Online Data Summaries for Semantic Mapping and Anomaly Detection with Autonomous Underwater Vehicles

Jeffrey W. Kaeli[1] and Hanumant Singh[1]

***Abstract*— Autonomous underwater vehicle (AUV) operations are inherently bandwidth limited but increasingly data intensive. This leads to large latencies between the capture of image data and the time at which operators are able to make informed decisions using the results of a survey. As AUV endurance and reliability continue to improve, there is a greater need for real-time on-board data processing capabilities. In this paper, we apply online data summary techniques to optical and acoustical imagery collected by AUVs, then show how they can be used to both create low-bandwidth semantic maps and to detect anomalies on the seafloor.**

I. INTRODUCTION

Seventy percent of the Earth's surface is covered by water, below which lie diverse ecosystems, rare geological formations, important archeological sites, and a wealth of natural resources. Understanding and quantifying these areas presents unique challenges for the robotic imaging platforms required to access such remote locations. Low-bandwidth acoustic communications prevent the transmission of images in real-time, while the large volumes of data collected often exceed the practical limits of exhaustive human analysis. As a result, the paradigm of underwater exploration has a high *latency of understanding* between the capture of image data and the time at which operators are able to gain a visual understanding of the survey environment.

A robotic vehicle capturing one optical still image every few seconds can easily generate thousands of images within a matter of hours. This sheer volume of data presents a formidable obstacle to any individual attempting to gain an understanding of the survey environment. Often, when a vehicle operator obtains a dataset for the first time, their instinct is to quickly scan thumbnails of the images for any that "pop out." While this can be useful to detect obvious anomalies, it is not necessarily the best or fastest way to obtain images that "represent" the data in a meaningful way. In this paper, we explore the use of online summaries [1], [2], [3] as a framework for both mapping and anomaly detection by maintaining a small subset of the images that exceed some threshold of novelty when they are first encountered.

II. RELATED WORK

A. Underwater Communications

Without a physical link to the surface, AUVs rely on acoustic signals to communicate with shipboard operators. These channels have very limited bandwidth with throughput on the order of tens of bytes per second depending on range, packet size, other uses of the channel (for instance, navigation sensors), and latencies due to the speed of sound in water [4], [5]. While much higher data rates have been achieved using underwater optical modems for vehicle control [6] and two-way communication [7], these systems are limited to ranges on the order of 100 meters and are inadequate for long-range communication [8]. In the absence of mission-time operator feedback, an AUV must either navigate along a preprogrammed course or use the data it collects to alter its behavior. Examples of the latter, termed adaptive mission planning, include detecting mines so potential targets can be re-surveyed in higher-resolution [9] and using chemical sensors to trace plumes back to their source [10], [11]. The overarching implication is that, with the exception of low-bandwidth status messages, data collected by an AUV is not seen by operators until after the mission is completed and the vehicle recovered.

B. Clustering Data

Clustering can be viewed as an unsupervised compression strategy that allows multidimensional data to be quantized to one of several discrete distributions by defining a distance metric between samples and minimizing some measure of that distance. We can think of each image or section of an image as a data point characterized by some distribution of features, such as a quantized descriptor (which itself could have been obtained through clustering). One of the most well-known clustering algorithms is the K-means algorithm which seeks to find a set of cluster centers that minimize the within-class distances between each cluster center and the members of its representative class [12]. While this method has been extremely useful in generating texton dictionaries for texture analysis [13], [14], the fact that the cluster centers are not guaranteed to occur at a data point makes mapping back to a single representative image for each class difficult. A similar algorithm, k-medoids, only considers data points as potential cluster centers, and is more useful for generating representative images. Both of these methods require the number of cluster to be set *a priori*.

Other methods seek to determine the number of clusters based on the natural structure of the data. Affinity propagation accomplishes this by picking "exemplars" that are suggested by nearby data points [15] and has found use in building texton vocabularies [16]. Hierarchical methods have also been used to learn objects [17], scenes [18], and underwater habitats [19] based on topic models using Latent Dirichlet Allocation (LDA) [20]. However, a drawback of

[1]Jeffrey W. Kaeli and Hanumant Singh are with the Woods Hole Oceanographic Institution, Woods Hole, MA 02543, USA {jkaeli,hsingh}@whoi.edu

all methods mentioned thus far is that they operate upon a static dataset. This "offline" approach is ill-suited to real-time robotic imaging because it offers no way to characterize the dataset until after all the data has been collected.

Clustering data in an "online" fashion provides two important benefits. Firstly, it allows data to be processed continuously throughout the mission, reducing the overall computational load. Secondly, at any point in time it provides a summary of the imagery captured thus far by the vehicle. A drawback to online methods is that they offer less guarantees of stability and are ultimately dependent upon the order in which images are presented to the algorithm [21]. The worst-case scenario for online approaches would be for the most extreme data points to occur first, followed by interior points which become poorly represented. Fortunately, natural underwater environments are highly redundant with habitat domains that persist across many frames. One possible approach uses incremental clustering of topic models using LDA [3]. We are particularly interested in recent work on navigation summaries [1], [2] which operate on the concept of "surprise."

C. Surprise-Based Summaries

An event can be said to be "surprising" because it happens unexpectedly. The idea of what is expected can be modeled as a probability distribution over a set of variables and considered as *prior* knowledge about the world. When a novel event occurs, it augments this body of knowledge and creates a slightly different *posterior* knowledge of the world. If the amount of knowledge added by any single event is large enough, that event can be said to be unexpected and thus is "surprising."

This concept has been formalized in a Bayesian framework as the difference between the posterior and prior models of the world [22]. For measuring this difference, the Kullback-Leibler divergence, or relative entropy, was shown to correlate with an attraction of human attention,

$$d_{KL}(p \parallel q) = \sum_x p(x) log \frac{p(x)}{q(x)} \tag{1}$$

$$\Phi = \sum_i d(p \parallel n_i) \tag{2}$$

where $p(x)$ is the posterior model, $q(x)$ is the prior model, and x is some observed variable over which distributions can be computed. Rather than modeling the prior knowledge Π^- as a single distribution $P(F)$ over a set of features F, we follow [1] and model it over each member of summary set $\boldsymbol{S}$ containing M members.

$$\Pi^- = \{P(F|S_1), \cdots P(R|S_M)\} \tag{3}$$

The posterior knowledge Π^+ is simply the union of prior knowledge with the new observation Z

$$\Pi^+ = \{P(F|S_1), \cdots P(R|S_M), P(F|Z)\} \tag{4}$$

The set theoretic surprise ξ can be defined as the Hausdorff distance between the posterior and prior distribution using the KL divergence as a distance metric [1]. The Hausdorff metric is a measure of the distance between two sets based on the greatest possible difference between one point in the first set to the nearest point on the other sets. Since the prior and posterior sets differ only by Z, the surprise can be simply expressed as the KL distance between observation Z and the nearest summary image in $\boldsymbol{S}$.

When a new observation's surprise exceeds a threshold, it is added to the summary set. The threshold is generally set as the lowest value of surprise in the current summary. That member of the old summary set with the lowest surprise is then removed and replaced by the new observation, and the surprise threshold set to the next least-surprising member of the summary set. In this manner, a temporally global summary of the images is maintained at all times [1].

III. SEMANTIC MAPPING

A. Modified Online Summaries

Our interest in creating online summaries is motivated by advances in image compression and acoustic communications that facilitate the transmission of images during a mission [23]. Summary images transmitted during a mission can serve as map bases such that each non-summary image is indexed as belonging to one of the summary set types. In this way, a low-bandwidth semantic map can be created to give an operator a fast, high-level understanding of the survey environment while a mission is still underway.

There are several drawbacks to existing approaches that make them ill-suited for picking which images to transmit during a mission. First, the summary represents a dynamic set of images, so there is no guarantee that an image that is transmitted will remain a member of the summary set throughout the rest of the mission. Second, simply transmitting images based on the highest "surprise" value can result in a handful of "outlier" images that are not representative of the dominant habitats in a survey. Lastly, if our goal is to use these summary images as the bases for building a semantic map to spatially characterize the survey environment, we need a means of reliably classifying non-summary images online as well.

Our first modification is to represent each non-summary image with a member of the summary set. Assuming that we have navigation data available to be transmitted as well, we can combine these representations with the approximate vehicle position to create spatial coverage maps based on the summary set. Intuitively, a non-summary image should be best represented by the summary image that is most similar. Representing a non-summary image by its nearest neighboring summary in this way can be thought of as minimizing the surprise one would have when looking through all the non-summary images represented by a given summary image.

We next must determine which summary images to transmit. Obviously, it is desirable to transmit the first image as soon as possible to minimize the latency of understanding for the operator. However, early in the mission the surprise

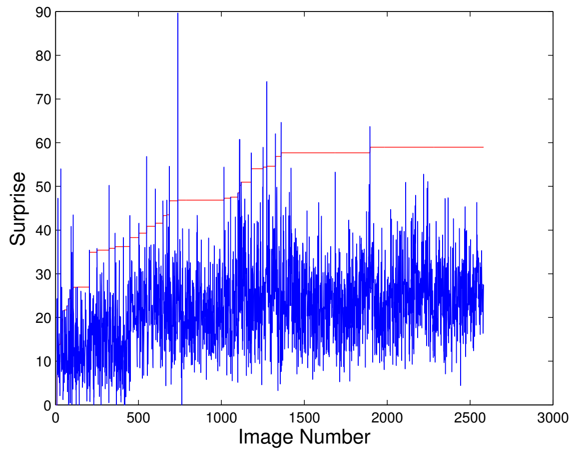

Fig. 1. Surprise as a function of image number. The threshold of surprise grows as more images are incorporated into the summary set.

threshold grows rapidly as the algorithm determines which images best represent the data. Thus, we wait until the surprise threshold does not change for a specified number of images, implying that the vehicle is imaging consistent terrain that could be represented well by a single image.

For subsequent images, we assume that the vehicle will be ready to transmit another image after a set number of frames. We would like to choose a summary image that is different enough from the previously transmitted summary images while at the same time representative of enough non-summary images to make it a worthwhile choice for a map basis. This can be formulated to minimize the surprise one would have when looking through the other summary images. We are effectively choosing a summary subset within the summary set. However, simply choosing the summary image that minimizes this surprise does not guarantee that it represents enough non-summary images to make it a useful basis for the map. Hence, we select the summary set that both minimizes the Haussdorff distance when the summary set is partitioned into subsets as well as represents enough non-summary images to exceed a given threshold.

Selecting good summary images to transmit is important because these images will be used to represent the entire dataset for the duration of the mission. Furthermore, this means that, as new summary images are added to the summary set, previously transmitted summary images should not be removed from the summary set given the high cost of transmitting an image. Subsequently, after a summary image is transmitted, it becomes "static," as opposed to the other "dynamic" summary images.

Online summary methods do not require distances to be recomputed for all existing data points when new data appears which is one quality that makes them attractive for power-limited underwater robots. Thus, when a new summary is added to the set, we would rather not lose the information we have gained by simply removing the least-surprising summary image and the non-summary images that it represents. Instead, we propose to merge it with the nearest summary image so that it and its non-summary images all become non-summary images represented by the nearest summary image.

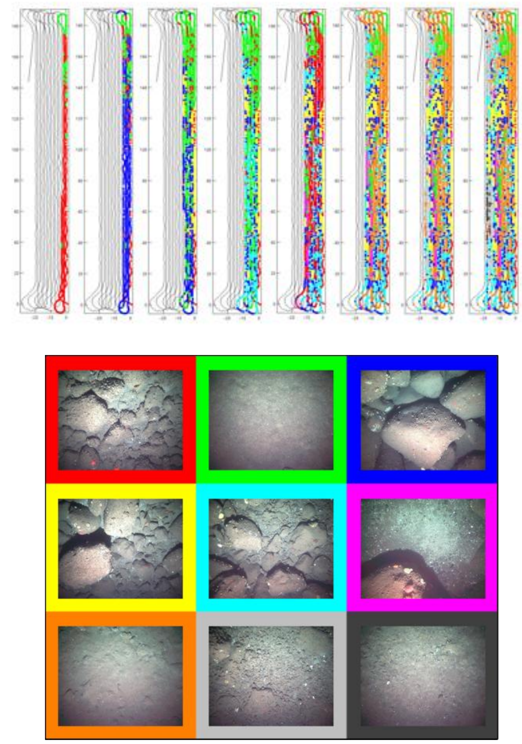

Fig. 2. Semantic maps created after each subsequent image is transmitted (top) with summary images and respective color codes (bottom).

B. Generating Semantic Maps

We implemented this new approach on a 2800 image dataset collected by the SeaBED AUV [24] in 2003 in the Stellwagen Marine Sanctuary. The survey consisted of multiple track lines over various habitats composed of boulders, rubble, sand, and mud. For each image, we computed 1000 keypoints, accumulated a histogram of oriented gradients around each keypoint, and quantized each to one of 14 binary patterns [25]. A global histogram was then computed for the entire image and used as the model distribution $P(F)$. Considering that images are captured every 3 seconds, the total mission time to capture 2800 images is over 2 hours. With the current state of the art in acoustic image transmission being approximately one full-resolution 1-megapixel image every 15 minutes [23], we estimated that about 8 images could be transmitted back within the course of a mission. We set the summary set size to twice that to allow for some flexibility.

The summary set is initialized with the first 16 images and their corresponding surprise values are set to the smallest surprise measured relative to the set of images before it. Progress then continues throughout the rest of the data until the surprise threshold is exceeded by a novel image. When this happens, the novel surprising image is incorporated into

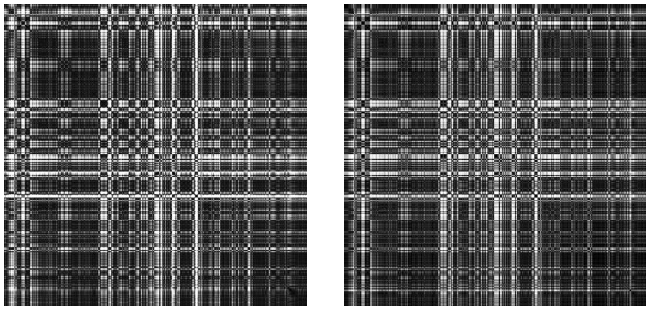

Fig. 3. Surprise values between all images (left) and symmetric surprise between all images (right) using only the surprise values from their representative summary set.

the summary set, the least surprising image removed, and the surprise threshold augmented to the new lowest surprise value within the set as previously described. Figure 1 plots the surprise value and threshold throughout the course of the mission. As more of the environment is surveyed, the more surprising a new image must be to become incorporated into the summary set.

Figure 2 shows the resulting progressive semantic maps created after each subsequent image and corresponding data are transmitted. The first image (red) was transmitted when the surprise threshold stabilized after 147 images. Each subsequent transmitted image was chosen after 300 frames had elapsed, simulating a realistic 15 minute transmission interval [23]. The first map is based on the first (red) and second (green) images, the second on the first three, and so on, until all 9 images are used.

To show that our overall approach preserves distance information, we plot the surprise distance between all 3000+ images in Figure 3. At left, distances have been calculated between each image. At right, the distances for each image have been replaced by their representative summary image's distances. Remarkably, the structure within the dataset is preserved quite well given the almost 30,000:1 compression ratio.

Some of these classes are similar and the operator may wish to merge them for visual clarity. In Figure 4 the 9 transmitted images have been heuristically merged into 5 distinct classes: (from top to bottom at right) sand, piled boulders, lone boulders in sand, mud, and rubble. From the complete mosaic and the bathymetric map, it is clear that the piled boulders correspond to the tops of ridges. Depths in the bathymetric map range from 60 meters (warmer hues) to 70 meters (colder hues). Between these ridges are sandy areas, all of which are bordered by mud and smaller rubble.

This level of dataset understanding would be extremely valuable for an operator to possess during a mission. For instance, if the boulder fields were of particular interest to a scientist, the vehicle could be issued a redirect command to resurvey that area at higher resolution. Conversely, if a particular substrate of interest is not being imaged, the mission can be terminated and the vehicle recovered and relocated to another area. Furthermore, upon recovery of

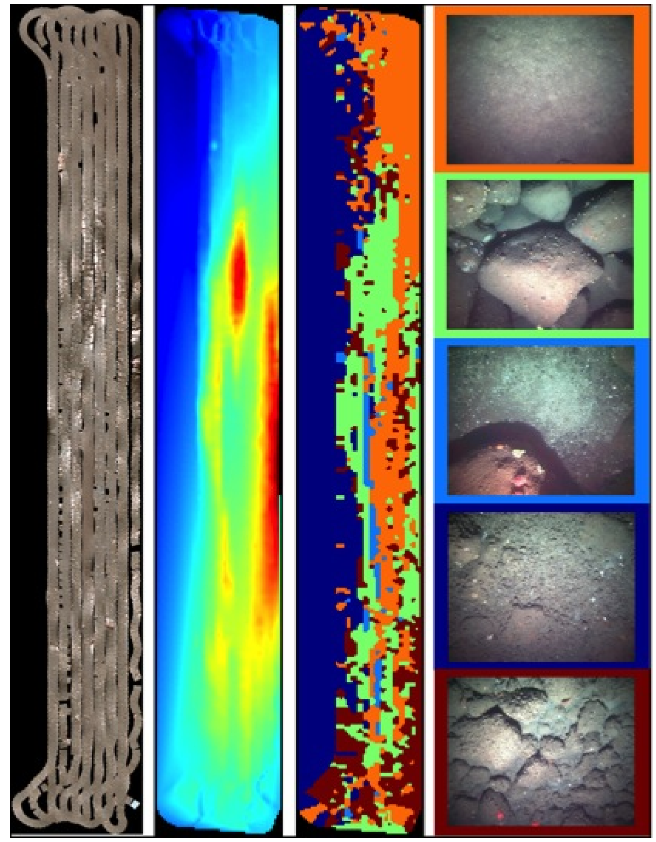

Fig. 4. Photomosaic (left) and bathymetry (middle left) of the entire mission. The final semantic map (middle right) using 9 images which have been heuristically merged into 5 distinct classes (right) and color coded.

the vehicle, the operator has a fully classified dataset with additional summary images as well. The non-summary images represented by each summary images can be browsed to check the class validity. Several randomly selected non-summary images have been chosen from each of the 5 summary sets in Figure 4 and are shown in Figure 5.

We have described modifications which enable us to select summary images to transmit that characterize the diversity in the dataset and will not change as additional summary images are added and merged. After the first image is transmitted and received, an operator has an initial understanding of the survey environment. After the second image is transmitted and received, additional scalar data containing navigation and classification information can be compressed and transmitted as well, providing the operator with ample information to begin to construct a spatial map of the survey environment. The classification masks exhibit high redundancy and covariance so they can be compressed at high rates. These data can be transmitted using very little bandwidth with the techniques presented in [26] and [23].

IV. ANOMALY DETECTION

A. Semantic Mapping and Anomaly Detection

Semantic mapping and anomaly detection can be viewed as complementary problems. In the previous section we introduced a method to generate semantic maps by using

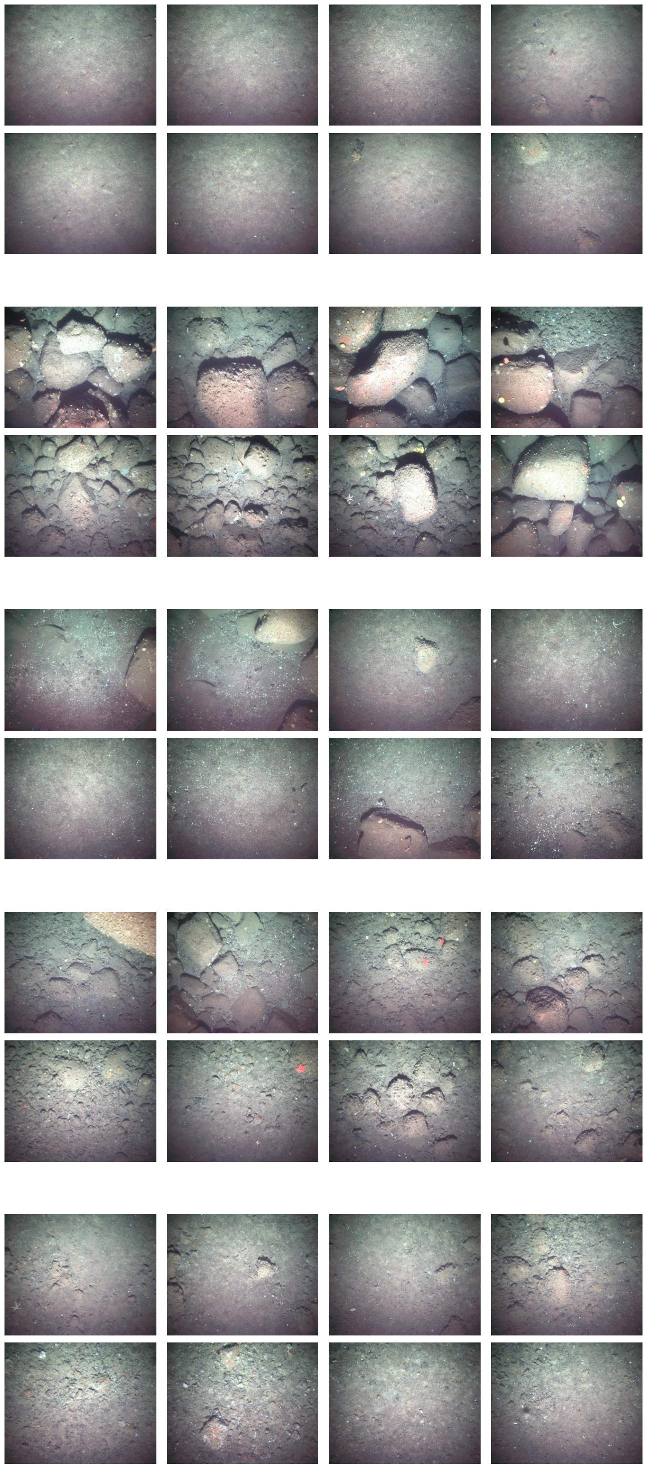

Fig. 5. Example imagery from each of the five heuristically merged classes.

modified online summaries to select a set of images that best represents the survey environment. While this approach selects for images that are “surprising,” it is biased against “anomalies,” or surprising images which serve poorly as summary images because few or no non-summary images are similar to them. If we instead retain these images in the summary set and how “surprising” they were when first

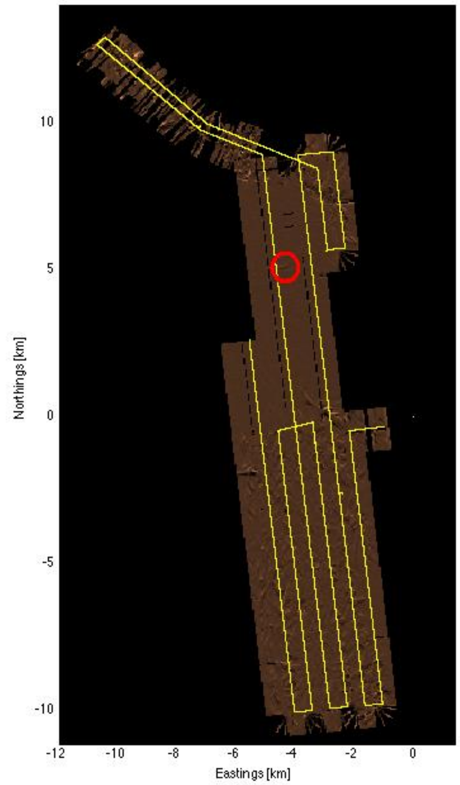

Fig. 6. Side scan sonar mosaic of the survey. Track lines are shown in yellow, with the vehicle starting at (0,0). The known location of a downed airplane is indicated by the red circle.

encountered, we can build up a map of how well the vehicle “understands” the survey environment. Such a map could inform a robot making adaptive multi-resolution surveys of a region, suggesting a “closer look” at regions which it does not “understand” as well as others and avoiding areas which fit well into its current model of the world.

We demonstrate a framework for this approach using part of a dataset collected by a 120 KHz side scan sonar from a REMUS 600 AUV at 70m altitude. At this altitude, the swatch width is approximately 700m to each side. Each return is resampled to an approximately square grid given the distance traveled between pings, in this case 1.5m, to account for the slant range. Intensity distortions of the port and starboard channels are corrected for using a running mean updated at each ping. Altitudes that were not between 65m - 75m were ignored. This can be seen in the northwest corner of the mosaic in Figure 6 where the AUV had difficulty maintaining a constant altitude over the rugged terrain.

Consecutive “good” pings are stored in a buffer until there are enough to form an image tile. Histograms of oriented gradients were computed in a dense fashion across each tile. Each histogram is treated as a distribution and quantized according to a summary of previously observed histograms. Each quantized pattern is then added to its corresponding map bin based on the navigational information of the AUV. Once a map bin has not been augmented after a ping, the distribution of features is also quantized according

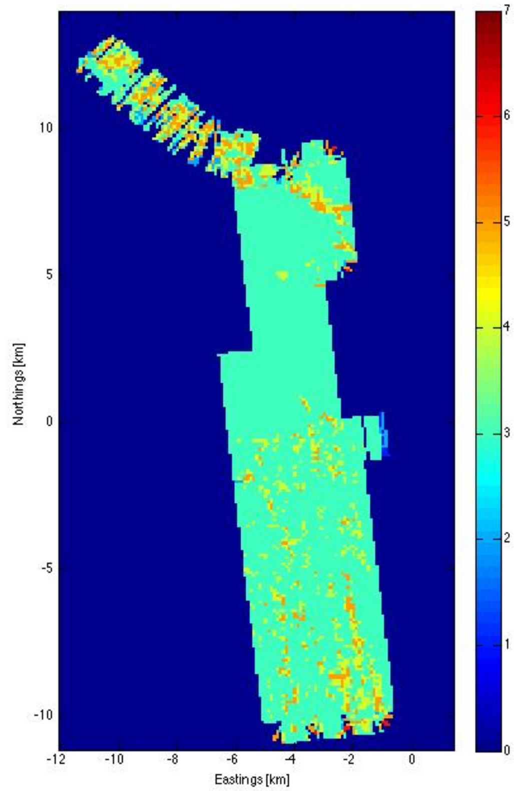

Fig. 7. Semantic map of the survey environment.

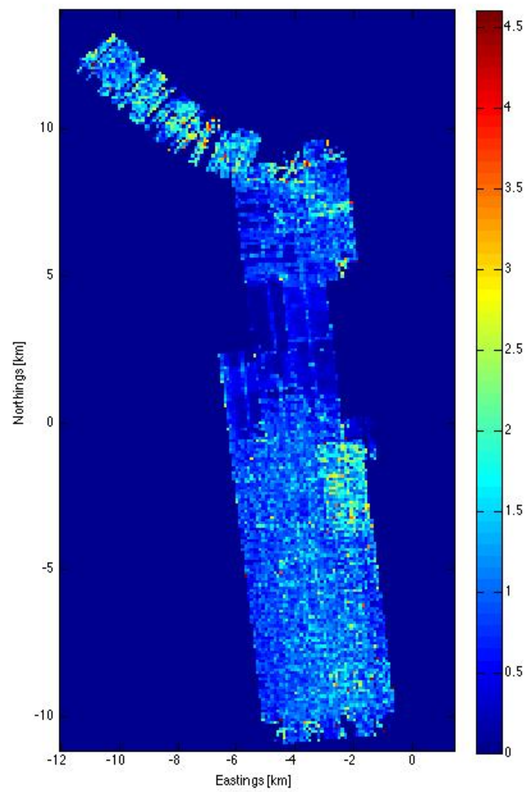

Fig. 8. Maximum surprise for each map bin. Note how values are low in the flat basin in the middle of the transect and higher in rocky aras and at the airplane. The higher values at the beginning of the survey are the inital surprise of the robot exploring somewhere for the first time.

to a summary of previously observed bins. This two-level approach allows both new "texton" [27], [13], [28] features and new substrate "texture" classes to be "learned" as new regions are explored.

Figure 7 shows the semantic map generated by the algorithm for the same mosaic shown in Figure 6. Map bins are in 100m spacings. The majority of the bins seem to correspond well with flatter areas in the mosaic, where several other classes correspond with rocky areas. A downed airplane in the survey has been classified as similar to rocky terrain. Figure 8 shows the maximum surprise value recorded at each bin. Lower values of surprise can be thought of as where the AUV "understands" what it is observing. It is not surprising that the AUV understands the flatter areas better than the rocky areas. It is also interesting to note that the AUV is surprised to encounter the airplane relative to the surrounding terrain.

V. DISCUSSION

This work makes contributions to the field of autonomous underwater robotics by describing a framework that can be used to reduce the "latency of understanding," or the time delay between when an image is captured and when it is finally "understood" by an operator. This latency is propagated from two sources: first, from the low-bandwidth of the acoustic communication channel which greatly restricts the throughput of data; second, from the large volume of image data that must be analyzed. The second source has been addressed by numerous automated classification algorithms designed to annotate image data in an offline post-processed sense. The first source has been addressed by recent compression work allowing a small set of images to be transmitted over the course of a mission. We have addressed both of these sources by describing a lightweight framework designed to run in real time aboard a robotic vehicle that can produce environmental maps based on a subset of summary images and infer areas that merit additional study.

In Figures 7 and 8 it can be seen that the alrorithm is very sensitive to noisy data. While we have taken care to only use pings captured within an acceptable range of altitudes, additional ping filtering and quality control would likely produce cleaner results. Furthermore, one would intuitively hope that a downed airplane would be classified as an anomaly apart from rocky areas. We are continuing to explore the use of more descriptive features as well as possibly adding a third summary layer that learns the distribution of classified map bins.

While optical imagery offers excellent resolution at close range, the high attenuation of light in water makes it less ideal as a large area search sensor. Acoustic imaging sensors such as side scan sonars are much better suited to mapping large areas. In one possible paradigm, a side scan sonar could be used for a high altitude search to characterize large areas and detect anomalies which are subsequently imaged at lower altitudes by a camera. Another area where these approaches could be useful is the selective offloading of data after a mission.

Existing techniques approach the visual summary problem strictly as a visual summary problem; we approach it from a compression standpoint in two contexts. Firstly, a robot vehicle's ability to communicate a high-level understanding of its environment given the limitations of acoustic modems. Secondly, a robot vehicle's ability to understand its environment and resurvey areas it doesn't understand. Our work represents an enhancement of the capabilities of robotic vehicles both to explore new environments and to improve the quality of operator involvement during vehicle missions.

ACKNOWLEDGMENT

We are grateful to Clay Kunz for his generous help in generating photomosaics of the data. We are also grateful to Mike Purcell for providing the sidescan sonar imagery.

REFERENCES

[1] Y. Girdhar and G. Dudek, "Onsum: A system for generating online navigation summaries," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 746–751.

[2] ——, "Efficient on-line data summarization using extremum summaries," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 3490–3496.

[3] R. Paul, D. Rus, and P. Newman, "How was your day? online visual workspace summaries using incremental clustering in topic space," in *Robotics and Automation (ICRA), 2012 IEEE International Conference on*. IEEE, 2012, pp. 4058–4065.

[4] L. Freitag, M. Grund, S. Singh, J. Partan, P. Koski, and K. Ball, "The whoi micro-modem: an acoustic communications and navigation system for multiple platforms," in *OCEANS, 2005. Proceedings of MTS/IEEE*. IEEE, 2005, pp. 1086–1092.

[5] M. Stojanovic, "Recent advances in high-speed underwater acoustic communications," *Oceanic Engineering, IEEE Journal of*, vol. 21, no. 2, pp. 125–136, 1996.

[6] M. Doniec, C. Detweiler, I. Vasilescu, and D. Rus, "Using optical communication for remote underwater robot operation," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 4017–4022.

[7] M. Doniec and D. Rus, "Bidirectional optical communication with aquaoptical ii," in *Communication Systems (ICCS), 2010 IEEE International Conference on*. IEEE, 2010, pp. 390–394.

[8] N. Farr, A. Bowen, J. Ware, C. Pontbriand, and M. Tivey, "An integrated, underwater optical/acoustic communications system," in *OCEANS 2010 IEEE-Sydney*. IEEE, 2010, pp. 1–6.

[9] L. Freitag, M. Grund, C. von Alt, R. Stokey, and T. Austin, "A shallow water acoustic network for mine countermeasures operations with autonomous underwater vehicles," *Underwater Defense Technology (UDT)*, 2005.

[10] J. A. Farrell, S. Pang, and W. Li, "Chemical plume tracing via an autonomous underwater vehicle," *Oceanic Engineering, IEEE Journal of*, vol. 30, no. 2, pp. 428–442, 2005.

[11] M. V. Jakuba, "Stochastic mapping for chemical plume source localization with application to autonomous hydrothermal vent discovery," Ph.D. dissertation, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, 2007.

[12] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd ed. John Wiley & Sons, 2001.

[13] T. Leung and J. Malik, "Representing and recognizing the visual appearance of materials using three-dimensional textons," *International Journal of Computer Vision*, vol. 43, pp. 29–44, 2001, 10.1023/A:1011126920638. [Online]. Available: http://dx.doi.org/10.1023/A:1011126920638

[14] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images," *International Journal of Computer Vision*, vol. 62, pp. 61–81, 2005, 10.1023/B:VISI.0000046589.39864.ee. [Online]. Available: http://dx.doi.org/10.1023/B:VISI.0000046589.39864.ee

[15] B. J. Frey and D. Dueck, "Clustering by passing messages between data points," *science*, vol. 315, no. 5814, pp. 972–976, 2007.

[16] N. C. Loomis, "Computational imaging and automated identification for aqueous environments," Ph.D. dissertation, MIT/WHOI Joint Program in Oceanography / Applied Ocean Science & Engineering, 2011.

[17] J. Sivic, B. Russell, A. Zisserman, W. Freeman, and A. Efros, "Unsupervised discovery of visual object class hierarchies," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, june 2008, pp. 1 –8.

[18] L. Fei-Fei and P. Perona, "A bayesian hierarchical model for learning natural scene categories," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 2. IEEE, 2005, pp. 524–531.

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

[20] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003. [Online]. Available: http://dl.acm.org/citation.cfm?id=944919.944937

[21] S. Zhong, "Efficient online spherical k-means clustering," in *Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on*, vol. 5. IEEE, 2005, pp. 3180–3185.

[22] L. Itti and P. F. Baldi, "Bayesian surprise attracts human attention," in *Advances in neural information processing systems*, 2005, pp. 547–554.

[23] C. Murphy, "Progressively communicating rish telemetry from autonmous underwater vehicles via relays," Ph.D. dissertation, MIT/WHOI Joint Program in Oceanography / Applied Oceans Science and Engineering, 2012.

[24] H. Singh, R. Eustice, C. Roman, and O. Pizarro, "The seabed auv–a platform for high resolution imaging," *Unmanned Underwater Vehicle Showcase*, 2002.

[25] J. W. Kaeli, "Computational strategies for understanding underwater optical image datasets," Ph.D. dissertation, Massachusetts Institute of Technology and Woods Hole Oceanographic Institution, 2013.

[26] T. E. Schneider, "Advances in integrating autonomy with acoustic communications for intelligent networks of marine robots," Ph.D. dissertation, Massachusetts Institute of Technology, 2013.

[27] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, pp. 91–97, 1981.

[28] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 31, no. 11, pp. 2032 –2047, nov. 2009.