Seafloor Image Compression with Large Tilesize Vector Quantization

Chris Murphy
Deep Submergence Laboratory
Woods Hole Oceanographic Institution
Woods Hole, MA
Email: cmurphy@whoi.edu

Robert Y. Wang
CSAIL
Massachusetts Institute of Technology
Cambridge, MA
Email: rywang@mit.edu

Dr. Hanumant Singh
Deep Submergence Laboratory
Woods Hole Oceanographic Institution
Woods Hole, MA
Email: hanu@whoi.edu

***Abstract*— Autonomous Underwater Vehicles (AUVs) often communicate with scientists on the surface over an unreliable acoustic channel. The challenges of operating in deep waters, over long distances, and with surface ship noise amount to a communication channel with a very low effective bandwidth. This restriction makes transmission of images, even highly compressed images, quite difficult. We present an image compression algorithm designed to convey the gist of an image to surface operators in a very small number of bytes. Our technique divides a large existing database of underwater images into 'tiles', and uses these to reconstruct an approximation to new underwater images from a similar domain. We achieve significantly higher compression ratios than conventional image compression techniques, such as JPEG or SPIHT, while still being able to provide useful visual feedback to the surface.**

I. Introduction

For those who study the seafloor, underwater robots provide a crucial pair of eyes in an otherwise forbidding environment. Robots allowing scientists from all disciplines to capture images in environments as diverse as lively Puerto Rican coral reefs, historically valuable shipwrecks, or the barren volcanic seafloor beneath Arctic ice caps. Tethered underwater robots operate near a surface ship and can transmit data to waiting scientists on the surface while receiving power and commands in return. Remotely Operated Vehicles (ROVs) range from shallow-water commercial models to deep-sea research vehicles like the JASON II [1], [2], and enable human operators to explore the depths almost as if they were there. As live imagery is received on video monitors, scientists can begin to form hypotheses from new observations, develop plans for upcoming dives, and even alter the plan for what remains of the current dive.

This contrasts strongly with typical Autonomous Underwater Vehicle (AUV) missions, where the lack of a physical tether can cut data throughput by seven orders of magnitude or more. A tethered vehicle communicating with single-mode fiber optics can easily support gigabits of data per second, yet autonomous vehicles typically communicate over acoustic links at speeds of tens or hundreds of bits per second. The lack of a tether, however, allows AUVs to reach areas that are inaccessible to ROVs, such as under Arctic ice sheets [3]. The lack of information on the surface means that AUVs are typically preprogrammed as a set of waypoints, leaving few mission planning decisions to be made while underwater.

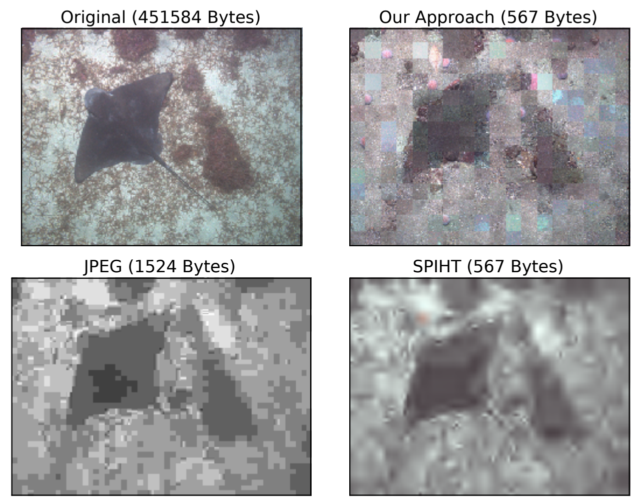

Fig. 1. A comparison of our method used to compress a 432x336 image of a ray, versus standard JPEG compression and the SPIHT wavelet coder.

We seek a way to involve surface scientists in the high-level decision-making processes for AUVs. Providing the operator with contextual data, beyond the basic location and health information currently translated, is essential to this goal. AUVs designed and optimized for seafloor photographic mapping can easily capture high dynamic range images every 3-5 seconds. Transmitting each of these images at full resolution to the surface over an acoustic link is not currently practical, but we, and others, have had success transmitting highly compressed versions of individual images. Automatically identifying which images should be sent up poses a difficult and task-specific computer vision challenge.

This paper presents a method for encoding "high-resolution thumbnails" of seafloor imagery, designed to give the operator a rough idea of the contents of each transmitted image. This method provides extremely high compression of seafloor imagery which, unlike traditional image compression methods, exploits the high levels of inter-dive redundancy. The method is inspired by the vector quantization literature, but differentiates itself from past work by using large quantization vectors, or 'tiles', much like an image mosaic. Compression artifacts are reduced during decoding, resulting in a caricature of the original image. While a naïve implementation requires a time-consuming and computationally difficult search, several modifications to increase suitability for subsea operation are proposed. Results are shown for images captured during two AUV dives in the Channel Islands National Marine Sanctuary. The first dive took place over the submerged wreck of a World War II aircraft shown in Figure 2, and the second over nearby rocky fish habitat as shown in Figure 3.

Fig. 2. The submerged wreck of an Avenger torpedo bomber, lost in the Channel Islands National Marine Sanctuary and Park. Images courtesy NOAA Northwest Fisheries Science Center, mosaic by Chris Murphy, Clay Kunz and Hanumant Singh.

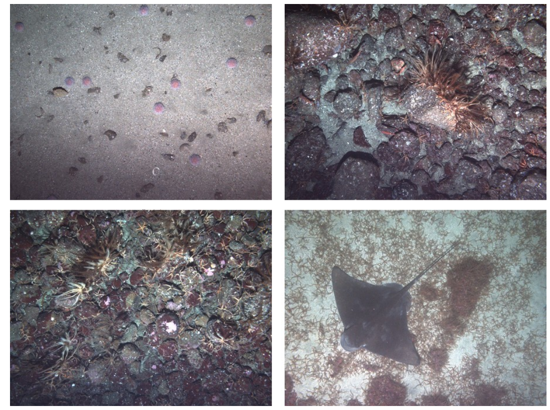

Fig. 3. Example images captured during an AUV dive in (generally) rocky habitat.

II. Related Work

A. Underwater Communications

the ocean environment presents numerous challenges to acoustic communication, including low available bandwidth and large propagation delays [4]. These challenges are made worse by operating over long distances [5] and by environmental conditions such as seafloor makeup and water depth. AUV and surface ship noise transmit directly into the channel, further exacerbating the problem. As a result, use of long-range underwater communication is characterized by extremely low effective bandwidth, high latency, and frequent packet loss.

To accomodate the peculiarities of the medium, channel coding methods with high rates of error-correction are typically employed. While underwater acoustic communications has achieved rates up to hundreds of kilobits per second [6], reliable acoustic communications over long distances currently requires the use of low-rate communications with high error tolerance, such as frequency-hopping frequency shift keying (FH-FSK) or highly error-corrected phase shift keying (PSK). In addition, AUVs may rely on acoustic navigation schemes such as LBL [7] or USBL. Since the ocean is a shared broadcast medium, time-multiplexing of the channel for navigation or communication with other vehicles may be required, which lowers effective bit-rates further. The WHOI Micro-Modem, used by Seabed-class AUVs, uses low frequency bandwidth to allow for multiple senders and receivers. It is capable of sending one 256-bit FH-FSK packet in slightly over 3 seconds, or one 1536-bit error-tolerant PSK packet in slightly over 6 seconds, delivering an effective bit-rate between 80 and 256 bits per second. Commercially available options from Teledyne-Benthos and LinkQuest advertise 80-360 bits per second for environments with harsh multi-path. Advances in coding theory bring increased bitrates, but there is always a tradeoff between enhanced reliability and higher bitrates. Summarizing data for transmission at ultra low bit-rates, especially when time-multiplexed with acoustic navigation methods, presents a significant hurdle. As a result, current telemetry is often quite limited.

B. Acoustic Telemetry

During many AUV deployments, a surface operator monitors simple vehicle telemetry to track the AUV and watch for anomalies. Perhaps the most widely used standard for this telemetry is the Compact Control Language (CCL) [8], which defines a number of standardized methods for encoding vehicle state and health, as well as samples of bathymetry, salinity, and other data [9]. In addition, several 256-bit packets for AUV to AUV, and AUV to surface-ship, communications are specified. While CCL is adequate for transmitting individual datapoints from an AUV, or for transmitting basic commands to an AUV, it is not particularly efficient. CCL relies only upon quantization to provide compression and makes no use of the inherent correlation between successive samples from most instruments.

There has been extensive experimentation with the transmission of still and video imagery over relatively high bandwidth (1-10kbps) acoustic tethers. In 1992, researchers from NEC presented a system for transmitting low-resolution compressed images from the Shinkai 6500 submersible [10]. Researchers at WHOI have developed high speed prototype acoustic tethers capable of transmitting video [11]. In addition, Hoag, Ingle et al. have extensively studied the application of wavelet compression techniques to underwater images [12] and video sequences [13]. Craig Sayers, and others at the University of Pennsylvania, developed techniques for selecting specific frames and 'regions of interest' from a video sequence that best describe an ROV manipulator and environment state, and transmitted these regions to surface operators over a 10000 bit per second acoustic tether as JPEG images [14]. These techniques don't generalize well to the ultra-low bandwidth situations we are confronted with, though we compare our technique to both JPEG and wavelet compression.

III. Method

Our approach decomposes each image into a grid of tiles, and encodes each tile as the index of a visually or semantically similar, but previously captured, image tile. We are inspired by the widely used JPEG standard, which encodes images as a grid of 8x8 patches. During JPEG encoding, each 8x8 patch is converted into a frequency domain representation by projecting it onto a series of cosine basis functions, using the Discrete Cosine Transform (DCT). The resulting coefficients are quantized and compressed using run length and Huffman encoding. While our technique is also a patch-based approach, our patches are larger and instead of quantizing the coefficients, we use the coefficients to look up the most similar entry in a database of tiles taken from other related images – a form of vector quantization.

Vector quantization represents an arbitrary point in high-dimensional space by the index of the nearest point in a preselected set, or library, of basis points. Widely used in audio compression and early video compression standards, vector quantization has been applied to image compression by decomposing an image into small (e.g. 4x4) vectors, and encoding using vectors from a previously selected basis set. Our approach uses tiles (or vectors) much larger in size – up to 40 pixels square. Our library of precaptured tiles is generated from previous dives in nearby or similar areas, and may contain tens or hundreds of thousands of tiles. Large image databases have been used extensively in computer vision for applications such as object recognition [15], semantic image completion [16], and image geolocation [17]. Torralba and colleagues have also explored ways of accelerating nearest neighbor image searching in a large database [18]. To the best of our knowledge, such an approach has been used to construct "photomosaic" puzzles and to similar ends, but not for image compression in this manner.

The tile library is shared between the AUV and the receiver prior to the dive. Tile similarity can be calculated using a variety of methods; we found L_2 distance to be effective, though computationally challenging for an embedded system. We therefore use Principle Components Analysis (PCA) to accelerate tile comparisons, and identify the best match for each source tile. Tile indices are transmitted to the surface, where indexes are transformed back into tiles, and a caricature, or 'high-resolution thumbnail', of the original image is generated. We present two distinct methods, based upon gradient-preserving blur Poisson editing techniques and image quilting, for reducing compression artifacts in these reconstructed images.

A. Tile Database

The first step of our technique is therefore to construct a database of tiles which describe the space of images we intend to encode. On N images of consistent dimension $w_i \times h_i$, we impose a grid with square dimension $w_t \times h_t$. Computing a tile database given an image set amounts to sampling sub images along a uniform grid. The image set is ideally of the same type of scene (e.g. coral reefs) as the images to be compressed. For many applications, such a training set is readily available. For instance, on an expedition consisting of several dives, an image sequence captured on a dive on the first day can be used to build a database used on subsequent days. In practice, we construct sample databases upwards of 100,000 tiles, depending on tile and image dimensions. The final size of a single compressed image is governed by Equation 1.

$$\frac{\text{bits}}{\text{pixel}} = \overbrace{\frac{1}{w_t \times h_t}}^{\text{Tiles per pixel}} \times \overbrace{\left\lceil \log_2 \Big(\underbrace{N \times \left\lfloor \frac{w_i}{w_t} \right\rfloor \times \left\lfloor \frac{h_i}{h_t} \right\rfloor}_{\text{Tiles in library}} \Big) \right\rceil}^{\text{Bits per tile}} \quad (1)$$

In addition to a large database of tiles, we require a notion of similarity or distance between tiles. This metric is used to determine the nearest neighbors to a query tile. We are interested in metrics that can capture the gist of a tile while being fast to evaluate. Ideally, we would like to use the L_2 distance metric across all of the pixels in an image so that the nearest neighbor selected minimizes the root mean square (RMS) reconstruction error. However, the L_2 distance is too expensive to use due to the high dimensionality of the tiles, and the large library size. Instead we approximate the L_2 distance using PCA by projecting each tile into a low-dimensional space, allowing encoding on commodity hardware on board an AUV. We could use any metric that captures the gist of a tile while being relatively fast to evaluate. A SIFT descriptor [19] or other feature-based approach would be effective, and could result in visually different but still semantically correct tile matches.

We use Principal Components Analysis (PCA) to accelerate the evaluation of the L_2 distance. Specifically, we use PCA to optimally compress (in the L_2 sense) our set of high dimensional image tiles into low dimensional feature vectors. We proceed as follows:

First we flatten each RGB color image tile x with dimension $n \times m$ into $3nm \times 1$ vectors. We pack these vectors side-by-side into a matrix X of size $3nm \times N$ where N is the size of our training set. We mean center this matrix by subtracting the mean column from each column.

$$X_m(i,j) = X(i,j) - \bar{X}_j$$

We then use the Nonlinear Iterative Partial Leaset Squares (NIPALS) algorithm to compute a basis $P = [p_1, p_2, \ldots, p_k]$ such that,

$$X_m = PT + E$$

where $T \in \mathbb{R}^{k \times N}$ is a matrix containing the low-dimensional feature vector representation of our training set and E is the residual.

A new tile query y can be projected on to the principal components P to generate a compressed representation $t(y) = P^T y$ comparable with the low-dimensional representation of the training set. The result is that the L_2 distance metric can be efficiently evaluated by comparing only a few components of low-dimensional vectors instead of taking the pixel-wise difference of two image tiles:

$$\begin{aligned} d(x_1, x_2) &= \|x_1 - x_2\|_2 \\ &\approx \|Pt(x_1) - Pt(x_2)\|_2 \\ &= \|P\| \, \|t(x_1) - t(x_2)\|_2 \\ &= \|t(x_1) - t(x_2)\|_2 \end{aligned}$$

As the number of components approaches the dimensionality of the data set, the PCA distance metric is guaranteed to converge to the L_2 distance metric. The results of approximating a sample tile with increasing numbers of PCA components are shown in Figure 4. Finally, the nearest image tile indices are packed bitwise, and transmitted over the acoustic link.

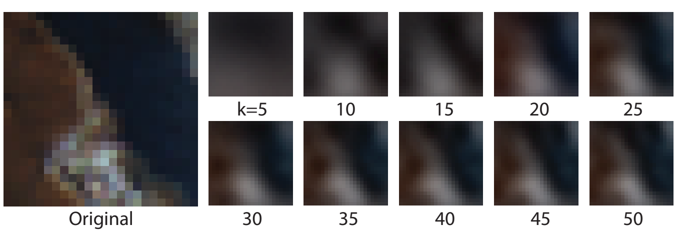

Fig. 4. Shown above is a visual depiction of the reconstruction fidelity of a sample tile as we increase the number of PCA components.

B. Block Artifact Reduction

Due to the tile-based nature of our compression algorithm and the constraints of extremely low bit rates, the images have noticable "block artifacts" along the grid structure. We can ameliorate this issue with post-processing. We explored two techniques: a gradient-preserving blur and an image quilting technique. These post processing techniques are performed by the receiver, typically a surface ship where computation power is less restricted, after an image has been received. While both techniques reduced the blocking artifacts, we relied upon the gradient preserving blur as the results were more visually appealing and only minimally affected the texture detail.

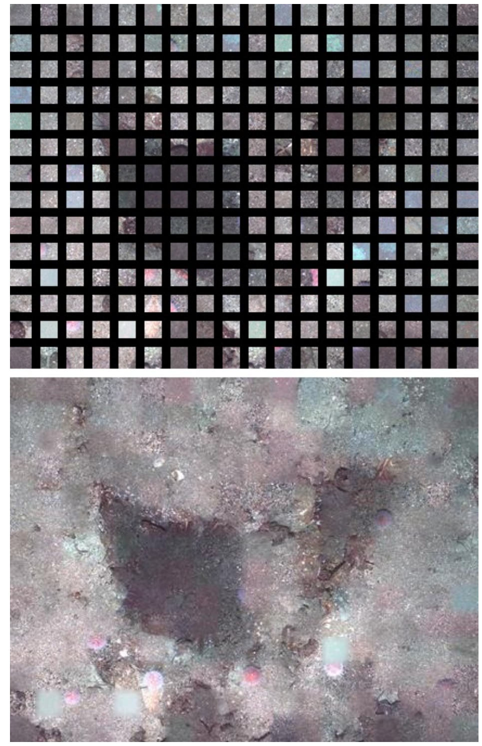

Fig. 5. Our gradient-preserving blur technique reduces the block artifacts over the indicated grid.

1) Gradient-Preserving Blur: One approach to reducing the block artifacts is to blur the image along the grid. However, this produces an equally disturbing artifact where the pixels along the grid are noticably too smooth. Ideally, we would like to smooth the less noticable low-frequency content while preserving sharp boundaries. We achieve this technique by applying a blur that preserves image gradients. We follow the Poisson image editing technique as described by Perez and colleagues [20].

Poisson image editing interpolates a region of an image $f \in \Omega$ constrained at the boundary $\partial\Omega$ with Direchlet boundary conditions f^*. The interpolated region is made to follow a vector guidance field $\mathbf{v}$ which in our case are composed of the image gradients of each tile (see Figure 6).

$$\min_f \iint_\Omega \|\triangledown f - \mathbf{v}\|^2 \text{ with } f|_{\partial\Omega} = f^*|_{\partial\Omega}$$

We discretize the above equation along the pixel grid. Let f_p be the value at pixel p and $\langle p, q \rangle$ be the set of neighboring pixel pairs.

$$\min_{f|_\Omega} \sum_{\langle p,q \rangle \cap \Omega \neq \emptyset} (f_p - f_q - v_{pq})^2 \text{ with } f_p = f_p^*, \text{ for all } p \in \partial\Omega$$

For all $\langle p, q \rangle$, we set

$$v_{pq} = \begin{cases} f_p - f_q & \text{if tile}(p) = \text{tile}(q), \\ 0 & \text{otherwise.} \end{cases}$$

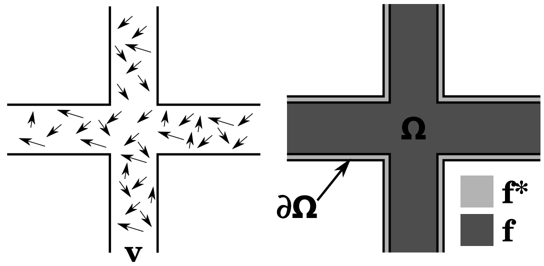

Fig. 6. To reduce block artifacts without oversmoothing, the gradients $\mathbf{v}$ of each tile are preserved while blurring the region between tiles Ω. We constrain the blur to respect the boundary $\partial\Omega$.

That is, we respect all image gradients as best we can except those between two pixels of from different tiles. The minimization above is given as the solution to a system of sparse linear equation. We solve these equation separately for each color channel using direct LU factorization. The results are shown in Figure 5. Our implementation takes several seconds per image to compute on a modern laptop processor.

2) Image Quilting: The results of our compression are reminiscent of texture synthesis literature. We are, in some ways, attempting to identify and resynthesize different substrate textures based upon our previously captured tiles. Thus inspired, we tested Efros and Freeman's image quilting [21] method as a way of reducing the blocking artifacts. In [21], the match fidelity between neighboring tiles is improved by constraining neighbors to agree with each other within some error tolerance. While such an approach could be incorporated into our technique, we currently have no such constraint. Image tiles are presently selected solely on the quality of their match to the image being compressed. The results of performing this method on a single image are shown in Figure 7. The white lines in the top image indicate the location of the minimum-error cut.

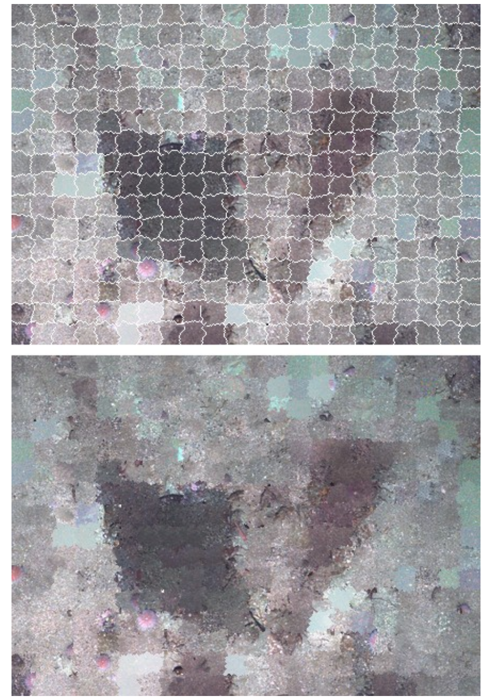

Fig. 7. Blocking artifact reduction with the minimum error boundary cut. At top, the minimum error boundaries are highlighted; when adjacent tiles are mated along these cuts, the bottom image is the result.

IV. Results

As shown in Figure 1, standard JPEG compression does not perform well at very low bitrates. The JPEG standard highly quantizes color information, meaning that at low bitrates color information is almost entirely gone, or badly distorted. High-resolution details are also entirely lost, leaving large blotches of different brightnesses, with few discernable features. The severe losses in coding quality observed in JPEG images when compressed to more than 0.15–0.20 bits per pixel (40:1) are well documented [12]. More modern image compression methods of course exist, including wavelet-based methods such as SPIHT, but all trade off texture information at high compression rates. We hoped to preserve texture information in our resulting thumbnails, while still being able to discern objects of interest.

We demonstrate our method over a set of images collected by a SeaBED AUV [22] during dives near Channel Islands National Park, off the California coast. The first dive was performed over the wreck of a World War II era fighter plane, shown in Figure 2. 150 randomly selected images from this dive were used to calculate PCA features as described previously. Images from the dive were then converted to tiles, and each tile was projected onto the generated PCA basis. The top PCA features that were generated are shown in Figure 8.

The results were saved as the tile library, shared by both the encoder and decoder. The second dive was over nearby rocky habitat, with the goal of studying fishery health. A selection of interesting images (ie - not purely sand) were hand picked from those captured during the dive. Each image was encoded, by projecting tiles onto the PCA basis and matching to those tiles acquired in the first dive. The results of encoding using this method are shown in Figure 9. As the source image resolution (448x336) was not an integer multiple of all tile sizes, the bottom and right edges were trimmed in some images.

The results of our trials are difficult to quantify, as appropriate metrics are largely application dependent. Still, as a method for obtaining an arbitrarily small representation of an image, it holds promise. To begin to quantify the encoding efficiency, we calculated the average per-pixel RMS error for twelve images encoded with a variety of tile sizes, and tile library sizes. The results, shown in Figure 10, suggest that larger tile sizes are more efficient at encoding an image to the same number of bits. This suggests that pursuing a large library of large tiles may hold promise.

V. Conclusions

In this paper we have presented a method for obtaining very high compression levels from imagery, scalable to arbitrarily

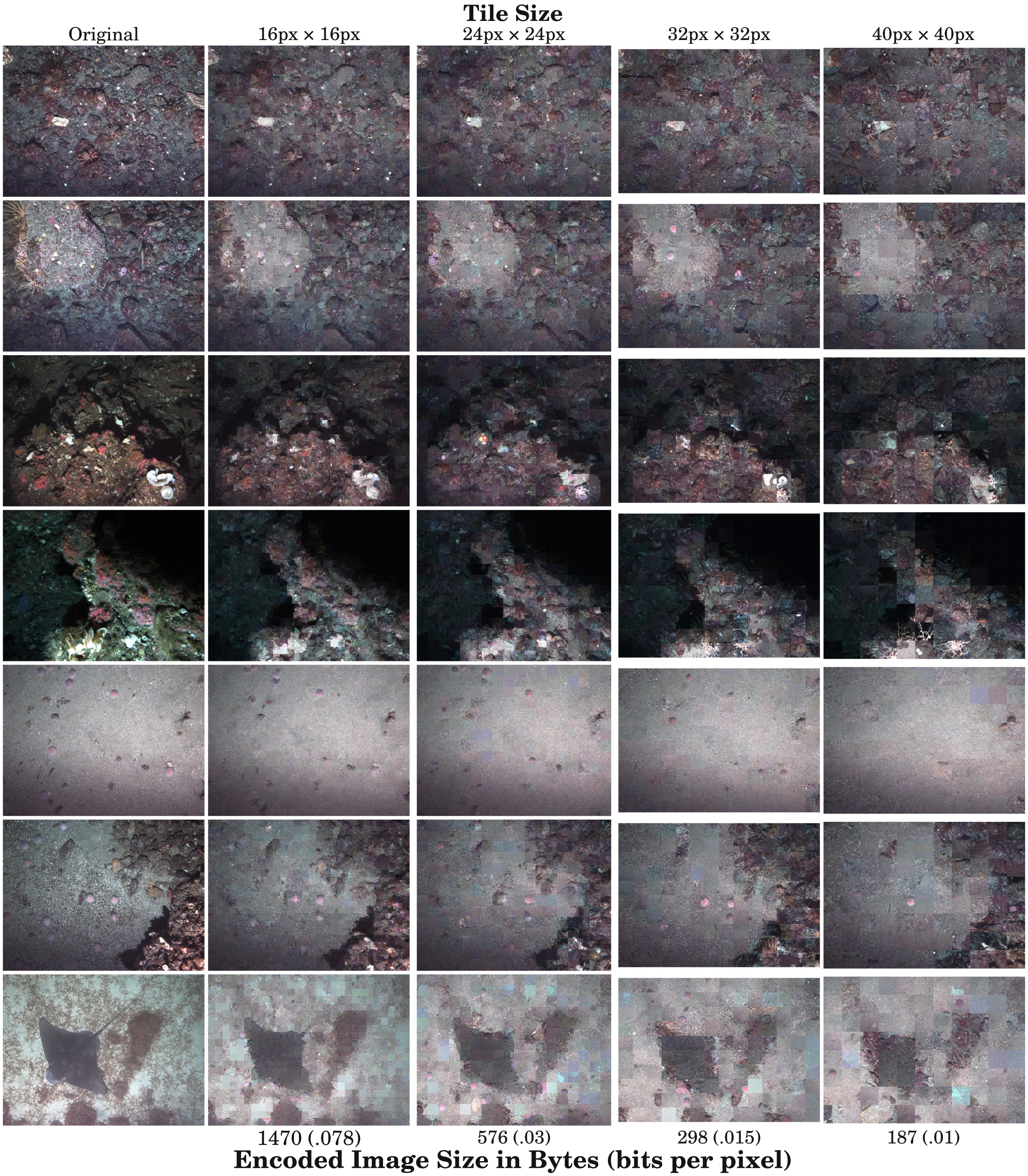

Fig. 9. Results of encoding a selection of images using this method.

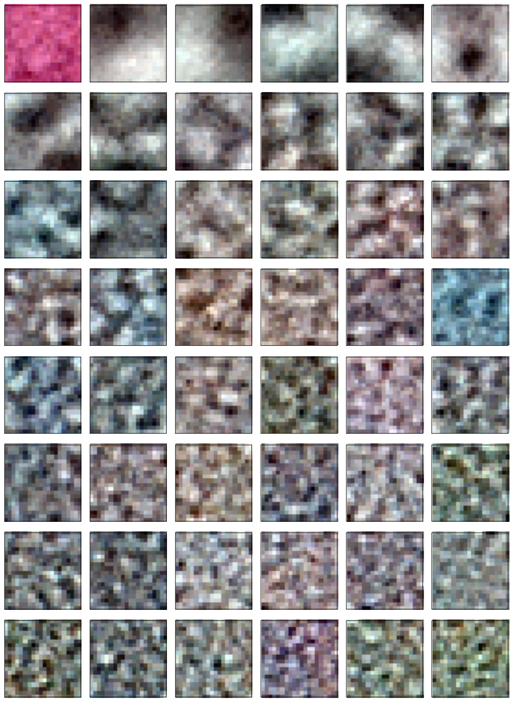

Fig. 8. Shown above are the first forty-eight components of the basis we computed from our training set, working from left to right and then down the page.

low bandwidth. Seafloor imagery is highly repetitive; we present one way that image compression can exploit that fact. While the compression is less effective on detailed objects in the imagery, like fish or manmade items, the method seems well suited to more repetitive areas of images, such as substrate.

VI. Future Work

During library generation, we currently do not attempt to discard any duplicate or similar tiles from the library. This is likely wasteful; careful pruning of the tile library could reduce the number of bits necessary to represent a tile. This is similar to the approach taken in normal vector quantization, though actual tiles would continue to be used rather than summarized centroids.

To compute tile similarity, we used the L_2 distance metric. Were more semantically meaningful information can be obtained from a tile, such as substrate classification or identified objects, image encoding could be performed based on generating a semantically similar image, rather than simply a visually similar image. This approach suffers from a lack of generality however, as any encoding would likely be mission-specific. Pizarro et al. have had luck applying [23] Bag-of-Words models to seafloor imagery; a similar approach may prove fruitful.

Acknowledgment

The authors would like to thank the *National Science Foundation Center for Subsurface Sensing and Imaging* (CenSSIS ERC). This work was funded by the CenSSIS ERC through grant EEC-99868321.

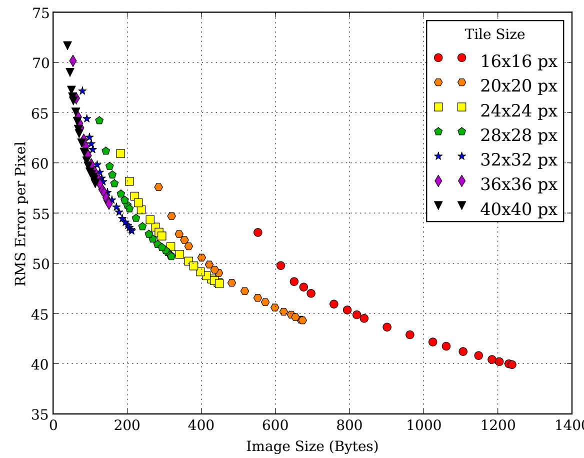

Fig. 10. As the size of the tile library grows, so does the number of bits necessary to represent each tile. This data suggests that the library size is less of a concern than the size of the tiles, however.

References

[1] R. D. Ballard, D. R. Yoerger, W. K. Stewart, and A. Bowen, "ARGO/JASON: a remotely operated survey and sampling system for full-ocean depth," in *OCEANS, 1991. Proceedings of MTS/IEEE*, 1991, pp. 71–75.

[2] B. Elder, A. Bowen, M. Heintz, M. Naiman, C. Taylor, W. Seller, J. Howland, and L. Whitcomb, "Jason 2: a review of capabilities," in *EOS: Trans. Amer. Geophysical Union Fall Meeting Supplement*, 2003, abstract.

[3] J. Bellingham, M. Deffenbaugh, J. J. Leonard, and J. Catipovic, "Arctic under-ice survey operations," *Unmanned Systems*, vol. 12, pp. 24–29, 1994.

[4] I. F. Akyildiz, D. Pompili, and T. Melodia, "Underwater Acoustic Sensor Networks: Research Challenges," in *Ad Hoc Networks*. Elsevier, Mar. 2005, vol. 3, no. 3, pp. 257 – 279.

[5] M. Stojanović, "On the relationship between capacity and distance in an underwater acoustic communication channel," in *SIGMOBILE Mobile Computing and Communications Review*, vol. 11, no. 4, Oct. 2007, pp. 34 – 43.

[6] ——, "Recent advances in highspeed underwater acoustic communications," *IEEE Journal of Oceanic Engineering*, vol. 21, no. 2, pp. 125–136, Apr. 1996.

[7] P. H. Milne, *Underwater Acoustic Positioning Systems*. Houston, TX: Gulf Publishing Co., 1983.

[8] R. Stokey, L. Freitag, and M. Grund, "A Compact Control Language for auv acoustic communication," *Oceans 2005 - Europe*, vol. 2, pp. 1133–1137, June 2005.

[9] R. P. Stokey, "A compact control language for autonomous underwater vehicles," Woods Hole Oceanographic Institution," Specification, Apr. 2005, http://acomms.whoi.edu/40x%20Specifications/401100%20Compact%20Control%20Language/CCL%20April%202005%20Public%20Release%201.0.pdf.

[10] M. Suzuki, T. Sasaki, and T. Tsuchiya, "Digital acoustic image transmission system for deep-sea research submersible," *Oceans 1992, Proceedings of the MTS/IEEE*, vol. 2, pp. 567–570, Oct 1992.

[11] C. Pelekanakis, M. Stojanovic, and L. Freitag, "High rate acoustic link for underwater video transmission," *OCEANS 2003. Proceedings*, vol. 2, pp. 1091–1097, Sep 2003.

[12] D. F. Hoag and V. K. Ingle, "Underwater image compression using the wavelet transform," in *OCEANS '94. Proceedings*, vol. 2, Sep 1994, pp. 533–537.

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

[14] C. Sayers, A. Lai, and R. Paul, "Visual imagery for subsea teleprogramming," in *Proc. IEEE Robotics and Automation Conference*, 2005.

[15] A. Torralba, R. Fergus, W. Freeman, and C. MIT, "80 Million Tiny Images: A Large Data Set for Nonparametric Object and Scene Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 11, pp. 1958–1970, 2008.

[16] J. Hays and A. A. Efros, "Scene completion using millions of photographs," in *SIGGRAPH '07: ACM SIGGRAPH 2007 papers*. New York, NY, USA: ACM, 2007, p. 4.

[17] J. Hays and A. Efros, "IM2GPS: estimating geographic information from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008*, 2008, pp. 1–8.

[18] A. Torralba, R. Fergus, and Y. Weiss, "Small codes and large image databases for recognition," in *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008*, 2008, pp. 1–8.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[20] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," *ACM Trans. Graph.*, vol. 22, no. 3, pp. 313–318, 2003.

[21] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *28th annual conference on Computer graphics and interactive techniques (SIGGRAPH)*, 2001, pp. 341–346.

[22] H. Singh, A. Can, R. Eustice, S. Lerner, N. McPhee, O. Pizarro, and C. Roman, "Seabed AUV offers new platform for high-resolution imaging," *EOS, Transactions of the AGU*, vol. 85, no. 31, pp. 289,294–295, Aug. 2004.

[23] O. Pizarro, S. Williams, and J. Colquhoun, "Topic-based habitat classification using visual data," in *IEEE OCEANS 2009 Europe*, may 2009, pp. 1–8.