attached to a 10 m pole.

B. Signals used in the experiments

The input digital video had 129 frames, playing at 15 frames/sec. The frame size was 144 x 176 pixels, with resolution of 8 bits/pixel. A representative frame is shown in Fig. 4. Without compression, this video would need a transmission rate in excess of 3 Mbps.

Using the simple compression method described in Sec.III, various compression ratios can be achieved depending upon the desired image quality. Fig. 5 shows one of the frames obtained with an average compression ratio of 40, i.e., with 0.2 bits/pixel. With this compression ratio, the bit rate needed to sustain a frame rate of f frames/sec is roughly $R = f \cdot 5$ kbps. Thus, for a frame rate of 15 frames/sec, the bit rate of 75 kbps is needed.

Transmission was organized in packets of fixed duration. Each packet contained 3958 data symbols, and an additional block of training data, whose minimum size depends on the adaptive filtering method used by the receiver. The training sequence was generated as a pseudo-random binary sequence, mapped into the same signal constellation as the rest of the data block. Each data block was preceded by a synchronization probe and a guard time whose duration need not be longer than the expected multipath spread. The design parameters used in the experiment, 500 training symbols and 50 ms guard time, were chosen with a larger-than-necessary safety margin. Transmission rate was 25000 symbols/sec for all modulation methods considered. For all the data packets, peak transmission power was equal. Note that the resulting average symbol energy, and consequently the detection SNR, is lower for a higher level modulation method.

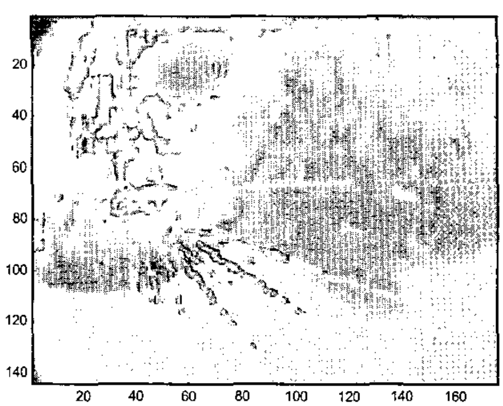

Fig. 4. An image from the original video sequence. Frame size is 144 x 176 pixels, 8 bits/pixel.

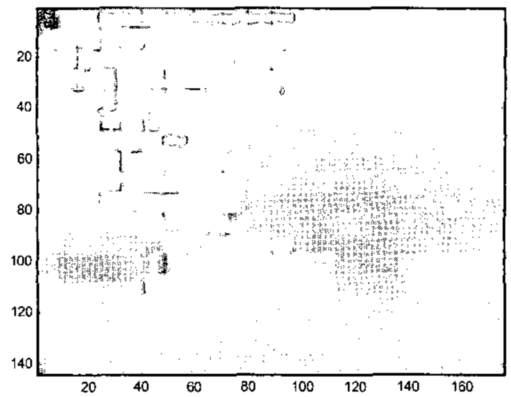

Fig. 5. The compressed image. Average compression ratio (over 129 frames) was 40, resulting in 0.2 bits/pixel.

C. Results

The short vertical channel chosen for these preliminary tests proved to have very little distortion, allowing for excellent signal detection using all modulation formats. Fig. 6 shows the results of linear equalization using an 11-tap adaptive filter operating under the LMS algorithm. Shown in the left and right-hand columns are the scatter plots before and after equalization. While linear adaptive equalization succeeded in errorless detection of all data packets with modulation levels up to 16, it encountered occasional difficulties with 32 and 64-QAM signals, where sensitivity to phase jitter and residual intersymbol interference at lower SNR limited its performance. In these cases, a decision-feedback equalizer, with the same number of taps (6 feedforward and 5 feedback), operating under an RLS algorithm and with integrated phase tracking, provided errorless performance in all cases investigated. Fig. 7 shows the receiver performance in this case. Shown in the figure is the received signal in time, the estimated channel response, and the scatter plots before and after equalization.

The results of signal processing presented above show excellent performance of all modulation methods. We note that the bit rates achieved with 8-PSK, 16-QAM, 32-QAM and 64-QAM are 75 kbps, 100 kbps, 125 kbps and 150 kbps, respectively. These rates suffice for real-time transmission of the 144 x 176 frame sequence at 15 frames/sec, even if still image compression is used. The quality of received video is identical to that of compressed (Fig. 5) as there are no bit errors in detection.

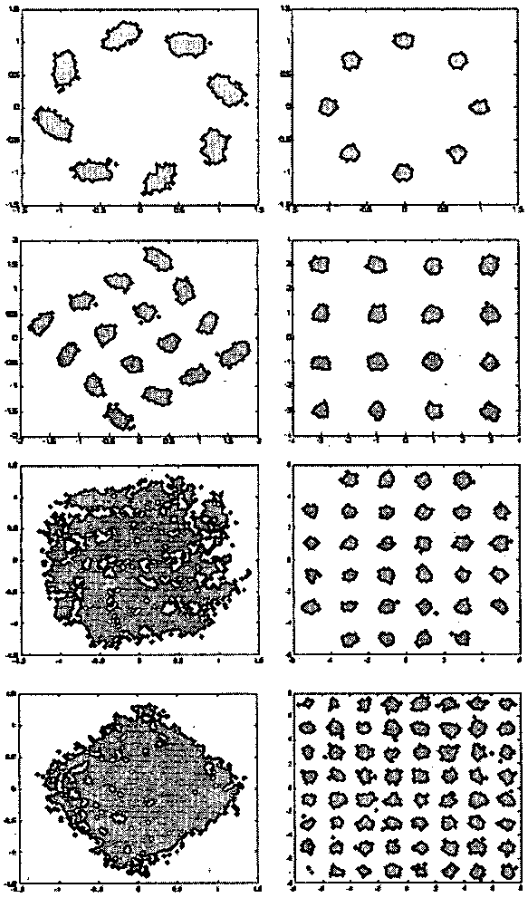

Fig. 6. Scatter plots of 8-PSK, 16-QAM, 32-QAM and 64-QAM, before and after equalization using an 11 tap linear LMS equalizer.

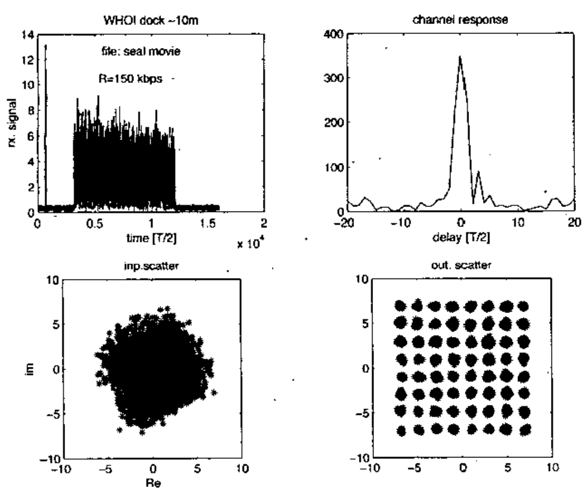

Fig. 7. Results of 64-QAM data processing using an RLS DFE(6,5) with integrated phase tracking.

V. SUMMARY AND FUTURE WORK

Two factors are instrumental in enabling real-time video transmission over wireless acoustic underwater channels: (1) efficient data compression and (2) bandwidth-efficient modulation. The focus of this paper was on the latter class of methods. A system employing variable rate M-QAM techniques was designed and applied to the experimental data transmitted over a short vertical channel. Excellent results were achieved at bit rates up to 150 kbps, using modulation methods with bandwidth efficiency as high as 6 bits/sec/Hz. This rate is sufficient to support real-time transmission of compressed video.

Future work in this area will focus on investigating the applicability of high-level modulation methods to longer-range channels with controlled degree of mobility between the transmitter and receiver. In addition, signal constellation shapes other than rectangular QAM will be investigated for possibly lower sensitivity to channel distortions. Finally, video compression techniques based on motion compensation will be integrated into the system.

REFERENCES

[1] M.Stojanovic, J.Catipovic and J.Proakis, "Phase coherent digital communications for underwater acoustic channels," IEEE J. Oceanic Eng., vol.19, pp.100-111, Jan. 1994.

[2] D.Kilfoyle and A.Baggeroer, "The state of the art in underwater acoustic telemetry," IEEE J. Oceanic Eng., vol. 25, pp. 4-27, Jan 2000.

[3] L.Freitag, M.Johnson and D.Frye, "High-rate acoustic communications for ocean observatories -- performance testing over a 3000 m vertical path," in Proc. IEEE Oceans'00 Conference, Providence, RI, Sept. 2000.

[4] H.Kwon, M.Venkatramam and N.Nasrabadi, "Very low bit-rate video coding using variable block-size entropy-constrained residual vector quantizers," IEEE J. Select. Areas Commun., vol.15, pp. 1714-1725., Dec.1997.

[5] M.Suzuki and T.Sasaki, "Digital acoustic image transmission system for deep sea research submersible,'' in Proc. IEEE Oceans'92 Conference, Newport, RI, Sept. 1992, pp. 567-570.

[6] G.Ayela and J.Coudeville, "TIVA: A long range, high baud rate image/data acoustic transmission system for underwater applications", Proc. Underwater Defense Technol. Conf., Paris, France, 1991.

[7] J.Gomes, V.Barosso, G.Ayela and P.Coince, "An overview of the ASIMOV Acoustic communication system," in Proc. IEEE Oceans'00 Conference, Providence, RI, Sept. 2000.

[8] J.Kojima, T.Ura, H.Ando and K.Asakawa, "High-speed acoustic data link transmitting moving pictures for autonomous underwater vehicles", in Proc. IEEE Intl. Symp. on Underwater Technology, 2002, pp. 278-283.

[9] D.Hoag, V.Ingle and R.Gaudette, "Low-bit-rate coding of underwater video using wavelet-based compression algorithms," IEEE J. Oceanic Eng., vol. 22, pp. 393-400, April 1997.

[10] S.Negahdaripour and A.Khamene, "Motion-based compession of underwater video imagery for the operations of unmanned submersible vehicles", Computer Vision and Image Understanding, Academic Press, 2000.

[11] J.Barbosa and V. Barroso, "Very low bit-rate for acoustic transmission of subsea images of hydrothermal vents," in Proc. IEEE Oceans'00, Providence, RI, Sept. 2000.

[12] V.Bhaskaran and K.Konstantinides, "Image and Video Compression Standards," 2nd Edition, Kluwer Academic Publishers, 2003.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Darryl Symonds and Mr. Harry Maxfield of RD Instruments for supplying the transducers used in the experiments.