Robust optical flow estimation using underwater color images

Shahriar Negahdaripour and Hossein Madjidi
Underwater Vision & Imaging Lab
Electrical and Computer Engineering Department
University of Miami
Coral Gables, FL 33124-0640

***ABSTRACT*: Accurate and efficient optical flow estimation is a major step in many computational vision problems, including 3-D mapping applications. While processing of gray scale image sequences has been the main approach, some authors have recently investigated and reported the use of color images. The main advantage in the use of color images is to exploit redundant information at each pixel, namely 3 channels, as potential cues for visual motion. Moreover, several spectral-dependent factors including medium attenuation may justify the use of color cues in underwater, calling for a physics-based approach to optical flow analysis.**

This paper explores a simplified underwater image formation model 1) to provide some preliminary insight on the information encoded in various color channels, 2) draw some general conclusions on the use of color and its various representations, and 3) verify if the results are consistent with the experiments with 7 sets of coral reef video recoded in relatively clear shallow waters of the Bahamas. These results suggest that optical flow algorithms based on the HSV representation can provide improved localization and accuracy.

I. BACKGROUND

The perceived motion of image brightness patterns, the so-called optical flow [7], is the primary visual cue for 2-D and 3-D reconstruction from time-varying imagery. Recent theoretical developments of uncalibrated stereovision have provided a framework for the unified treatment of the motion and stereo problems, that is, the SFM and SFS problems [3]. In this framework, we can interpret stereo disparity, and thus compute it, as the optical flow (image displacement) between the left and right images. Alternatively, optical flow is the disparity map in a stereo system (with a differential baseline).

Typical obstacles in the accurate computation of optical flow include uncontrolled lighting and shadows, insufficient or weak texture, occlusion boundaries, camera noise, image under-sampling, limited camera field of view, etc. Addressing these and other issues has been responsible for continued research on the robust computation of optical flow; e.g., relaxing the brightness constancy assumption [16], oriented Gabor filters [1], treatment of multiple objects and discontinuities [2], etc. Additional challenges in processing the underwater video imagery arise from the unique imaging conditions and natural scene characteristics: illumination from artificial finite-energy sources that are installed on and move with submersible platforms in deep-water operations, medium attenuation, complex topographies of natural objects, etc.

Robustness with respect to various sources of noise and error is typically addressed by carrying out the calculations over small local regions of each pixel; explicitly addressed by various formulations imposing the smoothness constraint. However, a critical issue, when dealing with arbitrary surface topographies, is the localization accuracy which can be compromised by excessive smoothing. In fact, most such algorithms tradeoff accuracy with respect to high-frequency variations in the optical flow, due to sharp surface variations, against immunity to image noise, digitization error, inaccuracy in gradient estimation by the application of finite difference methods, etc. Loss of high-frequency variations defeats the purpose of high-resolution close-range seafloor imaging for 2-D and 3-D mapping, as well as other application that rely on the estimation of accurate 2-D and 3-D motion.

This study explores the robust computation of optical flow by the application of multiple constraints at each pixel, based on the information in color imagery. More precisely, instead of utilizing several constraints from a local region, we would exploit constraints from the three channels at each pixel, which typically may have different and independent noise characteristics. As a result, integration of information from 3 channels can provide more robust results even when the visual cues may be correlated. Moreover, this enables us to take advantage of information redundancy without the need to compromise localization accuracy. Although it has been concluded that there is little advantage in the use of color imagery in land applications, several spectral-dependent factors including medium attenuation, may justify the use of color imagery in underwater. In particular, the important issues investigated in this paper can be summarized as follows:

- Theoretical analysis of the brightness constancy assumption in different channels of color images in the presence of medium attenuation: A simplified underwater image model is employed and the conclusions are then compared with the same results obtained from a variety of real images. This analysis provides an insight on the irradiance component in each channel and how this may vary with the movement of the camera;

- Computation of optical flow based on the different color representations, as well as gray-scale (intensity) channel, and comparison among the results;

- Evaluation of robustness in optical flow computation by varying the estimation window size.

II. RELATED WORK

0-933957-31-9

A. Optical Flow Computation from Color Imagery

The brightness constancy model (BCM)

$$e_x \delta x + e_y \delta y + e_t = 0, \qquad (1)$$

a constraint between the optical flow (δx,δy) and spatio-temporal gradient (e_x,e_y,e_t) of time-varying imagery $e(x, y, t)$, has been used extensively in the estimation of the image motion/disparity field for terrestrial applications [1]. However, this constraint alone is not sufficient to uniquely compute optical flow due to the aperture problem [6]; mathematically, we have one constraint equation in 2 unknowns, per pixel. Ideally, we need 2 or more independent constraints at or around a small enough region around each pixel. For example, one solution is to formulate a least-square problem by utilizing the constraints from N pixels within a local region, to compute the displacement of the center pixel [8]. This ambiguity is resolved in the estimation of disparities in a calibrated stereo system, as the flow is constrained by the epipolar geometry [5]. Still, the LSE formulation is suitable to reduce sensitivity to camera noise, image quantization noise, quantization error in the estimation of image gradients by finite difference, etc. The region size establishes the trade-off between noise immunity and loss of accuracy (localization) due to excessive smoothing.

While only the image irradiance at each point can be obtained from a gray-scale image, a color or multi-spectral image provides information on narrower bands within the visible and invisible spectrum. Therefore, these images can potentially 1) improve the solution of the optical flow problem, 2) overcome the aperture problem and (or) 3) reduce the region size in the LSE formulation (we now have cN constraints, where c is the number of channels). With color video, we can reduce the region size by 3 (to improve localization) for the same number of constraints in the LSE formulation.

If $R(x, y, t)$, $L(\lambda)$, and $S_i(\lambda)$ denote the spectral reflectance of the imaged surface at pixel (x, y) and time t, the spectral power distribution of the incident light radiation, and the spectral response function of sensor i, the reflectance model of the image can be written

$$E_i(x, y, t) = \int_\Lambda L(\lambda) R(x, y, t, \lambda) S_i(\lambda) d\lambda, \qquad (2)$$

From this, we can derive an analytical solution for optical flow based on the BCM model, with a minimum two different sensor sensitivity functions $S_i(\lambda),\ i = 1, 2$. More generally, n sensors give

$$\begin{array}{l} e_x^1 u + e_y^1 v + e_t^1 = 0 \\ \vdots \\ e_x^n u + e_y^n v + e_t^n = 0 \end{array} \qquad (3)$$

Use of the multi-spectral images for optical flow computation was proposed in the 80's, but no result was reported [13, 14, 20]. Improved optical flow computation was demonstrated in later work using two sensors, gray-scale CCD and IR cameras, from the visible and infrared spectra [10]. More recently, the use of different color spaces (RGB, HSV, and rgb) has been compared, showing that the normalized "rgb" typically provide the best results for general motions [4].

B. Underwater Image Modeling

In relatively clear shallow waters (with natural lighting) where scattering may be ignored, the image irradiance of a surface point $P = [X, Y, Z]$ may be written [21]

$$E_I(P, \lambda) = i_o \rho(X, Y) r(\lambda) e^{-c(\lambda) R_c}, \qquad (4)$$

where $R_c = \sqrt{X^2 + Y^2 + Z^2}$ is the distance of object point from the camera, $0 \leq \rho(X, Y) \leq 1$ is the surface albedo, $r(\lambda)$ is the reflectance map, $c(\lambda)$ is water attenuation, and i_o is the source intensity.

C. Generalized Dynamic Image Model

In underwater imagery, both shallow water and deep sea applications, radiometric variations are induced by medium attenuation, surface wave effects (in shallow water) and the source motion (in deep sea operations), leading to the violation of the BCM. The generalized dynamic image model (GDIM), with a multiplier field δm,

$$e_x \delta x + e_y \delta y + e_t - e \delta m = 0, \qquad (5)$$

has been shown as a viable constraint for the robust computation of 2-D optical flow and 3-D camera motion from underwater gray-scale video [16, 17, 18]. The new term $e\delta m$ accounts for the brightness variation of a scene point in the two views, but adds another unknown for a total of three at each pixel.

III. THEORETICAL EVALUATION OF BRIGHTNESS VARIATION IN COLOR OPTICAL FLOW

In processing color images, BCM or GDIM may be appropriate for a particular channel. For example, we may employ the HSV representation and assume "hue and saturation constancy" in the left and right (before and after motion) views, but allow radiometric variation δm_V in the intensity channel. This means that we have three constraints in three unknowns $(\delta x, \delta y, \delta m_V)$ at each pixel. Where the channel gradients are independent, we can theoretically determine the three unknowns at each point, without needing to utilize information from the neighboring pixels. Alternatively, we may explore application of the GDIM for each of the R,G, and B (or "rgb") channels. However, we now have 5 unknowns with three equations at each pixel.

In order to evaluate the significance of δm term for each channel, a theoretical analysis is undertaken, by utilizing the model in (4). Though this is a simplified model, the findings can still provide us with some useful general conclusions. In particular, it is speculated (which remains to be explored) that scattering effects may vary slowly with motion, and thus remain relatively steady over nearby viewing positions.

Replacing the brightness function ($E_I(P, \lambda)$) into (2), we can write

$$\begin{aligned} E_i(x, y, t) &= \int_\Lambda E_I(P, \lambda) S_i(\lambda) d\lambda \\ &= \int_\Lambda i_o \rho(X, Y) r(\lambda) e^{-c(\lambda) R_c} S_i(\lambda) d\lambda, \\ &= i_o \rho(X, Y) \int_\Lambda r(\lambda) e^{-c(\lambda) R_c} S_i(\lambda) d\lambda \end{aligned} \qquad (6)$$

where $i \in \{R, G, B\}$ for RGB representation. Now, δm is the rate of brightness (color) that can be computed from the time derivative of (6):

$$\frac{dE_i}{dt} = \sum_{j=1}^{3} \left(\frac{\partial E_i}{\partial U_j} \right) \left(\frac{dU_j}{dt} \right), \qquad (7)$$

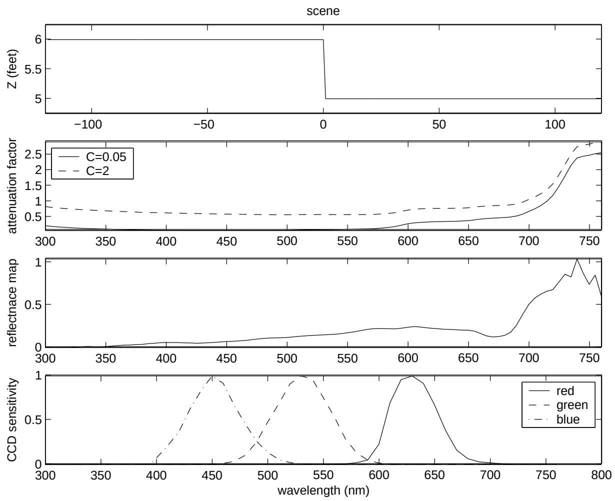

Figure 1: Top to bottom: scene profile, attenuation factor with 2 levels of turbidity, reflectance map, and the spectral sensitivity of the assumed camera for analytical study.

where $U_j = \{X(t), Y(t), Z(t)\}$. We have

$$\begin{aligned}\frac{\partial E_i}{\partial U_j} &= \left(\frac{\partial E_i}{\partial \rho}\right)\left(\frac{d\rho}{dU_j}\right) + \left(\frac{\partial E_i}{\partial R_c}\right)\left(\frac{dR_c}{dU_j}\right)\\ &= \left(\frac{E_i}{\rho}\right)\left(\frac{d\rho}{dU_j}\right) - \frac{\int_\Lambda c(\lambda) r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}{\int_\Lambda r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda} E_i \frac{U_j}{R_c}\\ &= \left(\frac{1}{\rho}\frac{d\rho}{dU_j} - \frac{\int_\Lambda c(\lambda) r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}{\int_\Lambda r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}\frac{U_j}{R_c}\right) E_i.\end{aligned} \tag{8}$$

Applying these partial derivatives to (7), we come up with δm for channel E_i

$$\begin{aligned}\delta m_i = \frac{1}{E_i}\frac{dE_i}{dt} = \frac{1}{\rho}\left(\frac{d\rho}{dX}\frac{dX}{dt} + \frac{d\rho}{dY}\frac{dY}{dt}\right) -\\ \frac{\int_\Lambda c(\lambda) r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}{R_c \int_\Lambda r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}\left(X\frac{dX}{dt} + Y\frac{dY}{dt} + Z\frac{dZ}{dt}\right).\end{aligned} \tag{9}$$

The color representation is assumed to be RGB, but the formulation can also be modified for the other representations, HSV, normalized RGB (rgb), etc.

In order to quantify these results, a scene model, medium spectral attenuation function, reflectance map, sensor spectral sensitivity function, and camera motion are necessary. To simplify the analysis, we consider only one single scan line of the image, corresponding to a one-dimensional surface profile. This is sufficient to explore how a step change on the surface affects the visual information in various channels. Furthermore, we assume that the camera moves forward toward the scene, corresponding to an irradiance gain over the entire scan line. Now, as we are interested in the variation in brightness/color, we simplify (1) for this case:

$$\delta m_i = -\left(\frac{Z \cdot t_Z}{R_c}\right)\frac{\int_\Lambda c(\lambda) r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}{\int_\Lambda r(\lambda) e^{-c(\lambda)R_c} S_i(\lambda) d\lambda}, \tag{10}$$

where t_Z is the forward motion component.

A. Computer Simulation

Without loss of generality, we let $Z_0 = 6'$ as initial distance to the scene with a 1' step change; see fig. 1a. The parameters are chosen such that the forward motion induces a maximum of 4-pixel displacement at the extreme ends.

To assess the magnitude of the multiplier term, we also need to know the attenuation factor, the spectral reflectance map and the spectral sensitivity of camera for different channels. In real world, these parameters depend on the environment and the camera that is used. Here, we refer to published data for underwater environments, underwater objects, and a specific camera. The attenuation factor, $c(\lambda)$, is computed based on an empirical model given in [15] (fig. 1b),

$$c(\lambda) = c_w(\lambda) + [c(490) - c_w(490)][1.563 - 1.149 \times 10^{-3}\lambda] \tag{11}$$

where c_w is the attenuation coefficient for pure sea water, $c(490) = 0.39C^{0.57}$, and C is directly proportional to the level of water turbidity. The spectral reflectance map $r(\lambda)$ is assumed based on the measurements that have been made for the coral reef objects with a DiveSpec [11] (fig. 1c), and the spectral sensitivity of the camera is that of Kodak DCS200 (fig. 1d).

It is noted that the mathematical expression of δm can be derived for other color representations with minor changes. Accordingly, we compute the brightness/color variation factor for RGB, HSV and rgV [1] representations, and for two type of waters, low turbidity $C = 0.05$ and high turbidity $C = 2.0$. Now, though the scattering effects may become significant for a higher turbidity, resulting in deviations in the assumed model, it is still informative to investigate how sensitive are the model and the general conclusions to the turbidity level.

The results, depicted in fig. 2, and table 1, lead to these conclusions:

- RGB representation:
  1. δm_G and δm_B vary more significantly than δm_R, with variation in turbidity (one order of magnitude);
  2. Depth discontinuities is more evident in the red channel (R), but becomes less significant with increased turbidity.
- HSV representation;
  1. δm_H is not significant, compared to δm_S and δm_V;
  2. δm_H and δm_S are relatively independent of turbidity level, but δm_V increases with turbidity;
  3. Scene discontinuity is more evident in the hue and saturation channels.
- rgV representation:
  1. δm_g is insignificant;
  2. δm_r and δm_g are relatively independent of turbidity level;
  3. Scene discontinuity is more evident in g channel.

[1] We have chosen rgV in place of commonly used rg, to include the intensity channel corresponding to gray-scale imagery.

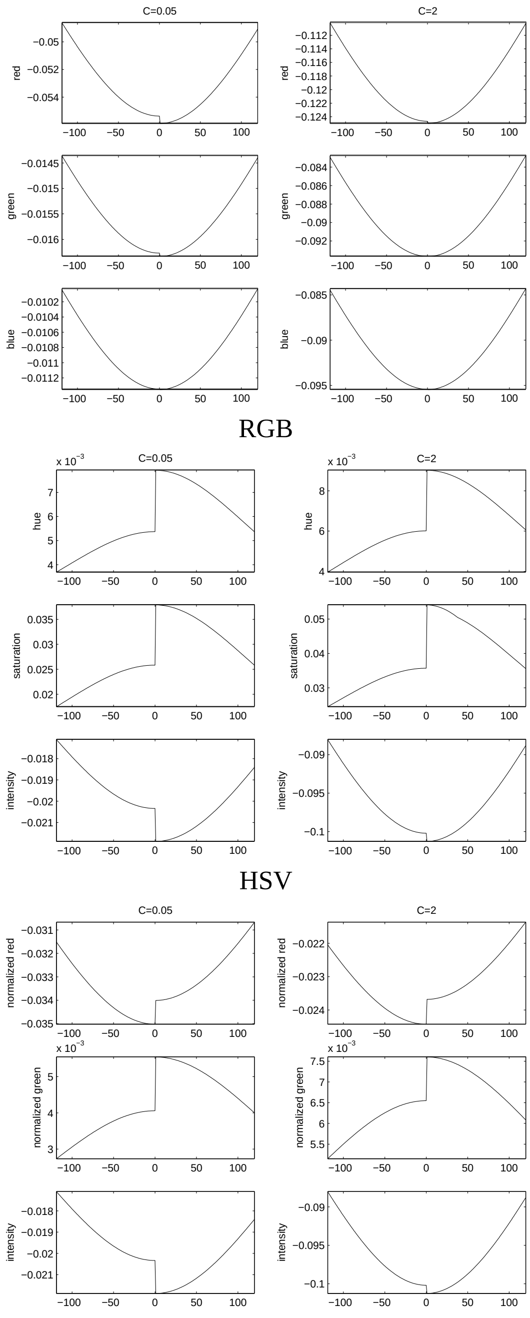

Figure 2: Simulation results; image brightness of various channels for RGB, HSV and rgV representation with low turbidity (left column) and high turbidity (right column).

	C=0.05			C=2.0		
	$\mu_{\delta m_1}$ $\sigma_{\delta m_1}$	$\mu_{\delta m_2}$ $\sigma_{\delta m_2}$	$\mu_{\delta m_3}$ $\sigma_{\delta m_3}$	$\mu_{\delta m_1}$ $\sigma_{\delta m_1}$	$\mu_{\delta m_2}$ $\sigma_{\delta m_2}$	$\mu_{\delta m_3}$ $\sigma_{\delta m_3}$
RGB	-0.05 0.00	-0.02 0.00	-0.01 0.00	-0.12 0.00	-0.09 0.00	-0.09 0.00
HSV	0.01 0.00	0.03 0.01	-0.02 0.00	0.01 0.00	0.04 0.01	-0.10 0.00
rgV	-0.03 0.00	0.00 0.00	-0.02 0.00	-0.02 0.00	0.01 0.00	-0.10 0.00

Table 1: Mean and std of radiometric field δm in the simulation study.

The simulation results allows us to quantify δm, average size of δm in each channel for the three representations, and the effects of depth discontinuity or sharp variations. In the next section, we apply the generalized dynamic image model in (5) in estimating the optical flow and δm for 7 pairs of real underwater images with varying conditions of the medium, illumination and scene structure. The results of these computations are compared to those based on the model in (9).

B. Real Data

The optical flow and δm were computed for a consecutive pair from 7 sets of video images from shallow coral reef in the Bahamas; sample images depicted in table 2. Care was taken in choosing pairs with a dominantly forward motion, in consistency with the theoretical analysis. Table 2 summarizes the result of the estimation process in the form of mean and standard deviation of δm over the image.

Analyzing the results, the following observations can be made:

- δm_R,δm_G, and δm_B channels are almost the same;
- δm_H is not significant;
- δm_g is smaller than δm_r and both are insignificant compared to δm_V.

Comparing tables 1 and 2, discrepancies are noted (and expected) in that the simulated data uses a specific camera model, target reflectance pattern, medium conditions, camera motion and a simple scene geometry. However, the analytical evaluation of the brightness variation factor δm for different channels provides a rough estimate (order) of the magnitude for real data. One can also see that δm_H is smaller than the other two for HSV representation, as concluded from analytical study (Table 1). One may expect to arrive at more consistent conclusions with (1) a more complete underwater image formation model in (4), e.g., including the scattering effects [19], and (2) comparable values of attenuation coefficient, spectral reflectance map and sensor spectral sensitivity to those for real data.

We can compare the tabulated results for the simulated and real data more directly in fig. 3, where the dashed line shows $\mu_{\delta m}$ for the simulation case, while the solid lines represent $\mu_{\delta m} \pm 2.5\sigma_{\delta m}$ for various real data sets. As can be seen, the analytical value falls within the range for real data, providing a reason order of magnitude estimate.

IV. EMPIRICAL EVALUATION OF COLOR OPTICAL FLOW

image	color rep.	$\mu_{\delta m_1}$ / $\sigma_{\delta m_1}$	$\mu_{\delta m_2}$ / $\sigma_{\delta m_2}$	$\mu_{\delta m_3}$ / $\sigma_{\delta m_3}$
image 1 	RGB	-0.013	-0.012	-0.016
		0.011	0.008	0.012
	HSV	-0.004	-0.021	-0.009
		0.013	0.021	0.008
	rgV	-0.000	-0.001	-0.012
		0.004	0.003	0.009
image 2 	RGB	-0.013	-0.011	-0.014
		0.011	0.007	0.009
	HSV	-0.004	-0.023	-0.009
		0.012	0.026	0.007
	rgV	0.001	-0.000	-0.012
		0.005	0.004	0.008
image 3 	RGB	-0.033	-0.029	-0.034
		0.017	0.011	0.014
	HSV	-0.009	-0.027	-0.027
		0.022	0.029	0.011
	rgV	0.001	-0.000	-0.031
		0.006	0.004	0.013
image 4 	RGB	-0.012	-0.006	-0.008
		0.018	0.015	0.018
	HSV	-0.002	-0.026	-0.005
		0.031	0.049	0.015
	rgV	-0.005	0.001	-0.007
		0.006	0.005	0.016
image 5 	RGB	-0.028	-0.011	-0.016
		0.023	0.015	0.015
	HSV	-0.003	-0.017	-0.009
		0.009	0.023	0.015
	rgV	-0.005	0.000	-0.014
		0.009	0.005	0.015
image 6 	RGB	-0.047	-0.031	-0.032
		0.045	0.027	0.027
	HSV	-0.001	-0.028	-0.027
		0.009	0.024	0.026
	rgV	-0.001	-0.003	-0.032
		0.016	0.006	0.029
image 7 	RGB	-0.044	-0.019	-0.023
		0.027	0.024	0.027
	HSV	0.001	-0.016	-0.013
		0.005	0.017	0.017
	rgV	-0.008	-0.001	-0.024
		0.012	0.005	0.026

Table 2: Mean and std over the entire image of the radiometric fields (δm_1, δm_2 and δm_3) for each data set.

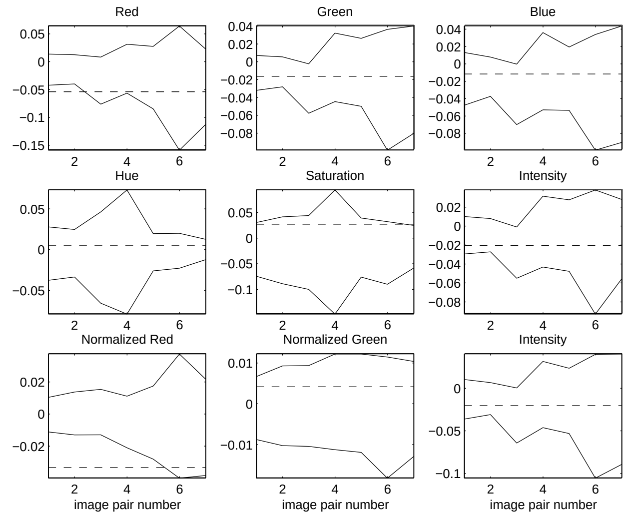

Figure 3: Comparison of simulated scene and real images. The dashed line shows $\mu_{\delta m}$ of each channel from analytical model. The solid lines represent $\mu_{\delta m} \pm 2.5\sigma_{\delta m}$ for various data sets.

So far, 1) we have utilized a model for the brightness variation δm for the channels in various color representations, to estimate the significance of this factor in optical flow estimation from color imagery; investigated consistency with real data in terms of the (order of) magnitude. The balance of the paper deals with the optical flow, and its robust computation. The same 7 pairs of real images are used to compare the performance of optical flow algorithm for color and gray-scale images, and various color representations. In the absence of ground-truth for optical flow components, the emphasis is on the *robustness*, as becomes evident. More precisely, the size of the computation window [9], controlling the imposed degree of smoothness in the optical flow field, becomes of great importance. The main question is: What is the optical flow variations of a representative pixel whose surrounding points have almost the same depth and motion, as we vary the window size? Therefore, the behavior of the optical flow for varying window sizes is used to assess the accuracy and stability of each method. In particular, does a particular method stands out that generally gives robust estimates? And if so, can we expect accurate enough results to achieve good localization by reducing the window size (for maintaining the high frequency variations over regions with sharp topographical changes)? In addressing these questions, we first select a representative point (marked by a small square in Table 2: One that has 1) a relatively similar flow compared to its neighborhood; 2) roughly the same robustness characteristics as the average taken over the entire image. The optical flow is estimated for different representations of color images and for intensity channel (gray-scale). The window size is varied from 2x2 to 11x11, where the upper limit is established by analyzing the data for convergence (see below).

Figure 4 shows a summary of the results of optical flow computations: Column 1 shows the magnitude $\sqrt{\delta x^2 + \delta y^2}$ of the optical flow of the selected point for various window sizes. The result has been derived for gray-scale, RGB, HSV, and rgV representations. As one notes, in all but one case, the estimates converge to a similar solution for size 11– chosen

data set	μ_{gray}	μ_{RGB}	μ_{HSV}	μ_{rgV}
1	1.03	0.85	0.79	0.85
2	0.57	0.47	0.35	0.45
3	0.89	0.70	0.54	0.71
4	0.61	0.51	0.41	0.52
5	1.01	0.83	0.75	0.79
6	1.20	0.80	0.73	0.70
7	0.93	0.79	0.54	0.74

Table 3: For the 7 data sets, average normalized difference in optical flow magnitudes for window sizes 2 and 11 (scaled by average flow over the image in each case) have been calculated from the histogram of the errors given in the last column of fig. 4.

as the limit for this study. However, while the performance of various methods are comparable for larger window sizes (reduced localization accuracy), the accuracy for smaller sizes is of special interest. Among these methods, results for gray-scale imagery are generally in the lower end of accuracy and robustness scales. Among various color representations, HSV stands out and thus is compared to performance with gray scale imagery in subsequent comparisons. The middle 2 columns is the estimated optical flow, with window size 2x2, for the HSV and gray-scale images. The uniformity of the flow is immediately obvious by visual inspection. Finally, the last column is the histogram over the entire image of the difference in the estimated optical flow magnitudes between window sizes 2 and 11. In other words, *if we assume the "converged" value for size 11 to be "ground truth"*, this represents the histogram of the error magnitudes. Again, a higher/lower peaks of HSV/gray at lower error magnitudes verify the earlier results. To summarize these results across various images, we have calculated both the error of the selected point and the mean error over the entire image, for the 7 data sets. Since the optical flow magnitudes are different for various images, the results have been normalized with respect to the average flow magnitude of each case. These normalized average errors are produced in table 3. The normalized errors of the selected point for various methods is given in fig. 5 (left), augmented with additional results for methods based on HS and rg channels (that have been used in some earlier studies [4]).

Finally, we have repeated the same computations, in each case, for consecutive frames over a portion of the video where the camera translation is parallel to the scene. In this case, the optical flow is (almost) constant over a small region with uniform depth. The difference between the estimated flows for window sizes 2x2 and 11x11 have been computed, averaged over region and normalized by the average flow magnitude of each case. This is given in fig. 5 (right).

IV. SUMMARY

This paper presents an investigation on the computation of optical flow from gray-scale and color imagery. A simplified physics-based model of underwater image formation–accounting for medium attenuation in shallow water, is the basis of a theoretical analysis to determine radiometric variations in various color channels. In particular, the goal was to establish if certain channels of various color representations may satisfy the brightness constancy assumption commonly used in the processing of land images [1]. Next, robustness and accuracy of optical flow was tackled by analyzing the results from 7 data sets, for various window sizes. Robustness for small window sizes translates to more efficient processing and better localization accuracy over surface regions with sharp topographical variations. The results of this study suggest the advantage in the use of underwater color imagery for optical flow computations, and particularly the HSV representation appears to provide superior estimates. Estimates for the smallest window size of 2x2 was evaluated in details. This provides 12 constraints (3 channels for 4 points) on a maximum of 5 unknowns, 2 optical flow components and δm for the 3 channels. The unknowns may be reduced by eliminating δm_H, which was shown to be insignificant based on the analytical and empirical results.

Acknowledgement

This article is based upon work supported by the Office of Naval Research (ONR) under Grant No. N000140310074. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the ONR.

References

[1] Baron, J.L., D.J. Fleet, and S.S. Beauchemin, "Performance of optical flow techniques," *IJCV*, Vol 12(1), 1994 (also TR-299, Dept. Comp. Science, University Western Ontario, March, 1992).

[2] Black, M., and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comp. Vision & Image Und.*, Vol 63(1), 1996.

[3] Faugeras, O., "What can be seen in three dimensions with an uncalibrated stereo rig?," *European Conf. in Comp. Vision*, 1992.

[4] Golland, P., "Use of Color for optical flow," M.S. Thesis, Computer Science Dept., Israel Institute of Technology, 1995.

[5] Hartley, R., and A. Zisserman, "Multiple View Geometry," Cambridge Univ. Press 2000.

[6] Hildreth, E. "Measurement of Visual Motion," MIT Press, Cambride, MA, 1980.

[7] Horn, B.K.P. and Schunck, B.G., "Determining Optical Flow," *Artif. Intelligence*, Vol 17, 1981.

[8] Kearney, J.K., W.B. Thompson, and D.L. Boley, "Optical flow estimation: An Error Analysis of Gradient-Based Methods with Local Optimization," *IEEE PAMI*, Vol 9, No 2, March, 87.

[9] Lucas, B.D. and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," *Proc. IJCAI*, 1981, 674-679.

[10] Markandey, V., B.E.Flinchbaugh, "Multispectral Constraints for Optical Flow Computation," *ICCV*, 1990.

[11] Mazel, C. H., "Diver-operated instrument for in situ measurement of spectral fluorescence and reflectance in benthic marine organisms and substrates," Optical Engineering 36: 2612-2617.

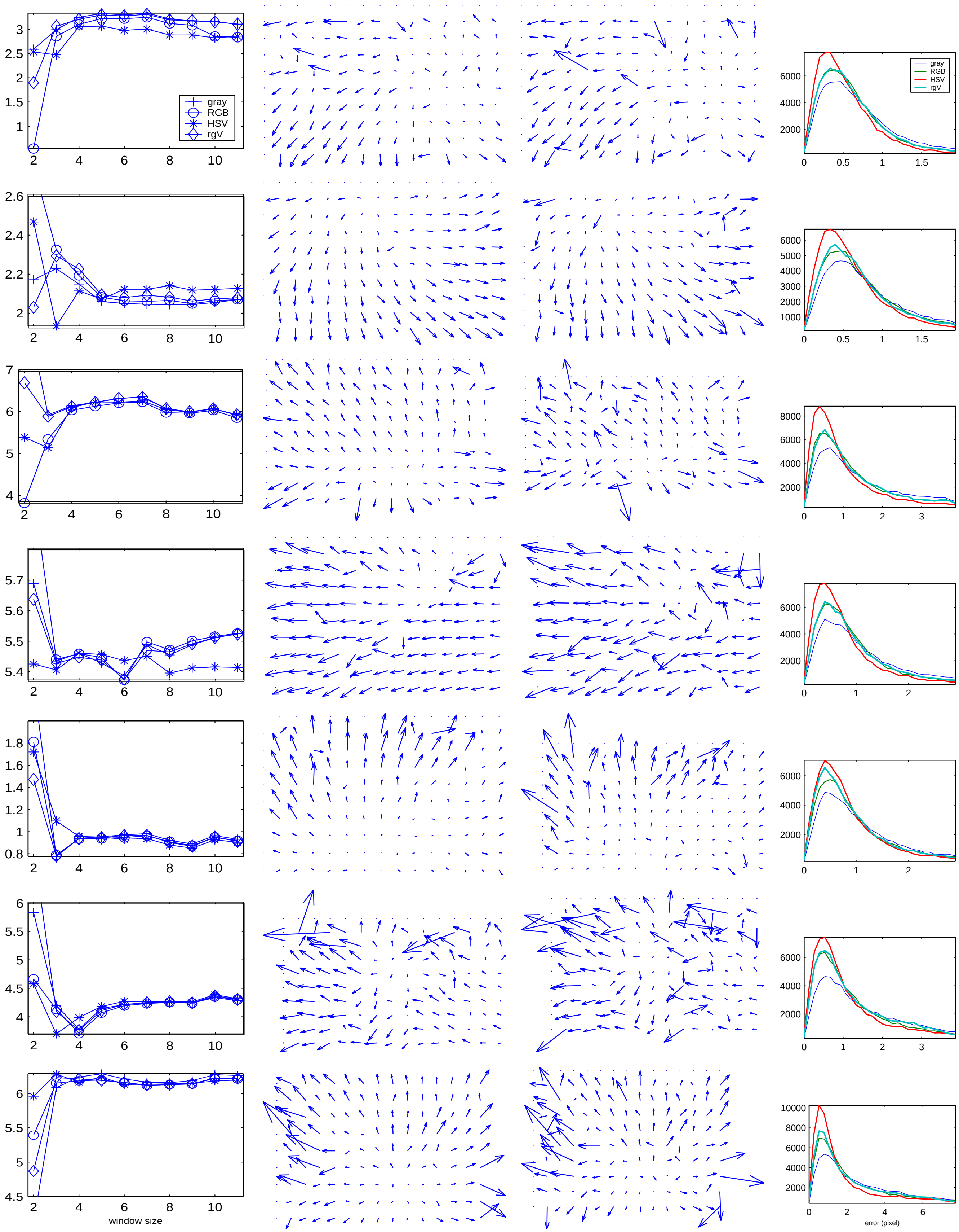

Figure 4: The result of optical flow computation on 7 pairs of real images including: optical flow magnitude for different window sizes (1^{st} column), optical flow field with window size of 2 based on HSV (2^{nd} column), and the same based on just intensity (3^{rd} column), histogram of the the difference between optical flow magnitudes for window size of 11 and 2 computed for all the points (4^{th} column).

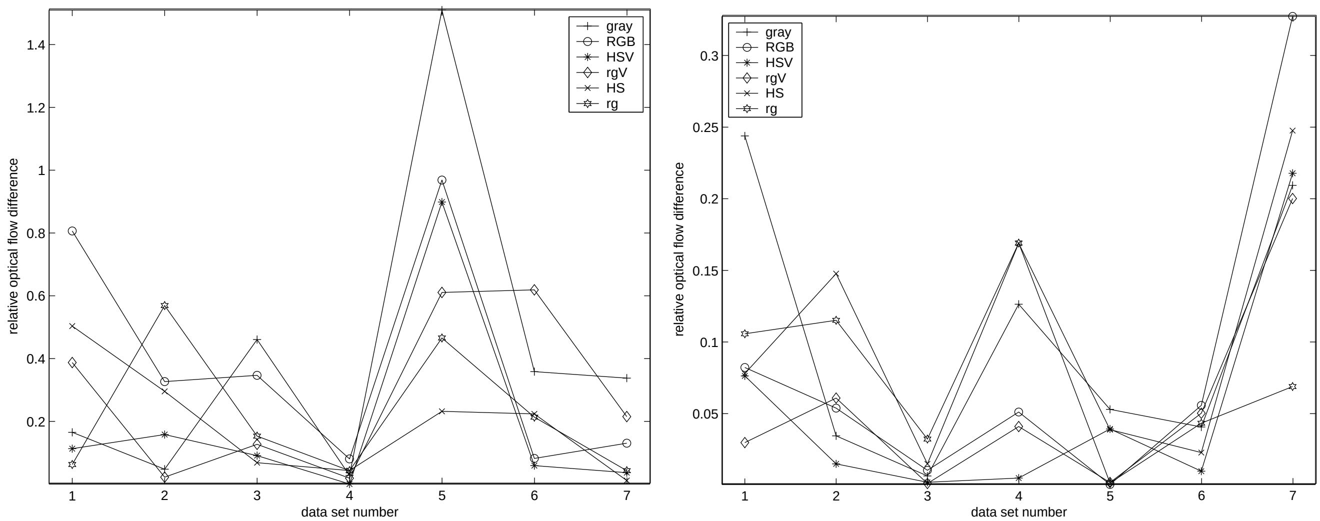

Figure 5: For 7 sample images, normalized difference between optical flow magnitudes for window sizes 2 and 11 have been computed for various color representations. Left is the result at the single representative point for a segment of the video that corresponds to forward camera motion. Right gives the average of the same error over a region with relatively uniform flow (for a part of the video with motion parallel to the scene).

[12] McGlamery, B.L., "Computer analysis and simulation of underwater camera system performance," SIO Ref. 75-2, Scripps Inst. of Oceanography, UCSD, January, 1975 (Also in A computer model for underwater camera systems, Ocean Optics VI, Proceedings of SPIE, Monterey, California, October, 1979.

[13] Mitiche, A., "Computation of optical flow and rigid motions," *Proc. Workshop on Computer Vision: Representation and Control*, 1984.

[14] Mitiche, A., Y.F. Wang, and J.K. Aggrawal, "Experiments in computating pptical flow with the gradient-based, multiconstraint methods," *Pattern Recognition*, Vol. 20(2), 1988.

[15] Mobley, C., "Light and Water, Radiative Transfer in Natural Waters," *Academic Press*, 1994.

[16] Negahdaripour, S., "Revised Interpretation of Optical Flow; Integration of Radiometric and Geometric Cues in Optical Flow," *IEEE Trans. Pattern Analysis & Machine Intelligence*, Vol 20, No 9, September, 1998.

[17] Negahdaripour, S., X. Xu, L. Jin, "Direct estimation of motion from sea floor images for automatic station-keeping of submersible platforms," *IEEE Journal of Oceanic Engineering*, July, 1999.

[18] Negahdaripour, S., X. Xu, "Mosaic-based positioning and improved motion estimation methods for autonomous navigation," *IEEE Journal Oceanic Engineering*, Vol 27 (1), January, 2002.

[19] Negahdaripour, S., H. Zhang, and X. Han, "Investigation of Photometric Stereo Method for 3-D Shape Recovery from Underwater Imagery," Proceedings of MTS/IEEE Oceans '02 Conference, Biloxi, Mississippi, October 29-31, 2002.

[20] Wohn, K., L.S.Davis, and P.Thrift, "Motion estimation based on multiple local constraints and nonlinear smoothings," *Pattern Recognition*, Vol 16 (6), 1983.

[21] Zhang, S., and S. Negahdaripour, "3-D shape recovery of planar and curved surfaces from shading cues in underwater images," *IEEE Trans. Oceanic Eng.*, Vol 27 (1), January, 2002.