Design and Analysis of a High-Rate Acoustic Link for Underwater Video Transmission

by

Konstantinos Pelekanakis

Diploma, Electrical Engineering (2001)
Technical University of Crete, Greece

Submitted to the Department of Ocean Engineering
In Partial Fulfillment of the Requirements for the Degree of
Master of Science in Ocean Engineering
at the
Massachusetts Institute of Technology
June 2004

Signature of Author...

<div align="right">

Department of Ocean Engineering
May 7, 2004

</div>

Certified by ....

<div align="right">

Dr. Milica Stojanovic
Principal Scientist, MIT Sea Grant College Program
Thesis Supervisor

</div>

Accepted by ....

<div align="right">

Michael Triantafyllou
Professor of Ocean Engineering
Chairman, Department Committee on Graduate Studies
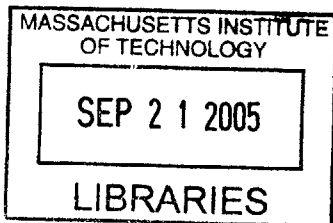
</div>

Design and Analysis of a High-Rate Acoustic Link for Underwater Video Transmission

by

Konstantinos Pelekanakis

Submitted to the Department of Ocean Engineering

on May 7, 2004 in Partial Fulfillment of the Requirements for

the Degree of Master of Science in Ocean Engineering

ABSTRACT

A high bit rate acoustic link for underwater video transmission is examined. Currently, encoding standards support 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 at 30 kbps over a deep water channel.

The key to bridging the remaining gap between the bit-rate needed for video transmission and that supported by the acoustic channel lies in two approaches: use of efficient image/video compression algorithms and use of high-level bandwidth-efficient modulation methods. An experimental system, based on discrete cosine transform (DCT) and Huffman entropy coding for image compression, and variable rate M-ary quadrature amplitude modulation (QAM) was implemented. Phase-coherent equalization is accomplished by joint operation of a decision feedback equalizer (DFE) and a second order phase locked loop (PLL). System performance is demonstrated experimentally, using transmission rate of 25000 symbols/sec at a carrier frequency of 75 kHz over a 10 m vertical path. Excellent results were obtained, thus demonstrating bit rates as high as 150 kbps, which are sufficient for real-time transmission of compressed video.

As an alternative to conventional QAM signaling, whose high-level constellations are sensitive to phase distortions induced by the channel, M-ary differential amplitude and phase shift keying (DAPSK) was used. DAPSK does not require explicit carrier phase synchronization at the receiver, but instead relies on simple differentially coherent detection. Receiver processing includes a linear equalizer whose coefficients are adjusted using a modified linear least square (LMS) algorithm. Simulation results confirm good performance of the differentially coherent equalization scheme employed.

Thesis Supervisor: Milica Stojanovic
Title: Principal Scientist, MIT Sea Grant College Program

Table of contents

List of Figures

List of Tables

Acknowledgements

"Give me where to stand, and I will move the earth."

Archimedes, 300 B.C.

Chapter 1

1 Introduction

1.1 Motivation

In the recent years, there has been an increasing interest in underwater acoustic (UWA) communications brought by the growing need for wireless transmission of information for applications such as pollution monitoring, remote control of submerged devices and collection of scientific data recorded in moored stations without retrieving the instruments. 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. The majority of today's underwater imagery is obtained by transmitting signals via optical cables. However, cables limit the operational range and maneuvering capability. Advances in UWA communications make the transmission of wireless video signals possible, thus leading to the vision of an autonomous oceanographic

sampling network (Figure 1), a collection of bottom-mounted instruments and free-swimming devices that provides remote and almost instantaneous access to oceanographic data.

Figure 1: Autonomous Oceanographic Sampling Network [1].

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 so that the achievable bit rate is limited only by the system bandwidth.

Video transmission involves large amount of data, which must be compressed prior to transmission in order to reduce the bit rate that is presented to the acoustic channel. 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 available system bandwidth. Currently available commercial acoustic modems provide transmission rates up to several kbps. For example, the Benthos acoustic modem operates at rates between 0.1 kbps and 2.4 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.

1.2 Coherent detection for underwater acoustic channels

The bit rate available over an UWA channel is limited by the system bandwidth. For example, long range systems operating over several tens of kilometers are confined to few kHz of bandwidth, medium range systems operating over several kilometers are limited to tens of kHz of bandwidth, while only short range systems operating over several tens of meters may have more than a hundred kHz. Within this limited bandwidth, the signal is subject to distortions caused by multipath propagation. While in radio channels a typical multipath spread spans a few symbol intervals, the UWA channels suffer from extended multipath propagation which usually increases with range, and, depending on the signaling rate, causes the intersymbol interference to span up to several tens of symbol intervals.

Consequently, bandwidth efficiency of candidate systems becomes an important issue in UWA communications. FSK (frequency shift keying) modulation, which has been favored for underwater channels due to its robustness to phase distortion, inefficiently uses the bandwidth, and, consequently, is not suitable for high data rate applications. To fully utilize the limited system bandwidth and increase the speed of data transmission, bandwidth-efficient signaling based on high-level PSK and QAM modulation methods must be employed.

Phase-coherent detection systems, necessary for PSK/QAM modulations, are based on adaptive equalization and synchronization techniques, which have been used in the majority of high-rate acoustic communication links [2],[3]. A representative system based on phase-coherent modulation operates over 60 m range at a carrier frequency of 1 MHz using 125 kHz bandwidth and achieving data rate of 500 kbps [4]. This system is used for communication with an undersea robot which performs maintenance of a submerged platform. A 16QAM modulation scheme was used and the performance is aided by an adaptive LMS equalizer. The work by Stojanovic et al. [5] is based on an effective algorithm that combines decision feedback equalization with a second-order PLL. Transmissions were made in long-range deep water (110 nautical miles, 660 b/s), long-range shallow water (48 nautical miles, 1kbps) and short-range shallow water (2 nautical range, 10kbps). This technique also forms 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 [6]. In the last two applications, the modulation methods used were 4PSK or 8PSK. Thus, the maximum bandwidth efficiency of 3 b/s/Hz was achieved.

1.3 Differentially coherent detection

The implementation and operation of a coherent receiver is, in general, a complex process involving fine tuning of the PLL parameters that change from one acoustic channel to another. In many applications, simplicity and robustness of implementation take precedence over achieving the best possible performance. In these cases, differential detection, which obviates the need for explicit phase synchronization offers an attractive alternative to coherent detection. Aside from implementation simplicity, differentially coherent detection may offer a more robust performance with high-level signal constellations.

As an alternative to bandwidth-efficient modulation methods, a multilevel differential modulation technique, the so-called Differential Amplitude and Phase Shift Keying (DAPSK) is investigated. DAPSK encodes information in the signal phase and amplitude relative to the previous symbol, a strategy that leads to simple receiver structure, at the expense of higher bit error rate.

The problem of concurrent differentially coherent detection and equalization necessary for multipath channels is addressed in the early work by Sehier and Kaleh [7], later by Masoomzadeh-Farad and Pasupathy [8] and more recently by Schober et al. [9]-[11]. These methods have not been considered previously for underwater channels. Below, a brief summary of the existing solutions is given.

In [7], a linear equalizer followed by the differentially coherent detector is employed. The tap coefficients of the equalizer are adjusted to minimize the minimum square error (MSE) obtained after differential detection. The system of equations to solve is of 3rd order, as opposed to coherent reception, preventing direct application of stochastic gradient-based

algorithms such as LMS. To circumvent this difficulty, the tap coefficients are considered to be constant from one symbol interval (iteration) to another. Then, the system of equations becomes linear yielding a modified LMS algorithm. Effectively, the equalizer is applied to the input signal pre-multiplied by the complex conjugate of the previously estimated symbol, which accounts for differentially coherent phase compensation prior to equalization.

In [8] two receivers were implemented: one based on linear equalization and another based on nonlinear equalization both compatible with DPSK signals. For linear equalizers, a modified recursive least squares (RLS) algorithm was considered using the same approach of linearizing the post-detection error as in [7]. For severe channel degradations (i.e. for channels with spectral nulls), the authors proposed to perform differentially detection first, followed by a DFE. However, this approach generates additional nonlinear distortion and is not optimal.

Scober et al. considered both linear equalization [9], [11] and nonlinear equalization schemes [10]. The nonlinear scheme, which is compatible with DPSK signals, includes a DFE followed by a differential demodulator and is optimal because maximum-likelihood sequence estimation (MLSE) is applied to the input signal containing linear ISI; thus, it can approach the performance of a coherent MMSE-DFE arbitrarily close. The linear scheme of [9], which is suitable for M-ary DPSK, approaches a coherent linear MMSE equalizer and the linear scheme of [11], which is suitable for M-ary DAPSK, approaches a zero-forcing equalizer. The improvement stems from generating a reference symbol at the output of the equalizer to track the carrier-phase of the received signal prior to demodulation. The reference symbol is generated from N previously detected symbols and as N tends to infinity, optimum carrier reference is obtained. Hence, these methods approach the performance of coherent detection. When the reference symbol is generated from a single previous observation ($N=1$), the linear

equalizer [9] reduces to that of [7]. For an efficient recursive adaptation of the equalizer coefficients, a modified LMS and a modified RLS are developed.

1.4 Compression techniques for low bit-rate image/video coding

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 quantized, using, in the simplest form, a scalar quantizer. Vector quantization can be applied instead to groups of coefficients, to ultimately

13

provide better compression. Finally, the quantized levels are encoded, using an efficient method such as entropy coding in which 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. This approach is analogous to Differential Pulse Code Modulation (DPCM) used for voice transmission, and many of the same ideas that were originally developed for voice transmission apply to video as well.

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.

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 bandlimited 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 [12] describes a compression scheme that transmits 144x176 pixel images, with 8 bits/pixel and 10 frames/sec 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 bandlimited channels, all but the most recent experimental underwater systems rely on encoding of still images using JPEG principles. JPEG is a standard method for image compression that is based on DCT, scalar quantization and Huffman coding.

1.5 Experimental systems

The first system to demonstrate image transmission over a vertical path was developed in Japan [13]. The JPEG standard DCT was used to encode 256x256 pixel still images with 2 bits per pixel. Transmission of about one frame per 10 seconds was achieved using 4-DPSK at 16 kbps. The remarkable results obtained with this system included a video of a slowly-moving crab, 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 [14]. This system was also based on the JPEG standard and used binary DPSK for transmission at 19 kbps.

More recently, an image transmission system has been developed in a Portuguese effort called ASIMOV [15]. 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 of about 2 frames/sec is accomplished at 30 kbps using an 8PSK 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 [16], employs 4PSK, 8PSK and 16QAM signals with 40 KHz bandwidth to achieve transmission at up to 128 kbps. The system uses 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.

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 [17] 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 256x256 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 144x176 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 [18]. Although it attains approximately the same compression ratio (100:1) as in [17], 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 128x128 pixel frames and 30 frames/sec, the resulting bit rates needed to support real-time video transmission is of 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 [19], 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.

1.6 Thesis overview

From the above discussion, it is evident that even after massive compression, video signals require bit rates that are relatively high according to the current standards of underwater acoustic modems. The key to achieving video transmission over the band-limited underwater channels lies in two approaches:

1. Efficient video compression.

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 thesis is on the second approach.

Higher-level modulation methods efficiently use a fixed transmission bandwidth by allocating several bits of information to a single transmitted symbol. In this manner, several bits per second are transmitted per Hz of occupied bandwidth. The modulation methods that can achieve this come from the class of M-ary QAM and M-ary PSK methods, which require coherent detection at the receiver. For example, an 8PSK symbol carries 3 bits of information, while a 16QAM symbol carries 4 bits of information.

Over horizontal channels, where multipath propagation causes strong interference, these methods also require sophisticated signal processing for equalization. For this reason, their

application so far has mostly focused on the basic, binary and quaternary PSK methods. In contrast, the vertical path represents a better quality communication channel, which can be expected to support higher-level modulation schemes, thus making full utilization of the bandwidth-efficient signaling. However, this fact has not been exploited in the existing systems that are often limited to 8PSK signals. In comparison, some telephone channels and microwave radio links today use modulations with 256 levels.

The focus of this work is on the design and experimental demonstration of high-level bandwidth-efficient modulation methods with bandwidth efficiency greater than 3 bits/s/Hz. As the number of bits per symbol increases, the performance, measured in terms of bit error probability, degrades. However, the high signal to noise ratio of a shallow or moderately deep vertical channel and the modest bit error probability requirement of image transmission offer favorable conditions for a trade-off between bandwidth-efficiency and performance. The major concern in using a high-level signal constellation is its sensitivity to phase jitter and time-varying channel response. While the vertical channel exhibits little multipath, vertical motion due to waves causes substantial Doppler effects [20]. These effects are evident both in the carrier phase variation and in pulse compression/dilation. The frequency offset caused by vertical motion cannot be assumed constant as the direction of the motion changes in time.

In choosing the modulation method, it is not only the constellation size that plays a role in the signal design, but also the constellation shape. The conventional shapes that are easy to implement are the circular shape of PSK and the square shape of QAM methods. In addition to these shapes, there are signal constellations that are potentially more robust to the channel impairments. One such class of constellations is the DAPSK (star shaped QAM), in which concentric circles contain PSK-shaped signal points. These signals have been found to

provide improved performance in situations with difficult phase tracking and signal level variation [21]. The improvement is based on differential signal encoding, in both amplitude and phase, which permits differentially coherent demodulation, thus eliminating the need for explicit phase tracking and gain control.

In this thesis, a system employing variable rate modulation based on coherent (M-ary QAM) or noncoherent (M-ary DAPSK) detection was implemented in software and tested over a 10 meter vertical path using transducers with 25 kHz bandwidth and center frequency of 75 kHz. The system was used to transmit an underwater video clip, encoded as a sequence of still images using standard JPEG coding. Packet synchronization is achieved by matched-filtering to a 28-bit Barker sequence. Receiver processing includes two options:

1. Phase synchronization and channel equalization using jointly optimized decision-feedback equalizer and decision-directed phase estimator. The weights of the equalizer are adjusted by an adaptive algorithm (RLS or LMS) and second order digital PLL.

2. Differentially coherent detection with linear channel equalization using modified LMS algorithm.

All the equalizers were implemented as fractionally spaced to eliminate the need for explicit symbol delay tracking [5]. Excellent results were obtained, thus demonstrating bit rates as high as 150 kbps, which are sufficient for real-time transmission of quality compressed video.

The thesis is organized as follows. In Chapter 1, an overview is given of the existing work on coherent and differentially coherent equalization as well as their applications in underwater video transmission over underwater acoustic channels. Chapter 2 presents the theoretical background of the JPEG standard, of the modulation techniques employed at the transmitter and of the signal processing employed at the receiver. Chapter 3 describes the experimental

20

system. Results of data processing are given in Chapter 4. Finally, Chapter 5 summarizes

conclusions and directions for future work.

Chapter 2

2 Theory

2.1 JPEG image compression

The goal of a still-image-coding algorithm is to exploit spatial redundancy among neighboring pixels thereby reducing the size of the 2-D data set. A popular approach is transform domain coding in which the original data is projected onto a set of basis functions such that a large number of the transform coefficients contain little energy. By zeroing out the coefficients deemed insignificant and efficiently coding the remaining coefficients, the original data set becomes compressed. Video compression can be regarded as an extension to the still image compression problem since digital video is composed of a sequence of images. The main difference between the two problems is that video may be compressed at much higher rates since temporal correlation exists among pixels across the frames of the sequence.

A common image compression standard is JPEG. According to this standard, an image is partitioned into 8x8 non-overlapping pixel blocks form left to right and top to bottom. Each

block is DCT coded, and the 64 transform coefficients are quantized to the desired quality. The quantized coefficients are entropy-coded and output as part of the compressed image data. Figure 2 illustrates the JPEG compression algorithm. Each 8-bit sample is level shifted by subtracting 128 before being DCT coded. The 64 DCT coefficients are then uniformly quantized according to the step size of a given quantization matrix. Entropy coding is accomplished in two stages. The first stage is the translation of the quantized DCT coefficients into an intermediate set of *symbols*. In the second stage, each symbol is divided into two parts; a variable length code (VLC) is assigned to the first part (*symbol 1*), followed by a binary representation of the amplitude for the second part (*symbol 2*).
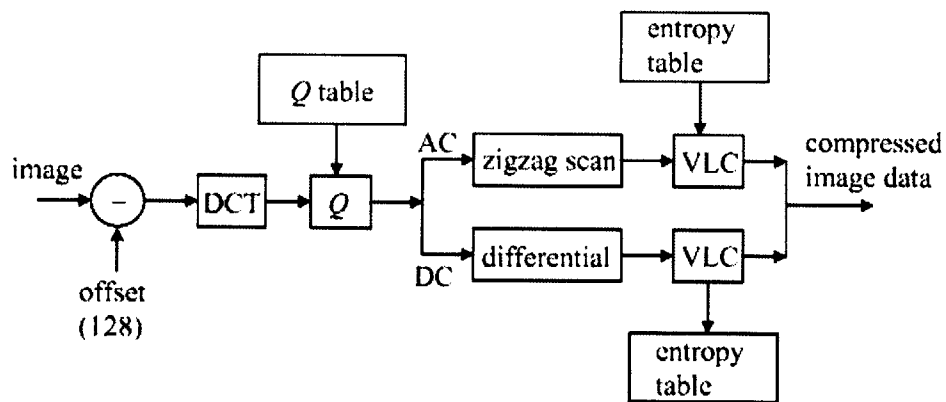
Figure 2: Block diagram of a JPEG encoder [22].

2.1.1 Data compression using DCT

The DCT is an orthogonal transform with basis sequences that are cosines; thus it is closely related to the DFT. In the development of the DFT, a periodic sequence is first formed by appending an infinite number of shifted replicas of the finite length sequence $x[n]$ in such a way so that $x[n]$ can be uniquely recovered. Similarly, since cosines are periodic and even functions, the DCT corresponds to forming a periodic and even sequence from $x[n]$ in such a way so that $x[n]$ can be uniquely recovered. Since there are many ways do to this, the DCT has many definitions. The most commonly used is:

$$X_{DCT}[k] = 2\sum_{n=0}^{N-1} x[n]\cos(\frac{\pi k(2n+1)}{2N}), \quad 0 \le k \le N-1$$

$$x[n] = \frac{1}{N}\sum_{k=0}^{N-1} \beta[k]X_{DCT}[k]\cos(\frac{\pi k(2n+1)}{2N}), \quad 0 \le n \le N-1, \quad \beta[k] = \begin{cases} 1/2, & k=0 \\ 1, & 1 \le k \le N-1 \end{cases} \tag{1}$$

For a real sequence $x[n]$, $0 \le n \le N-1$, the DCT coefficients $X_{DCT}[k]$ are real numbers while the DFT coefficients are, in general, complex, providing a significant advantage of the DCT method in the case of data compression algorithms. The relationship between DCT and DFT is developed in [23]:

$$X_{DCT}[k] = 2\operatorname{Re}\{X_{DFT}[k]e^{-j\pi k/(2N)}\}, \quad k = 0,...,N-1 \tag{2}$$

where $X_{DFT}[k]$ is the 2N-point DFT of the N-point sequence $x[n]$ defined as follows

$$X_{DFT}[k] = \sum_{n=0}^{N-1} x[n]e^{-j2\pi kn/(2N)}, \quad k = 0,...,2N-1 \tag{3}$$

The DCT is preferred to DFT in many data compression algorithms because of its "energy compaction" property. With respect to the DFT, the DCT of a finite length sequence often has its coefficients highly concentrated at low indices. This means that one can force to zero many

24

high-indexed coefficients without a significant impact on the energy of the signal. The "energy compaction" property can be intuitively justified by observing that the periodic, symmetric expansion of a finite-length signal often results in a smoother signal than its simple periodic expansion producing low energy content at high frequencies.

Exploiting the close relationship between the DCT and the DFT, fast algorithms such as the Fast Fourier Transform (FFT) can be used to compute the DFT in equation (2) making the DCT a very efficient method for data compression. In image compression, the 2-D DCT is used which is a separable transform, meaning the transform can be performed on each dimension independently. Therefore, one can first apply the 1-D DCT to the rows and then to the columns of a pixel matrix or vice versa. This property significantly reduces the complexity of the implementation.

2.1.2 Quantization

According to Figure 2 above, the quantization block is used to limit the range of the transform coefficients. This block gives the compression gain for the system; however, it introduces loss at the same time. There are many quantization schemes used in the lossy image compression. The JPEG recommends a quantizer similar to the optimal uniform quantizer, but designed specifically for the block-based DCT coefficients. Each frequency in the transformed block has a specific quantization level. These levels are derived from experimental results based on many images in order to give the best visual quality for a given bit rate. Table 1 contains the recommended quantization levels for the luminescence

component according to the JPEG standard. The top left corner corresponds to the quantization level for the DC component. The bottom right hand corner is the quantization level for the highest frequency component of the transform coefficients in 2-D.

If the elements of the quantization table are represented by $Q(u,v)$, then a quantized DCT coefficient with the horizontal and the vertical spatial frequencies of u and v, $F^q(u,v)$, is given by

$$F^q(u,v) = \left\lfloor \frac{F(u,v)}{Q(u,v)} \right\rfloor \qquad (4)$$

where $F(u,v)$ is the transform coefficient value prior to quantization. At the decoder the quantized coefficients are inverse quantized by

$$F^Q(u,v) = F^q(u,v) \times Q(u,v) \qquad (5)$$

to reconstruct the quantized coefficients.

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	13	16	24	40	57	69	56
14	17	22	29	51	87	80	62
18	22	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Table 1: Luminance quantization table [22].

2.1.3 Entropy coding

The entropy coder translates values or symbols into a string of binary digits. The encoder used in this project is the run-length Huffman encoder configured for JPEG. The run-length Huffman coder combines regular Huffman coding with the modifications of the quantized coefficients to achieve higher compression. This coder separates the coefficients into two parts: DC (commonly referred to as (0,0) coefficient) and the 63 AC coefficients. Each type is coded using a different method as shown in Figure 2. The general concepts are described below.

The DC coefficients have a very large range. For an 8-bit per pixel image, the DC coefficient lies in the interval [-2047, 2047]. This range requires 12 bits to represent. However, the value of the DC coefficients of the adjacent blocks are close to each other because most images are smooth and do not change abruptly. The DC coder uses differential pulse code modulation (DPCM) to first compute the difference between adjacent blocks, and then encodes the difference, as can be seen in Figure 3. The separate treatment of DC coefficients exploits the correlation between the DC values of adjacent blocks. Also, the DC coefficients are encoded more efficiently as they typically contain the largest portion of the image energy. The Huffman codeword for the DC coefficients can be separated into two parts: the category and the residue. The category (CAT) is determined by the range of magnitude of the DIFF (see Figure 3) and is then variable length coded using Huffman coding. The residue is determined by the sign and the amplitude of the DIFF. The CAT after being variable length coded is appended with additional bits to specify the actual DIFF values

(amplitude). The Huffman table for the DC coefficients of the DCT JPEG can be found in [22].
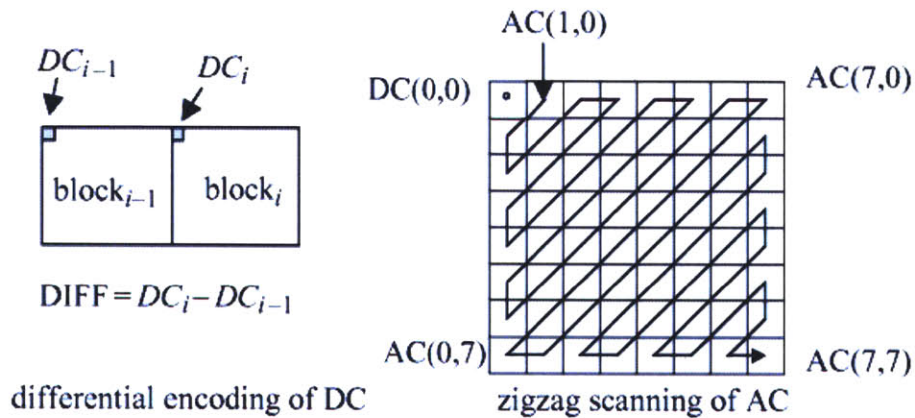

Figure 3: Encoding and zigzag scanning [22].

As explained previously, the quantized AC coefficients of the DCT are likely to be zeros at high frequencies. The run-length coder counts the maximum number of zeros before encountering a non-zero coefficient. The scanning method used for the AC coefficients is called zigzag scan and is illustrated in Figure 3. The maximum number of zero-run code word is 16. Each non-zero coefficient is represented in combination with the "run-length" (consecutive number) of zero-valued coefficients, which precede it in the zigzag sequence. Each such run-length/non-zero-coefficient combination is presented by a pair of symbols:

<div align="center">

Symbol-1   Symbol-2

(RUNLENGTH, CAT)   (AMPLITUDE)

</div>

Symbol-1 represents two pieces of information, RUNLEGTH and CAT. Symbol-2 represents a single piece of information, which is simply the amplitude of the non-zero coefficient. RUNLEGTH is the number of the consecutive zero-valued coefficients in the zigzag sequence preceding the non-zero coefficient being represented. CAT is determined by the range of magnitude of the amplitude. The end of block (EOB) is used if the rest of the scanning coefficients are zeroes. The run-length Huffman table coder for the AC coefficients can be found in [22].

2.2 Underwater acoustic channels

Sound propagation through sea water is affected by many factors. Vertical gradient of sound speed causes refraction of rays, sometimes in a way that prevents reception in certain locations known as shadow zones. Also, the signal is subject to random fluctuating Doppler shifts and spread because of surface and internal wave motion. In addition, there exist multiple propagation paths from transmitter to receiver in most underwater propagation geometries. In deep water these paths are described as rays or waveguide modes and in shallow water as reflections from discrete scatterers. Received signal fluctuations occur from medium fluctuations along a single path (micromultipath) and time-variant interference between several propagation paths (macromultipath) [20]. The mechanisms that cause distortions to signals transmitted through the UWA channel are summarized below.

2.2.1 Single-Path Fluctuation

From the viewpoint of building a robust communication receiver, the questions of primary interest are the types of ocean acoustic channels with fluctuation levels sufficiently low to permit coherent signaling in order to achieve greater data throughput. Low fluctuation geometries can be determined by investigating regions where the Rytov approximation for weak scattering holds [20]. The Rytov method yields perturbation log-intensity and log-phase variance parameters:

$$i \equiv \ln \frac{I_c + I_i}{I_c} \qquad (6)$$

where I_c is the intensity of the nonfluctuating field component and I_i is the fluctuating (noncoherent) field intensity. For weak fluctuations $I_c \gg I_i$, so that

$$i \cong \frac{I_i}{I_c} = \frac{1}{\gamma^2} \qquad (7)$$

where γ is a measure of acoustic fluctuating intensity defined as the ratio of coherent to noncoherent field intensity. It is customarily accepted that the Rytov approximation is valid if $i^2 < 0.3$ which is also close to the phase tracking limit for the coherent phase-locked loop [20]. The Rytov approximation validity region depends on the channel geometry and there is a strong experimental evidence of rapid system degradation with increasing transmission angle from the vertical [20]. For instance, the deep-water vertical path is amenable to coherent modulation and several coherent systems currently exist [6], [13].

2.2.2 Multipath

The mechanism of multipath formation depends on the channel geometry and also on the frequency of transmitted signals. Understanding of these mechanisms is based on the theory and models of sound propagation. Figure 4 illustrates the different sound propagation paths in deep water. Two principal mechanisms of multipath formation are reflections from the boundaries (bottom, surface and any objects in the water) and ray bending. In shallow water propagation largely occur in surface-bottom bounces in addition to a *direct path* and reflections from other boundaries, if any. Trapping of sound rays in the *mixed layer* through repeated surface reflection and upward refraction due to the slightly positive sound speed gradient creates the so-called *surface duct*. In deep water, there is a region in the water column where sound speed first decreases with depth to a minimum value and then increases with depth. At that depth sound is trapped in the *SOFAR (SOund Fixing And Ranging) channel* waveguide and can travel extreme distances without bouncing on the bottom or the surface. For steeper transmission angles sound is not ducted in the sound channel but is refracted downward and then back up again to the surface focusing in one region. Typical length of the so-formed *convergence zone* is 60-100 km.

Multipath propagation gives rise to extended temporal signal spreading in the ocean. The SOFAR channel has characteristic multipath spread from tens of milliseconds to several seconds. A commonly encountered multipath spread of 10 msec in medium-range shallow water channels causes the ISI to extend over 100 symbols if 10 kHz is the transmitted signal bandwidth. At short ranges, reflections from objects and channel boundaries dominate the

multipath; the problem becomes geometric specific and no generic solutions are available [20].
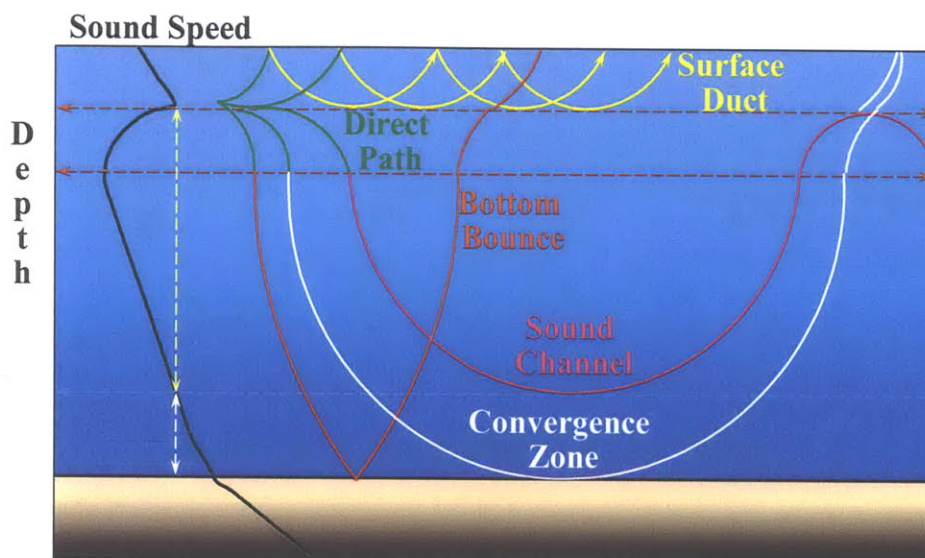

Figure 4: Sound propagation paths in deep water.

2.2.3 Temporal variability

A time-varying multipath channel is characterized by the time spread T_m and the Doppler spread B_d. T_m is the time over which most of the channel impulse response dies off, and B_d is the increase in the frequency content of a received signal over that which was transmitted. The channel coherence time is that time over which the impulse response of the channel remains essentially constant. A simple approximation to the channel coherence time is

$T_{coh} \approx 1/B_d$. The channel spread factor is defined as the product $T_m \cdot B_d$. These values must be compared with the corresponding signal parameters, T the symbol duration and $W \approx 1/T$ the signal bandwidth. Channels which do not introduce time dispersion are those for which $T \gg T_m$. Channels which are slowly varying are those for which $B_d \ll W$. In a mobile communication system relative motion between the transmitter and the receiver introduces additional Doppler spreading and shifting.

The time and frequency dispersion of the UWA channel affects the selection of signaling rate. Signaling at a low rate eliminates ISI but allows for high channel variation from one symbol interval to another. Signaling at a high rate results in ISI but also results in slower channel variation over the symbol duration. This allows an adaptive receiver to efficiently track the channel on a symbol to symbol basis. Thus, there is a trade-off in the choice of the signaling rate for a specific channel.

The trade-off in the choice of the signaling rate may be illustrated in the following case. In shallow water, the dominant mechanism leading to Doppler spreading is the surface scattering caused by waves. In [24], the estimated Doppler spread for telemetry systems is expressed as

$$B_w = 2f_w \left[1 + \frac{4\pi f_0 \cos\theta_0}{c} h_w \right] \tag{8}$$

where w is the wind speed, f_w is the wave frequency, f_0 is the carrier frequency, θ_0 is the incident grazing angle, c is speed of sound and h_w is the r.m.s. wave height. Doppler spread can be estimated using the following empirical formulas which relate the frequency and the wave height to the wind speed:

$$f_w = 2/w$$
$$h_w = 0.005w^{2.5}$$
(9)

For example, consider the case where $B_w=10$ Hz and the signal bandwidth is $1/T=1$ kHz. Then, the normalized Doppler spread is $B_w \cdot T=10^{-2}$ which approaches the limiting value for reliable coherent detection [3]. However, increasing the bandwidth reduces $B_w \cdot T$; thus, phase coherent detection is assured, provided a method for dealing with the resulting ISI.

2.2.4 Ambient noise

Ambient noise strongly affects the received SNR. The power spectral density of ambient noise usually decays by 20 dB/octave, thus it dictates the minimum frequency limit of the transmitted signal bandwidth. Ambient noise level depends on the site of operation, as well as on the presence of equipment, people, and marine organisms. Air bubbles generated by waves or rain are the main cause of ambient noise in the 10-20 kHz band [20].

2.2.5 Sound Attenuation

Attenuation of the received signal level is affected by range, frequency and channel waveguide properties. Having the range fixed, the rule of thumb to find the carrier frequency of transmission is to use the following equation: **absorption rate × range = 10 dB** [25]. In Figure 5, the absorption rate (dB/km) versus frequency in seawater is illustrated. Clearly the

absorption rate increases with frequency; hence it dictates the maximum frequency limit of the transmitted signal bandwidth.
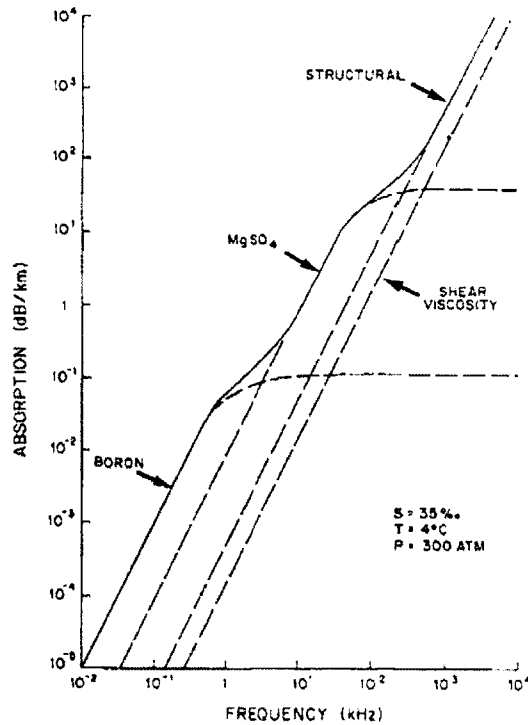

Figure 5: Absorption rate of acoustic energy in seawater [25].

Moreover, transmission loss depends on the channel waveguide properties. There are *shadow zones*, i.e. regions where data telemetry is infeasible due to the excessive transmission loss induced by energy refraction. Figure 6 shows the shadow zone caused by the splitting of the limiting ray (defines the deepest ray that remains in the *mixed layer* and the shallowest that escapes to the *thermocline*) at the sonic layer depth (SLD). All rays transmitted at smaller angles remain in the *mixed layer* and all rays transmitted at greater angles escape to the

35

thermocline. The depth at which the limiting ray reaches vertex speed is defined as the SLD. Theoretically, no sound should enter the shadow zone, but some does due to scattering and diffraction.
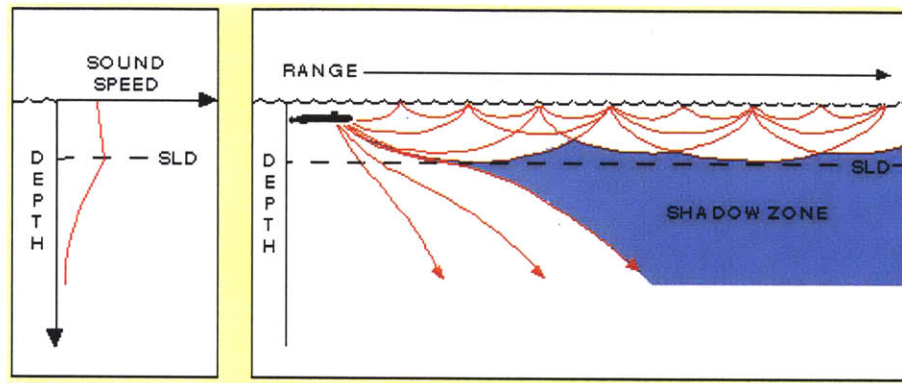

Figure 6: Shadow zone due to surface ducting

2.3 Bandwidth-efficient vs. power-efficient modulation methods

The choice of a signaling method in a communication system depends on the available bandwidth and SNR. A meaningful comparison of modulation methods is based on the normalized data rate (or bandwidth efficiency) R_b/W, which is measured in bits per second per Hertz of occupied bandwidth. In a typical communication system the data arrives at a rate of R_b bits/sec and is converted into a sequence of symbols chosen from the symbol constellation of size M. Thus, each symbol carries $log_2 M$ of information. The symbol rate is $R=R_b/log_2 M$

36

symbols/sec and the symbol interval is $T=1/R$. The required bandwidth is $W \approx 1/T$. The bandwidth efficiency for M-ary PSK or M-ary QAM modulation is $R_b/W = log_2 M$. Note that if a higher-level symbol constellation is used, for fixed bandwidth, the bit rate is increased.

Orthogonal signaling, e.g. FSK, behaves differently as bandwidth increases. If the M orthogonal signals are constructed by means of orthogonal carriers with minimum frequency separation $1/2T$ for orthogonality ($1/T$ for noncoherent detection), the bandwidth required for transmission is

$$W = M\frac{1}{2T} = \frac{M}{2\log_2 M}R_b \qquad (10)$$

so as M increases the bandwidth-efficiency decreases.

Figure 7 shows the graph of the normalized data rate R_b/W (bits/sec/Hz) versus SNR per bit for PAM, PSK and QAM and orthogonal signals for the AWGN channel and for the case in which the error probability is 10^{-5}. Note that in the case of PAM, PSK and QAM, R_b/W increases for increasing M at the cost of higher SNR per bit. Consequently, these modulation methods are appropriate for bandwidth-limited systems when $R_b/W>1$ is desired, provided that there is sufficient SNR to accommodate greater M. In contrast, M-ary orthogonal signals are appropriate for systems operating when $R_b/W<1$. These signals are efficient for power-limited channels (low SNR regimes) provided that there is enough bandwidth to accommodate the large number of signals. As $M\to\infty$, the error probability can be made as small as desired provided that that the SNR/bit>-1.6 dB. This is the result of Shannon's capacity theorem [26]. Figure 7 illustrates the normalized channel capacity of the bandlimited AWGN channel which is expressed as

$$\frac{C}{W} = \log_2\left(1 + \frac{C}{W}\frac{E_b}{N_0}\right)$$ (11)

where E_b is the energy per bit and N_o is the noise power spectral density. The ratio C/W

represents the highest achievable bit rate to bandwidth ratio. Therefore, it serves as the upper

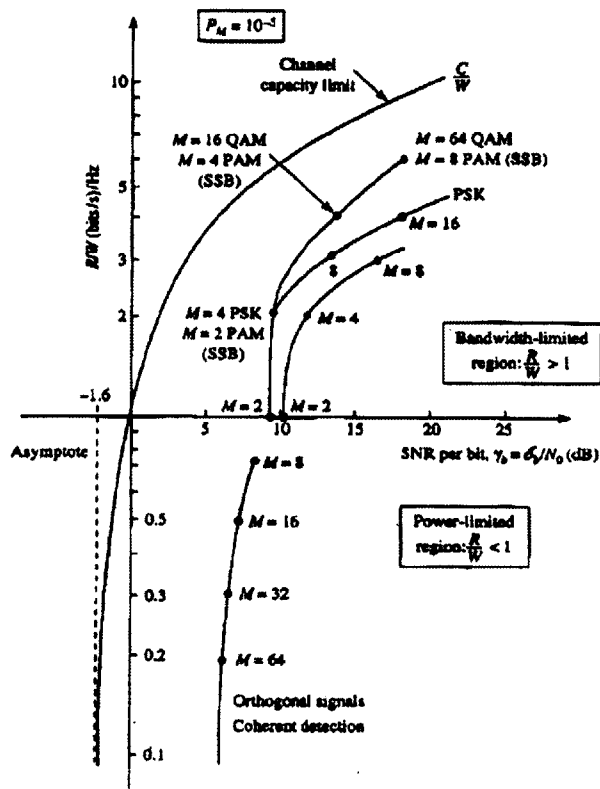bound on the bandwidth efficiency of any type of modulation.

Figure 7: Bandwidth efficiency as a function of SNR/bit (for constant P_e=10-5) [26].

38

2.4 Joint adaptive equalization and phase estimation for coherent detection

The vertical underwater channel is benign in comparison to the horizontal channel, primarily due to the lack of boundary interaction. While vertical channels exhibit little time dispersion, horizontal channels suffer from extended multipath propagation that may cause the ISI to span up to several tens of symbol intervals. Moreover, the received signal undergoes carrier offset due to the relative motion of the source and receiver. In addition, the channel impulse response varies with time, causing amplitude and phase distortion.

A solution to the ISI problem is to design a receiver that employs an equalizer which compensates or at least reduces ISI in the received signal. The optimum, from the probability of error viewpoint equalizer is based on the maximum-likelihood (ML) sequence detection [26]. The major shortcoming of this structure is its complexity, which grows exponentially with the length of the delay dispersion, making it impractical for high symbol rates when the channel response spans more than a few symbol intervals.

A suboptimum type of channel equalizer which circumvents the complexity problems is an adaptive decision feedback equalizer (DFE). Conventional systems perform carrier phase recovery by using a PLL structure and channel equalization separately. However, this approach may not be well suited for a rapidly changing medium because residual phase fluctuations impair the performance of the equalizer and may cause the problem of equalizer tap rotation. Therefore, joint carrier phase estimation and adaptive equalization has been developed and shown to be an effective solution for many underwater channels [5]. The receiver structure is shown in Figure 8.
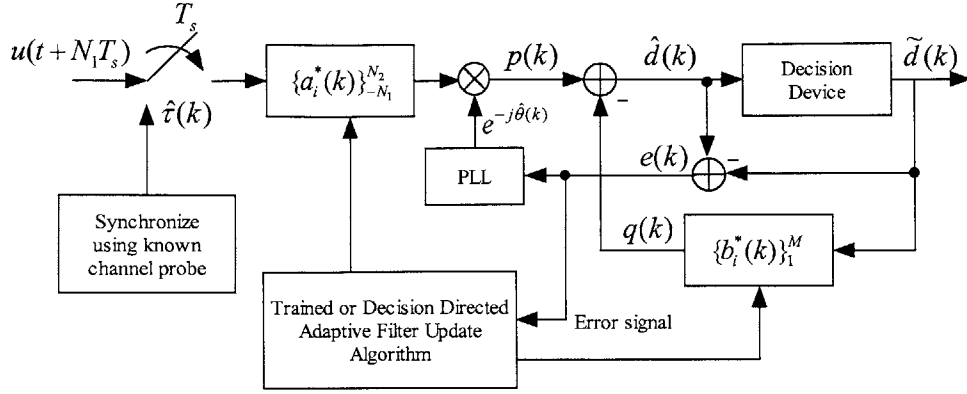
Figure 8: Receiver structure for joint equalization and synchronization based on [5].

2.4.1 Receiver algorithm

The received signal, after being brought to baseband and lowpass filtered, is frame synchronized by matched filtering to a known channel probe. Then, the received signal can be modeled as

$$u(t) = \sum_k d(k)h(t - kT - \tau)e^{j\theta(t)} + w(t) \qquad (12)$$

where $\{d(k)\}$ are the transmitted symbols taken from an M-ary constellation, $1/T$ is the symbol rate, $h(t)$ includes the physical channel and any transmit and receive filtering and the delay uncertainty τ is within one symbol interval after coarse synchronization in time. The carrier phase $\theta(t)$ can be modeled as a sum of three terms: constant phase offset, variable Doppler shift and random phase jitter. $w(t)$ is white Gaussian noise.

Past and future symbols cause the ISI, so a delay of N_1 sampling intervals is introduced at the receiver. The $N=N_1+N_2+1$ tap weights of the feedforward linear transversal filter are

arranged in a row vector $\boldsymbol{a}$ and the input vector at the n-th time instance stored in the feedforward filter is given as

$$u(k) = [u(kT + N_1 T_s + \tau)...u(kT - N_2 T_s + \tau)]^T .$$ (13)

where $T_s = T/2$ (Nyquist rate). The output of the feedforward filter is produced once per symbol interval yielding:

$$p(k) = a(k)^H u(k) e^{-j\hat{\theta}(k)} .$$ (14)

where ^{H} denotes conjugate transpose. The feedback filter has tap weights $\boldsymbol{b}$ and operates on the sequence of M previously detected symbols collected in the vector

$$\tilde{\boldsymbol{d}}(k) = [\tilde{d}_{k-1}...\tilde{d}_{k-M}]^T .$$ (15)

The output of the feedback filter which estimates the residual ISI is expressed as

$$q(k) = b(k)^H \tilde{d}(k)$$ (16)

This estimate is subtracted from the output of the feedforward filter to produce the overall estimate of the data symbol $d(k)$ which is:

$$\hat{d}(k) = p(k) - q(k)$$ (17)

By quantizing the estimate $\hat{d}(k)$ to the nearest symbol value, the decision $\tilde{d}(k)$ is formed.

In the MSE criterion, the receiver parameters are adjusted to minimize the mean square value of the error:

$$e(k) = d(k) - \hat{d}(k) .$$ (18)

In the decision-directed mode, $d(k)$ should be replaced by $\tilde{d}(k)$. The equations that give the gradient of the MSE with respect to all the receiver parameters are given in [5] and are reproduced below for convenience:

$$\frac{\partial E(|e(k)|^2)}{\partial a} = -2E\left(\boldsymbol{u}(k)e^*(k)\right)e^{-j\hat{\theta}}$$

$$\frac{\partial E(|e(k)|^2)}{\partial \boldsymbol{b}} = 2E\left(\tilde{\boldsymbol{d}}(k)e^*(k)\right)$$

$$\frac{\partial E(|e(k)|^2)}{\partial \hat{\theta}} = -2\,\mathrm{Im}\left\{E\left(p(k)(d(k)+q(k))^*\right)\right\}$$ (19)

$$\frac{\partial E(|e(k)|^2)}{\partial \hat{\tau}} = -2\,\mathrm{Re}\left\{E\left(\dot{p}(k)e^*(k)\right)\right\}$$

where

$$\dot{p}(k) = \boldsymbol{a}(k)^H \dot{\boldsymbol{u}}(k)e^{-j\hat{\theta}(k)} \quad,$$ (20)

and $\dot{\boldsymbol{u}}(k)$ is the time derivative of the received signal. Setting the gradients equal to zero

results in the set of equations whose solution represents the jointly optimal solution of channel

equalization and phase synchronization. Since the optimal values of the receiver parameters

are time varying we seek to obtain a solution to the system of equations in a recursive manner.

To obtain the necessary tracking of the carrier phase estimate, a second order update equation

is used based on the analogy of the digital PLL but using the expression of the gradient of the

MSE with respect to the carrier phase in (19). The update equations for the carrier phase

estimate are:

$$\hat{\theta}(k+1) = \hat{\theta}(k) + K_1\Phi(k) + K_2\sum_{i=0}^{k}\Phi(i)$$

$$\Phi(k) = \mathrm{Im}\left\{p(k)(d(k)+q(k))^*\right\}$$ (21)

where K_1 and K_2 are the proportional and integral tracking constants. A good rule of thumb is

$K_2=K_1/10$. Two commonly used adaptive algorithms that produce the vector coefficients

update equations are the least mean squares (LMS) algorithm and the recursive least squares

(RLS) algorithm.

Update equations for the LMS algorithm

The update equations for the LMS algorithm are derived in [27]:

$$c(k+1) = c(k) + \mu \cdot e^*(k)U(k), \quad k = 0,1,\ldots \quad (22)$$

where

$$c(k) = \begin{bmatrix} a(k) \\ -b(k) \end{bmatrix} \quad (23)$$

is the composite weight vector, $e(k)$ is given in (18) and

$$U(k) = \begin{bmatrix} u(k)e^{-j\hat{\theta}(k)} \\ d(k) \end{bmatrix} \quad (24)$$

is the composite data vector. μ is a step size parameter of the LMS algorithm. The initial convergence rate is strongly influenced by the channel spectral characteristics, which are related to the eigenvalues of the received signal covariance matrix. If the eigenvalue ratio $\lambda_{max}/\lambda_{min}$ (condition number) is close to unity the equalizer converges to its optimum tap coefficients relatively fast. On the other hand, if the channel exhibits poor spectral characteristics, the eigenvalue ratio $\lambda_{max}/\lambda_{min} \gg 1$ and hence the convergence rate of the LMS algorithm will be slow. A good rule of thumb to determine μ is by the following equation [26]:

$$\mu = \frac{0.5}{(N_1 + N_2 + 1 + M)P_R} \quad (25)$$

where P_R denotes the received signal-plus-noise power, which can be estimated from the received signal.

43

Update equations for the RLS algorithm

The price paid for the simplicity of the LMS algorithm is the slow convergence. Besides allowing for shorter training periods, fast convergence is desired on rapidly changing channels since it enables the receiver to make full use of only temporarily present multipath components. The update equations for the RLS algorithm are derived in [27]:

$$K(k) = \frac{\lambda^{-1} \boldsymbol{R}^{-1}(k-1)\boldsymbol{U}(k)}{1 + \lambda^{-1}\boldsymbol{U}^{H}(k)\boldsymbol{R}^{-1}(k-1)\boldsymbol{U}(k)}$$

$$a(k) = d(k) - \boldsymbol{c}^{H}(k-1)\boldsymbol{U}(k) \qquad\qquad k = 0,1,... \qquad (26)$$

$$c(k) = c(k-1) + K(k)a^{*}(k)$$

$$\boldsymbol{R}^{-1}(k) = \lambda^{-1}\boldsymbol{R}^{-1}(k-1) - \lambda^{-1}K(k)\boldsymbol{U}^{H}(k)\boldsymbol{R}^{-1}(k-1)$$

where *c(k)* and *U(k)* are given in (23) and (24) respectively and the correlation matrix *R(k)* is defined as

$$R(k) = \sum_{i=1}^{k} \lambda^{k-i} U(i)U^{H}(i) \qquad\qquad (27)$$

The applicability of the RLS algorithm requires the initialization of the recursion by choosing $\boldsymbol{R}^{-1}(0)=\delta\cdot\boldsymbol{I}$ (δ is a small positive constant) that assures the nonsingularity of the correlation matrix. The use of the "forgetting factor" λ is intended to ensure that data in the distant past are "forgotten" in order to accommodate the time variations of the channel. λ is a positive constant close to, but less than 1.

An adaptive algorithm has two sources of misadjustment error. The first one is the estimation noise of correlation matrix that arises from computing estimates based on windowed data sequences. The second one is lag noise caused by the reaction time of the adaptation algorithm to changes in its environment. Lag error is minimized by rapidly

44

discounting the past observations and computing estimates predominantly based on recent data. These two errors place conflicting demands on the adaptation algorithm. The estimation noise is reduced by choosing λ close to 1; however, in the nonstationary situation, the lag error is minimized by choosing a small forgetting factor. The optimization of the combination of these two errors requires a trade-off in the choice of λ.

2.5 Differential Amplitude Phase Shift Keying (DAPSK) modulation and detection

In this section, we discuss a multilevel differential modulation technique, namely the DAPSK technique. The reason for considering this technique is that differential modulation techniques appear robust against carrier-phase variations of the received signal without requiring any carrier recovery. Moreover, a differential modulation technique doesn't require any explicit knowledge about the channel properties in the differential demodulation process if flat fading is assumed. It requires, however, that the channel remains relatively constant over two consecutive symbol intervals [28].

In its generalized form, the transmitter uses a star constellation with N concentric amplitude rings, each containing M phasors and sequentially encodes information onto the *amplitude ratio* and *phase difference* between current and previously transmitted symbols. We refer to this scheme as *NxM* DAPSK. The transmitted bit stream is first grouped into a sequence of real-valued pairs $\{(\Delta\beta_i, \Delta\gamma_i), i = ...,0, 1, 2, ...\}$ in which $\Delta\beta_i \in \{0, 1, ..., N-1\}$ and $\Delta\gamma_i \in \{0, 1, ..., M-1\}$. All $N \cdot M$ possible pairs are assumed equally likely and Gray coding is used for binary

representations of $\Delta\beta_i$ and $\Delta\gamma_i$. The pairs $(\Delta\beta_i, \Delta\gamma_i)$ are then used to determine the *amplitude ratio* and *phase difference* between current and previous transmitted symbols. With $\alpha > 1$, the ith transmitted complex-valued symbol s_i is given by

$$s_i = \lambda a^{\beta_i} e^{j(2\pi/M)\gamma_i} \tag{28}$$

where

$$\lambda = \sqrt{N(\alpha^2 - 1)/(\alpha^{2N} - 1)} \tag{29}$$

is used to normalize the signal constellation to unit energy. In equation (28), β_i and γ_i indicate the amplitude and phase levels of s_i respectively, and are related to $(\Delta\beta_i, \Delta\gamma_i)$ by

$$\beta_i = (\beta_{i-1} + \Delta\beta_i) \bmod N$$

$$\gamma_i = (\gamma_{i-1} + \Delta\gamma_i) \bmod M \tag{30}$$

Note that the *NxM* DAPSK with $N=1$ corresponds to the *M-ary* DPSK modulation.

Assuming a flat fading situation, the effect of the channel is analytically described by a complex multiplication:

$$r_i = s_i h_i \tag{31}$$

where r_i is the received i-th data symbol. For a QAM modulation/demodulation technique, an estimation of the channel transfer factor h_i would be necessary; however in the differential demodulation process, the receiver forms the ratio between r_i and r_{i-1}:

$$\hat{d}_i = \frac{r_i}{r_{i-1}} = \frac{s_i h_i}{s_{i-1} h_{i-1}} = \Delta s_i \frac{h_i}{h_{i-1}} \tag{32}$$

where

$$\Delta s_i = a^{(\beta_i - \beta_{i-1})} e^{j(2\pi/M)(\gamma_i - \gamma_{i-1})} \tag{33}$$

If we assume that the channel changes negligibly over two consecutive symbols, i.e., $h_i \approx h_{i-1}$, no channel estimation is necessary, thus reducing the complexity of the receiver. The differential demodulation (32) yields $\hat{d}_i \approx \Delta s_i$. The phase and amplitude reference in the demodulation process is known by the previously received state r_{i-1}; therefore, we can decode the transmitted bits by observing Δs_i.

As an example, consider the 4x16DAPSK with four different amplitude levels and 16 different phase states. The symbol constellation is shown in Figure 9 below. It consists of four concentric circles with increasing amplitude values λ, λa, λa^2, λa^3. On each circle there are 16 different phase states. The pair $(\Delta\beta_i, \Delta\gamma_i)$ is calculated from the 6-bit binary input sequence $(c_1, c_2, ..., c_6)$. The information of the first four bits $(c_1, ..., c_4)$ is mapped to one of the $2^4 = 16$ possible phase states positioned on a single ring. Thus, $\Delta\gamma_i \in \{0, 1, ..., 15\}$ and adjacent phase states in the constellation diagram have a phase difference of $2\pi/16$. The information remaining in the two bits c_5 and c_6 is used for the transition among the possible amplitude levels in the constellation. Since there are $2^2 = 4$ possible amplitude states, then $\Delta\beta_i \in \{0, 1, 2, 3\}$.

In order to construct the symbol constellation shown in Figure 9 the transition between the different phase and amplitude states is calculated from (30) where $N=4$ and $M=16$. Although the phase transitions are directly applied as in 16DPSK, the amplitude transitions among the four rings must be applied in a cyclic manner. From (33), observe that $(\beta_i - \beta_{i-1}) \in \{-3, -2, -1, 0, 1, 2, 3\}$ so seven different amplitude ratios can occur related to seven possible transitions among the four different amplitude rings in the constellation diagram. Table 2 illustrates the

procedure. The amplitude $|\Delta s_i|$ is calculated by the two bits c_5 and c_6 and the amplitude $|s_{i-1}|$ of the previously transmitted symbol.
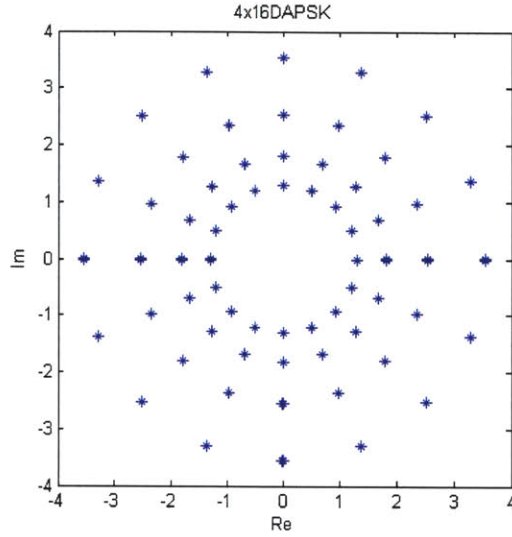

Figure 9: Constellation diagram of the 4x16DAPSK modulation scheme.

$|\Delta s_i|$	Information bits c_6 and c_5			
	00	01	11	10
$|s_{i-1}| = \lambda$	1	α	α^2	α^3
$|s_{i-1}| = \lambda\alpha$	1	α	α^2	$1/\alpha$
$|s_{i-1}| = \lambda\alpha^2$	1	α	$1/\alpha^2$	$1/\alpha$
$|s_{i-1}| = \lambda\alpha^3$	1	$1/\alpha^3$	$1/\alpha^2$	$1/\alpha$

Table 2: Generation of the amplitude $|\Delta s_i|$ given $|s_{i-1}|$ and the information bits c_5 and c_6.

The amplitude factor a between two adjacent amplitude circles is a system parameter. In [28], the amplitude factor has been optimized for the AWGN channel in order to minimize the bit error rate and has been calculated as $a=1.4$. The amplitude factor characterizes the distance between the amplitude rings and additionally influences the average power of all complex states on the constellation. Note that the two properties have opposite effect on the resulting bit error rate.

In the differential demodulation process the binary information is obtained from the complex state $\hat{d}_i$ (see equation (33)). For the phase decision, only the phase of $\hat{d}_i$ is of interest. Thus, the complex plane is divided into 16 sectors corresponding to the 16 possible values ($2\pi m/16$, $m=0$, ..., 15) of the phase of Δs_i and the estimated phase is quantized into the closest possible value. The two bits c_5 and c_6 are recovered from the magnitude of $\hat{d}_i$. The decision rule is the following:

if $s_{i-1} = 1$ then:

$$
|\Delta s_i| = \begin{cases}
1, & \text{if } |d_i| \le (1+a)/2 \\
a, & \text{if } |d_i| > (1+a)/2 \quad \text{and } |\Delta s_i| \le (a+a^2)/2 \\
a^2, & \text{if } |d_i| > (a+a^2)/2 \quad \text{and } |\Delta s_i| \le (a^2+a^3)/2 \\
a^3, & \text{if } |d_i| > (a^2+a^3)/2
\end{cases}
$$

if $s_{i-1} = a$ then:

$$
|\Delta s_i| = \begin{cases}
a^{-1}, & \text{if } |d_i| \le (a^{-1}+1)/2 \\
1, & \text{if } |d_i| > (a^{-1}+1)/2 \quad \text{and } |\Delta s_i| \le (1+a)/2 \\
a, & \text{if } |d_i| > (1+a)/2 \quad \text{and } |\Delta s_i| \le (a+a^2)/2 \\
a^2, & \text{if } |d_i| > (a+a^2)/2
\end{cases}
$$

if $s_{i-1} = a^2$ then:

$$|\Delta s_i| = \begin{cases} a^{-2}, & \text{if } |d_i| \le (a^{-2}+a^{-1})/2 \\ a^{-1}, & \text{if } |d_i| \succ (a^{-2}+a^{-1})/2 \quad \text{and } |\Delta s_i| \le (a^{-1}+1)/2 \\ 1, & \text{if } |d_i| \succ (a^{-1}+1)/2 \quad \text{and } |\Delta s_i| \le (1+a)/2 \\ a, & \text{if } |d_i| \succ (1+a)/2 \end{cases}$$

(34)

if $s_{i-1} = a^3$ then:

$$|\Delta s_i| = \begin{cases} a^{-3}, & \text{if } |d_i| \le (a^{-3}+a^{-2})/2 \\ a^{-2}, & \text{if } |d_i| \succ (a^{-3}+a^{-2})/2 \quad \text{and } |\Delta s_i| \le (a^{-2}+a^{-1})/2 \\ a^{-1}, & \text{if } |d_i| \succ (a^{-2}+a^{-1})/2 \quad \text{and } |\Delta s_i| \le (a^{-1}+1)/2 \\ 1, & \text{if } |d_i| \succ (a^{-1}+1)/2 \end{cases}$$

2.6 Adaptive equalization for differentially coherent detection

DFE is an efficient, practical approach for the transmission of high speed data over slowly varying, fading multipath channels. For slow channel variations, the tap weights of a conventional adaptive equalizer remove the phase difference between the transmitter and receiver clocks and the carrier phase doesn't have to be recovered. For fast channel variations, however, the conventional DFE cannot compensate for carrier phase. One approach to solve this problem is to incorporate a PLL at the receiver which is a complex process involving many ancillary functions associated with the carrier synchronization loop. An easier way to mitigate carrier phase variations is to employ noncoherent signaling such as DAPSK and a corresponding receiver based on differentially coherent detection. This choice reduces the computational complexity at the receiver at the expense of higher probability of error.

50

The combination of a linear equalizer and a decision-aided differential detector that works for both N-DPSK and NxM-DAPSK has been proposed in [11]. Using the equalizer output signal, a reference symbol is generated recursively or nonrecursively. For adaptation of the equalizer coefficients a modified LMS algorithm is implemented. Figure 10 illustrates the receiver structure employed for differentially coherent detection [11].
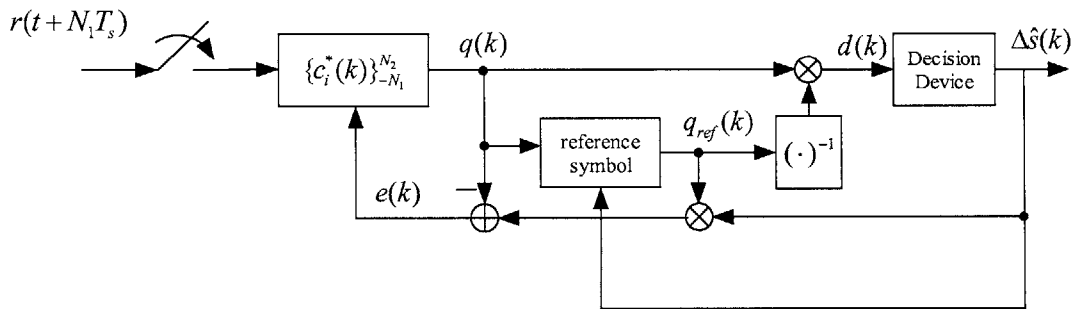

Figure 10: Block diagram of the differentially coherent receiver based on [11]

2.6.1 Receiver algorithm

In this section, the differentially coherent equalizer of Schober et al [11] is described. The DAPSK symbols $s(k)$ (see (28)) are transmitted over the UWA channel and the baseband equivalent received signal at the input of the fractionally spaced equalizer is modeled as

$$u(t) = \sum_k s(k)h(t - kT)e^{j\theta(t)} + w(t) \qquad (35)$$

where T is the symbol interval, $h(t)$ is the combined impulse response of the cascade of transmit filter, channel, and receiver input filter. The unknown carrier phase is denoted as $\theta(t)$. $w(t)$ is white Gaussian noise.

Past and future symbols cause ISI so a delay of N_1 sampling intervals is introduced at the receiver. The $N = N_1 + N_2 + 1$ tap weights of the linear transversal filter are arranged in a row vector c and the input vector at the n-th time instance stored in the equalizer is given as

$$u(k) = [u(kT + N_1 T_s) \ldots u(kT - N_2 T_s)]^T \qquad (36)$$

where $T_s = T/2$. It is assumed that proper time alignment has been achieved so that this signal contains significant contributions from the data symbol $s(k)$. The output of the filter is produced once per symbol interval yielding:

$$q(k) = c(k)^H u(k). \qquad (37)$$

The next stage of the receiver is a decision-aided differential detector which determines the estimate $\Delta \hat{s}(k)$ for the transmitted symbol $\Delta s(k) = s(k)/s(k-1)$. To do so, a reference symbol is generated recursively as

$$q_{ref}(k-1) = \frac{q(k-1) + W \dfrac{q_{ref}(k-2)}{\Delta \hat{s}^*(k-1)}}{1 + \dfrac{W}{|\Delta \hat{s}(k-1)|^2}} \qquad (38)$$

where W, $W \geq 0$ is a design parameter. Alternatively, the reference symbol may be generated nonrecursively as:

$$q_{ref}(k-1) = \frac{\sum_{v=1}^{m-1} q(k-v) \prod_{\mu=1}^{v-1} \frac{1}{\Delta \hat{s}^*(k-\mu)}}{\sum_{v=1}^{m-1} \prod_{\mu=1}^{v-1} \frac{1}{|\Delta \hat{s}(k-\mu)|^2}} \qquad (39)$$

where $m \geq 2$ is the number of equalizer output symbols used to determine $\Delta\hat{s}(k)$. Note that for $m=2$, $q_{ref}(k-1)=q(k-1)$, i.e., the previous equalizer output symbol is employed to provide a phase and amplitude reference for the current equalizer output $q(k)$. This reference is used for conventional differential detection [7]. Since $q(k-1)$ is a noisy reference, the performance of the resulting noncoherent receiver is worse than that of a coherent receiver that assumes perfect knowledge of the amplitude and phase reference. In order to alleviate this problem, $m \geq 2$ equalizer output symbols may be used to improve the phase and amplitude reference. The decision variable for estimation of $\Delta\hat{s}(k)$ is given by

$$d(k) = \frac{q(k)}{q_{ref}(k-1)} \qquad (40)$$

and the decision rules are described in Section 2.5.

The equalizer minimizes the variance of the post-detection error:

$$e(k) = \Delta\hat{s}(k)q_{ref}(k-1) - q(k) \; . \qquad (41)$$

This error is used to determine the equalizer vector c according to the MMSE criterion. However, an unconstrained minimization of $E\{|e(k)|^2\}$ yields the all zero vector $c=0$. Therefore a unity energy constraint is introduced, i.e., $c^H c=1$. An efficient solution can be obtained using the stochastic gradient approach. The adaptive gradient algorithm is obtained from

$$\hat{c}(k+1) = c(k) - \mu \frac{\partial}{\partial c^*(k)} |e(k)|^2 \ . \tag{42}$$

and

$$c(k+1) = \frac{\hat{c}(k+1)}{\left\| \hat{c}(k+1) \right\|} \tag{43}$$

where μ is the adaptation step size parameter and is equal to the step size parameter of the conventional LMS algorithm. Equation (43) ensures that the unity energy constraint is satisfied. Note that $q_{ref}(k-1)$ depends only on $c(k-v)$, $v \geq 0$ but not on $c(k)$. Therefore, it has to be treated like a constant when differentiating with respect to $c(k)$. Hence, (42) is rewritten to

$$\hat{c}(k+1) = c(k) + \mu \cdot e^*(k) u(k) \tag{44}$$

Note that $e^*(k)r(k)$ is not a function of θ therefore the resulting modified LMS algorithm is noncoherent.

Chapter 3

3 Experimental description

3.1 Experiment setup

The experiment took place at the Woods Hole Oceanographic Institution (WHOI) dock. Two electroacoustic transducers were used. One acted as the sound projector and the other as the hydrophone. The transducers have a 6° conical beampattern providing approximately 25 kHz of usable bandwidth centered at 75 kHz. Figure 11 illustrates the two RD Instruments Long Ranger transducers used for the experiment.

Figure 12 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. Also at the lower end, a ROS 20/20 underwater camera was attached providing the input video frames to the system for transmission. The gray-scale video had the following specifications: 15 frames/sec, 144x176 pixels/frame and 8 bits/pixel.

Figure 11: Transducers with pressure housing.
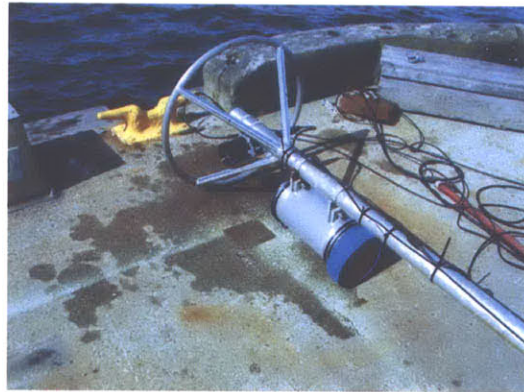
Figure 12: The transducers mounted on a 10m pole.

All the signal processing was implemented in Matlab®, running on two laptop computers. Both of the laptops were connected with the transducers through 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 laptop connected with the sound projector had also an Imperx VCE-B5A01 frame grabber which was used to acquire the frames from the camera and then to deliver them to the routine responsible for data compression. Figure 13 shows the experimental set-up and illustrates the hardware interconnections.
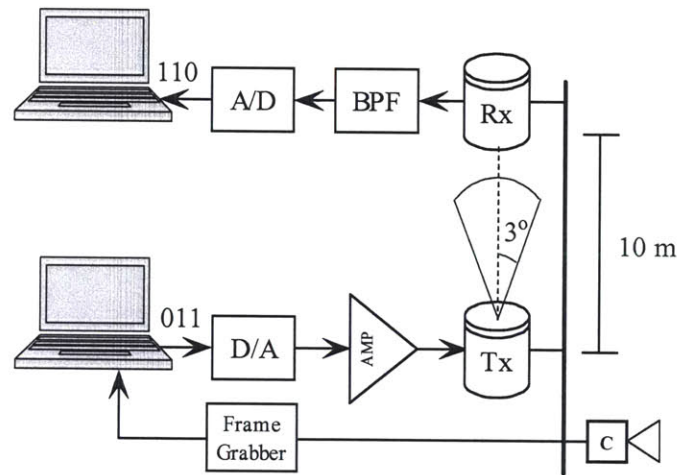
Figure 13: Experiment setup.

3.2 System description

Transmission was organized in packets of fixed duration. Each packet contained 4158 data symbols, and an additional block of training data. 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 which was a 28-bit Barker code 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. Figure 14 illustrates how the symbols are organized into a packet. Sequential packet transmission was employed so as to provide periodic frame synchronization and retraining for the equalizer. The transmission rate was 25000 symbols/sec for all the modulation methods considered. For all the data packets, average transmission power per bit was equal. The passband sampling rate was 200 kHz (8 samples/symbol).

probe signal	50 ms	training data 500 symbols	data 4158 symbols	50 ms
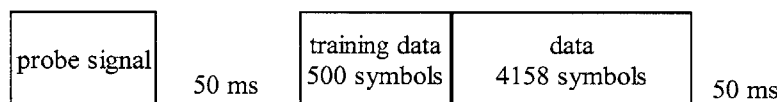
Figure 14: Signal framing.

The system block diagram is shown in Figure 15. At the input to the system are 8-bit gray-scale 144x176-pixel video frames, which are first compressed using a selected method. In the current implementation, frames in the video sequence are compressed individually, by applying the JPEG algorithm. The resulting bit stream is mapped into the symbols of desired signal constellation. The following have been implemented:

1. 8PSK, 16QAM, 32QAM and 64QAM for use with coherent detection

2. 4x4DAPSK, 2x8DAPSK, 4x8DAPSK, 2x16DAPSK and 4x16DAPSK for use with
    differentially coherent detection

The signal constellations for the first and second modulation methods are illustrated in Figure 16 and Figure 17 respectively. After addition of training data and organizing transmission in packets, transmitter filtering is performed using a 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 A/D conversion is shifted to baseband, low-pass filtered and down-sampled to 2 samples/symbol. Packet coarse synchronization is achieved by matched-filtering to the 28-bit Barker sequence. Adaptive filtering by a $T/2$ fractionally-spaced equalizer is used. For coherent detection, joint decision feedback equalization and phase tracking is employed. The adaptive equalization algorithm is a combination of RLS or LMS and a second order PLL. For differentially coherent detection, linear equalization is implemented using a modified LMS algorithm. The detected data symbols are finally converted to bits and passed on to the video decoder.
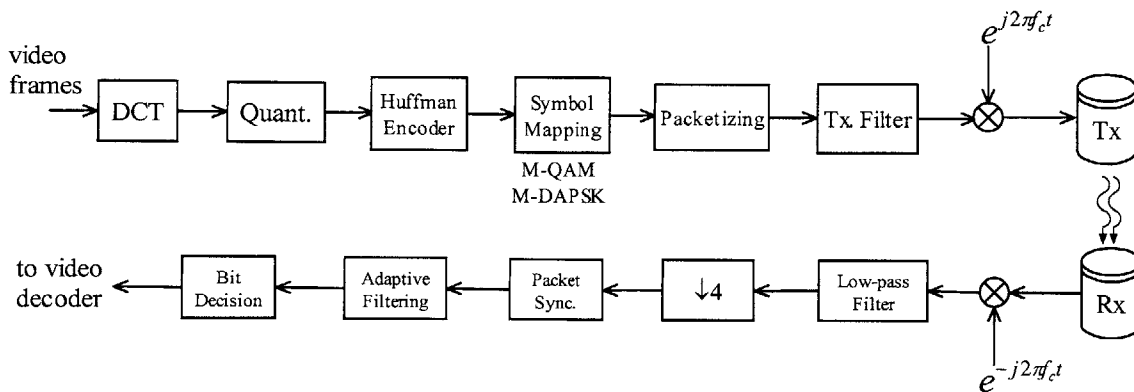
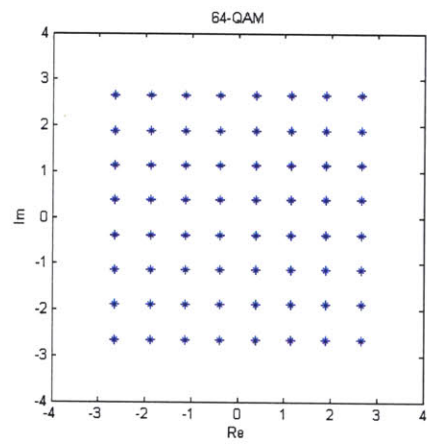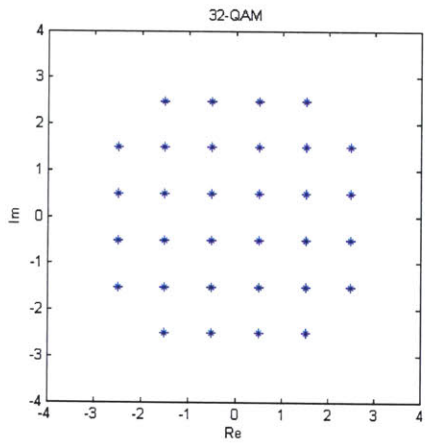Figure 15: Block diagram of the transmitter and the receiver.
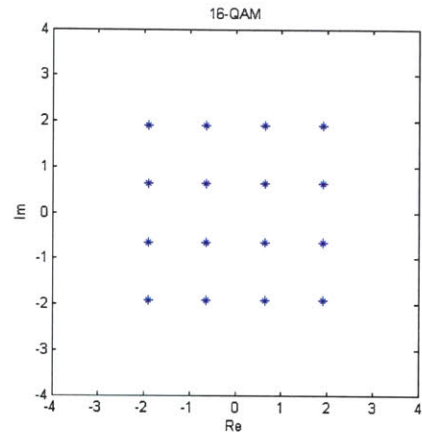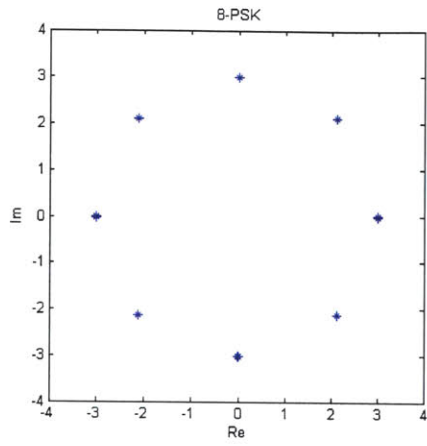
59

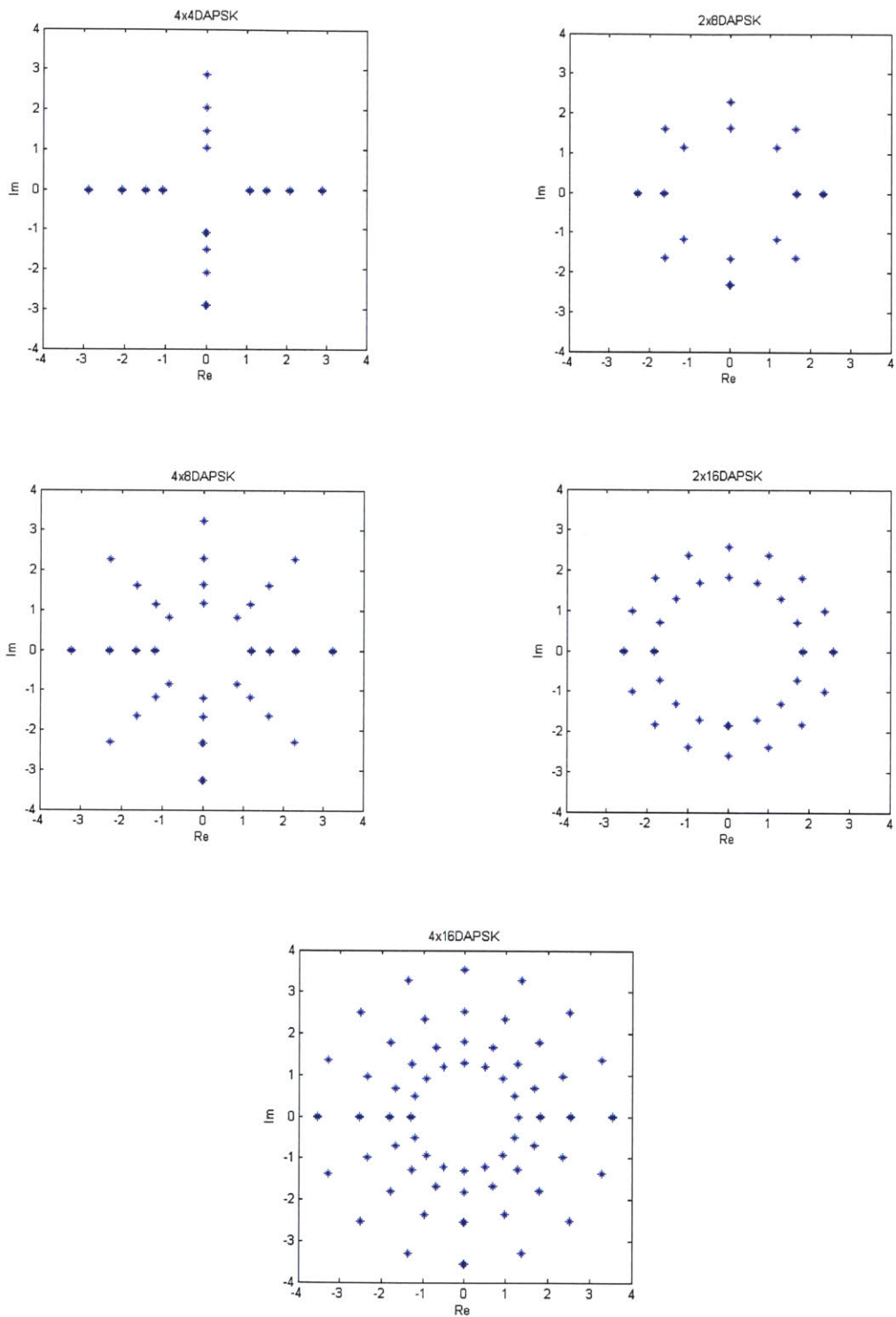**Figure 16: Signal constellations based on coherent detection.**

Figure 17: Signal constellations based on differentialyl coherent detection.

Chapter 4

4 Results

4.1 Experimental results for phase coherent detection

Experimental data obtained from the experiment conducted at WHOI were processed employing the coherent receiver described in Section 2.4. 8PSK, 16QAM, 32QAM and 64QAM modulation methods were used. In all the modulation schemes, the DFE had 10 $T/2$-spaced feedforward taps and two feedback taps. The PLL tracking constants were $K_1=10^{-3}$ and $K_2=K_1/10$. The forgetting factor of the RLS algorithm was chosen by trial to be $\lambda=0.999$. The step size μ of the LMS algorithm is given in (25).

The short vertical channel chosen for the tests proved to have very little distortion (negligible multipath), allowing for excellent signal detection using all modulation formats. Figure 18 - Figure 21 present the results for all the modulation methods employed. Shown in the left and right-hand columns are the scatter plots before and after equalization respectively. In addition, the mean square error for each adaptive algorithm (RLS/LMS) is given. Note that in all cases RLS converges faster than LMS. No errors were observed in any of the packets.

Table 3 summarizes the achieved bit rates and bandwidth efficiency for the modulation schemes employed. Note that for the 64QAM case, the bandwidth efficiency of 6 bits/sec/Hz was achieved, corresponding to the rate of 150 kbps.
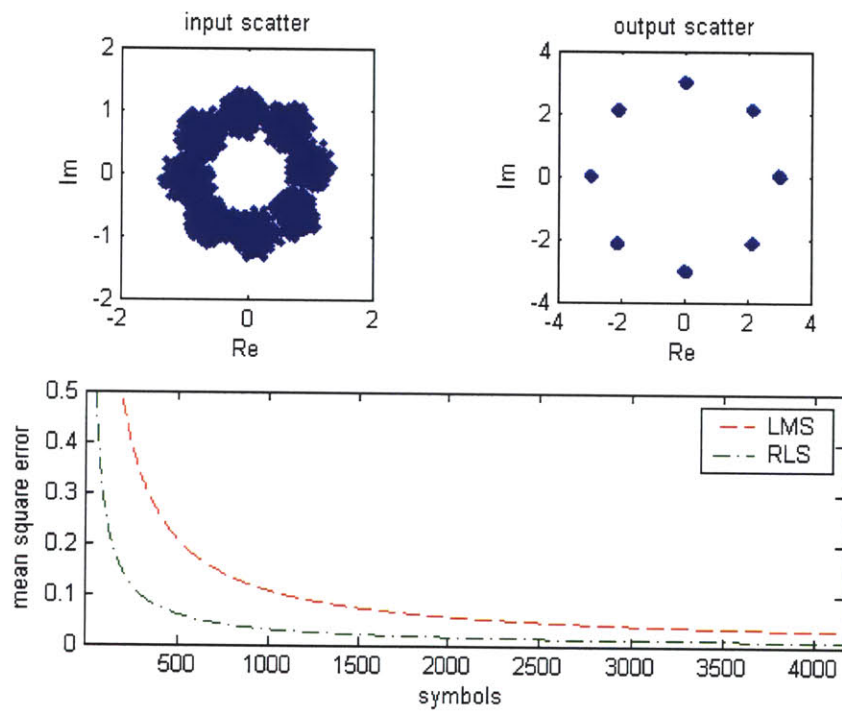

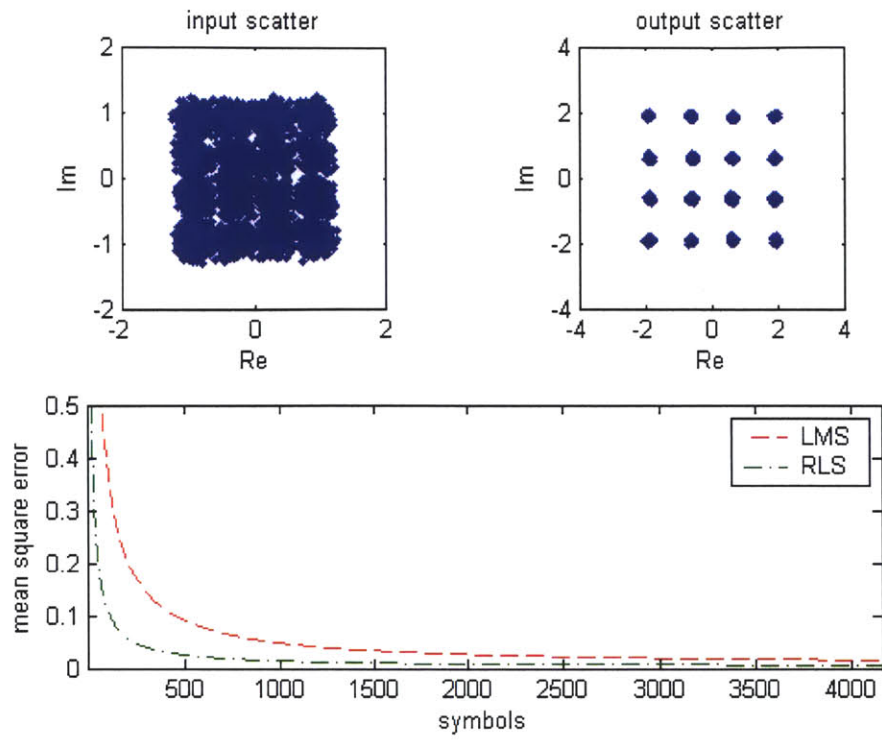**Figure 18: 8PSK input/output scatter. Input SNR/bit = 43 dB.**

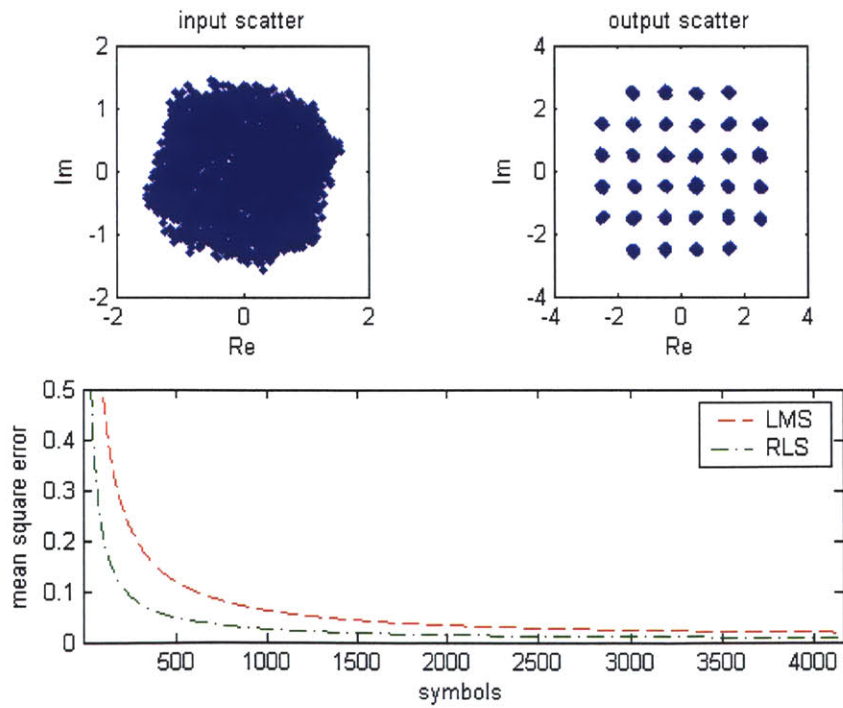**Figure 19: 16QAM input/output scatter. Input SNR/bit = 42 dB.**

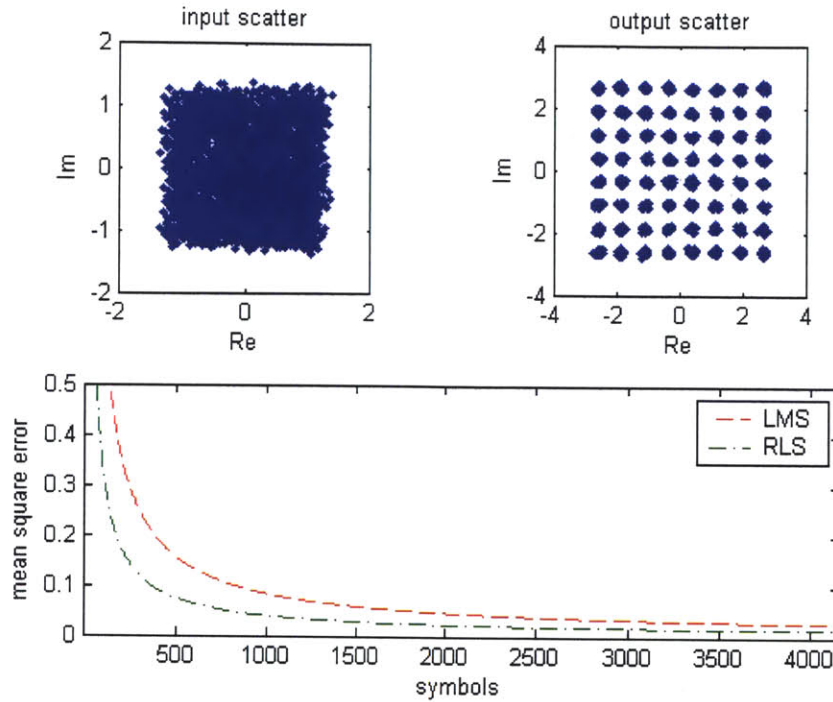**Figure 20: 32QAM input/output scatter. Input SNR/bit = 43 dB.**

Figure 21: 64QAM input/output scatter. Input SNR/bit = 41dB.

The transmitted video had the following specifications: 15 frames/sec, 144x176 pixels/frame and 8 bits/pixel. If no compression were used an acoustic link of 3,041 kbps would be needed to support transmission of the selected video clip. Employing JPEG compression at each individual frame, the required rate was reduced to 101 kbps (30:1 compression ratio). Therefore, in the current implementation, bandwidth efficiency of ≥3 bit/sec/Hz, i.e. 16QAM, 32QAM or 64QAM, is adequate in order to support real time underwater video transmission. Note also that the rates summarized in Table 3 suffice for real-time transmission if 64 kbps video encoding is employed which is a typical rate for the video encoding standards MPEG-4, H.263.

Modulation method	Bit rate	Bandwidth efficiency
8PSK	75 kbps	3 bits/sec/Hz
16QAM	100 kbps	4 bits/sec/Hz
32QAM	125 kbps	5 bits/sec/Hz
64QAM	150 kbps	6 bits/sec/Hz

Table 3: Achieved bit rate and bandwidth efficiency for all modulation methods.

The 144x176 sample frame was extracted from an underwater clip. The frame exhibits low contrast and detail inherent to underwater imagery. Figure 22 illustrates the original image frame and three JPEG compressed images produced from the original one. The various compression ratios were achieved by zeroing high spatial frequency DCT coefficients (AC coefficients). Figure 22 (b) is generated without zeroing any AC coefficients (pure JPEG compression), Figure 22 (c) is produced by zeroing all the AC coefficients with absolute value less than 1 and Figure 22 (d) is produced by zeroing all the AC coefficients with absolute value less than 2. Note that as the threshold is increased, the compression ratio increases at the expense of visual integrity. The "blockiness" distortion inherent to JPEG compression is due to the processing of an image by 8x8 subblocks and the zeroing of many AC coefficients produces rough pixel transitions across the boundaries of the subblocks. The results demonstrate that the JPEG algorithm yields higher quality compressed images for bit-rates of 0.2 bits/pixel and up. However, in a very low bit-rate scenario ($\leq$ 0.2 bits/pixel) the quality of the JPEG reconstructed image drops off dramatically.
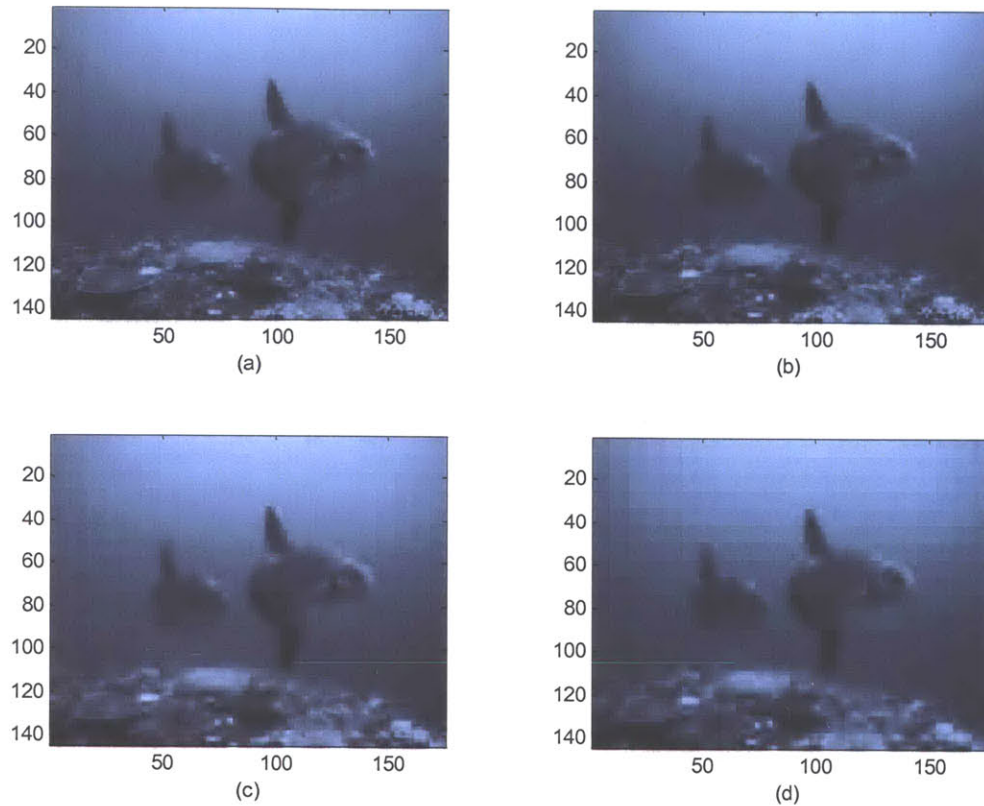
Figure 22: (a) original frame, (b) 16:1 compression ratio, (c) 22:1 compression ratio and (d) 30:1 compression ratio.

4.2 Simulation results for differentially coherent detection

In this section the performance of the linear equalizer followed by a reference-aided differentially detector is evaluated for 4x4DAPSK, 2x8DAPSK, 4x8DAPSK, 2x16DAPSK and 4x16DAPSK by computer simulations. For all signal constellations, the amplitude factor between two adjacent rings was $\alpha=1.4$. The packets transmitted in the experiment are used to conduct the simulations. All the packets had a training sequence of 500 symbols. The

differentially coherent equalizer had 10 *T/2*-spaced taps and the tap coefficients are adjusted adaptively by the modified LMS algorithm described in Section 2.6. In every modulation method the reference symbol was generated recursively by setting the parameter *W=1*.

Figure 23 - Figure 27 show the bit error probability versus E_b/N_o for all the modulation formats. In every case, the differentially coherent equalizer is compared with a coherent linear *T/2*-spaced equalizer with the same number of taps (ten taps). The tap coefficients of the coherent equalizer were adjusted adaptively by joint operation of conventional LMS algorithm and second-order PLL. The proportional carrier phase tracking constant was $K_1=10^{-2}$ and the integral tracking constant was $K_2=10^{-3}$. The step parameter μ of the LMS algorithm was the same for both equalizers. Furthermore, in every figure three family plots have been considered. The blue family represents zero frequency offset, the green family represents $f_d \cdot T=10^{-3}$ frequency offset and the red family represents $f_d \cdot T=2 \cdot 10^{-3}$ frequency offset.

In all cases, the differentially coherent equalizer approached the coherent equalizer for zero frequency shift. In addition, 2x8DAPSK and 2x16DAPSK signals performed satisfactory for $f_d \cdot T=10^{-3}$. However, for $f_d \cdot T=2 \cdot 10^{-3}$ the performance dropped off dramatically, especially for 4x4DAPSK, 4x8DAPSK and 4x16DAPSK, indicating that the equalizer could not tolerate frequency offset greater than $f_d \cdot T=10^{-3}$. On the other hand, the coherent equalizer provided excellent performance for all modulation methods for all frequency shifts. Such performance is made possible by inclusion of decision directed phase estimation.
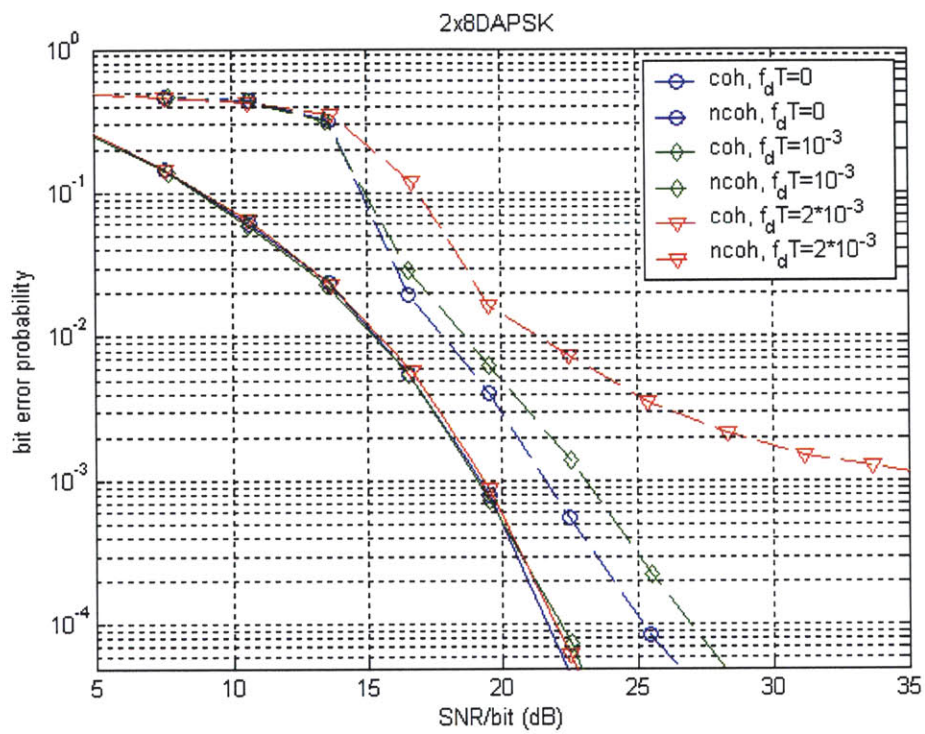
Figure 23: BER versus $10\log_{10}(E_b/N_o)$ for the coherent and noncoherent 2x8DAPSK receiver.
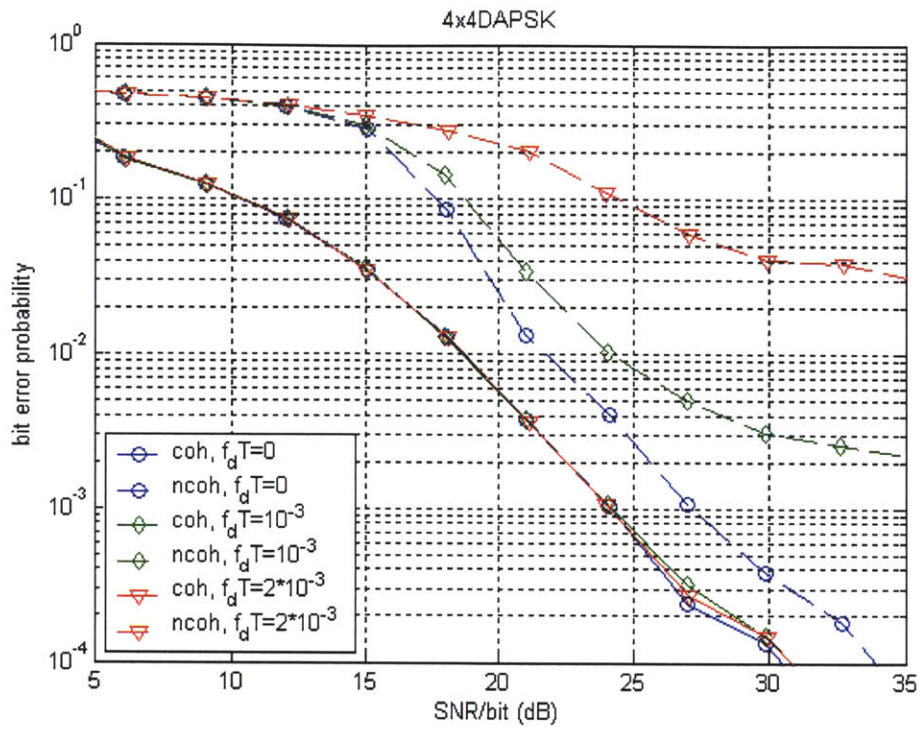
Figure 24: BER versus $10\log_{10}(E_b/N_o)$ for the coherent and noncoherent 4x4DAPSK receiver.
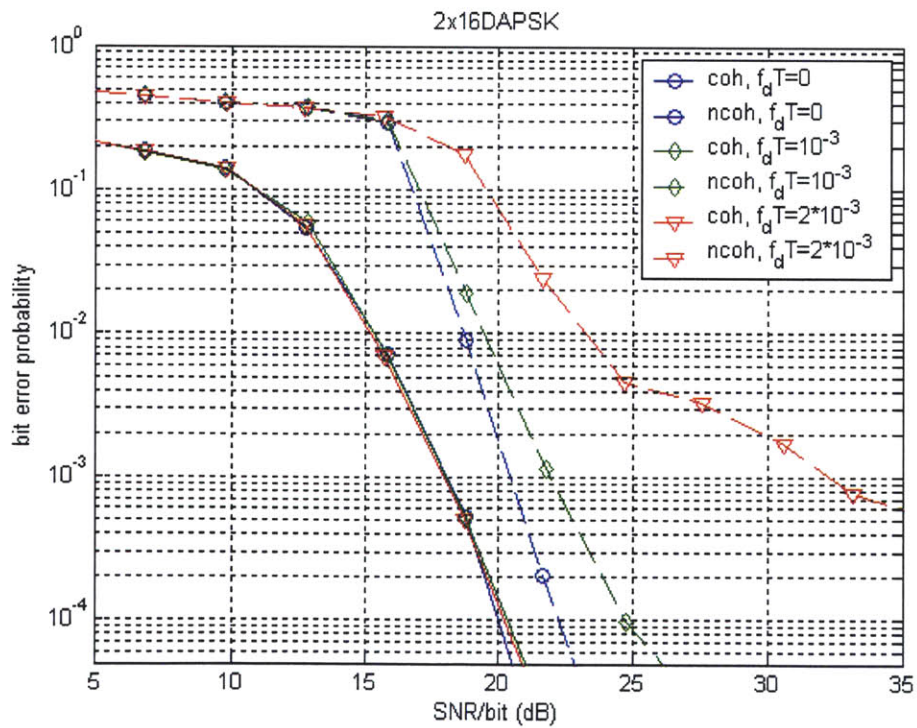

Figure 25: BER versus $10\log_{10}(E_b/N_o)$ for the coherent and noncoherent 2x16DAPSK receiver.
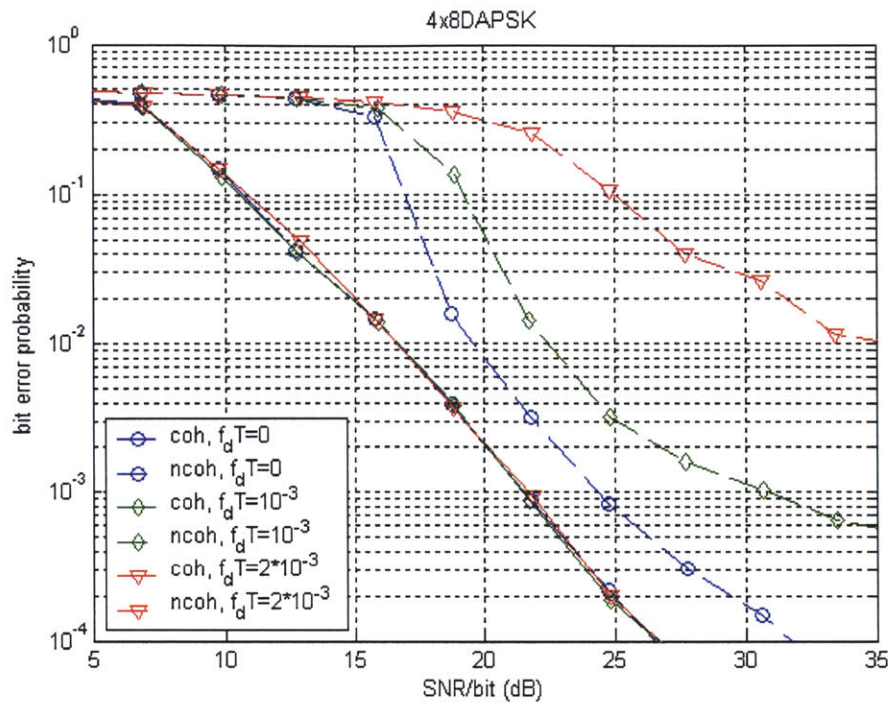
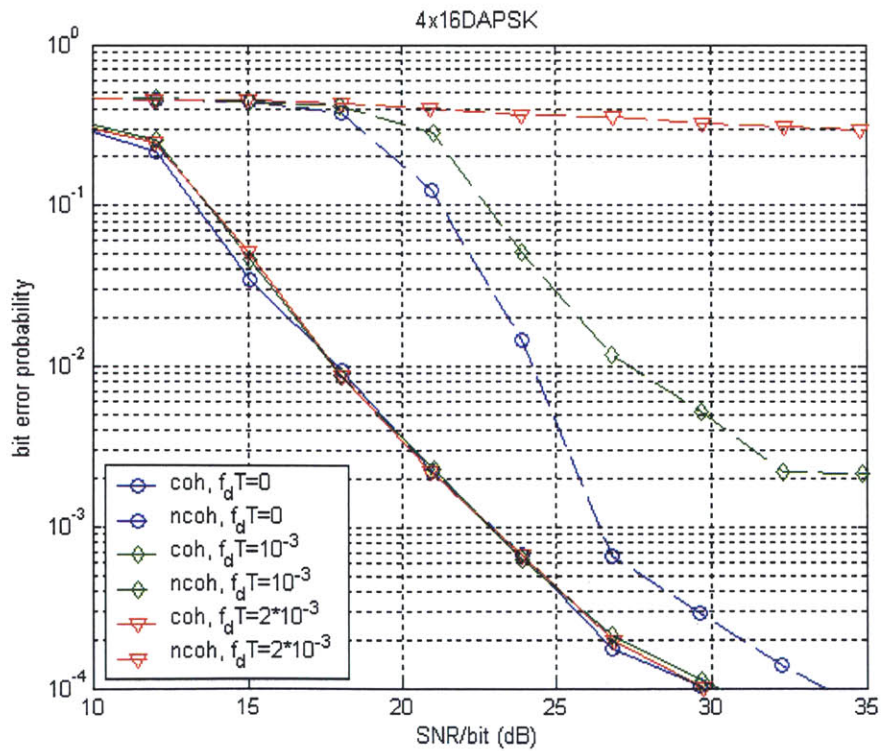**Figure 26: BER versus $10\log_{10}(E_b/N_o)$ for the coherent and noncoherent 4x8DAPSK receiver.**

Figure 27: BER versus $10\log_{10}(E_b/N_o)$ for the coherent and noncoherent 4x16DAPSK receiver.

Chapter 5

5 Conclusion

5.1 Accomplishments

A high rate acoustic link for underwater video transmission was implemented. Image encoding was accomplished using the JPEG DCT, scalar quantization and run-length Huffman encoding. Transmitter processing includes signal constellations based on:

1. Coherent detection, i.e., 8PSK, 16QAM, 32QAM, 64QAM.

2. Differentially coherent detection, i.e., 2x8DAPSK, 4x4DAPSK, 2x16DAPSK, 4x8DAPSK and 4x16DAPSK.

Receiver processing includes:

1. Jointly optimized DFE and decision-directed PLL based on [5].

2. Linear equalizer followed by reference-aided differentially coherent detection based on [11].

System performance was demonstrated experimentally, using 25000 symbols/sec at a carrier frequency of 75 kHz over a 10 m vertical path. The coherent receiver showed excellent

performance for all constellations thus achieving bit rates as high as 150 kbps, which are sufficient for real-time transmission of compressed video.

Computer simulations compared the differentially coherent equalizer with a linear coherent LMS equalizer. The fist presents satisfactory results for moderate frequency offsets $(f_d T = 10^{-3})$.

5.2 Future work

The eventual goal of our project is to compress underwater video data onboard an AUV and transmit it acoustically to the surface. Satisfactory results of preliminary tests reported in this thesis serve as an encouragement to perform experiments in deeper water. Two directions for future development are the following. First, a more sophisticated video encoder should be used which exploits motion compensation and prediction. Second, work should continue on high-level bandwidth-efficient modulation/detection methods. In particular, decision-feedback equalization methods suitable for DAPSK signals in the presence of nonnegligible frequency offset should be developed to make these methods applicable to underwater channels other than vertical channel which has little multipath. Moreover, these methods should be combined with array processing, which is known to be the single most effective means of providing the high SNR necessary for detection of high-level signal constellations.

Bibliography

[1] http://orca.rsmas.miami.edu/~majumdar/.

[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] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *IEEE J. Oceanic Eng.*, vol. 21, pp. 125-136, 1996.

[4] A. Kaya and S. Yauchi, "An acoustic communication system for subsea robot," *in Proc. OCEANS'89*, pp. 765-770, Seattle, Washington, Oct. 1989.

[5] 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.

[6] 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.

[7] P. Sehier and G. K. Kaleh, "Adaptive equalizer for differentially coherent receiver," *IEE Proc.*, vol. 137, pp. 9-12, Feb. 1990.

[8] A. Masoomzadeh-Farad and S. Pasupathy, "Nonlinear equalization of multipath fading channels with noncoherent demodulation," *IEEE J. Select. Areas Commun.*, vol. 14, pp. 512-520, April 1996.

[9] R. Scober, W.H. Gerstacker and J. B. Huber, "Adaptive linear equalization combined with noncoherent detection for MDPSK signals," *IEEE Trans. Commun.*, vol. 48, pp. 733-738, May 2000.

[10] R. Scober and W.H. Gerstacker, "Adaptive noncoherent DFE for MDPSK signals transmitted over ISI channels," *IEEE Trans. Commun.*, vol. 48, pp. 1128-1140, July 2000.

[11] R. Scober, W.H. Gerstacker and J. B. Huber, "Adaptive noncoherent linear minimum ISI equalization for MDPSK and MADPSK signals," *IEEE Trans. Sig. Proc.*, vol. 49, pp. 2018-2030, Sept 2001.

[12] 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.

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

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] J. Capitovic, "Performance limitations in underwater acoustic telemetry", *IEEE J. Oceanic Eng.* vol. OE-15, pp. 205-216, July 1990.

[21] W. T. Webb, L. Hanzo, and R. Steele, "Bandwidth efficient QAM schemes for Rayleigh fading channels," *Proc. Inst. Elect. Eng.*, vol. 138, pp. 169-175, June 1991.

[22] V.Bhaskaran and K.Konstantinides, Image and Video Compression Standards, 2nd edition, Kluwer Academic Publishers, 2003.

[23] Alan V. Oppenheim, Ronald W. Schafer, Discrete Time Signal Processing, 2nd edition, Prentice Hall, 1999.

[24] R. Coates, Underwater Acoustic Systems, New York: Willey, 1989.

[25] I. Dyer, "Subject notes for MIT course 13.851: Fundamentals and applications of underwater sound," Dept. of Ocean Eng., MIT, Cambridge, MA.

[26] J. G. Proakis, Digital Communications, New York: McGraw-Hill, 2001.

[27] S. Haykin, Adaptive Filter Theory, New Jersey: Prentice Hall, 1991.

[28] V. Engels, H. Rohling, "Multilevel differential modulation techniques (64-DAPSK) for multicarrier transmission systems," *European Trans. Telecommun.*, vol. 6, pp. 633-640, 1995.