A Study on the SPIHT Image Coding Technique for Underwater Acoustic Communications

Beatrice Tomasi[⋆], Laura Toni[†], Paolo Casari[§⋆], James Preisig[‡], Michele Zorzi[§⋆]

[⋆]DEI, University of Padova — Padova, Italy
[§]CFR, Consorzio Ferrara Ricerche — Ferrara, Italy
[†]IIT, Italian Institute of Technology — Genova, Italy
[‡]AOPE, Woods Hole Oceanographic Institution — Woods Hole, MA, USA

{tomasibe,casarip,zorzi}@dei.unipd.it laura.toni@iit.it jpreisig@whoi.edu

ABSTRACT

In this paper, we consider the transmission of progressively encoded images over underwater acoustic links, where transmitted symbols are protected by Forward Error Correction (FEC). The allocation of redundancy is performed according to both a Basic and a Multiple Description (MD)-like technique. The performance of this system is analyzed in terms of the resulting Peak Signal-to-Noise Ratio (PSNR) as image packets are transmitted over the measured realizations of acoustic channel impulse responses. These measurements were taken near the coast of Martha's Vineyard during October 2008, in different environmental conditions.

In the results, we quantify the performance improvement of the multiple description (MD) technique compared to the Basic allocation, thus suggesting its suitability for the transmission of images in the underwater acoustic scenario.

Categories and Subject Descriptors

C.2.0 [**Communication/Networking and Information Technology**]: General—*Data communications*; I.6.6 [**Simulation and Modeling**]: Simulation Output Analysis

General Terms

Measurement, Performance

Keywords

Underwater acoustic communications, source coding, image coding, simulation

1. INTRODUCTION

The acquisition and transmission of images is one of the most appealing applications envisioned for autonomous underwater networks, and at the same time one of the most overlooked so far [1,2]. Nevertheless, the feasibility of real-time underwater imagery would open up the way to new explorations of the marine environment, paving the way to discoveries in the fields of marine biology, ecology and geology and being of help in oil spill relief operations, and in the detection of underwater mines, to name a few.

The efficient encoding and transmission of images over terrestrial radio links has been widely studied [3, 4]; however, as underwater acoustic links are different from their terrestrial counterparts, it is not clear whether the technologies and communication models employed for the latter can be applied to the former as well [5]. In fact, the typically high error rates, high path loss, and low channel capacity in underwater channels make the purpose of high quality image transmission even more challenging than in terrestrial radio links [6, 7]. The first works on underwater image transmission were mainly focused on studying coding techniques aimed at compressing the large amounts of information gathered by Autonomous Underwater Vehicles (AUVs) during their missions. For example, in [8] a wavelet-based video coding technique was proposed to compress underwater video sequences.

More recent works investigating source coding techniques are [7, 9, 10]. In [9], a hybrid global-local motion-compensated scheme has been proposed, which leverages on the main features of underwater video sequences (e.g., primarily static scenes with small moving objects, low contrast due to the artificial light conditions, etc.). Although the authors demonstrated that the proposed compression technique provides a satisfactory visual quality while maintaining a high compression efficiency, realistic underwater transmissions have not been considered in the simulation tests. A simpler source compression scheme has been investigated in [7], where the Set Partitioning In Hierarchical Trees (SPIHT) source coding scheme [11] has been applied to underwater transmissions. The SPIHT encoded bitstream, divided in consecutive chunks, has been transmitted over underwater links. Results have demonstrated that SPIHT is an efficient compression technique for data and image transmission, but still suffers from high end-to-end delays. However, the authors only consider a basic version of SPIHT, where packets are retransmitted if lost, which is an inherently inefficient hybrid Automatic Repeat reQuest (ARQ) strategy [12], and where channel coding is assumed without any FEC optimization based on the progressive nature of the source bitstream. End-to-end delays might be considerably reduced by adopting more sophisticated transmission schemes, which make the system more robust to channel impairments via a different management of redundancy, leading to a considerable reduction of the number of retransmissions. For example, the choice of the redundancy per chunk should be strictly correlated to the priority level of the source bitstream, by adopting an unequal error protection (UEP) strategy (i.e., a more reliable transmission for more important bits).

WUWNet'11, Dec. 1 - 2, 2011, Seattle, Washington, USA

This is the approach presented in this paper. Specifically, we propose a comparative study of FEC allocation strategies for the transmission of SPIHT source coded bitstreams in underwater acoustic links. We focus on the SPIHT scheme because it is inherently designed for progressive image encoding (whereby the quality of received images depends on the number of correctly received packets), and for the capability to manage redundancy in a way that is convenient to communications and that can be adapted to actual channel requirements (see also a description of SPIHT in Section 2). To the best of our knowledge, this is the first study on rate allocation schemes for the transmission of progressively coded source information over underwater channels.

In the following, a symmetric n-channel FEC-based MD [13], able to exploit the progressive nature of the encoded source, is compared to a baseline algorithm [7]. For both algorithms, we propose a forward error correction (FEC) profile optimization, which minimizes an objective cost function (e.g., the expected distortion) under the constraints that the total number of bits to be transmitted is fixed, and that the channel state information (CSI) available at the transmitter is limited. We remark that due to the large underwater propagation delays and to the usually lower channel coherence time, any instantaneous CSI fed back to the transmitter would be likely outdated. Therefore, the FEC profile optimization will take into account only the average signal-to-noise ratio (SNR), assuming that no information is available on instantaneous fading fluctuations and packet correlation.

The results, illustrated in Section 3, demonstrate how suitable the MD technique is for underwater scenarios, especially in the low SNR regime.

2. SYSTEM DESCRIPTION

2.1 Progressive Source Coding

Progressive/embedded coding methods (e.g., SPIHT), which have been deeply investigated for terrestrial radio links, are appealing for underwater communications as well. First, the quality of the decoded information is proportional to the amount of received information. This means that the more the bits adopted for decoding, the higher the quality of the reconstructed bitstream. Second, the decoding process can be stopped as soon as a target bit budget (i.e., the total number of bits) or a target distortion objective is met, and the quality will be the best possible for that bit budget.

Third, given an encoded bitstream, the image can be reconstructed from different numbers of bits processed by the decoder, providing several qualities and several compression rates with just one coding process. All these features make progressive coding methods highly suitable for heterogeneous networks (in terms of channel bandwidths, mobility levels of terminals, end-user quality-of-service requirements), or for transmissions with limited CSI. However, progressive encoding is generally error-sensitive, as an error generally makes the subsequent bits useless. Thus, the placement and amount of FEC within the bitstream should be carefully designed, e.g., by adapting the amount of FEC to the importance of bits within the SPIHT stream.

2.2 Overview of FEC Allocations for SPIHT

In this section we describe FEC allocation techniques, aimed at assigning the most convenient protection to the encoded bitstream when limited CSI is available. Neither instantaneous channel fluctuations nor the information on fading correlation will be taken into account for the design of the allocation/coding schemes. In both algorithms, we assume that the SPIHT encoded information is transmitted in a resource block (RB) of fixed size, characterized by N_t packets, each consisting of L_p modulated symbols.

2.2.1 *The Basic Allocation*

In the basic algorithm, depicted in Fig. 1, the encoded bitstream is sequentially ordered, and allocated into packets. Each packet is protected in the time domain by a concatenation of Reed-Solomon (RS) and cyclic redundancy check (CRC) channel coding. We assume erasure channels, which means that, at the receiver side, if the CRC check fails (i.e., if the RS decoder is unable to correct all errors affecting the received packet), the packet is lost. Losing a packet makes all subsequent packets useless, as decoding a progressively encoded stream requires in-order processing of transmitted data. It follows that every source bit is more important than subsequent ones, and requires UEP, i.e., the assignment of a greater amount of FEC to the first bits in the bitstream, while leaving later (thus less important) chunks less protected. In Fig. 1, for example, the first information symbol is more protected than the subsequent ones, and the probability of losing the first packet will be lower than that of losing the other ones.

We assume that each image will be allocated and transmitted in a RB with fixed dimensions, such that by increasing the level of protection introduced by the channel coding, the information bits transmitted per RB will decrease. Fig. 2(a) denotes the RB generated by the basic allocation, and the notation adopted in the following: each image is allocated into N_t packets, each one consisting of L_p modulated symbols. The goal of the basic allocation algorithm is to optimize the right amount of FEC per packet, able to find the best tradeoff between redundancy and reliability, as it can be evaluated through the following problem formulation.

We assign a given channel code rate $r_j, j = 1, \ldots, N_\mathrm{t}$ to each packet of the RB, and we denote by $\boldsymbol{r} = [r_1, r_2, \ldots, r_{N_\mathrm{t}}]$ the rate vector for the channel codes in the time domain. In view of the previous discussion, any packet j is more important than all subsequent ones; therefore, we impose $r_j \leq r_{j+1}$. When all the first j packets are received, but the $(j+1)$-th packet is lost, the decoded image quality will be $D\left[R_j(\boldsymbol{r})\right]$, where $D[\cdot]$ is the rate-distortion curve of the employed source coding scheme [14], and R_j is the received bit budget (i.e., the total number of bits) for the information symbols. This received bit budget is computed as $R_j(\boldsymbol{r}) = \sum_{l=1}^{j} c_l(r_l)$, where c_l is the number of source information bits in the l-th packet. Given the source code rate-distortion curve $D\left[R_j(\boldsymbol{r})\right]$, and the conditional packet loss probability P_j, which is the probability of losing the $(j+1)$-th packet given that packets $1, \ldots, j$ are correctly received, we can evaluate the best channel code rate vector $\boldsymbol{r}^\star$ as the one that minimizes the expected distortion $E[D]$, i.e.,

$$E_{\boldsymbol{r}^\star}[D] = \min_{\boldsymbol{r}} \left\{ \sum_{j=1}^{N_\mathrm{t}} D[R_j(\boldsymbol{r})] P_j(\boldsymbol{r}) + D_0 P_0(\boldsymbol{r}) \right\} \qquad (1)$$

subject to a constraint on the overall bit budget

$$\sum_{j=1}^{N_\mathrm{t}} c_j / r_j \leq B_\mathrm{tot} \,,$$

where B_tot is the total bit budget of the RB, and D_0 corresponds to the distortion when no packets are received, i.e., when the decoder must reconstruct the message without being able to use any of the transmitted information. For a gray scale still image, this typically means reconstructing an entirely gray image. We use the iterative procedure described in [15] to solve the optimization problem (1).

In an ideal system, in which the instantaneous SNR experienced by each set of packets is available in the feedback channel, the Basic Allocation algorithm would be able to assign the right amount

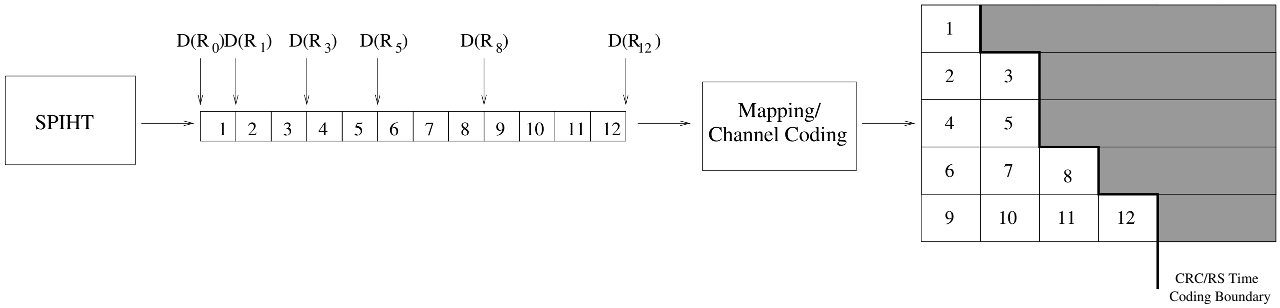
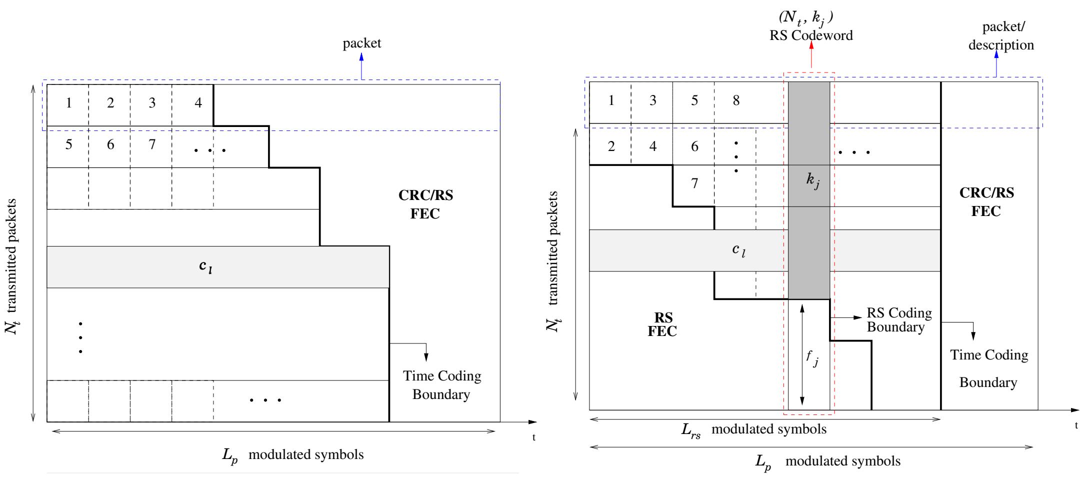


Figure 1: Description of the Basic Allocation scheme. The dark shaded area to the right of the CRC/RS Time Coding Boundary line represents the bits for CRC and Reed-Solomon coding applied to each packet. The unshaded area represents information symbols.



(a) Basic Allocation of the number of information bits

(b) MD-based Allocation

Figure 2: Resource block for progressive image transmission.

of FEC to each transmitted packet. In particular, since each RB experiences its own SNR vector, the optimization problem in (1) should be re-optimized for each RB. Thus, the Basic Allocation algorithm leads to the best performance if the SNR experienced by the packets is provided in the feedback channels, driving the system to be extremely sensitive to non-ideal or outdated channel estimations. To show this sensitivity, we compare ideal systems with realistic scenarios, in which the mean SNR per image is considered as information in the feedback channel.

This allocation technique is considered as a baseline, and has been already tested in underwater scenarios [7]. However, the allocation in [7] is based on a heuristic approach (by transmitting more redundancy when the channel is bad), whereas, here, we compute the optimized allocation, in terms of image distortion, by solving a minimization problem. This approach has not yet been investigated for the underwater acoustic scenario.

2.2.2 *The MD-Like Allocation*

We now illustrate the general mechanism for converting an embedded bitstream from a source encoder into multiple descriptions in which contiguous information symbols are spread across multiple packets. The basic idea of MD source coding is to split the information into multiple bitstreams approximately equally important, each of which can be independently decoded. In the MD-Like allocation technique [13], which is a n-channel FEC-based MD coding, contiguous information symbols are spread across N_t multiple packets (also called "descriptions"), instead of being assigned to the same packet.

We now show how an embedded bitstream, generated from a progressive source coding, can be allocated and opportunistically protected in a RB to be transmitted over underwater channels. With the final goal of improving the reliability of the bitstream transmission, we assume to protect the encoded image with a bidimensional channel code (product code). Fig. 3 depicts the detailed steps of the MD-Like technique: i) allocation and vertical channel coding, ii) horizontal channel coding.

In the first step, the bitstream is allocated across packets/ descriptions. The information is protected using systematic maximum distance separable (MDS) RS codes across packets, with a level of protection depending on the relative importance of the information symbols. This means that for a (N_t, k_l) codeword, where k_l is the number of source symbols to encode the l-th codeword, the FEC level is $f_l = N_t - k_l$ and the source information is decoded if at least k_l of N_t descriptions are received. Due to the progressive nature of the source information, we assume $f_l \geq f_{l+1}$: the more

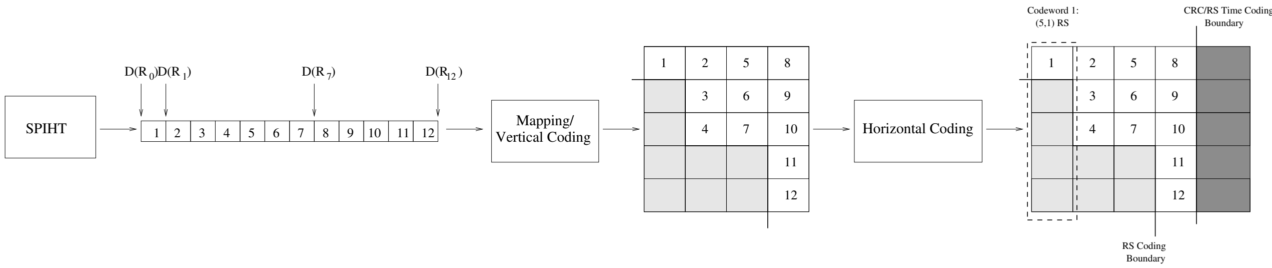


Figure 3: Description of the MD-Like Allocation scheme. The dark shaded area to the right of the CRC/RS Time Coding Boundary line represents the bits for CRC and Reed-Solomon coding applied to each packet (horizontal channel coding). The lighter shaded area under the RS coding boundary staircase represents Reed-Solomon parity symbols across the packets (vertical channel coding). The unshaded area represents information symbols. For each (N_t, k_l) RS codeword, k_l information bits are encoded into N_t total symbols.

important the information bits, the higher the protection. For example, in Fig. 3, the MD-Like allocation leads to the construction of 4 codewords. The first one is a $(5, 1)$ systematic MDS code (codeword 1) in which erasure of any four descriptions still allows to reconstruct information symbol 1 and to achieve a delivered quality equal to $D(R_1)$. It is worth noting that, due to independent descriptions, the MD-Like scheme only requires to know how many and not which packets have been received. Thus, all the packets have roughly the same importance, and, in the optimization process, the second order statistics of the channel will be a sufficient information for the FEC design, as described in the following. Moreover, due to the allocation of the information symbols across packets, rather than the sequential one as in the Basic Allocation, the first description will not contain all the most important source bits. So, the loss of the first packet will not jeopardize the reconstruction of the first source bits. In the example of Fig. 3, if the first description is lost, the delivered quality would be equal to $D(R_7)$, while in the Basic allocation it would be $D(R_0)$, since no information would be available for the source decoding.

A second step of the algorithm consists in concatenating CRC and RS codes to each description (horizontal coding), in order to enhance the system robustness to channel fluctuations (both fast and slow fading). Since each description is approximately equally important, an equal error protection (EEP) profile is employed.

Fig. 2(b) illustrates the proposed RB for the transmission of an embedded bitstreams, depicting the notation adopted in the following. Due to the UEP across the packets, a staircase boundary can be observed as RS coding line, while a vertical boundary represents the CRC/RS coding line. The L_{rs} symbols on the left of the vertical boundary are the RS codeword across packets, while those on the right represent the CRC/RS parity symbols. In the proposed allocation algorithm, we will be interested in evaluating the packet loss probability mass function $\mathbb{P}_{N_t}(g), g \in [0, N_t]$, which is the probability of losing any g out of N_t total packets in the RB. When $j = N_t - g$ packets are received, any l-th RS codeword across the packets with $c_l \leq j$ can be correctly decoded, and the total bit budget for the source information becomes

$$R_j = \sum_{l: c_l \leq j} c_l \times B_{\text{rs}} \tag{2}$$

where B_{rs} is the number of bits per RS symbol. In this MD-Like allocation technique, both FEC profiles of the product code need to be designed for the minimization of the expected distortion, based on the following problem formulation

$$E_{r^\star \boldsymbol{f}^\star}[D] = \min_{\{r, \boldsymbol{f}\}} \left\{ \sum_{l=1}^{N_t} D\left[R_l(r, f_l)\right] \mathbb{P}_{N_t}(N_t - l) + D_0 \mathbb{P}_{N_t}(0) \right\} \tag{3}$$

where $\boldsymbol{f} = [f_1, f_2, \ldots, f_{L_{\text{rs}}}]$ is the code rate vector for the RS code across the packets.

Unlike the Basic Allocation scheme, the MD-Like optimization does not depend on the instantaneous SNR vector. The reason can be found in two main factors: i) an EEP is assumed in the horizontal time coding, so the mean SNR is a sufficient information to set the right amount of FEC in the time domain coding. ii) the UEP profile of the vertical coding is optimized based on the packet loss PMF, which requires to know how many and not which descriptions are lost. Thus, from the FEC optimization process, the second order statistics of the channel is required, and not the instantaneous information. This means that, first, the FEC optimized scheme will be suitable for all the RBs transmitted over channels with the same second order statistics. Second, the optimization scheme is robust to non-ideal or outdated CSI.

2.3 Channel Model

We now address the correlation between the channel model and the UEP profiles. In [16], the MD-Like allocation scheme has been proposed for transmission of progressive images over OFDM networks. The authors considered a frequency-selective environment and deduced that the level of diversity offered by the channel in the frequency domain is one of the main factors which affect the UEP profile design. In our work, we aim at studying the possible effects of the channel characteristics on the FEC optimization.

We assume that packets are sequentially transmitted in time over underwater channels. We denote the coherence time as Δ_t, and we consider that K packets span Δ_t. Subdividing the N_t packets into blocks of K descriptions, for this theoretical analysis, we assume that packets in different blocks are considered to fade independently, while packets in the same block experience identical fades.

Knowing the coherence time of the channel, the Basic allocation algorithm can be slightly simplified. In particular, we might assume that a constant level of FEC is assigned to packets in the same block. Assuming that the packets in the subblock are correlated, one level of FEC will offer the right amount of protection to all the K transmitted packets. For the extreme case of $K = N_t$, all the packets will be correlated, and the UEP might degenerate to an EEP. The opposite case is when $K = 1$, and all the packets might experience a different channel code rate.

When the MD-Like algorithm is considered, the correlation between packets affects the packet loss PMF, which is an important parameter in the FEC optimization. In particular, a two-state Markov model, which corresponds to a Gilbert-Elliot channel, characterizes the channel by a “good" state (G) and a “bad" state (B), the probability that k packets are lost for a channel with K correlated packet is [16]

$$\mathbb{P}_K(k) = \mathbb{P}_K(k|G)P(G) + \mathbb{P}_K(k|B)P(B) \quad (4)$$

where $\mathbb{P}_K(k|G)$ ($\mathbb{P}_K(k|B)$) is the probability of having k corrupted packets when the channel is in the “good" (“bad") state, and $P(G)$ ($P(B)$) is the probability of being in the “good" (“bad") state. Assuming that the subblocks are mutually independent, the PMF of losing g packets out of N_t can be derived from the PMF of each subblocks as follows

$$\mathbb{P}_{N_t}(g) = \sum_{k_1,k_2,\ldots,k_L:k_1+k_2+\ldots+k_L=g} \mathbb{P}_K(k_1)\cdot\ldots\cdot\mathbb{P}_K(k_L) \quad (5)$$

where $L = N_t/K$ is the number of independent subblocks per RB, and we assumed that all the packet loss PMFs per subblock are the same, since the subblocks have the same statistics.

From the previous analysis, we deduce that the packet loss PMF of the RB depends on the coherence time of the channel, and it is determined by the number of combinations for g and k_j. For high coherence time values ($K = N_t$), due to the limited number of combinations in losing g packets, $\mathbb{P}_{N_t}(g)$ is mostly determined by the PMFs of the individual sub-channels. This means that, when most of the packets are correlated, there is a substantial probability of losing or receiving most of the packets, whereas the probability of losing only a small number of descriptions is negligible. Given this PMF, there is no level of diversity offered by the channel (or degrees of freedom), which means that in most cases the RB will be entirely received or lost. Therefore, in order to create a level of priority in the RB, a sharp UEP is required in the vertical coding.

For high L values (i.e., for low K values), by the Central Limit Theorem, $\mathbb{P}_{N_t}(g)$ approaches a normal distribution, leading to a high (low) probability of losing half of the (all or zero) descriptions. For example, when $K = 1$ all the packets will be mutually independent. Since the probability of losing most of the packets is negligible, there is no need to assign a lot of FEC in the first codewords in the vertical coding. Therefore, the system can increase the throughput, keeping the FEC value in the RS codeword low even for the codewords with high priority. This would drive the system choosing an EEP profile as the best option.

2.4 Description of the Acoustic Data Set

The experiment, during which the data set was collected, was conducted on October 2008 at the Martha’s Vineyard Coastal Observatory (MVCO) operated by the Woods Hole Oceanographic Institution [17]. The data set is particularly interesting because such environmental conditions as surface waves and wind speed and direction varied significantly over the duration of the experiment. While the collected data sets are from only one deployment, the results about the performance of the considered image transmission techniques can be extended to other scenarios with similar geometries. The scenario consists of a single transmitter and four receiving systems, deployed at different directions and distances. In particular, systems S3 and S5 were placed 200 m and 1000 m Southeast of the transmitter, respectively, while S4 and S6 were similarly deployed in the Southwestern direction. The devices were deployed in shallow water conditions (the water depth was around $15\,\mathrm{m}$ almost uniformly), near Martha’s Vineyard Island from October 18 to October 27, 2008. The temperature was almost constant on the water column. The transmitter central frequency was $11.5\,\mathrm{kHz}$. In these conditions, the most important factor causing the time-variability of the acoustic propagation is the surface condition.

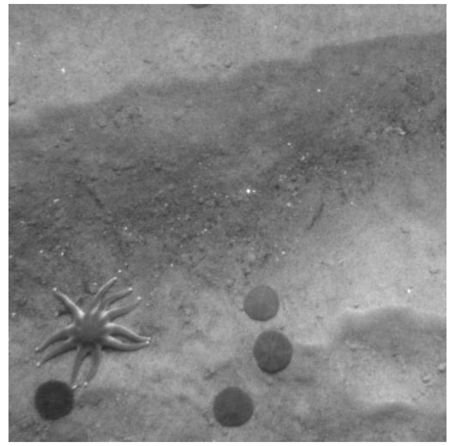


Figure 4: This picture, representing the sea floor, was collected by an AUV, and is an example of a typical underwater image.

In the simulations, we will consider the data received at system S3. The transmitted signals, from which we estimate the channel impulse response, consist of multiple repetitions of a 4095-symbol binary maximum length sequence transmitted at a symbol rate of 6.5 kbps. A transmission three minutes in duration was made once every two hours. A maximum length sequence is a particular pseudo-noise signal that is spectrally flat. Thanks to this property, it is possible to estimate the channel impulse response by computing the correlation between the transmitted signal and the received signal. We compute the channel estimate over segments of the received signal that are 400 symbols long (which corresponds to 60 ms). After each estimation we shift the window by 100 symbols, which corresponds to 15 ms, resulting in an estimate every 15 ms over windows of 60 ms. The data were collected at high SNR in order to allow an accurate estimation of the channel impulse response.

In this work we use these estimates as underwater acoustic channel realizations. Given the availability of a channel estimate every $15\,\mathrm{ms}$, by assuming a symbol duration of almost $2\,\mathrm{ms}$ and by assuming that the channel is constant over the interval between two estimates, we know the channel coefficients at every byte in a packet. An image consists of 153 packets, each 512 bits long. Every packet lasts almost $1\,\mathrm{s}$ so that all channel estimates over a three-minute period are used as channel realizations during the transmission of an image. In Section 3, more details on how the channel estimates were processed to compute the statistics of the performance will be provided.

3. SIMULATION RESULTS

3.1 Simulation Setup

We run the simulations on the 512×512 gray-scale image in Figure 4. The image was encoded using the SPIHT [14] algorithm to produce an embedded bitstream with a compression rate of 0.3. The serial bitstream was allocated into a RB with $N_t = 153$ packets, each consisting of 512 binary phase-shift keying (BPSK) modulated symbols, for a total budget of 9830 bytes. In the MD-Like allocation algorithm, we used MDS RS codes across the descriptions and there were 8 bits per RS symbol, that is $B_{\text{rs}} = 8$. As the available coding rates used in the optimization problems (1) and (3), we consider $\{0.7, 0.8, 0.9\}$.

As a representative metric, we compute the PSNR[1] of the image versus the SNR averaged over an image. In particular, we consider the low-high SNR range between 5 to $22\,\text{dB}$. By processing the data set described in Section 2.4, we obtain a collection of three-minute long sequences of consecutive instantaneous SNRs. We then compute the sequences of channel gains from the sequences of estimated SNRs. The simulated communication system consists of the source encoder, a modulator, the channel with additive noise, the demodulator and the decoder. During a simulation run, we multiply the transmitted byte stream by the corresponding portion of the channel gain sequence and we add noise. The result of this latter operation is the received data stream, which is then demodulated and decoded in order to assess the communication performance. In the simulation we consider all the inferred channel gains for each noise power level. This approach allows us to study the system behavior in different SNR regimes, while taking into account the channel gain fluctuations that occur in realistic acoustic propagation scenarios and heavily affect the communication performance.

For both the basic and MD schemes, we compare a realistic system (labeled as “shadowing"), in which the transmitter only knows the mean SNR per image, with a genie-aided system (labeled as “ideal"), in which the transmitter has a perfect knowledge of the instantaneous CSI, which is used in the optimization process. This comparison aims at quantifying, in terms of performance loss, how the allocation schemes are affected by a limited CSI. Moreover, we perform an analytical and a simulation study for both allocation techniques.

We compute analytically the packet loss probability mass function and the PSNR, by assuming independent and identically distributed (i.i.d.) packets and by precomputing the packet error rate (PER), from the data set. On the other hand, we obtain the simulation results by feeding the system with the channel realizations, and simulating realistic transmission of the images. It is worth noting that both the analytical and the simulation curves assume the same PER when the allocation is performed. By comparing the simulation and the theoretical curves, we show whether or not it is accurate to assume independence at the byte as well as the packet level.

3.2 Results

Figure 5 shows the computed FEC profiles, for both the Basic and the MD schemes, for two different average SNRs at system S3. The x-axis and y-axis indicate the number of information bytes and the packet transmission index, respectively. We recall here that an image is encoded into 153 packets, each of 64 bytes. Depending on the allocated FEC, each packet has a certain amount of redundancy. For this reason, the y-axis is limited between 1 and 153, whereas the x-axis can be between 1 and 64. The solid lines correspond to the profile for a middle SNR regime, whereas the dotted (with the diamond) lines represent the profiles at a high SNR regime. Black and blue colors indicate the MD and the Basic allocation schemes, respectively. While the Basic allocation scheme adds FEC to each packet, the MD-like technique introduces both vertical and horizontal redundancy. The double level of FEC introduced in the MD allocation technique increases the transmission reliability, at the price of a reduction of transmitted source bits. However, a reduction of transmitted source rate is not necessarily translated into a decreasing quality of the reconstructed image. In

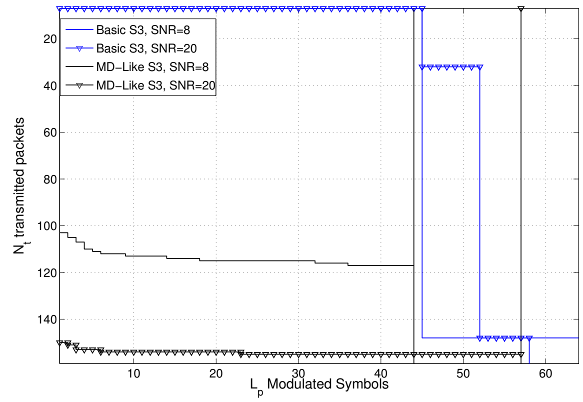


Figure 5: FEC profiles for Basic (in blue) and MD (in black) allocation schemes. The dotted (with the diamond) and solid lines correspond to $20\ \text{dB}$ and $8\ \text{dB}$ SNRs, respectively.

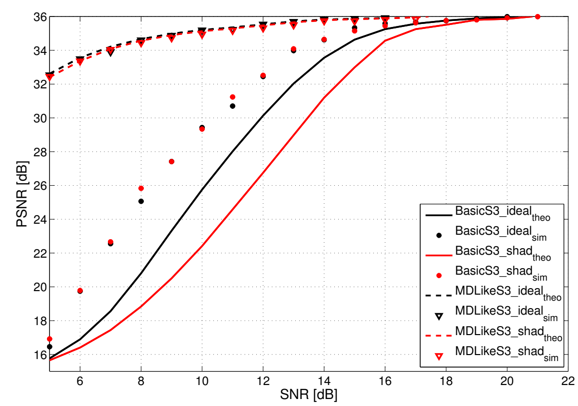


Figure 6: PSNR vs SNR for system S3 at $200\,\text{m}$ Southeast.

particular, the reconstructed image quality increases accordingly to the correctly received information bits. Thus, the best FEC profile is not the one which maximizes the transmitted source rate, but the one which maximizes the correctly received bits. This tradeoff between reliability and transmission rate is investigated in the following figure. As expected, for both the allocation schemes, the best FEC profile is reduced with the increasing of the mean SNR.

Figure 6 shows the results in terms of the received image quality, PSNR. We remind that the PSNR represents the tradeoff between transmission source rate and reliability. For this reason, we define as the best FEC profile, the one which maximizes the PSNR. As expected, the MD technique outperforms the Basic allocation scheme, especially in the low-medium SNR regime $(5 - 20\,\text{dB})$, which will be the operational SNR regime for typical underwater acoustic systems. The double channel coding in the MD technique reduces the transmitted source information, but increases, at the same time, the reliability of the transmission. This means that the MD allocation technique is able to find the best tradeoff between transmitted source rate and reliability. Therefore for future implementation, if memory is not a problem, the MD technique should be preferred to the basic one. As expected, the greater the SNR, the higher the transmission reliability, and the lower the required FEC in the RB.

[1] $\text{PSNR} = 10\log_{10}\left(255^2/E[D]\right)$.

(a) Basic allocation scheme, SNR=6 dB, PSNR=19.7 dB.

(b) Basic allocation scheme, SNR=12 dB, PSNR=32.3 dB.

(c) MD-Like allocation scheme, SNR=6 dB, PSNR=33.4 dB.

(d) MD-Like allocation scheme, SNR=12 dB, PSNR=35.6 dB.

Figure 7: Image quality for system S3 at $200\,\mathrm{m}$ Southeast. Both MD and Basic allocation schemes are considered for transmissions over $6\ \mathrm{dB}$ and $12\ \mathrm{dB}$ SNRs channels.

Figure 6 refers to system S3, which is at $200\,\mathrm{m}$ from the transmitter. We can observe that, for the Basic allocation technique, the theoretical ideal and shadowing curves are separated by $1-3\,\mathrm{dB}$ in the medium SNR range ($[8-15]\,\mathrm{dB}$) and by $0.5-1\,\mathrm{dB}$ in the low SNR regime ($[5-7]\,\mathrm{dB}$). This behavior represents the performance degradation due to a limited CSI in the shadowing curves, which relies on the mean, rather than the instantaneous, SNR information, when independent and identically distributed (i.i.d.) errors over a packet are assumed. From the simulated results, it can be observed that the theoretical evaluation is a pessimistic estimate of the system performance. Even though the optimum theoretical FEC profile underestimates the channel quality, it actually over protects the system, thus reducing the gap between ideal and shadowing curves. This means that, for the considered receiver, the mean SNR is a sufficient information for the design of the FEC profile in the basic technique.

For the MD allocation, the performance degradation is lower than $2\,\mathrm{dB}$ for both theoretical and simulated curves, proving the robustness of the algorithm to limited CSI. This is justified by the adoption of FEC protection in the time domain and across packets. From these results we can conclude that, for the considered receiver, a faster feedback would not give any PSNR performance improving to both the Basic and the MD allocation techniques.

Finally, this results in terms of reconstructed images given in Fig. 7, for systems with $6\ \mathrm{dB}$ and $12\ \mathrm{dB}$ SNRs, provide some subjective comparisons. Here, the advantage of the MD allocation scheme can be observed by comparing the reconstructed images with the original one (Fig. 4). The four reconstructed images in Fig. 7 represent four different simulated results, already provided in terms of PSNR score in Fig. 6. For both the allocation schemes, it can be visually observed that increasing the SNR value improves the quality of the reconstructed image. More important, for the same mean SNR value, it is noticeable the gain that the MD allocation scheme achieves compared to the Basic one. The gain is remarkable in very challenging environments, such as SNR=6 dB, in which the double FEC protection of the MD allocation schemes makes it possible to reconstruct the image at a more than decent quality, Fig. 7(c). On the contrary, the Basic allocation scheme does not provide sufficient protection against channel impairments for systems with SNR=6 dB, see Fig. 7(a).

To conclude, this study aims at investigating and comparing the performance of two allocation techniques for the SPIHT algorithm.

Compared to the basic method, MD is found to be more suitable for underwater acoustic systems in terms of PSNR and robustness against channel variability. For both methods, the mean SNR was found to be a sufficient information for the FEC profile optimization. This can be justified by a limited variation of the considered channel, and it might not be verified in cases of fast fading transmissions. Moreover, the Basic allocation technique turns out to be sensitive to the correlation between packets, leading to a substantial difference between simulated and theoretical curves.

4. CONCLUSION

In this work we have performed a simulation study in order to compare the performance of two allocation techniques for the SPIHT image coding in case of underwater acoustic communications. We have shown how much the MD allocation outperforms the basic allocation and we have verified the i.i.d. assumption for the error process at both the byte and the packet level. Moreover we have seen, in case of fast channel fluctuations, how robust the techniques are to imperfect CSI.

Acknowledgment

The authors would like to thank Dr. Hanumant Singh for providing the underwater images. This work has been supported in part by the Italian Institute of Technology within the Project SEED framework (NAUTILUS project), by ONR under Grants no. N00014-05-10085 and N00014-10-10422 and by the European Commission under the Seventh Framework Programme (Grant Agreement 258359 – CLAM).

5. REFERENCES

[1] R. M.Eustice, O. Pizarro, H. Singh, "Visually Augmented Navigation for Autonomous Underwater Vehicles," in *IEEE J. of Ocean. Eng.*, vol. 33, no. 2, pp. 103–122, Apr., 2008.

[2] J. Ribas, D. Sura, M. Stojanovic, "Underwater wireless video transmission for supervisory control and inspection using acoustic OFDM," in *Proc. IEEE OCEANS*, Seattle, USA, Sep., 2010.

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

[4] M. van Der Schaar, N. Shankar , "Cross-layer wireless multimedia transmission: challenges, principles, and new paradigms," in *IEEE Wireless Commun. Mag.*, vol. 12, no. 4, pp. 50–58, Aug., 2005.

[5] M. Stojanovic, J. Preisig, "Underwater acoustic communication channels: Propagation models and statistical characterization," in *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 84–89, Jan., 2009.

[6] K. Hobin, R. Annavajjala, P. Cosman, L. Milstein, "Source-channel rate optimization for progressive image transmission over block fading relay channels," in *IEEE Trans. Commun.*, vol. 58, no. 6, pp. 1631–1642, Jun., 2010.

[7] C. Murphy, H. Singh, "Wavelet Compression with Set Partitioning for Low Bandwidth Telemetry from AUVs," in *Proc. WuWNET'10*, Woods Hole, USA, Oct., 2010.

[8] D.F. Hoag, V.K. Ingle, R.J. Gaudette, "Low-bit-rate coding of underwater video using wavelet-based compression algorithms," in *IEEE J. of Ocean. Eng.*, vol. 22, no. 2, pp. 393–400, Apr., 1997.

[9] Q.-Z. Li, W. Bing, W.-J. Wang, G. Xiao-ling, "An Efficient Underwater Video Compression Algorithm for Underwater Acoustic Channel Transmission," in *Proc. Intern. Conf. on Commun. and Mob. Comp.*, Jan., 2009.

[10] Q.-Z. Li, W.-J. Wang, "Low-bit-rate coding of underwater color image using improved wavelet difference reduction," in *IEEE J. of Visual. Commun. and Image Repres.*, vol. 21, no. 7, pp. 762–769, 2010.

[11] D. G. Sachs, A. Raghavan, K. Ramchandran, "Wireless Image Transmission Using Multiple-Description Based Concatenated Codes," in *Proc. of the Conf. on Data Compression*, Jan., 2000.

[12] B. Tomasi, P. Casari, L. Badia, M. Zorzi, "A Study of Incremental Redundancy Hybrid ARQ over Markov Channel Models Derived from Experimental Data," in *Proc. WuWNET'10*, Woods Hole, USA, Oct., 2010.

[13] L. Toni, Y. S. Chan, P. C. Cosman, L. B. Milstein, "Channel Coding for Progressive Images in a 2-D Time-Frequency OFDM Block With Channel Estimation Errors," in *IEEE Trans. Image Process.*, vol. 18, no. 11, pp. 2476–2490, Nov., 2009.

[14] A. Said and W. A. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," in *IEEE Trans. Circuits Syst. Video Technol.*, vol. 6, pp. 243–249, Jun., 1996.

[15] A. E. Mohr, E. A. Riskin, R. E. Ladner, "Unequal loss protection: graceful degradation of image quality over packet erasure channels through forward error correction," in *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 819–828, Jun., 2000.

[16] Y. S. Chan, P. C. Cosman, L. B. Milstein, "A cross-layer diversity technique for multi-carrier OFDM multimedia networks," in *IEEE J. Trans. on Image Process.*, vol. 15, no. 4, pp. 833–847, Apr., 2006.

[17] www.whoi.edu/mvco.