Real-Time Video Mosaicking of the Ocean Floor

Richard L. Marks * Stephen M. Rock † Michael J. Lee ‡

Abstract This research proposes a method for the creation of real-time video mosaics of the ocean floor. New vision processing hardware is the enabling technology which makes real-time video mosaicking conceivable. A complete system has been developed to create mosaics from a single camera source. The mosaicking system was tested using a remotely operated vehicle. The test was targeted to determine the capabilities of the system in an actual scientific marine setting. The chief concerns were whether marine snow and nonuniform lighting of the field of view would hinder mosaic creation. Several mosaics of portions of a brachiopod field were successfully created.

I. INTRODUCTION

Visual mosaicking of the ocean floor is useful to many applications. Scientists use mosaics to gain an understanding of ocean biology and geology. Mosaics are also used to help visualize ship wreckage. Just as satellite photography has helped to give a macro understanding of the Earth's terrain, so has video mosaicking helped with the underwater terrain. Because of the limited visual range underwater (due to light attenuation), video mosaicking is necessary in order to gain this macro view.

The ability to produce video mosaics in real time offers several additional advantages. Because intermediate mosaic results can be observed while images are being gathered for the mosaic, steps can be taken to ensure that no gaps occur between images. Areas of particular interest can be determined immediately and be given special attention. Real-time video mosaicking can also be used to construct visual maps to aid navigation of underwater vehicles.

II. BACKGROUND

Video mosaicking has been conducted for many years. Manual mosaicking consisted of actually laying pictures side by side and taping them together [1]. Recently, computer-aided image processing has enabled the digital merging of stored images. This has allowed mosaics to be created as a single, large image. However, merging these images manually is quite tedious and time-consuming.

Recent research has addressed the idea of automatic mosaic creation. UNH proposed to create mosaics from images taken at precise positional coordinates [2]. However, the inaccuracies in position measurements may be large compared to the size of a video pixel. This would lead to unacceptable offset errors between images. Other researchers have explored panoramic scene representation, in which a visual history of a robot's motion is created[3]. However, this visual history is created solely for the purpose of robot navigation and does not correspond to a real-world visual mosaic of the traversed area.

Impressive automatic mosaics have been created at NASA Ames by using advanced image processing algorithms to match multiple interesting features in overlapping images. However, the computational requirements of these algorithms are extremely demanding, requiring minutes of supercomputer time. Other work has been done in automatic mosaicking of pairs of satellite images[4]. Examples are show for images with large translation, scaling, and rotation differences. These approaches are performed off-line by post-processing recorded data. Because there is no time constraint, high quality mosaics can be created by warping and adjusting the images to fine resolution. However, these approaches cannot guarantee that there will be no "gaps" in the data. Also, they do not provide real-time feedback of the macro view, which is often needed to decide what areas need to be explored more thoroughly.

The mosaicking system described below addresses all of these issues. It provides real-time mosaic feedback and guarantees no gaps (because the on-line vision processing determines when the next frame should be stored). The offsets for each image in the mosaic are determined directly from image correlation rather than from a position measurement from another sensor. In addition, it outputs a vector of pixel offsets for the optimal next image location which can be used as feedback for camera positioning.

*Doctoral Candidate, Department of Aeronautics and Astronautics, rlmarks@sun-valley.Stanford.Edu

†Associate Professor, Department of Aeronautics and Astronautics, rock@sun-valley.Stanford.Edu

‡Senior Research Engineer, Monterey Bay Aquarium Research Institute, lemi@hp850.Mbari.Org

III. THE VIDEO MOSAICKING SYSTEM

A. Visual Correlation

The key technology in our video mosaicking system is *visual correlation.* Visual correlation has been used in many applications. The most common uses are for the computation of optical flow and for stereo ranging. Optical flow is generated by correlating images taken separated in time by a fixed interval. Stereo ranging is generated by correlating images taken from different locations. For video mosaicking, a stored image is correlated with live camera images to compute the camera's positional offset in pixels.

1) Signum of Laplacian of Gaussian prefilter: Visual correlation of raw camera images has several problems. The correlation strength is sensitive to slight variations in the scene. Changes in orientation, scale, or lighting drastically lowers the correlation measurement quality. These problems are compounded in the ocean where marine snow, nonuniform lighting, and lack of contrast also degrade raw correlation measurements.

These problems can be overcome by filtering the images before correlation occurs. A filter with many useful properties is the two-dimensional spatial signum of Laplacian of Gaussian filter (Equation 1) [5] [6].

$$\mathbf{I}'(x,y) = \mathbf{sgn}[\bigtriangledown^2 * \frac{e^{\frac{-x^2-y^2}{2\sigma^2}}}{2\pi\sigma^2} * \mathbf{I}(x,y)] \qquad (1)$$

$\mathbf{I}(x,y)$ corresponds to the original image intensities, and $\mathbf{I}'(x,y)$ corresponds to the output intensities. The Gaussian convolution smoothes the image, acting like a low-pass filter. The frequency roll-off is adjusted by varying σ, the width of the Gaussian. Increasing the width reduces the amplitudes of high spatial frequencies in the image (such as marine snow). The signum of Laplacian portion of the filter converts the smoothed image into a binary image with black/white transitions occurring at intensity gradient local maxima. This reduces the effects of lighting biases and limited contrast. Figure 1 shows an example of a filtered image.

2) Special purpose hardware: Special purpose prefiltering and correlation hardware is used in the video mosaicking system. This hardware, obtained from Teleos Research, performs real-time signum of Laplacian of Gaussian filtering with a filter size of up to 40 by 40 pixels. It also provides 20,000 correlations of 100 by 100 pixel windows per second. A more thorough description of the capabilities of the Teleos system can be found in a paper on underwater object tracking[7]. This powerful hardware is the enabling technology which makes real-time video mosaicking feasible.

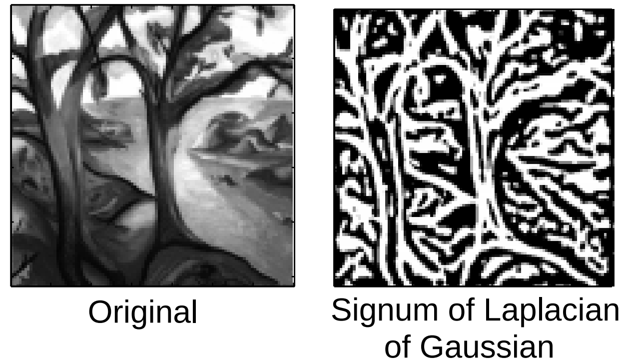

Figure 1: An image and the corresponding signum of Laplacian of Gaussian filtered image.

B. Mosaicking Algorithm

Figure 2 shows a schematic of the video mosaicking system. Images from a single camera are continuously digitized and filtered. Sections of the live image are correlated with sections of a stored image (which is also filtered). If the scene in the stored image is identical to the scene in the live image, maximum correlation will occur when a section in the live image is correlated with a section at the same location in the stored image. If the scene in the live image is slightly offset, however, maximum correlation will occur between sections which are offset (see Figure 3).

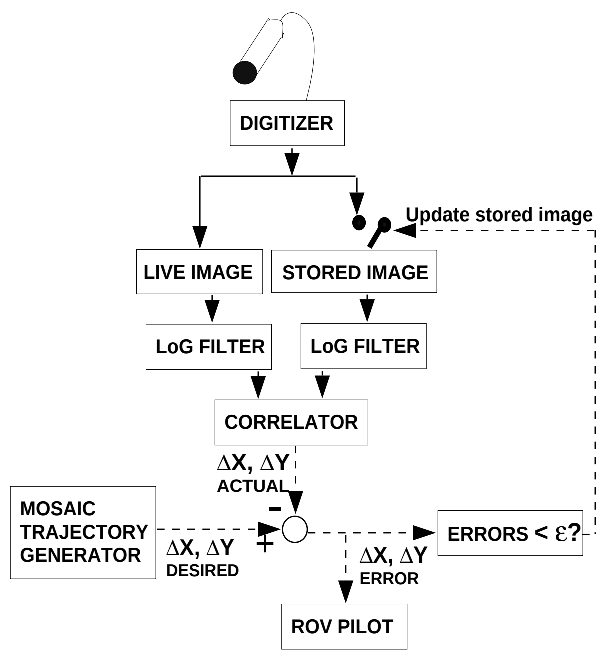

Figure 2: Block diagram for a single column mosaic.

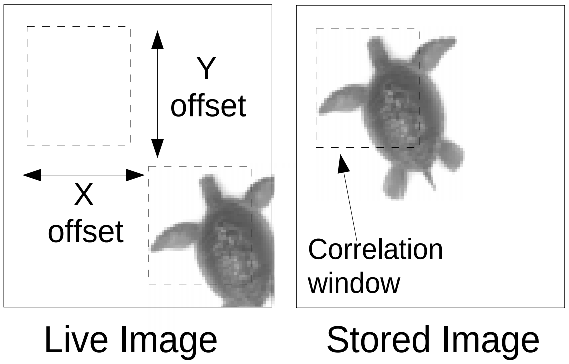

Figure 3: When the scenes in the images are offset, the best correlation occurs between windows which are also offset.

Thus, finding the corresponding sections which give maximum correlation determines the positional offsets of the live image from the stored image ($\Delta X, \Delta Y Actual$ in Figure 2).

Automatic mosaics are created by repeatedly storing images and knowing the positional relationships between their scenes. A new image is stored only when $\Delta X, \Delta Y Actual$ are within ϵ of some desired positional offsets ($\Delta X, \Delta Y Desired$). When a new image is stored, the live image scene matches the stored image scene, so $\Delta X, \Delta Y Actual$ are reset to zero, and new values for $\Delta X, \Delta Y Desired$ are obtained from the mosaic trajectory generator.

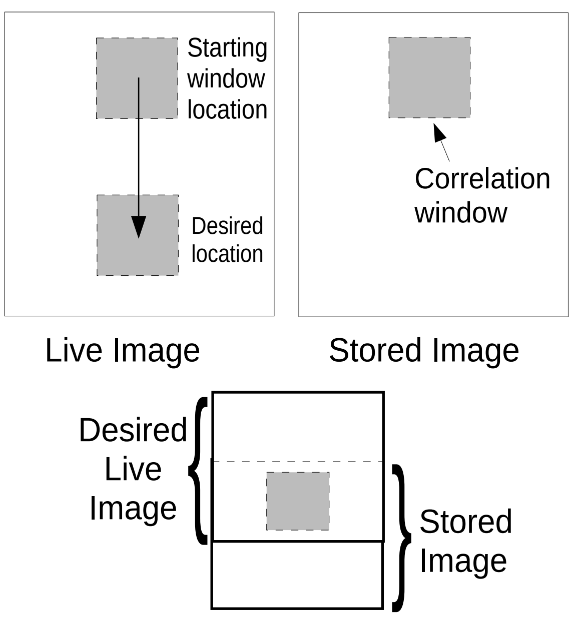

Figure 4: The desired live image scene is above the scene contained in the stored image.

C. Trajectory Generation

The trajectory generator shown in Figure 2 generates the desired offsets for the mosaicking system. Figure 4 shows an example of the desired offsets the trajectory generator could produce. For this example, the trajectory generator is attempting to move the scene in the live image upward such that the bottom of the scene in the live image scene will overlap with the top of the scene in the stored image. When this is achieved, the stored image will be replaced with the live image and the same process will be repeated. By following this strategy, a single-column mosaic can be generated in which every new picture added will belong above the previous picture in the mosaic. Similar strategies allow creation of single-column mosaics in the other direction and single-row mosaics.

IV. EXPERIMENTAL TESTING

A. Expedition on Ventana

The mosaicking system was tested in the ocean environment using *Ventana*, an ROV owned by the Monterey Bay Aquarium Research Institute. *Ventana* is operated via a tether from the support ship *Point Lobos*. A downward-pointing color camera was used to view the ocean floor, and a fiber-optic link through the tether carried the video to the ROV pilot control station on *Point Lobos*. Four lights arranged in a square pattern were mounted for bottom illumination.

The mosaicking system was situated in the pilot control station. It processed live video from Ventana and graphically displayed the offset errors ($\Delta X, \Delta Y Error$) to the pilot. The pilot controlled Ventana to zero the offset errors. The experimental tests were conducted with *Ventana* at a depth of 1000 feet, in Soquel Canyon just off the central Californian coast.

B. Test Plan

Several tests were performed to determine the ability of the mosaicking system to make mosaics of the ocean floor. The first test was the creation of a two-image mosaic to prove that correlation could be performed: 1) in the presence of marine snow, 2) with the four lights Ventana was carrying, and 3) on the type of scenes common to the brachiopod field. Following successful completion of the first test, multiple single column mosaics of different sizes were created of various areas of the brachiopod field. These mosaics tested the trajectory generation and the reliability of the correlation process.

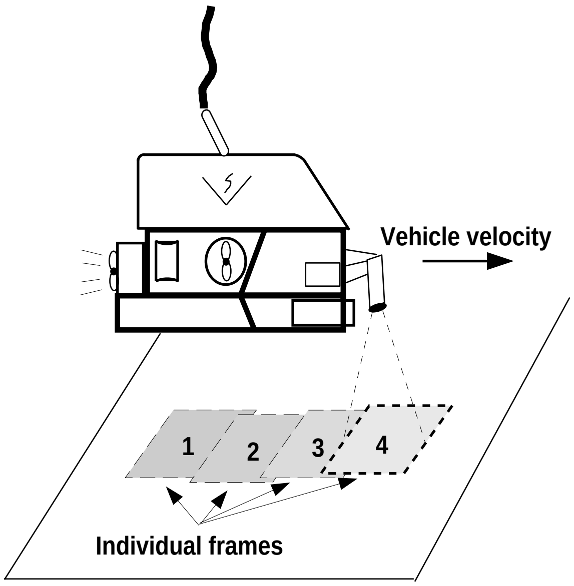

Figure 5: Creation of a single-column video mosaic of the ocean floor using Ventana.

C. Results

1) Two-image mosaic: A two-image mosaic was created on the first attempt (Figure 6).[1] The mosaicking system correctly measured the offset between the first and second pictures and was able to overlay the second picture in the proper location. However, the image quality of the mosaic was poor due to several effects. The most obvious effect came from the nonuniform lighting in the camera's field-of-view. The lights were shadowed by the vehicle's pan/tilt system and other protruding obstructions. This effect is evident at the bottom edge of the top image. The bottom portion of every image is darker than the top, and both sides are darker than the center. This effect is evident in all mosaics presented in this paper.

The high-frequency artifact that can be seen in the image is due to digitization problems. This problem was due to the fact that a color camera was used with a monochrome digitizer. The color subcarrier interfered with digitization process. A luminance-only copy of video recorded during the experiment did not display this artifact when digitized.

2) Single-column mosaics: Despite the noticeable problems in image quality of the two-image mosaic, the correlation system measured image offsets correctly. Several single-column mosaics were created (Figures 7, 9, and 10) . Due to considerations of vehicle control, only column mosaics were attempted (no row mosaics). Successful single-column mosaics were created travelling both forward and backward (although many more were attempted travelling forward, again for considerations of vehicle control). Mosaics ranging from two to twenty images were created. Figures 7, 9, and 10 demonstrate the variableness of the sections of ocean floor which were mosaicked.

[1]Note: the only post-processing performed on the mosaics is contrast enhancement. This was done to make objects in the mosaics easier to see and to make the mosaics more readily copiable.

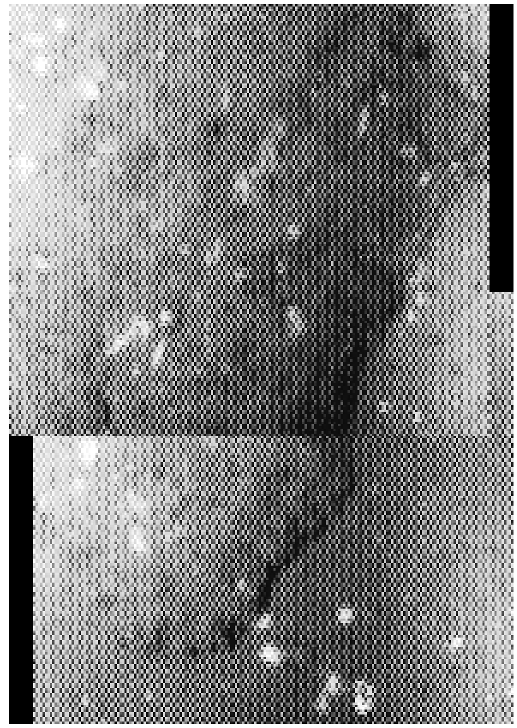

Figure 6: A two-image mosaic. Non-uniform lighting and color subcarrier digitization interference effects are evident.

D. Mosaic Analysis

1) First mosaic: Figure 7 shows an interesting mosaic of several rocks and a pair of fish (upper left and lower right). The mosaic has one noticeable error which occurs between the bottom image and its neighbor. This error is due to the fish at the right edge of the bottom image. One possible cause of the error is that the fish moved during the time the first and second images were snapped. Another more interesting possible cause is that the fish *appeared* to move with respect to the scene. Because the fish was above the ocean floor, it appeared to move with respect to the background as the camera moved over it. Figure 8 shows an exaggerated example of this. This well-known effect is used to accomplish stereo ranging by many vision processing systems.

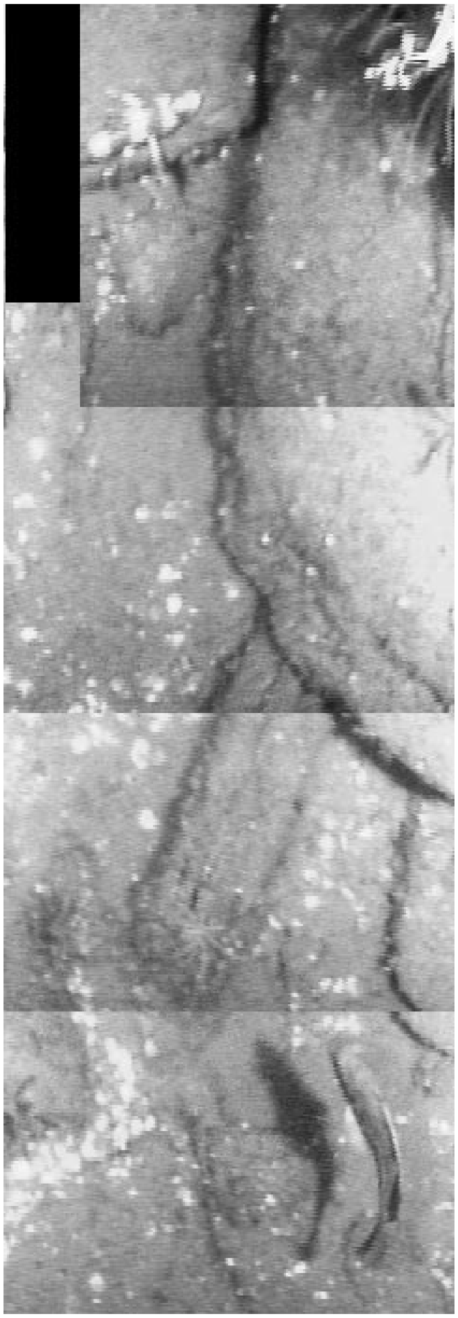

Figure 7: The ocean floor. The small (< 1 ft.) fish in the picture provide a scale.

The magnitude of the apparent motion is dependent on the field of view of the camera and the ratio of the altitude of the fish to the altitude of the vehicle. The mosaicking approach described earlier implicitly assumed this effect was negligible by assuming orthographic projection. The extent to which orthographic projection is a valid assumption is directly related to the achievable correctness of the mosaics. The assumption is valid for

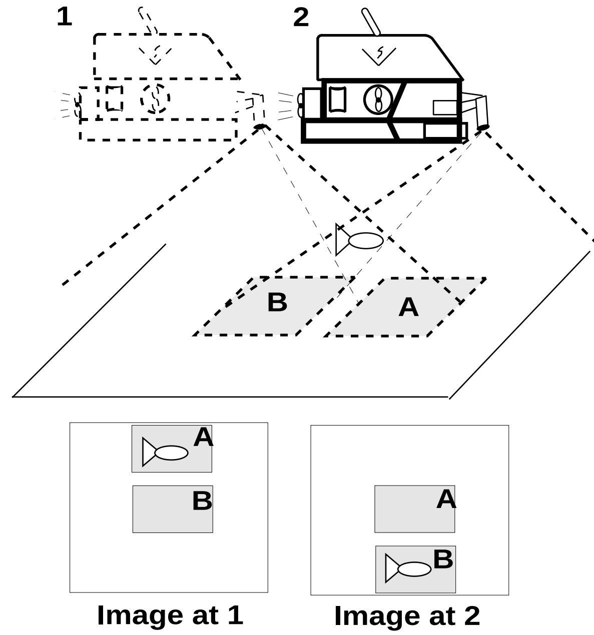

Figure 8: The fish appears over different portions of the ocean floor for images taken from different locations.

many cases and can always be made more reasonable by increasing the altitude of the vehicle (at the cost of lowering the mosaic resolution).

The situation of Figure 7 has the additional effect of the fish's shadow. Because the light source and camera were not coincident, the shadow moved in a manner different than both the fish and the floor. The magnitude of this effect is also dependent on the altitude ratio and camera field of view. In addition, it is dependent on the distance between the camera and the light source. Unlike the previous effect, it would still be present in true orthographic projection.

2) Second mosaic: Figure 9 demonstrates the extreme variation in lighting that was present in the mosaics. In the absence of sharp shadows, however, the mosaicking system was able to correlate perfectly. This demonstrates the effectiveness of the Laplacian portion of the filter which removes lighting biases and near-linear lighting variations. Many mosaics were created which would have been difficult with the naked eye due to the extreme lighting variation that was present.

Figure 9 also demonstrates the ability to correlate backgrounds which are of scientific interest. The brachiopod field shown has been studied by both biologists and geologists and is a typical portion of such fields.

3) Third mosaic: Figure 10 provides clear evidence that a side current was acting on the vehicle during the course of the mosaic. The steady drift to the right occurred in spite of the pilot's effort to keep the vehicle

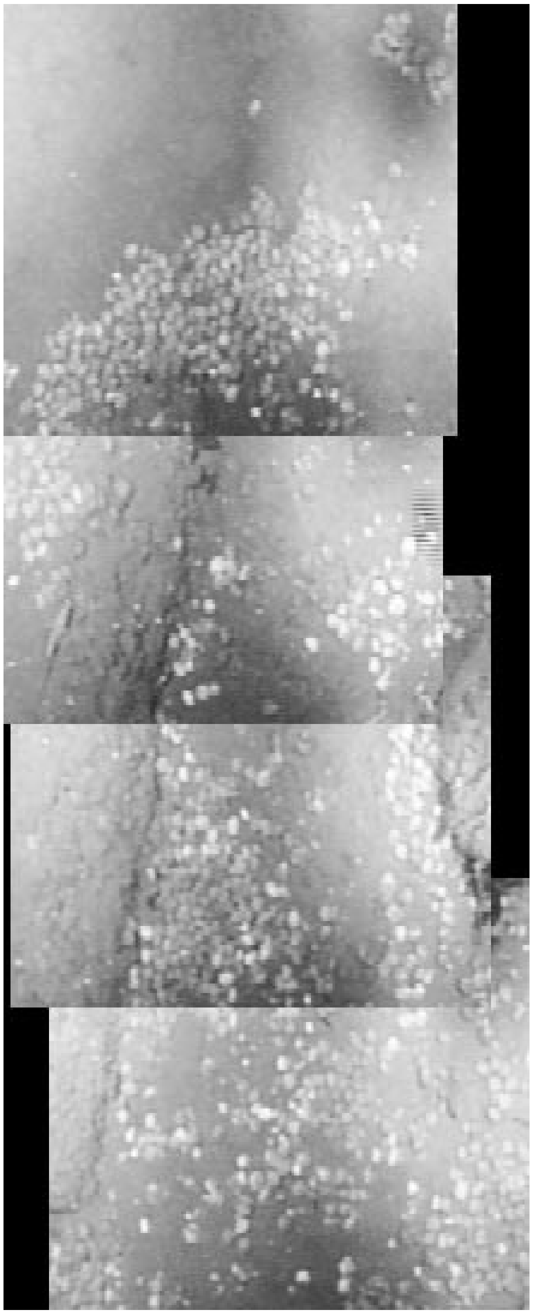

Figure 9: A field of brachipods.

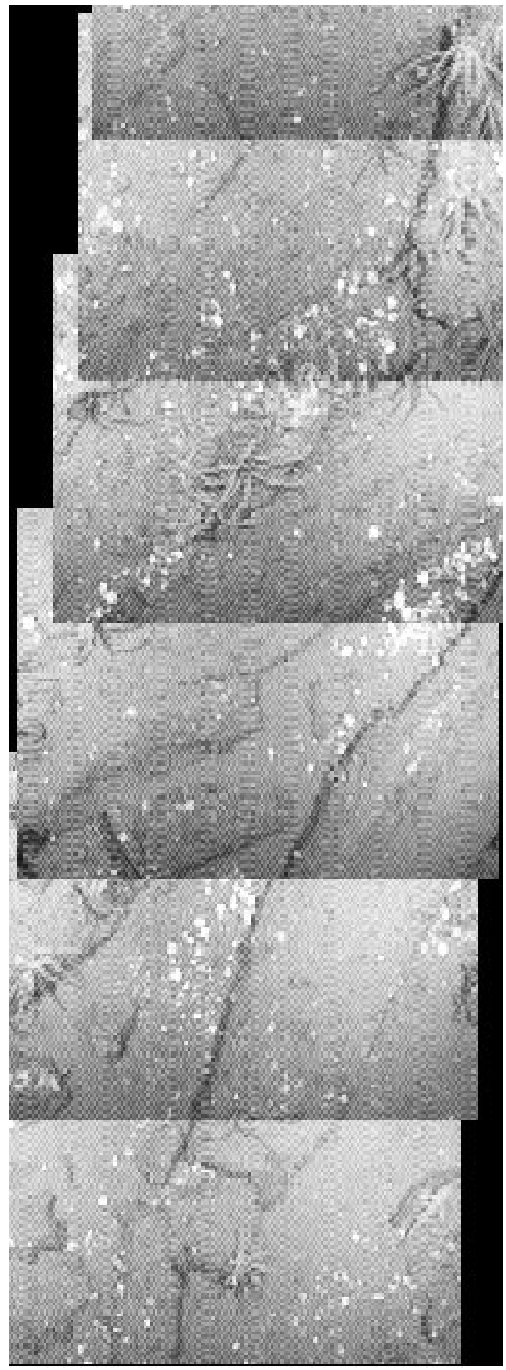

Figure 10: The strong current to the right is clearly visible.

moving straight ahead. The effect of such currents is to make mosaics of diagonal columns. Although this may not be noteworthy for single column mosaics, it could hinder the creation of multiple column mosaics. To create a second column correlated with the first, a similar flight path would be necessary. This would require more complexity in the mosaic trajectory generation. If the current was also time varying, correct trajectory generation would be extremely difficult.

V. CONCLUSIONS

An important conclusion drawn from the ocean tests is that a Laplacian of Gaussian prefilter can be used to accomplish image correlation in an actual ocean setting. In particular, the tests showed that correlation was robust despite the presence of marine snow and lighting variations (other than shadows). This conclusion is promising for future research in underwater visual sensing.

The test results also prove that it is feasible to create single-column mosaics of actual ocean floor settings *automatically* and in *real time*. The described approach to mosaicking yielded many successful mosaics of different types of ocean floor scenery. None of these mosaics contained visual gaps, and the mosaics were available in real time for inspection.

The single-column mosaicking success supports the prospect of creating multiple-column mosaics. Future research is planned addressing the strategies involved in creating large, muliple-column mosaics.

REFERENCES

[1] J. Pollio. Stereo-photographic mapping from submersibles. In *SPIE*, volume 12, pages 67–71. U.S. Naval Oceanographic Office, Deep Vehicles Branch, 1968.

[2] R. Haywood. Acquisition of a micro scale photographic survey using an autonomus submersible. In *IEEE Oceans*, volume 5, pages 1423–1426. University of New Hampshire, 1986.

[3] J. Zheng and S. Tsuji. Panoramic representation for route recognition by a mobile robot. *International Journal of Computer Vision*, 9(1):55–76, 1992.

[4] Q. Zheng and R. Chellappa. A computational vision approach to image registration. *IEEE Transactions on Image Processing*, 2(3):311–325, July 1993.

[5] H. Nishihara. Prism: a practical realtime imaging tereo matcher. In *SPIE*, pages 134–142, 1983.

[6] H. Nishihara. Practical realtime imaging stereo matcher. *Optical Engineering*, 23(5):536–545, October 1984.

[7] R. Marks, S. Rock, and M. Lee. Automatic Object Tracking for an Unmanned Underwater Vehicle using Real-Time Image Filtering and Correlation. In *Proceedings of IEEE Systems, Man, and Cybernetics*, France, October 1993. IEEE.