RECOVERY OF 3D INFORMATION FROM UNDERWATER IMAGES *

M. Arredondo, K. Lebart, D. Lane, A. Serrurier
Department of Computer and Electrical Engineering
Heriot Watt University
Riccarton EH14 4AS Edinburgh, Scotland, UK
{M.Arredondo,K.Lebart,D.M.Lane,as27}@hw.ac.uk

1 Abstract

The impact of trawling activity on the sea floor can be assessed by observing trawling marks, holes and mounds from video data. One of the key indicators used by scientists is the relief of the sea floor. Although the use of video cameras has become a standard practice in underwater missions, there is little technology to help marine scientists to analyze the data. The estimate of the optical flow provides a depth map of the seafloor towards the automation of the analysis of underwater videos, for the purpose of trawling impact assessment. We explore here the joint use of optical flow and feature tracking to improve the robustness of dense point correspondence, leading to a better depth map.

2 Introduction

Although the problems of trawling impacts on the marine environment were identified more than a century ago, they have only been a major concern in the last 10 years or so, with an increasing application of high level of scientific research. For a long time fishing has been blamed for loss of biodiversity through habitat modification and destruction and mechanisms of impact are now well understood. However, there is little or no information available on the scale of impacts because of the lack of wide area high resolution monitoring. With the increasing needs for efficient environmental management, scientists are being pressurized by managers to provide quick, reliable and usable information which they are presently unable to efficiently collect over wide areas [1]. Marine biologists currently log survey data manually, they have to watch long videos - several hours long - to identify and classify trawling marks according to their freshness and frequency; they also look for the frequency of mounds and holes on the seafloor.

In the context of the European project AMASON (Advanced Mapping with Sonar and Video) we have explored the joint use of feature tracking and optical flow to improve the robustness of dense point correspondence. We use the velocity of a set of tracked features and the knowledge of the motion of the vehicle (translating parallel to the seafloor with the camera pointing downwards) to impose a constraint on the solution of the optical flow. This robust estimate of the optical flow provides a better depth map of the seafloor towards the automation of trawling impact assessment from video. Other underwater tasks than can benefit from a robust optical flow estimate are: station keeping [13], object tracking [7] and navigation [4].

In the next section we will introduce the problem of recovering 3D information from a stream of images (referred in computer vision as the *structure from motion* problem). We present the proposed approach in Section 4 and show some results in Section 5. We conclude in Section 6 and give some directions of future work.

3 Recovering Depth

The *structure from motion problem* (SFM) is the recovery of a three dimensional (3D) structure from a sequence of two dimensional (2D) single frames. In SFM we want to recover 3D information about a collection of discrete structures (e.g. points or lines) from a 2D collection of such structures. Two dimensional images are formed from the 3D world by a projection process. SFM seeks to take the 2D information and recover the 3D original

*This work has been partially supported by the 5th Framework Programme of research of the European Community through the project AMASON (EVK3-CT-2001-00059).

information, inverting the effect of the projection process. The majority of the existing techniques decouple the solution in two steps and estimate motion first and then estimate structure.

Methods to recover motion can be classified into the discrete and differential approaches, depending on whether they use feature correspondences or the positions and apparent velocities (optical flow) of objects. While tracking a limited number of features through a sequence of images might be "easier" than estimating a dense correspondence, in the underwater environment it is harder to find strong visual features (e.g. distinct corners, edges or contours of objects). Tracking techniques have been successfully applied on underwater environments [11]; however, estimating the depth of a set of sparse features on the seafloor would not be suitable for the purpose of identifying trawling marks.

On the other hand, dense correspondence methods provide a dense structure of the scene. However, estimating optical flow accurately is not a trivial task and has been an important subject of research on the Computer Vision community for many years. Optical flow estimation may fail when the brightness constancy assumption, presupposed for the calculation of the optical flow, does not hold [9]. This makes the dense correspondence from underwater images a very difficult problem, given that underwater scenes are usually illuminated with nonuniform lighting (e.g. due to highly directional lighting sources or attenuation of light as a function of distance) and that illumination can change drastically from frame to frame on a video sequence. Previous researchers have tried to overcome the illumination problem using different approaches: Negahdaripour and Yu proposed an optical flow method that allows the brightness on the scene to change [9], while Eustice et al. [3] used contrast limited adaptive histogram equalization [12] to deal with lighting artifacts of nonuniform illumination and low contrast.

3.1 Optical Flow

The motion field is the 2D vector field which is the perspective projection on the image plane of the 3D velocity field of a moving scene. From the information available from a sequence of images (the spatial and temporal variation of the brightness pattern) it is only possible to derive an estimate of the motion field, called optical flow [6]. During the past two decades many methods to estimate optical flow have been proposed. In this section we describe a gradient-based method to estimate the optical flow [8]. Let $I(\mathbf{x}, t)$ denote the continuous space-time intensity function of an image, where $\mathbf{x} = (x, y)^\top$. If the intensity remains practically constant along a motion trajectory, we have: $\frac{dI(\mathbf{x},t)}{dt} = 0$, where $\mathbf{x}$ varies by t according to the motion trajectory. This is a total derivative expression and denotes the rate of change of intensity along the motion trajectory. Using the chain rule of differentiation, it can be expressed (dropping the $(\mathbf{x}, t)$ for compactness) as:

Figure 1: Nonuniform lighting and abrupt illumination change are a problem to estimate reliable optical flow from underwater images.

$$\frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \frac{\partial I}{\partial t} = 0 \tag{1}$$

where $u = \frac{dx}{dt}$ and $v = \frac{dy}{dt}$ denote the components of the image velocity vector in terms of the continuous image coordinates, and the partial spatial derivatives of the image brightness are the components of the spatial gradient ∇I.

Expression (1) can be rewritten as the *image brightness constancy equation*:

$$\nabla I^\top \mathbf{v} + I_t = 0 \tag{2}$$

where I_t denotes partial differentiation with respect to time. Expression (2) provides one linear equation for two unknown components of the velocity; hence, further constraints are necessary to solve for $\mathbf{v}$. Perhaps the simplest constraint is to assume that the motion is the same on a small spatial neighborhood (known as the *local spatial smoothness constraint*). Using this assumption, the optical flow can be estimated [8] within a neighborhood Ω (of size $N \times N$) as the vector, $\mathbf{v}$, that minimizes:

$$\Psi[\mathbf{v}] = \sum_{\mathbf{p}_i \in \Omega} ([I_x \quad I_y]^\top \mathbf{v} + I_t)^2 \qquad (3)$$

The solution to this least squares problem is:

$$\mathbf{v} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b} \qquad (4)$$

where the i-th row of the $N^2 \times 2$ matrix $\mathbf{A}$ is the spatial image gradient evaluated at point $\mathbf{p}_i$, $\mathbf{b}$ is the N^2-dimensional vector of the partial temporal derivatives of the image brightness, evaluated at $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{N^2}$, and $\mathbf{v}$ is the optical flow at the center of neighborhood Ω.

4 Constrained Optical Flow

Researchers have used different assumptions and constraints in order to improve the robustness of the estimated optical flow. For example, Foroosh [5] proposed a *temporal smoothness constraint* which implies that either objects in the scene have constant velocities or that the sequence have a sufficiently dense temporal sampling, so that the inter-frame acceleration is negligible. Chunke and Oe [2] generate two new constraints by spatially deriving equation (1); using these equations in conjunction with the spatial smoothness constraint a solution is found.

We propose to use the velocity of a set of tracked points and the knowledge we have of the motion of the vehicle to constraint the solution of the optical flow. We assume that the vehicle is translating parallel to the seafloor, hence all the points in the image will be moving in the same direction. We can track some points to estimate the direction of the motion and then use this information to constraint the solution of the optical flow for the rest of the points on the scene.

We want to minimize the function:

$$f(u, v) = [I_x u + I_y v + I_t]^2 \qquad (5)$$

subject to the constraint

$$g(u, v) = v - mu = 0 \qquad (6)$$

where $m = \hat{v}/\hat{u}$ and $\hat{v}$, $\hat{u}$ denote the average magnitude of the components of the velocity estimated from a set of tracked points. To solve this constrained optimization problem we use lagrangian multipliers.

The Lagrangian function is given by:

$$\begin{aligned} L(u, v, \lambda) &= (I_x u + I_y v + I_t)^2 + \lambda(v - mu) \\ &= (I_x u + I_y v)^2 + 2(I_x u + I_y v) I_t \\ &\quad + I_t^2 + \lambda(v - mu) \\ &= u^2 I_x^2 + 2uv I_x I_y + v^2 I_y^2 \\ &\quad + 2u I_x I_t + 2v I_y I_t \\ &\quad + I_t^2 + \lambda(v - mu) \end{aligned}$$

The critical points of L are the solutions (u, v, λ) to the following system of equations:

$$\begin{aligned} \frac{\partial L}{\partial \lambda} &= 0 = \quad v - mu && (7) \\ \frac{\partial L}{\partial u} &= 0 = \quad 2u I_x^2 + 2v I_x I_y + 2 I_x I_t - \lambda m && (8) \\ \frac{\partial L}{\partial v} &= 0 = \quad 2v I_y^2 + 2u I_x I_y + 2 I_y I_t + \lambda && (9) \end{aligned}$$

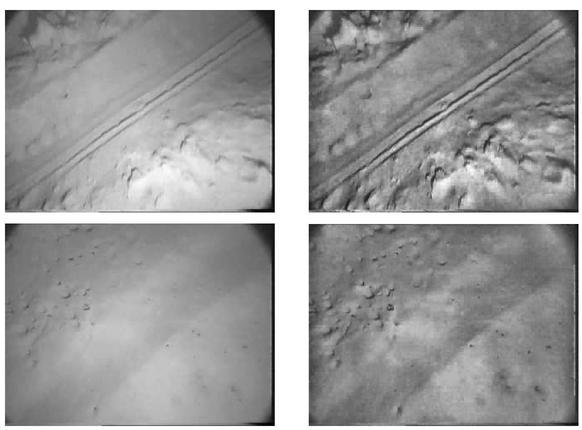
Using the local spatial smoothness constraint, each point on a small neighborhood Ω will provide two equations to estimate the velocity at the center of the patch. Using a neighborhood of size $N \times N$ the velocity at the center of the neighborhood is given by the solution of the system:

$$\overbrace{\begin{bmatrix} -m & 1 & 0 \\ 2I_{x_1}^2 & 2I_{x_1} I_{y_1} & -m \\ 2I_{x_1} I_{y_1} & 2I_{y_1}^2 & 1 \\ \vdots & \vdots & \vdots \\ 2I_{x_N}^2 & 2I_{x_N} I_{y_N} & -m \\ 2I_{x_N} I_{y_N} & 2I_{y_N}^2 & 1 \end{bmatrix}}^{\mathbf{A}} \begin{bmatrix} u \\ v \\ \lambda \end{bmatrix} = \underbrace{\begin{bmatrix} 0 \\ -2I_{x_1} I_{t_1} \\ -2I_{y_1} I_{t_1} \\ \vdots \\ -2I_{x_N} I_{t_N} \\ -2I_{y_N} I_{t_N} \end{bmatrix}}_{\mathbf{b}}$$

The least squares solution can be obtained using (4).

5 Results

We have implemented in MATLAB the gradient-based algorithms described in Section 1 and Section 4. To improve the quality of the images and minimize the effect of non-uniform and changing illumination on the estimation of the optical flow we followed the approach successfully applied by Eustice et al. [3]. We used contrast limited adaptive histogram equalization; with this technique the image is broken intro sub-regions. The optimal gray scale distribution is calculated for each of these sub-regions based upon its histogram and a previously determined transfer function, which is based upon the desired histogram of the sub-region. Following the results reported by Eustice et al. [3], we used a Raleigh distribution to equalize our underwater images (see Figure 2).



Original Images　　Equalized Images

Figure 2: Original underwater images and their equalized version, from which we estimated optical flow.

From the equalized images we estimated the velocity of a set of features in the image used the method proposed by Shi and Tomasi [10]. Using the average velocity of the tracked features we estimated the constrained optical flow. To further improve the estimated depth maps [1] we implemented a simple approach to combine the solution of multiple frames. We first estimated the region of the image that is common to the frames to combine by estimating the displacement between them. Once the common region is established we can average the magnitude of the velocity at each point. Of course, this simple approach causes some blurring on the final disparity map, given that not all the points move at the same velocity.

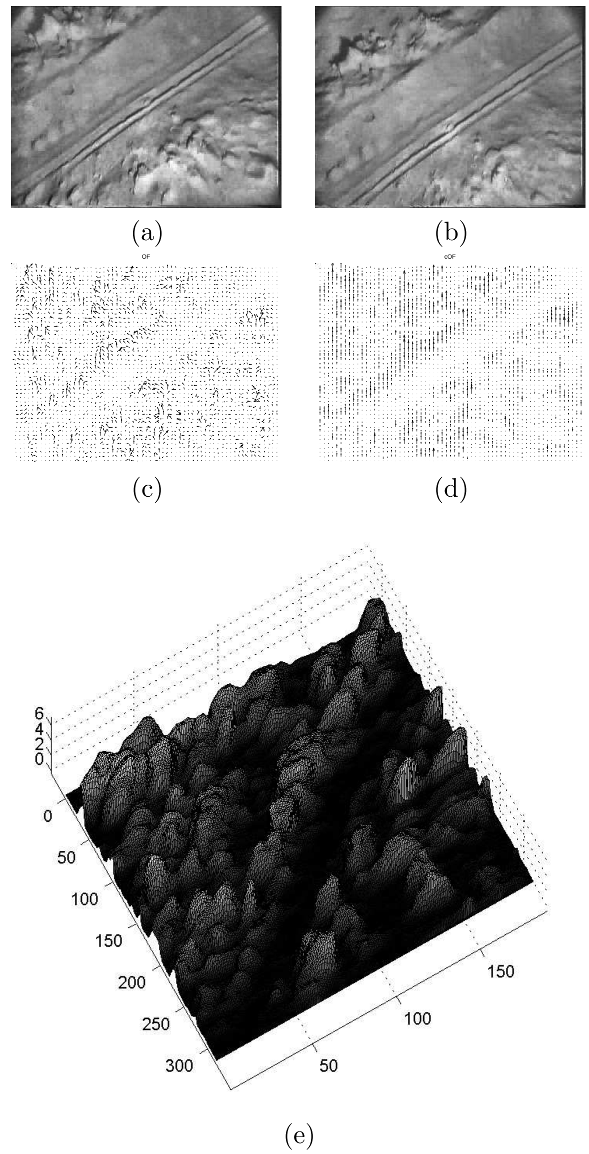


Figure 3: (a)(b) Images from the video sequence (c) Optical flow Lucas&Kanade (d) Optical flow proposed approach (e) Acumulated disparity map over four frames plotted as a surface.

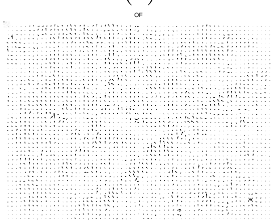
6 Conclusions

We have proposed a gradient-based method to estimate optical flow using information from tracked features to improve the reliability of the estimate. The approach has been tested with

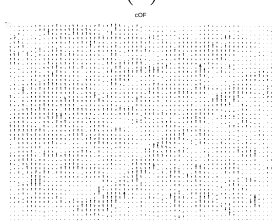
[1] The disparity map obtained from the optical flow gives a rough idea of the depth and structure of the scene: points closer to the camera move faster than points farther away.

(a)

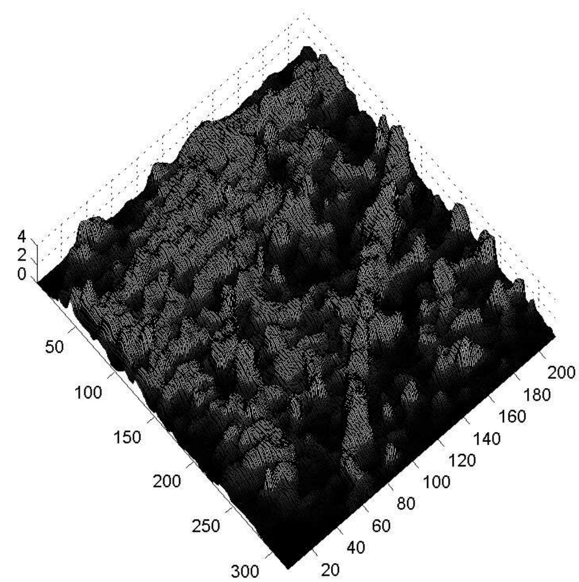
(b)



(c)



(d)



(e)

Figure 4: (a)(b) Images from the video sequence (c) Optical flow Lucas&Kanade (d) Optical flow proposed approach (e) Acumulated disparity map over four frames plotted as a surface.

several sequences of underwater images; the results are encouraging, showing a smoother optical flow compared to the traditional method proposed by Lucas and Kanade [8]. We are currently working on the development of algorithms for 3D feature extraction to test the disparity maps we obtain under a classification framework; this would allow us to judge the results quantitatively.

The constraint currently imposed for the solution of the optical flow allows only for a pure translational motion, which is a very restrictive case. We are working on a method that allows more degrees of freedom on the motion while still applying a strong geometric constraint using the differential epipolar equation.

7 Acknowledgments

We thank Chris Smith from the Institute of Marine Biology of Crete for providing us with the underwater video sequences. We also thank Stan Birchfield for the C implementation of the tracker and R. Eustice for the UWIT.

References

[1] Advanced Mapping with Sonar and Video (AMASON). Description of Work. European Commission Framework V RTD Project Proposal. 2001.

[2] Y. Chunke and S. Oe. *A new gradient-based optical flow method and its application to motion segmentation.* 26th Annual Confjerence of the IEEE Industrial Electronics Society, 2000. Vol. 2, pp. 1225-1230.

[3] R. Eustice, O. Pizarro, et al. *UWIT: Underwater Imaging Toolbox for Optical Image Processing and Mosaicking in MATLAB.* Proceedings of the Third Underwater Technology Symposium, 2002, Tokyo, Japan.

[4] Y. Fan and A. Balasuriya. *Optical flow based speed estimation in AUV target tracking.* MTS/IEEE Conference and Exhibition OCEANS 2001, Vol. 4, pp. 2377-2382.

[5] H. Foroosh. *A closed-form solution for optical flow by imposing temporal constraints.* Proceedings of the International Conference on Image Processing 2001, Vol. 3, pp. 656-659.

[6] B.K.P. Horn and B.G. Schunck. *Determining optical flow.* A. I., 1981, 17, pp. 185-203.

[7] D.M. Lane, M.J. Chantler and Dongyong Dai. *Robust tracking of multiple objects in sector-scan sonar image sequences using optical flow motion estimation* IEEE Journal of Oceanic Engineering, 1998, Vol. 23, Issue: 1, pp. 31-46.

[8] B.D. Lucas and T. Kanade. *An Iterative Image Registration Technique with an Application to Stereo Vision.* 7th Int. Joint Conf. on A. I., 1981, pp. 674-679.

[9] S. Negahdaripour and C. H. Yu. *A Generalized Brightness Change Model for Computing Optical Flow*. International Conference on Computer Vision 1993, pp. 2-11.

[10] J. Shi and C. Tomasi. *Good Features to Track*. IEEE Conference on Computer Vision and Pattern Recognition 1994, pp. 593-600.

[11] E. Trucco, Y. Petillot, I. Tena Ruiz, C. Plakas, D. M. Lane. *Feature tracking in video and sonar sub sea sequences with applications*. Computer Vision and Image Understanding, Vol. 79, Issue 1, July 2000.

[12] K. Zuiderveld. *Contrast Limited Adaptive Histogram Equalization*. Graphics Gems IV. P. Heckbert. Boston, Academic Press. IV: 474485, 1994.

[13] S. van der Zwaan and J. Santos-Victor. *Real-time vision-based station keeping for underwater robots*. MTS/IEEE Conference and Exhibition OCEANS, 2001. Vol. 2, pp. 1058-1065.