Determining Spatial Motion Directly from Normal Flow Field: A Comprehensive Treatment

Tak-Wai Hui and Ronald Chung

Department of Mech. and Automation Engineering,
The Chinese University of Hong Kong, Hong Kong
twhui1@mae.cuhk.edu.hk, rchung@cuhk.edu.hk

Abstract. Determining motion from a video of the imaged scene relative to the camera is important for various robotics tasks including visual control and autonomous navigation. The difficulty of the problem lies mainly in that the flow pattern directly observable in the video is generally not the full flow field induced by the motion, but only partial information of it, which is known as the normal flow field. A few methods collectively referred to as the direct methods have been proposed to determine the spatial motion from merely the normal flow field without ever interpolating the full flows. However, such methods generally have difficulty addressing the case of general motion. This work proposes a new direct method that uses two constraints: one related to the direction component of the normal flow field, and the other to the magnitude component, to determine motion. The first constraint presents itself as a system of linear inequalities to bind the motion parameters; the second one uses the rotation magnitude's globality to all image positions to constrain the motion parameters further. A two-stage iterative process in a coarse-to-fine framework is used to exploit the two constraints. Experimental results on benchmark data show that the new treatment can tackle even the case of general motion.

1 Introduction

A moving object or camera generally induces a certain apparent flow pattern in the acquired video. How the relative motion in space between the object and camera can be determined from the apparent flow is a classical problem whose solution has tremendous applications to autonomous navigation, visual control, robotics, human action understanding, and intelligent user interface.

Due to the well-known ambiguity between motion speed and object size-and-depth, from monocular video alone the translation magnitude of motion is generally not determinable and left as an overall arbitrary scale related to object depth. In other words, if we describe the spatial motion as consisting of a translation component $\mathbf{t}$ (as a 3D displacement vector) and a rotation component $\mathbf{w}$ (in rotation's angle-axis form), we are to determine the direction of $\mathbf{t}$ and the full $\mathbf{w}$. However, due to the familiar *aperture problem*, the full flow induced by the spatial motion at any image position is observable generally only partially, as a *normal flow*. This partial observability makes motion determination a challenge.

R. Koch et al. (Eds.): ACCV 2010 Workshops, Part I, LNCS 6468, pp. 23–32, 2011.

Classical solutions to motion determination are largely about establishing explicit motion correspondences across the image frames. One is the feature-based track, which tracks the distinct features in the image stream [1, 5]. Another track, originated from Horn and Schunck [10] as well as Lucas and Kanade [14], tracks practically all image positions. In [2], the multi-scale approach was extended by including the gradient constancy assumption to overcome the drawback of using the grey value constancy assumption. Higher order constancy assumptions had also been included into the variational model [15]. After estimating the full flow field, motion parameters are estimated in a subsequent stage. A method of using constraint lines to estimate the focus of expansion (FoE) was developed by [12]. However, the rotational component cannot be recovered. Linearity and divergence properties of the orthogonal cross-section of the projected flow fields were used in [7]. An iterative approach is used to de-rotate the image by feeding back the estimated rotational component in the next iterative cycle. A recent research work presents a linear formulation of the bilinear constraint [16]. It also pointed out that highly accurate estimate for full flow does not necessarily provide an accurate estimation of ego-motion because it also depends on the error characteristic within the estimated full flow field.

A few methods have been proposed to determine camera motion from the normal flow field directly. Such methods are often referred to as the direct methods. A classical direct method is [11], in which only camera motion having pure translation, pure rotation, or a motion with known rotation can be recovered. Another direct method that is based upon selection of image points that form global patterns in the image was developed by [8]. They transformed the parameter estimation problem to a series of boundary extraction problems. The method needs to determine boundaries between two sparsely labeled regions in the image domain. Another direct method utilizes a ψ-line searching algorithm to determine the direction of $\mathbf{t}$ [17]. However, only a limited number of normal flows could participate in the recovery. In a later work, the searching problem is formulated on the L-space [18]. More normal flows could contribute to the solution. Yet each set of affine parameters of r on each ψ-line of the L-space are still estimated by a limited number of normal flows. Both two methods require the application of minimization over unbounded decision variables.

The direct methods in the literature either cannot tackle the case of general motion in which both translation and rotation are present and unknown, or have to deal with the problem of identifying boundaries between image regions that are only sparsely labeled. In this paper, we describe a direct method that is free of such limitations. We propose the use of two constraints: one related to the direction component of the normal flows, and the other to the globality of the motion magnitude. While the former manifests as a system of linear inequalities that bind the motion parameters, the latter serves to constrain the motion parameters further by insisting that every image position must have a component of normal flow magnitude that is consistent with a global rotation magnitude. A two-stage iterative voting process, in a coarse-to-fine framework, is implemented

to determine the motion parameters. Experimental results on benchmark data show that the method is capable of recovering general motion.

2 The Apparent Flow Direction (AFD) Constraint

An earlier report[4] has laid down the foundation on how normal flow direction imposes a constraint on the spatial motion parameters. Here we provide a more geometric intuition of the constraint as well as a more complete formulation of it. In general, optical flow $\dot{\mathbf{p}}$ at any image position $\mathbf{p}$ is not directly observable from the image because of the well-known aperture problem. Only the projected component of the flow to the spatial intensity gradient at the position, by the name of normal flow $\dot{\mathbf{p}}_n$, is directly observable, and it is generally computed from the spatial and temporal derivatives of the intensity profile at $\mathbf{p}$.

The constraint is powerful but simple to state. If normal flow must be a projected component of full flow, the full flow must have an orientation no different from that of the observed normal flow by more than 90°. This restriction on the direction of the full flow in turn constrains the motion parameters ($\mathbf{t}$, $\mathbf{w}$). The constraint at $\mathbf{p}$ with non-zero full flow and normal flow can be formulated as:

$$\dot{\mathbf{p}} \cdot \dot{\mathbf{p}}_n = (\dot{\mathbf{p}} \cdot \mathbf{n})^2 > 0 \tag{1}$$

where $\mathbf{n}$ is a unit vector in the direction of the local intensity gradient at $\mathbf{p}$.

2.1 Preliminaries

Consider a camera that has the image plane placed at focal length f from its optical center. Define a camera-centered coordinate system C, which has the Z-axis pointing along the optical axis toward the imaged scene. Consider any object point $\mathbf{P} = (X,\ Y,\ Z)^T$ of the scene with respect to C. Suppose that its projection onto the image plane is $\mathbf{p} = (x,\ y)^T$. Under perspective projection,

$$\mathbf{P} \cong \widetilde{\mathbf{p}} \tag{2}$$

where $\cong$ denotes equality up to arbitrary nonzero scale, and $\widetilde{\mathbf{p}} = (x/f, y/f, 1)^T$ represents projective coordinates of $\mathbf{p}$ with f-normalized x and y components.

Suppose that the camera undergoes a general motion with a translation $\mathbf{t}$ and a rotation $\mathbf{w}$. The motion of the point $\mathbf{P}$ relative to frame C is:

$$\dot{\mathbf{P}} = -\mathbf{t} - \mathbf{w} \times \mathbf{P} \tag{3}$$

Suppose also that the normal flow $\dot{\mathbf{p}}_n$ of orientation γ in the image space is observed at image position $\mathbf{p}$. The motion parameters $\mathbf{t}$ and $\mathbf{w}$ must be such that they induce a full flow (at $\mathbf{p}$) that has a direction no different from γ by more than 90°. A few further algebraic manipulations will turn Equation (1) to:

$$AFD(\mathbf{p}, \gamma) : \mathbf{t} \cdot \mathbf{a}_t / Z - \mathbf{w} \cdot \mathbf{a}_w < 0 \tag{4}$$

where $\mathbf{a}_t = \widetilde{\mathbf{p}} \times (\sin\gamma, -\cos\gamma, 0)^T$, $\mathbf{a}_w = \mathbf{a}_t \times \widetilde{\mathbf{p}}$, and Z is the scene depth at $\mathbf{p}$. This is the constraint imposed by normal flow direction γ. It is expressed in a form more precise than that in [4]. We refer to it as the Apparent Flow Direction (AFD) constraint.

2.2 The Special Case: Pure Translation

Suppose the camera undergoes a pure translation $\mathbf{t}$ that has a component toward the imaged scene. All the optical flows should be pointing away from the focus of expansion (FoE). The full flow $\dot{\mathbf{p}}$ induced by the FoE could be any vector of orientation between (γ-90°) and (γ+90°) in the image. The case that the camera undergoes a pure translation that has a component away from the imaged scene is similar, except that the FoE is replaced by the focus of contraction (FoC).

By setting the $\mathbf{w}$-component to zero in the $AFD(\mathbf{p}, \gamma)$ constraint, the locus of $\mathbf{t}$'s direction in space (regardless of whether it represents an FoE or FoC) is

$$AFD_t(\mathbf{p}, \gamma) : \widehat{\mathbf{t}} \cdot \widehat{\mathbf{a}}_t < 0 \tag{5}$$

where $\widehat{\mathbf{t}}$ and $\widehat{\mathbf{a}}_t$ are unit vectors of $\mathbf{t}$ and $\mathbf{a}_t$ respectively. In essence, $AFD_t(\mathbf{p}, \gamma)$ is a linear inequality on the direction of $\mathbf{t}$, representing exactly half of the of the parameter space of $\mathbf{t}$'s direction.

2.3 The Special Case: Pure Rotation

The case of pure rotation is analogous to that of pure translation. Suppose the camera rotates about an axis $\mathbf{w}$ (in the right-hand manner) with an angular velocity given by the magnitude of $\mathbf{w}$. By setting the $\mathbf{t}$-component to zero in the $AFD(\mathbf{p}, \gamma)$ constraint, the locus of $\mathbf{w}$ imposed by $\mathbf{p}$ and γ can be expressed as:

$$AFD_w(\mathbf{p}, \gamma) : \widehat{\mathbf{w}} \cdot \widehat{\mathbf{a}}_w > 0 \tag{6}$$

where $\widehat{\mathbf{w}}$ and $\widehat{\mathbf{a}}_w$ are unit vectors of $\mathbf{w}$ and $\mathbf{a}_w$ respectively. The locus is again in the form of a linear inequality that binds the direction of $\mathbf{w}$.

2.4 Solving the System of Linear Inequalities for the Two Special Cases

Suppose there are m data points (image positions have observable normal flows) in the image. They will each give rise to an inequality described in either Equation (5) or Equation (6), about the directions of $\mathbf{t}$ and $\mathbf{w}$ respectively. Define $\mathbf{A}_t = [\widehat{\mathbf{a}}_t]_{m\times 3}, \mathbf{A}_w = [-\widehat{\mathbf{a}}_w]_{m\times 3}$, the entire set of inequalities can be expressed as either

$$\mathbf{A}_t\widehat{\mathbf{t}} < 0 \ \texttt{or} \ \mathbf{A}_w\widehat{\mathbf{w}} < 0 \tag{7}$$

The parameter space of $\mathbf{t}$ and $\mathbf{w}$ can be parameterized by the spherical coordinates $(\rho = 1, \phi, \theta)$. With a number of data points available, each supplying a different locus for $\mathbf{t}$ (or $\mathbf{w}$). Here we provide a geometric interpretation of the task, and supply an alternative solution mechanism that is computationally more efficient but also more accurate in its solution. The task in hand (as expressed by Equation (7)) is about seeking a 3-vector $\mathbf{n}$ (which is about either $\mathbf{t}$ or $\mathbf{w}$) that makes no acute angle with any row vector of a matrix $\mathbf{A}$ (which could be either $\mathbf{A}_t$ or $\mathbf{A}_w$, depending upon whether it is a case of pure translation or pure rotation). Notice that only the direction of $\mathbf{n}$, not its magnitude, is desired.

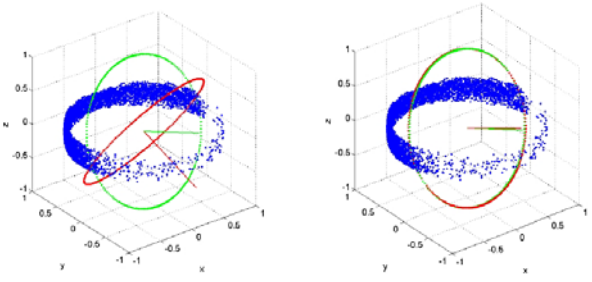

Fig. 1. Distribution of noisy $\hat{\mathbf{a}}_t$'s and the estimation of translation $\mathbf{t}$.(a) Original solution. (b) Solution using a resampling-based method. Red and green lines are the estimated and true directions of $\mathbf{t}$ respectively.

If the row vectors of $\mathbf{A}$ are viewed as radial vectors of a unit sphere in 3-space, the solution is about designating a pole $\mathbf{n}$ of the sphere, so that if the pole is regarded as the north pole, a maximal number of the above radial vectors lie in the sphere's south hemisphere. In this light, an alternative way of acquiring the solution to the inequality system could be formulated as an optimization problem: given $\mathbf{A}$, seek a unit 3-vector $\mathbf{n}$ such that

$$\mathbf{n} = \arg \min_{subject\ to\ \mathbf{n} \in \mathbf{R}^3, ||\mathbf{n}||=1} \sum (\widehat{\mathbf{a}}_i \cdot \mathbf{n}) \tag{8}$$

where $\mathbf{a}_i (i = 1, 2, ..., m)$ are normalized row-vectors of $\mathbf{A}$.

To avoid solution of local minimum, at the end of each minimization process we randomly re-sample the space of decision variables around the stabilized values, and use such sampled values to initialize the decision variables for re-running the minimization process. We iterate the process until random distribution of the stabilized decision values does not lower the objective function further. Figure 1 shows example results of recovering $\mathbf{t}$ on a synthetic dataset. The normal flows were corrupted by an additive white Gaussian noise of 5dB signal-to-noise ratio. The above resampling-based minimization mechanism was able to supply a direction of $\mathbf{t}$ (Figure 1(b)) much closer to the ground truth than the one without using resampling-based minimization (Figure 1(a)). The angular error in $\mathbf{t}$ was reduced from 52.9° to 1.6°. The case of pure rotation is similar. It is omitted here.

2.5 The Case of General Motion

When there are both translation and rotation in the motion, normal flows are affected by both $\mathbf{t}$ and $\mathbf{w}$ in unknown proportions. By observing Equation (4), it can be deduced that it cannot be $\mathbf{t} \cdot \mathbf{a}_t$ being positive and $\mathbf{w} \cdot \mathbf{a}_w$ being negative at the same time, or else the inequality expressed by Equation (4) will be violated. From the geometric perspective, and in particular the principle of vector addition, the same conclusion can be reached, as illustrated by Figure 2.

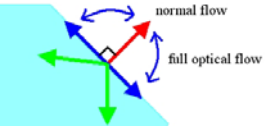

Fig. 2. Relationship between full flow and normal flow

Consider any image position $\mathbf{p}$ where the normal flow is observable. The direction γ of the normal flow introduces respective bounds, in the form of $AFD_t(\mathbf{p}, \gamma)$ and $AFD_w(\mathbf{p}, \gamma)$, to the directions of the vectors $\dot{\mathbf{p}}_t$ and $\dot{\mathbf{p}}_w$, which are optical flow components induced by the translation $\mathbf{t}$ and rotation $\mathbf{w}$ of the motion respectively. As the full flow must have an overall orientation between (γ-90°) and (γ+90°), it is impossible that both the two optical flow components $\dot{\mathbf{p}}_t$ and $\dot{\mathbf{p}}_w$ point to the shaded region shown in Figure 2. This imposes a constraint to ($\mathbf{t}$,$\mathbf{w}$), in the form of a logical OR operation over $AFD_t(\mathbf{p}, \gamma)$ and $AFD_w(\mathbf{p}, \gamma)$. More precisely, each of { $\mathbf{t}$,$\mathbf{w}$ } ought to span exactly half of their own direction space, and with ($\mathbf{t}$,$\mathbf{w}$) (their directions only not magnitudes) considered together, they span three-quarter of a 4-D parameter space. In other words, the AFD constraint allows each single normal flow to trim away one-quarter of ($\mathbf{t}$,$\mathbf{w}$)-direction-space from consideration.

3 The Apparent Flow Magnitude (AFM) Constraint

The AFD constraint is applicable regardless of whether the motion is a specific one or a general one. However, whether it alone could lead to a unique result of the motion parameters depend upon how general is the orientation distribution of the normal flow field, which is a function of the intensity-gradient distribution of the image. Our experimental results show that for some specific scenes, the AFD constraint could only restrict the motion parameters to a possible set of solutions, not a unique one. In this work we explore also how the magnitude information of the field can be brought in to narrow down the motion values further.

Consider the case of pure rotation. At any image position $\mathbf{p}$ the rotation $\mathbf{w}$ induces an optical flow $\dot{\mathbf{p}}_w$. This optical flow $\dot{\mathbf{p}}_w$ then projects to the local intensity gradient to form the observable normal flow $\dot{\mathbf{p}}_n$. If the angle between $\widehat{\dot{\mathbf{p}}}_w$ ($||\mathbf{w}||$-normalized $\dot{\mathbf{p}}_w$) and $\dot{\mathbf{p}}_n$ is α, we have

$$||\mathbf{w}|| = ||\dot{\mathbf{p}}_n||/(||\widehat{\dot{\mathbf{p}}}_w|| \cos\alpha) \tag{9}$$

In other words, any hypothesis of the rotation axis $\widehat{\mathbf{w}}$ will allow the rotation magnitude $||\mathbf{w}||$ to be determined from any image position $\mathbf{p}$ where normal flow is observable. Since this rotation magnitude is a global quantity for the entire normal flow field, it should have the same value from any particular normal flow data point. As a consequence, the consistency over $||\mathbf{w}||$ of the normal flow field could serve to confirm if the hypothesized rotation axis is correct or not. We refer

to this constraint on the motion parameters as the Apparent Flow Magnitude (AFM) constraint.

In our implementation, for any particular hypothesis of the rotation axis $\widehat{\mathbf{w}}$, the standard deviation (SD) of $||\mathbf{w}||$ computed from all the normal flow data points in the image space serves to indicate how likely the hypothesis is correct; an SD that is too large will make it justifiable to have the particular hypothesis of $\widehat{\mathbf{w}}$ ruled out. Notice that the AFM constraint also supplies the true value of $||\mathbf{w}||$ once the final answer on the rotation axis $\widehat{\mathbf{w}}$ is attained.

It is worth pointing out that, in Equation (9), an α close to 90° will amplify the uncertainty in estimating $||\mathbf{w}||$. We refer to such normal flow data points as the degenerate data points. It is necessary that in the above process such normal flow data points are excluded from computing the SD measure. Figure 3 shows the distribution of $||\mathbf{w}||$-estimation in a synthetic data experiment for a $\widehat{\mathbf{w}}$ that is only 0.1° away from the ground truth. Yet the SD of the distribution is still large, due mainly to the fact that the degenerate data points are not excluded from the estimation process. In the figure, those $||\mathbf{w}||$-estimates that are from the degenerate data points are marked with red circles. If such data points are excluded, the SD can be reduced from 0.183° to 0.0253°.

To extend the use of the AFM constraint to the case of general motion, we adopt a simple trick. For each particular set of $\left(\widehat{\mathbf{t}}, \widehat{\mathbf{w}}\right)$ (here we use $\widehat{\mathbf{t}}$ and $\widehat{\mathbf{w}}$, the normalized $\mathbf{t}$ and $\mathbf{w}$, to denote their directions), we find out which normal flow data point $\mathbf{p}$ has the intensity gradient direction $\widehat{\mathbf{u}} = \nabla I/||\nabla I||$ (where *I(x,y)* refers to the intensity distribution in the associated image) orthogonal to the optical flow component $\widehat{\dot{\mathbf{p}}}_t$ induced by the particular $\widehat{\mathbf{t}}$. At such $\mathbf{p}$'s, the optical flow $\dot{\mathbf{p}}$ and in turn the normal flow $\dot{\mathbf{p}}_n$ is solely governed by the optical flow component $\widehat{\dot{\mathbf{p}}}_w$ that is induced by $\widehat{\mathbf{w}}$. At such data points the above analysis could still apply. To summarize, by defining a set $\mathit{\Pi}_\perp\left(\widehat{\mathbf{t}}\right) = \left\{\mathbf{p} : \widehat{\dot{\mathbf{p}}}_t \perp \widehat{\mathbf{u}}(\mathbf{p})\right\}$, the AFM constraint for the case of general motion is expressible as:

$$AFM : \left(\widehat{\mathbf{t}}^T, \widehat{\mathbf{w}}^T\right)^T = \arg \min_{\widehat{\mathbf{t}},\widehat{\mathbf{w}} \in \mathbf{R}^3,\ ||\widehat{\mathbf{t}}||,||\widehat{\mathbf{w}}||=1 \text{ and } \mathbf{p} \in \mathit{\Pi}_\perp(\widehat{\mathbf{t}})} SD||\mathbf{w}(\mathbf{p})|| \qquad (10)$$

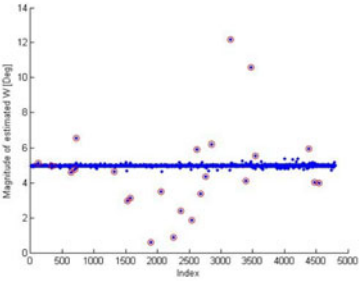

Fig. 3. The distribution of $||\mathbf{w}||$-estimates in a synthetic data experiment

4 Putting the Two Constraints Together

We adopt an iterative two-stage process to apply the AFD and AFM constraints alternately for determining motion. The first stage involves the use of AFD to constrain the directional components $\left(\widehat{\mathbf{t}}, \widehat{\mathbf{w}}\right)$ of the motion. Such a process could give rise to unique solution of the directional components, but not always. In case that it does not, the constrained $\left(\widehat{\mathbf{t}}, \widehat{\mathbf{w}}\right)$ will go through the second stage, which is about the application of the AFM constraint.

To speed up the processing and to avoid the effect of local minima, a coarse-to-fine strategy over the bounded $\left(\widehat{\mathbf{t}}, \widehat{\mathbf{w}}\right)$-parameter space is used. A set of motion vectors $\left(\widehat{\mathbf{t}}, \widehat{\mathbf{w}}\right)$ parameterized by spherical coordinates ϕ and θ are first generated using a coarse sampling resolution. In the first stage, the AFD constraint is applied to reject impossible combinations of $\left(\widehat{\mathbf{t}}, \widehat{\mathbf{w}}\right)$. For each of such motion vectors, the total number of normal flow data points that fulfill the AFD constraint is also recorded. Those motion vectors that have too few of such data points are rejected. The AFM constraint is applied to refine the solution further in the second stage. The motion vectors that do not have small enough SD of $||\mathbf{w}||$ are rejected. The magnitude of $\mathbf{w}$ is also estimated at the same time. The result is then carried forward to serve as seeds of new sampling points in the next iteration that adopts a slightly finer resolution. The iterations continue until motion values of enough precision are attained.

5 Experimental Results on Benchmark Data

Two sets of data are both benchmark data used in the literature, namely the Fountain sequence (FS) [16] and the very widely used Yosemite sequence (YS). Both datasets are about general motions. One latest result is that of the method presented in [16], which makes use of high-accuracy optical flow estimation as input flow (meaning that the method has to use smoothness constraint of some sort for interpolating the full flows from the apparent flows), namely the methods of Brox *et al.* [2] and Farnebäck [6]. To make our experiments results directly comparable to those results, the same pair of image frames was used in each experiment. The ground truth of the FS is $\mathbf{t}$ = (-0.6446 0.2179 2.4056)^T pixels and $\mathbf{w}$ = (-0.125 0.2 -0.125)^T deg/frame [16]. The ground truth of YS is $\mathbf{t}$ = (0 0.17 0.98)$^T\times$ 34.8 pixels and$\mathbf{w}$ = (1.33 9.31 1.62)$^T \times 10^{-2}$ deg/frame [9].

To determine the normal flow at any particular image position, we simply use the equality $\dot{\mathbf{p}}_n = -I_t \nabla I / ||\nabla I||^2$. The spatial ($\nabla I$) and temporal ($I_t$) derivatives are calculated by the 8-point method described by Horn and Schunck [10]. To ease the differentiation process, the image data were smoothed by 2D Gaussian filter beforehand. FS and YS had 16.46% and 26.59% detectable normal flows respectively. The iterative two-stage process, with a coarse-to-fine strategy, was then applied to each of the datasets to determine the direction $\hat{\mathbf{t}}$, the rotation axis $\hat{\mathbf{w}}$, and the rotation magnitude $||\mathbf{w}||$. Tables 1 and 2 show the results in

Table 1. Errors of motion estimation on the Fountain sequence

Technique	Ang. error, $\mathbf{t}$ [deg]	Absolute error, $\mathbf{w}$ [deg/frame] $\lvert\Delta w_x\rvert$	$\lvert\Delta w_y\rvert$	$\lvert\Delta w_z\rvert$
Raudies *et al.*[16] (Brox *et al.*(2D) [2])	4.395	0.001645	0.0286	0.02101
Raudies *et al.* [16] (Farnebäck [6], 100% density)	6.841	0.01521	0.05089	0.025
Raudies *et al.* [16] (Farnebäck [6], 25% density)	1.542	0.0008952	0.01349	0.003637
Our method	0.740115	0.0071692	0.010532	0.018084

Table 2. Errors of motion estimation on the Yosemite sequence without clouds

Technique	Ang. error, $\mathbf{t}$ [deg]	Absolute error, $\mathbf{w}$ [deg/frame] $\lvert\Delta w_x\rvert$	$\lvert\Delta w_y\rvert$	$\lvert\Delta w_z\rvert$
Raudies *et al.*[16] (Brox *et al.*(2D) [2])	4.893	0.02012	0.1187	0.1153
Raudies *et al.* [16] (Farnebäck [6], 100% density)	4.834	0.03922	0.00393	0.07636
Lourakis [13]	3.7	0.038732	0.028419	0.011516
Heeger *et al.* [9]	3.5	0.0568	0.0344	0.0807
Lourakis [12]	3.1			
Raudies *et al.* [16] (Farnebäck [6], 25% density)	1.208	0.007888	0.01178	0.02633
Raudies *et al.* [16] (Farnebäck[6], 25% density, RANSAC)	1.134	0.01261	0.008485	0.02849
Our method	0.988969	0.0037171	0.012006	0.025528

comparison with those reported in the literature. Our method achieved better result in recovering the translational direction as compared with the existing methods. On the recovery of rotation, overall speaking, the proposed method also achieved results comparable to or better than those of the existing methods. Our method performed slightly less than [16] on estimating $\mathbf{w}$, yet it is worth noticing that the method reported in [16] is a full flow-based one and requires to introduce the smoothness assumption somewhere in attaining the full flows, which could cause a problem for scenes that are not smooth enough.

6 Conclusion and Future Work

We have described two important constraints that allow the normal flow field to be used directly for motion determination. One constraint is related to the directional information of the normal flow field, and the other to its magnitude information. We have also outlined how the two constraints can be used alternately in an iterative fashion to determine the motion parameters. Notice also that the method, being a direct method itself, does not require the presence of

distinctly trackable features in the imaged scene, nor does it require to assume smoothness about the imaged scene for interpolating the full flows. Experimental results over benchmark datasets show that not only can the method recover general motion, it also has a performance comparable even to those of methods that have to bring in the smoothness assumption as well. Future work will address how multiple motions in the imaged scene can be detected, and how depth can be recovered from visual motion.

Acknowledgement. The work described was partially supported by the Chinese University of Hong Kong 2009-2010 Direct Grant (Project No. 2050468).

References

1. Armangue, X., Araujo, H., Salvi, J.: A review on egomotion by means of differential epipolar geometry applied to the movement of a mobile robot. Patten Recognition 36(12), 2927–2944 (2003)
2. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. In: Pajdla, T., Matas, J(G.) (eds.) ECCV 2004. LNCS, vol. 3024, pp. 25–36. Springer, Heidelberg (2004)
3. Bruss, A.R., Horn, B.K.P.: Passive navigation. In: Int'l Conf. on Computer Vision Graphic and Imagining, vol. 21, pp. 3–20 (1983)
4. Chung, R., Yuan, D.: Direct determination of camera motion from normal flows. In: Int'l Conf. on Visualization, Imaging, and Image Processing, pp. 153–157 (2009)
5. Cipolla, R., Okamoto, Y., Kuno, Y.: Robust structure from motion using motion parallax. In: Int'l Conf. Computer Vision, pp. 374–382 (1993)
6. Farnebäck, G.: Polynomial expansion for orientation and motion estimation. PhD thesis, Dept. of Electrical Engineering, Linköping University (2002)
7. Fejes, S., Davis, L.S.: What can projections of flow fields tell us about the visual motion. In: Int'l Conf. on Computer Vision, pp. 979–986 (1998)
8. Fermüller, C., Aloimonos, Y.: Qualitative egomotion. IJCV 15, 7–29 (1995)
9. Heeger, D.J., Jepson, A.D.: Subspace methods for recovering rigid motion I: algorithm and implementation. IJCV 7(2), 95–117 (1992)
10. Horn, B.K.P., Shunck, B.G.: Determining optical Flow. AI 17, 185–203 (1981)
11. Horn, B.K.P., Weldon, E.J.: Direct methods for recovering motion. IJCV 2, 51–76 (1988)
12. Lourakis, M.I.A.: Using constraint lines for estimating egomotion. In: Asian Conf. on Computer Vision, pp. 971–976 (2000)
13. Lourakis, M.I.A.: Egomotion estimation using quadruples of collinear image points. In: European Conf. on Computer Vision, pp. 834–848 (2000)
14. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision. In: Int'l Conf. on Computer Vision, pp. 933–938 (1981)
15. Papenberg, N., Bruhn, A., Brox, T., Didas, S., Weickert, J.: Highly accurate optic flow computation with theoretically justified warping. IJCV 67(2), 141–158 (2000)
16. Raudies, F., Neumann, H.: An efficient linear method for the estimation of egomotion from optical flow. In: Denzler, J., Notni, G., Süße, H. (eds.) DAGM. LNCS, vol. 5748, pp. 11–20. Springer, Heidelberg (2009)
17. Silva, C., Santos-Victor, J.: Robust egomotion estimation from the normal flow using search subspaces. IEEE Trans. PAMI 19(9), 1026–1034 (1997)
18. Silva, C., Santos-Victor, J.: Egomotion estimation on a topological space. In: Int'l Conf. on Pattern Recognition, pp. 64–66 (1998)