Dense Range Flow from Depth and Intensity Data

Hagen Spies, Bernd Jähne
Interdisciplinary Center for Scientific Computing
University of Heidelberg
INF 368, 69120 Heidelberg, Germany
{Hagen.Spies,Bernd.Jaehne}@iwr.uni-heidelberg.de

John L. Barron
Department of Computer Science
University of Western Ontario
London, Ontario, N6G 5B7 Canada
barron@csd.uwo.ca

Abstract

The combined use of intensity and depth information greatly helps in the estimation of the local 3D movements (range flow) of moving surfaces. We demonstrate how the two can be combined in both a local total least squares algorithm and in an iterative global variational technique. While the first assumes locally constant flow the second method relies on a smoothly varying flow field. The improvement achieved through incorporating intensity is illustrated qualitatively and quantitatively on synthetic and real test data.

1. Introduction

From a sequence of depth maps the 3D motion of the observed surface can be computed [7, 8], we denote this motion field *range flow*. When the depth maps are obtained from stereo, structured lighting, depth from X or active laser triangulation we typically have some intensity information as well. As this intensity is perfectly registered with the depth data it can be used to aid the range flow estimation. This contribution shows how intensity and depth data can be jointly used in a total least squares (TLS) estimation technique based on a locally constant flow model and in an iterative regularization scheme based on globally smooth flow.

The estimation of the 3D motion parameters (translation plus rotation) of either a moving sensor or a rigid object are what most previous work on range sequence analysis is concerned with [4, 6]. Towards this end intensity and depth have also been used together [1, 3]. The algorithms presented here estimate the local 3D displacements of not necessarily rigid bodies. This is related to previously reported model based range flow estimation on non-rigid surfaces [12, 9]. The 3D range flow can also be recovered from optical flow if other surface properties such as depth or correspondences are available [10].

2. Constraint Equations

We consider depth as a function of space and time $Z = Z(X,Y,T)$, from the total derivative with respect to time we obtain the range flow motion constraint equation (RFMC) [4, 12, 8]:

$$Z_X U + Z_Y V + W + Z_T = 0 \ . \tag{1}$$

Here indices denote partial derivatives. For the intensity data the well known brightness change constraint equation (BCCE) can be used [5]. As we do possess intensity values for each measured 3D point given in the sensors coordinate system this directly gives a constraint equation in the X and Y movements:

$$I_X U + I_Y V + I_T = 0 \ . \tag{2}$$

We now have two equations in three unknowns. Even if they are independent, which is not always the case, we still need to make further assumptions to solve for the range flow. In Sect. 3 it is shown how a TLS solution can be obtained when a locally constant flow field is assumed. Section 4 describes how a solution is obtained based on the premiss of a smoothly varying flow field.

2.1. Scaling the Constraints

When combining the two constraints on intensity (2) and depth (1) their relative weight has to be taken into account. In the range data we are working with, see Sect. 5.2, the intensity values are typically by about a factor of 100 larger than the depth values. Thus the former will dominate any numerical calculation. Therefore we choose to weight the two constraints by a factor $\beta^2 = \frac{<||\nabla Z||^2>}{<||\nabla I||^2>}$, computed from average intensity and depth gradient magnitudes calculated on a representative data set for the given application.

With this scaling, intensity and depth are now on an equal footing. However, this might not always be desirable. An additional factor, e.g. based on the noise levels, can easily be incorporated into β^2.

3. Local TLS Estimation

The RFMC (1) and the BCCE (2) are recast in a total least squares framework [7, 8]:

$$\begin{aligned} u_1 Z_X + u_2 Z_Y + u_3 + u_4 Z_T &= 0 \ , & (3) \\ u_1 I_X + u_2 I_Y + u_4 I_T &= 0 \ . & (4) \end{aligned}$$

The desired range flow is then found from:

$$\vec{f} = [U\ V\ W]^T = \frac{1}{u_4}[u_1\ u_2\ u_3]^T \ . \qquad (5)$$

In order to avoid the trivial solution we require the solution vector $\vec{u} = [u_1\ u_2\ u_3\ u_4]^T$ to be normalized. Incorporating this requirement by means of a Lagrangian multiplier λ the total least squares solution in an aperture A (typically 5×5) can be found from minimizing an energy functional:

$$\begin{aligned} J &= \int_A \left[(u_1 Z_X + u_2 Z_Y + u_3 + u_4 Z_T)^2 \right. \\ &+ \beta^2 (u_1 I_X + u_2 I_Y + u_4 I_T)^2 \\ &+ \left. \lambda \left(\vec{u}^T \vec{u} \right) \right] \mathrm{d}X \mathrm{d}Y \ \rightarrow \min \ . \qquad (6) \end{aligned}$$

An obvious extension to the presented approach is to extend the integration in (6) to the temporal domain. As that would increase the number of frames used this is currently not done. Setting the partial derivatives of J with respect to all four variables $u_{i=1\ldots4}$ to zero and moving those variables out of the integral (they are assumed to be locally constant) yields conditions on the choice of λ. It is easily shown that for these conditions to hold λ has to be an eigenvalue of the following extended structure tensor $\boldsymbol{F}$:

$$\begin{bmatrix} < Z_X Z_X > + \beta^2 < I_X I_X > & < Z_X Z_Y > + \beta^2 < I_X I_Y > \\ < Z_Y Z_X > + \beta^2 < I_X I_Y > & < Z_Y Z_Y > + \beta^2 < I_Y I_Y > \\ < Z_X > & < Z_Y > \\ < Z_T Z_X > + \beta^2 < I_X I_T > & < Z_T Z_Y > + \beta^2 < I_Y I_T > \end{bmatrix}$$

$$\begin{bmatrix} < Z_X > & < Z_X Z_T > + \beta^2 < I_X I_T > \\ < Z_Y > & < Z_Y Z_T > + \beta^2 < I_Y I_T > \\ < 1 > & < Z_T > \\ < Z_T > & < Z_T Z_T > + \beta^2 < I_T I_T > \end{bmatrix}$$

Here the integration is abbreviated as $\int \cdot = < \cdot >$, which becomes a convolution with a smoothing operator (typically Box or Binomial) in the discrete case. The minimum is given by the smallest eigenvalue and the range flow is then found from the corresponding eigenvector from equation (5). The magnitude of the smallest eigenvalue directly indicates the success of the fit, this can for instance be exploited to detect motion boundaries between differently moving objects [8].

3.1. The Aperture Problem

Neither the use of both depth (RFMC) and intensity (BCCE) nor the pooling of those constraints over an aperture A guarantees the presence of three linear independent constraints. Yet, such are necessary for a unique range flow estimation. If there are enough constraints available all but the smallest eigenvalue of $\boldsymbol{F}$ are bigger than zero. This allows for the recovery of *full flow* as described above.

Using only depth planar structures result in only one independent constraint, the solution with minimum norm has thus been called *plane flow* in this case. As linear structures yield only two constraints the corresponding minimum norm solution is termed *line flow*. In both cases the appropriate normal flow can be recovered from the eigenvectors corresponding to the non-vanishing eigenvalues [8].

For the combined depth and intensity analysis the same situations may occur and we will call the corresponding normal flows as before. One of the key points we want to emphasize is that the use of intensity information does enable the computation of denser full flow fields, see Sect. 5.

4. Global Smoothness

The computation of smoothly varying range flow presented here is an extension of the well known optical flow algorithm by Horn and Schunk [5]. Combining the RFMC (1), the BCCE (2) and a simple membrane model yields the following energy to be minimized:

$$\begin{aligned} J(U,V,W) &= \int_A \left[(Z_X U + Z_Y V + W + Z_T)^2 \right. \\ &\quad + \beta^2 (I_X U + I_Y V + I_T)^2 \\ &\quad + \left. \alpha^2 (\nabla U^2 + \nabla V^2 + \nabla W^2) \right] \mathrm{d}X \mathrm{d}Y \qquad (7) \end{aligned}$$

The influence of the smoothness term is regulated by the constant factor α^2. This time the integration area A may be the entire image or subsets that have been previously segmented. Clearly this approach fails if there is a motion discontinuity within A. In that case, an anisotropic smoothness term as used in recent optical flow algorithms, is certainly a better choice [11].

The minimum of the energy J in equation (7) is reached when the Euler-Lagrange equations are satisfied. If we replace the occurring Laplacian with the difference between an average and the central pixel $\Delta f = \bar{f} - f$ (here $\bar{\cdot}$ denotes the average, best computed with a Binomial) we obtain a set of linear equations:

$$\begin{bmatrix} \alpha^2 + Z_X^2 + \beta^2 I_X^2 & Z_X Z_Y + \beta^2 I_X I_Y & Z_X \\ Z_X Z_Y + \beta^2 I_X I_Y & \alpha^2 + Z_Y^2 + \beta^2 I_Y^2 & Z_Y \\ Z_X & Z_Y & \alpha^2 + 1 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix}$$

$$= \begin{bmatrix} \alpha^2 \bar{U} - Z_X Z_T - \beta^2 I_X I_T \\ \alpha^2 \bar{V} - Z_Y Z_T - \beta^2 I_Y I_T \\ \alpha^2 \bar{W} - Z_T \end{bmatrix} \ .$$

Let us denote the matrix as $\boldsymbol{A}$, then we can iteratively compute the solution at step n+1 from step n:

$$\begin{bmatrix} U^{n+1} \\ V^{n+1} \\ W^{n+1} \end{bmatrix} = \boldsymbol{A}^{-1} \begin{bmatrix} \alpha^2 \bar{U}^n - Z_X Z_T - \beta^2 I_X I_T \\ \alpha^2 \bar{V}^n - Z_Y Z_T - \beta^2 I_Y I_T \\ \alpha^2 \bar{W}^n - Z_T \end{bmatrix} \ . \qquad (8)$$

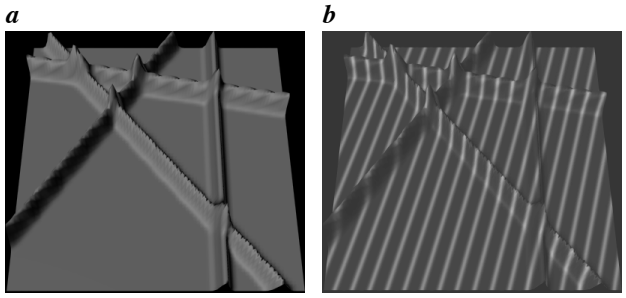

Figure 1. Rendered synthetic test sequence: a) depth only and b) with intensity texture.

It can be shown that $\boldsymbol{A}$ is always regular provided $\alpha^2 > 0$ [8]. Initialization is currently done with zero. We aim to show that the use of the intensity as well as the depth does provide good results after considerably fewer iterations.

5. Experiments

Apart from qualitative proof that the use of the intensity improves range flow estimation we also want to demonstrate the accuracy of the proposed methods quantitatively. To do so we report the relative error in the flow magnitude and the deviation between the correct 3D movement and the estimated displacement in degrees [7, 8].

5.1. Synthetic Data

To illustrate the effect of the additional intensity information we use the synthetic sequence shown in Fig. 1, the correct movement is $[0.66, -0.46, 0.34]^T$. Using depth maps alone ($\beta^2 = 0$) local TLS estimation only gives sparse full flow, Fig. 2a. The flat surface only allows the computation of plane flow perpendicular to it. On the lines any movement in the direction of the lines can not be resolved. When the linear intensity pattern (Fig. 1b) is taken into consideration the displacements along the lines can now be computed. Hence the full flow density increases from 7% to 38%, see Fig. 2b. Clearly this is not the case if intensity and depth lines coincide. On the planar surface the intensity allows for the computation normal to the intensity lines, yet one direction remains ambiguous. Both the relative error and the directional error remain very low on this ideal test case ($< 1\%$ and $< 1^\circ$).

The global smoothness algorithm can in principle compute the correct flow everywhere even in the depth only case at 5000 iterations. However, due to the large regions with only limited information such a large number of iterations is required. That the incorporation of the intensity yields superior results when equal iteration numbers are compared can be seen by studying Fig. 2c and d. In numbers the following results are obtained ($\beta^2 = 1;\ \alpha^2 = 10$):

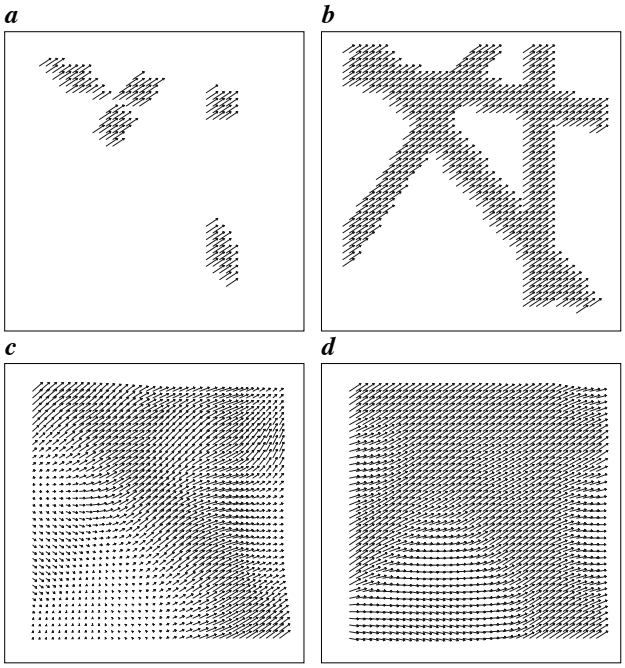

Figure 2. X-Y-range flow. TLS: a) depth only, b) with intensity. Global Smoothness (500 iterations): c) depth only, d) with intensity.

data	iterations	rel mag error [%]	dir error [°]
Z only	100	42.6 ± 20.6	40.4 ± 22.7
Z only	500	30.1 ± 21.5	24.0 ± 21.5
Z only	1000	23.0 ± 20.6	16.4 ± 18.9
Z only	5000	7.8 ± 11.5	3.9 ± 5.9
I and Z	100	18.3 ± 11.1	22.7 ± 16.9
I and Z	500	10.9 ± 10.7	11.2 ± 13.3
I and Z	1000	7.4 ± 9.0	6.6 ± 9.6
I and Z	5000	0.9 ± 1.6	0.5 ± 1.0

Clearly the combination converges considerably faster. However in the very end both methods provide excellent results.

5.2. Real Data

We present work on data captured with a *Biris* laser range finder [2], here the intensity is the brightness of the reflected laser beam. Thus we are actually looking at monochromatic images. Figure 3 shows our real test sequence, the correct movement of $[-0.16, 0.19, -0.32]^T$mm was achieved by placing the object (a toy tiger) on a system of linear positioners. Due to rapid depth changes there are some holes in the data, those were left out in the computation.

The X-Y components of the computed flow fields are shown in Fig. 4. The weight is chosen to be $\beta^2 = 0.05$, see Sect. 2.1. The increased density in the TLS case is apparent. Some numbers may illustrate the algorithms behavior:

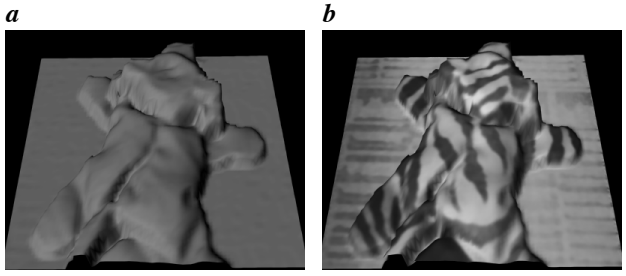

Figure 3. Real test sequence: a) depth only and b) with intensity texture.

data	density [%]	rel mag error [%]	dir error [°]
Z only	10.5	13.8 ± 15.4	12.7 ± 10.3
I and Z	59.0	7.9 ± 9.0	9.9 ± 8.6

Not only does the density increase but also the error decrease.

Figure 4c,d indicate that the combined use of intensity and depth in the global flow algorithm achieves a better estimate after the same number of iterations. Numerical results are ($\alpha^2 = 10$):

data	iterations	rel mag error [%]	dir error [°]
Z only	500	12.4 ± 7.6	13.5 ± 9.5
Z only	1000	9.0 ± 7.0	8.9 ± 6.8
Z only	5000	4.2 ± 4.3	5.0 ± 2.8
I and Z	500	7.5 ± 5.0	9.8 ± 5.8
I and Z	1000	4.8 ± 3.4	6.6 ± 4.3
I and Z	5000	4.7 ± 3.4	6.3 ± 4.0

Thus we conclude that using both data sets speeds up the calculation dramatically. Slightly better results are found from the depth alone. This is because the depth data is less noisy for this kind of sensor [1]. To get the best performance and speed the algorithm can be started with a high intensity influence (β^2) which is subsequently lowered to zero.

6. Conclusions

The integrated use of intensity and depth data could be shown to greatly improve 3D range flow estimation. Typically the additional intensity information helps to reduce the aperture problem, thus we can compute denser flow fields. The accuracy of the proposed methods was demonstrated on both synthetic and real test data. The methods are currently used to investigate the diurnal motion patterns of castor oil leaves [8].

References

[1] D. Ballard and O. Kimball. Rigid body motion from depth and optical flow. *CVGIP*, 22:95–115, 1983.

[1] If noise levels are given by σ_I^2 and σ_Z^2 for intensity and depth respectively we could update the influence factor as: $\beta^2 \rightarrow \frac{\sigma_Z^2}{\sigma_I^2}\beta^2$.

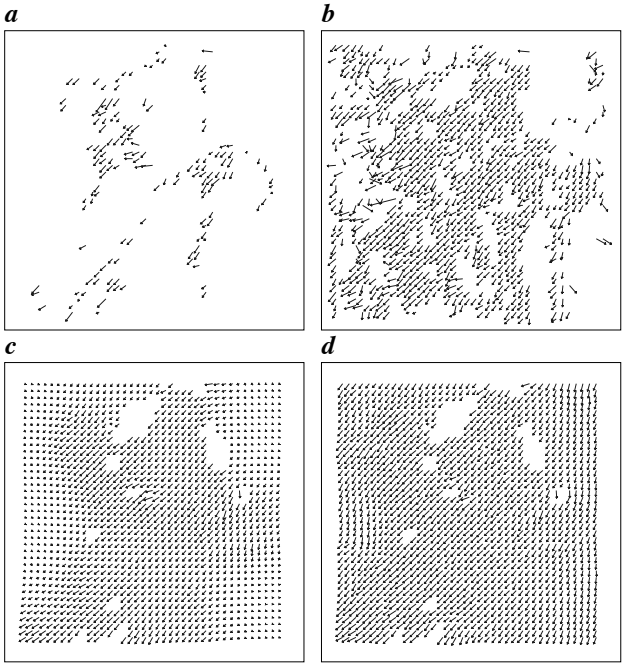

Figure 4. X-Y-range flow. TLS: a) depth only, b) with intensity. Global Smoothness (500 iterations): c) depth only, d) with intensity.

[2] J.-A. Beraldin, S. El-Hakim, and F. Blais. Performance evaluation of three active vision systems built at the national research council of canada. In *Conf. on Optical 3D Measurement Techniques III*, pages 352–361, Oct. 1995.

[3] M. Harville, A. Rahimi, T. Darrell, G. Gordon, and J. Woodfill. 3d pose tracking with linear depth and brightness constraints. In *ICCV '99*, pages 206–213, 1999.

[4] B. Horn and J. Harris. Rigid body motion from range image sequences. *CVGIP*, 53(1):1–13, Jan. 1991.

[5] B. Horn and B. Schunk. Determining optical flow. *Artificial Intelligence*, 17:185–204, 1981.

[6] B. Sabata and J. Aggarwal. Estimation of motion from a pair of range images: A review. *CVGIP*, 54(3):309–324, Nov. 1991.

[7] H. Spies, H. Haußecker, B. Jähne, and J. Barron. Differential range flow estimation. In *21.Symposium fur Mustererkennung, DAGM '1999*, pages 309–316, Sept. 1999.

[8] H. Spies, B. Jähne, and J. Barron. Regularised range flow. In *ECCV '2000*, June 2000.

[9] L. Tsap, D. Goldgof, and S. Sarkar. Model-based force-driven nonrigid motion recovery from sequences of range images without point correspondences. *Image and Vision Computing*, 17(14):997–1007, 1999.

[10] S. Vedula, S. Baker, P. Rander, R. Collins, and T. Kanade. Three-dimensional scene flow. In *ICCV '99*, pages 722–729, 1999.

[11] J. Weikert. On discontinuity-preserving optic flow. In *CVMR '98*, pages 115–122, 1998.

[12] M. Yamamoto, P. Boulanger, J. Beraldin, and M. Rioux. Direct estimation of range flow on deformable shape from a video rate range camera. *PAMI*, 15(1):82–89, Jan. 1993.