Image Registration Underwater for Fluid Flow Measurements and Mosaicking

Ryan Eustice, Hanumant Singh, Jonathan Howland

Department of Applied Ocean Physics and Engineering, MS #7
Woods Hole Oceanographic Institution
Woods Hole MA 02543
e-mail: ryan@whoi.edu

Abstract - *In this paper we present a unified framework of increasing complexity that handles image registration ranging from translation only motion between images, to a full eight parameter projective transformation. Our formulation deals specifically with the constraints that are peculiar to these applications underwater. To examine the role of translation only motion across images we look at video imagery from fluid flow across a flange at a hydrothermal vent site in Guaymas Basin in Mexico. We then show the extensions to this algorithm that are required to handle the case for 2D mosaicing, which involves translations, rotations, scale, and shear in an unstructured three dimensional underwater world. Finally, we show the effectiveness of these techniques on data acquired during photographic surveys of an ancient Roman shipwreck located in the Skerki Bank region in ~800m of water in the Mediterranean.*

Introduction

From its modest beginnings in the 1890s with the work of Louis Baton in the French Riviera, underwater photography has progressed to become an important tool in deep sea research[1]. With the loss of the nuclear powered submarine the USS Thresher in 1963 came a surge of interest in deep sea photographic equipment and techniques that has eventually led to the routine use of photomosaicking in an underwater context. The early techniques employed film and used rectifying lenses that reprojected images to allow the merging of multiple images into a composite mosaic. Numerous examples of the use of these techniques exist today including mosaics of the Monitor[2], the Thresher[2], and of geological sites of interest photographed during project Famous[3].

With the advent of digital computers and CCD cameras in the late 1980s, efforts at the Deep Submergence Laboratory (DSL) of the Woods Hole Oceanographic Institution (WHOI) have focused on using specialized image processing hardware for manipulating and processing digital images of underwater scenes to manually construct photomosaics[4][5]. Work on underwater photomosaicking is currently also being pursued at other institutions[6][7][8].

Independent of the work being carried out underwater, photomosaicking is routinely used for satellite imagery[9] as well as for mobile land robots[10][11][12][13]. Although these fields have reported significant advances, there are fundamental differences between these applications and the underwater domain. The chief differences are in the area of lighting and in the unstructured three dimensional nature of the underwater terrain. Thus, we focus in this paper on these differences and the techniques involved in addressing these differences to build large area photomosaics underwater.

2D Translation & Fluid Flow

Beginning with the simplest case, the following describes the mathematics behind image registration of pure translation only motion between images, set in the context of an interesting application. Photomosaicking techniques were used in the summer of 1999 to make volumetric flow rate measurements of a hydrothermal vent site in Guaymas Basin, Mexico; something which had never been done before. A clear graduated cylinder was modified to have a special funnel base which directed all of the expelled hydrothermal exhaust water through the cylinder. The special base also had a dye injection system which released a continuous stream of fluorescent dye into the funnel coloring the exhaust water inside the cylinder. The remotely operated vehicle (ROV) Jason, operated by DSL at WHOI, was used to deploy this graduated cylinder device at a hydrothermal vent site and then record the fluid flow imagery onto standard video at a frame rate of 30Hz.

The concept for extracting volumetric flow rate measurements at the vent lies in the assumption that between successive frames of video, the dye in the cylinder undergoes pure translation only motion. If this translation can be found, an average velocity
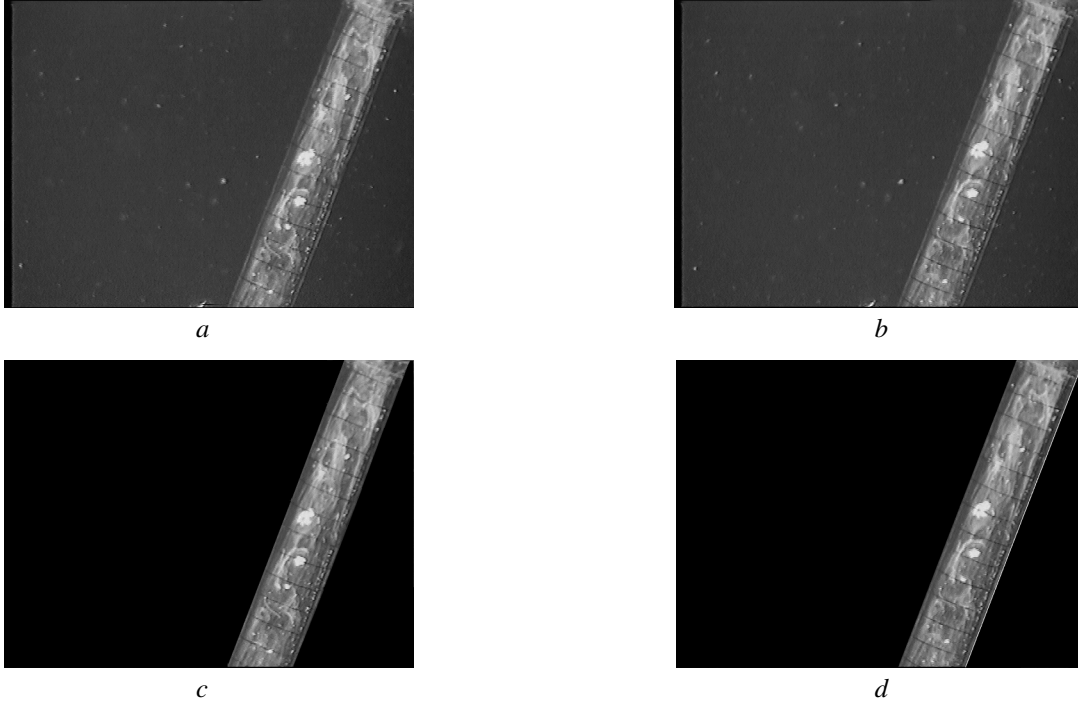
Figure 1. Flow visualization of exhaust flow of a hydrothermal vent in Guaymas Basin, Mexico. Images *a* and *b* show raw picutres of frames 4-01 and 4-06 in a sequence of video. Images *c* and *d* are the same two frames, but have the background removed. Notice the validity of the assumption of pure translation.

measurement of the fluid can be calculated from the time between successive frames. Multiplying the average velocity by the cylinder cross-sectional area produces an average volumetric flow rate.

To recover the fluid translation through Fourier transform phase correlation techniques, the first step was to reduce the backscatter noise present in the background of the image; this was accomplished by isolating the cylinder and dye from the rest of the image as shown in Figure 1. Clearly the assumption of translation only motion is valid as frame 4-06 is seen to be merely a shifted version of frame 4-01.

Thus, defining the image in frame 4-01 to be $x_1[n_1, n_2]$, the image in frame 4-06 can be written as:

$$x_2[n_1, n_2] = x_1[n_1 - m_1, n_2 - m_2] \qquad (1)$$

To find the delay m_1 and m_2, begin by noting the Fourier transforms, $\Im$, of $x_1[n_1, n_2]$ and $x_2[n_1, n_2]$:

$$\Im(x_1[n_1, n_2]) = |X_1(j\Omega_1, j\Omega_2)| e^{\theta_x(j\Omega_1, j\Omega_2)} \qquad (2)$$

$$\Im(x_2[n_1, n_2]) = \qquad (3)$$
$$|X_1(j\Omega_1, j\Omega_2)| e^{\theta_x(j\Omega_1, j\Omega_2)} e^{-jm_1\Omega_1} e^{-jm_2\Omega_2}$$

Multiplying $X_1(j\Omega_1, j\Omega_2)$ by the conjugate of $X_2(j\Omega_1, j\Omega_2)$ and then dividing by the magnitude of their respective product yields:

$$G(j\Omega_1, j\Omega_2) = \frac{X_1 \cdot X_2{}^*}{|X_1 \cdot X_2{}^*|} = e^{jm_1\Omega_1} e^{jm_2\Omega_2} \qquad (4)$$

Taking the inverse Fourier transform of $G(j\Omega_1, j\Omega_2)$ results in the spatial function $g[n_1, n_2]$

$$g[n_1, n_2] = \delta[n_1 - m_1, n_2 - m_2] \qquad (5)$$

from which the translation measurements are obtained from the coordinates of the unit sample. From here, the velocity and volumetric rates are easily calculated.

Photomosaicking

The process of photomosaicking involves aligning multiple images to one another and then merging these images into a seamless composite representation. The alignment of a pair of images may be accomplished by a

variety of techniques which include manual tie point location, spatial techniques, and Fourier transform based methods. A good survey of image registration techniques is that by Brown[14]. These techniques vary from being very labour intensive to completely automated with attendant trade-offs in ease of use and accuracy. In particular, we look at manual tie point location and spatial techniques and their roles in the context of underwater applications.

Manual Location of Tie Points

We should point out that several packages exist that provide an easy to use interface that allows manual photomosaicking. Operators are presented with views of the images to be mosaicked and then select common features across these images. A linear least squares solver package is typically employed to solve the over constrained set of equations that correspond to the manually selected tie points to estimate the transformation between the pair of images.

The residuals associated with the process of solving the over constrained set of equations provides a level of quality control over the process. Typically, operators can fine-tune their choice of tie points to reject those with high residuals to produce an aesthetically pleasing mosaic.

The techniques associated with manual mosaicking have additional significance in areas with unstructured hard-to-model relief.

Spatial Methods

Our spatial approach to photomosaicking mirrors that of Sawhney and Kumar[10], although it differs in important ways to deal with the peculiarities of the underwater environment.

Let I_1 ... I_N be the N images that we are trying to register.

If we consider an ideal coordinate system where a point p is given by $p = (x, y)$ and $p_i = (x_i, y_i)$ is the representation of that point in image I_i such that the point p is related to p_i by the mapping

$$p_i = P(p; A_i) \qquad (6)$$

where A_i is the 3D-to-2D or 2D-to-2D projection transformation that maps p to p_i; A_i may be a Helmert, affine, projective, or higher order transformation.

For the purposes of this paper we assume A_i to represent an eight parameter projective plane transformation

$$x_i = \frac{a_{11}^i x + a_{12}^i y + a_{13}^i}{a_{31}^i x + a_{32}^i y + a_{33}^i} \qquad (7)$$

$$y_i = \frac{a_{21}^i x + a_{22}^i y + a_{23}^i}{a_{31}^i x + a_{32}^i y + a_{33}^i} \qquad (8)$$

where $a^i{}_{11}$ through $a^i{}_{33}$ are the plane projective parameters that correspond to image I_i.

The correspondence between the points in the images and the coordinate transformation must both be accomplished simultaneously. The brightness constancy constraint can be used in this respect as it relates

$$I_i(p_i;(p, A_i)) = I_{i+1}(p_{i+1};(p, A_{i+1})) \qquad (9)$$

However, while this may serve as a good enough approximation in some cases, it is not generally applicable for underwater imagery. The light source moves with the camera on underwater imaging platforms and in the case for non-uniformly illuminated images, equeation (4) does not hold true. Thus we deviate from the formulation proposed by Sawhney and Kumar in two important ways.

1. We substitute adaptively histogram equalized imagery in place of the original images even though, due to aesthetic considerations, we may use the unequalized images for the mosaic.

2. We detect and discount the contribution due to shadows. As the light source moves, the shadows tend to also shift. Due to the low contrast nature of underwater imagery, techniques which rely on edge based operators are often fooled by this apparent motion across images.

Thus if we consider the ith histogram equalized image I_i^h and the subset S of the original image I such that the intensity in S corresponds to shadowed regions and their neighbors, then we can minimize the error criterion,

$$\underset{A_i, A_{i+1}}{min} \sum_{p \notin S} \frac{1}{M(p)} \sum_{i, i+1} (I_i^h(p_i) - \bar{I}(p))^2 \qquad (1)$$

This can be linearized

$$E((p_i;p);A) = \frac{1}{\sqrt{M(p)}}(I_i^h(p_i) - \bar{I}(p)) \qquad (10)$$

such that each iteration of the Levenberg-Marquadt optimization procedure solves the following linear sum of squares problem

$$\binom{min}{\delta A} \sum_{p \notin s} \sum_{i} (E(p_i;p) + \nabla E(p_i;p)[\delta A]) \qquad (11)$$

Our implementation uses a two step progressively complex model in combination with a laplacian of gaussian pyramid. We first estimate the translation parameters at a coarse level of the pyramid using a scheme similar to that used for fluid flow measurements and then use those values to estimate the parameters corresponding to the projective transformation.

The results of our techniques are highlighted in Figures 2 and 3. Figure 2 shows the general methodology we follow for the photomosaicking process. Figure 3 shows a complete photomosaick of a 4th century B.C. Roman shipwreck located in ~800m of water in the Mediterranean. This photomosaic is a composite of 180 images that were assembled in four strips that were then manually assembled into a composite photomosaic.

These and other results along with the implementation of this algorithm are discussed at length in [15].

Conclusions

In this paper we have attempted to demonstrate a unified framework to photomosaicking with a particular emphasis on the unique constraints of underwater imagery. The techniques presented are general building blocks towards a global underwater image registration suite. Areas of future and on-going research here at DSL which work towards this end goal are: color histogram equalization to produce color mosaics, understanding and quantifying the effects of an unstructured 3D environment on mosaics, and finally, 3D image reconstruction from a single camera on a moving underwater platform.
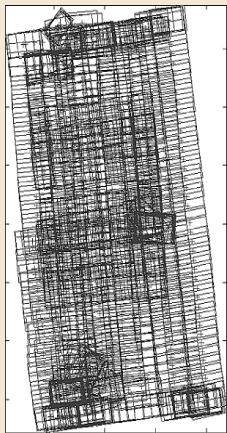
Step 1 *Vehicle operators conduct a carefully planned survey over the area of interest to ensure sufficient coverage and overlap in the imagery. Image footprints are projected on the area of interest to allow operators to choose individual images for use in the mosaic.*
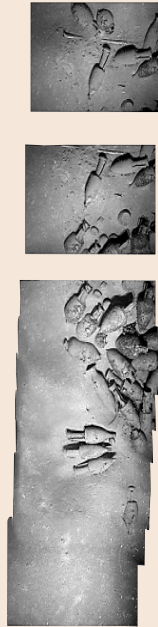
Step 2 *Individual images are processed to remove lighting and other artifacts. These images are then merged into single strip mosaics by indentifying common features in successive images.*
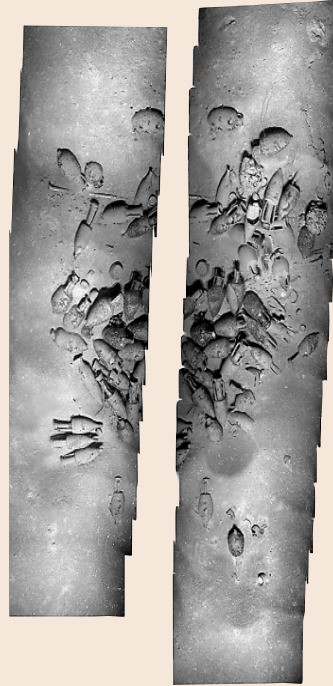
Step 3 *Individual strips are then mosaiced together using a technique similar to that used in Step 2.*

Step 4 *Ongoing work focuses on understanding and improving the quantitative nature of the mosaicing process.*

sk97.066

sk97.067

sk97.113

sk97.112

sk97.063

sk97.068

sk97.055

sk97.065

sk97.056

sk97.078

sk97.071

lead sheathing
pan tile

sk97.064

sk97.072

hand quern

sk97.070

sk97.061

sk97.069

sk97.052

sk97.079

sk97.079a

sk97.080

sk97.073

Dressel 1c

sk97.053

sk97.054

sk97.060

sk97.077

sk97.059

sk97.075

bilge pipe

sk97.062

sk97.076

sk97.051

sk97.074

sk97.050

sk97.058

Figure 3. Mosaic of a 4th Century B.C. Roman shipwreck located in ~800m of water in the Mediterranean; the mosaic was constructed from 180 images taken from the Jason ROV.

References

[1] Vine, A.C. Early History of Underwater Photography, *Oceanus*, vol. 18, no. 3, Spring, pp. 2-10, 1975.

[2] Patterson, R.B. Future Developments in Deep-Sea Imaging, *Oceanus*, vol. 18, no. 3, Spring, pp. 17-23, 1975.

[3] Ballard, B.D. Photography from a Submersible During Project Famous, *Oceanus*, vol. 18, no. 3, Spring, pp. 40-43, 1975.

[4] Sulanowska, M.M., Humphris, S.E., Howland, J.C. Detailed Analysis of the Surface Morphology of the Active TAG Hydrothermal Mound by Mosaicking of Digital Images, *EOS, Transactions of the American Geophysical Union 77,* p. 768, 1996.

[5] Howland, J.C., Singh, H., Marra, M., Potter, D. Digital Mosaicking of Underwater Imagery, *Sea Technology,* pp. 65-69, June, 1999.

[6] Gracias, N., Santos-Victor, J. Underwater Video Mosaics as Visual Navigation Maps, *Computer Vision and Image Understanding,* vol. 79, pp. 66-91, 2000.

[7] Negahdaripour, S., Xu, X., Khamene, A. Applications of Direct 3D Motion Estimation for Underwater Machine Vision Systems, *in Proceedings of the IEEE Oceans'98,* Nice, France, September, 1998.

[8] Marks, R., Rock, S., Lee, M. Real-Time Video Mosaicking of the Ocean Floor, *IEEE Journal of Ocean Engineering*, vol. 20, no. 3, pp. 229-241, 1995.

[9] Slama, C.C., Ed. *Manual of Photogrammetry.* American Society of Photogrammetry, Falls Church, VA, 1980.

[10] Sawhney, H.S., Kumar, R. True Multi-Image Alignment and Its Application to Mosaicing and Lens Distortion Correction, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, pp. 235-243, March, 1999.

[11] Mann, S., Picard, R.W. Virtual Bellows: Constructing High Quality Stills From Video, *ICIP*, 1994.

[12] Peleg, S., Herman, J. Panoramic Mosaics by Manifold Projection, *CVPR*, pp. 338-343, 1997.

[13] Szeliski, R. Image Mosaicing for Tele-Reality Applications, *IEEE Workshop Applications of Computer Vision*, pp. 44-53, 1994.

[14] Brown, L.G. A Survey of Image Registration Techniques, ACM Computing Surveys, vol. 24, no. 4, pp. 325-376, December, 1992.

[15] Singh, H., Howland, J., Pizarro, O., Eustice, R., Large Area Photomosaicking Underwater, *submitted to the IEEE Journal of Oceanic Engineering.*