A GLOBAL ANALYSIS OF OPTICAL SNOW FOR ARBITRARY CAMERA MOTIONS

Vincent Chapdelaine-Couture[1] and Sébastien Roy[2]

University of Montreal (DIRO), Montreal, Quebec, H3T 1J4 Canada
[1] **email: chapdelv@iro.umontreal.ca**
[2] **email: roys@iro.umontreal.ca**

Abstract

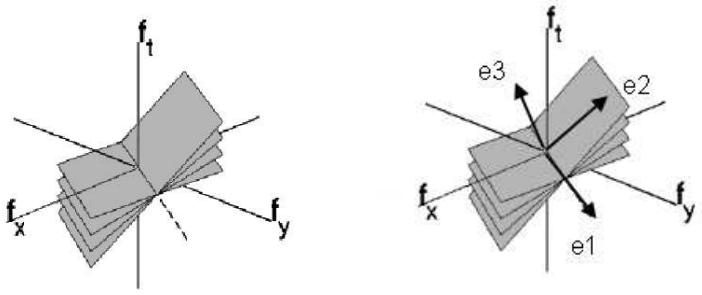
Optical snow, introduced in [4], is a new category of motion for highly cluttered scenes in which no spatial continuity can be assumed. Since no smoothness constraint can be imposed on the velocity field, traditional optical flow methods can no longer be used [1]. However, a model of optical snow has been proposed in [5] and algorithms based on this model were suggested using an analysis in the spatio-temporal frequency domain [5, 2]. This model assumes lateral motion and can be used to solve the 3D camera motion problem by decomposing sequences in sufficiently small patches [7]. We would like to use the same model to find arbitrary camera motions globally instead of using patches. In the present paper, we introduce a complementary model for purely non-lateral optical snow. The standard optical snow model and this complementary form could lead to a new global approach for solving the general egomotion problem. We show how non-lateral optical snow sequences can be rectified such that standard methods to analyze optical snow can be applied. The effectiveness of the method is shown for both real and synthetic sequences.
Keywords*: Optical snow, Egomotion, Fourier transform, Motion analysis, Optical flow*

1 Introduction

Most studies assume a unique velocity at each point in the visual field [1]. This assumption is only valid if the depth map is continuous. If an observer moves in a 3D highly cluttered scene, a forest for instance, this assumption no longer holds; branches and leaves at many depths cause discontinuities in the motion field. Such cases can be solved by a human observer [10]. However, these scenes are hard to solve since feature points cannot be tracked and since traditional optical flow methods cannot be expected to recover an image velocity.

Recently, [4] introduced a new category of movement called optical snow which generalizes optical flow by abandoning assumptions of spatial continuity. A model to analyse optical snow induced by all one-parameter set of velocities has been proposed in [5]. Taking the whole sequence in consideration, this model allows for *lateral observer motions* only [5]. Since the image velocity field can locally be approximated as a sum of two fields, a parallel field due to camera translation and a constant field due to camera rotation [9, 7], it was showed in [7] that this model could be used to recover 3D egomotion by subdividing the image sequence in sufficiently small patches.

We suggest a more global approach without the use of patches. In this paper, we analyze purely *non-lateral observer motions* and present a method to rectify these sequences such that standard optical snow methods can be applied. Finally, we give some experimental results.



(a) (b)

Figure 1: (a) Bowtie signature in the frequency domain (b) Eigenbasis of bowtie signature

2 Previous works

2.1 Optical Snow

The model of optical snow defined in [4, 5] is an extension of the motion plane property [11] which states that an image pattern translating with uniform image velocity produces a plane of energy in the frequency domain. Formally, let $I(x, y, t)$ be a time varying image. If an image patch is translating by (v_x, v_y), we know from [3] that this velocity is constrained by

$$v_x \frac{\partial I}{\partial x} + v_y \frac{\partial I}{\partial y} + \frac{\partial I}{\partial t} = 0 \tag{1}$$

This constraint, transposed in the Fourier domain, yields the following equation:

$$-2\pi * (v_x f_x + v_y f_y + f_t) * \hat{I}(f_x, f_y, f_t) = 0 \tag{2}$$

where $\hat{I}(f_x, f_y, f_t)$ is the Fourier transform of $I(f_x, f_y, f_t)$. As noted in [5], equation 2 implies that all frequencies $\hat{I}(x, y, t) \neq 0$ lie on the plane

$$v_x f_x + v_y f_y + f_t = 0. \tag{3}$$

This model was extended in [4] for the case in which there is a one-parameter set of velocities within an image region, i.e. where velocities vary according the following equation:

$$(v_x, v_y) = (u_x + \alpha t_x, u_y + \alpha t_y) \tag{4}$$

where u_x, u_y, t_x, t_y are constants and α depends on the visible depth at point (x,y) [5].
Substituting Eq. 4 into Eq. 3 yield a family of planes in the frequency domain,

$$(u_x + \alpha t_x) f_x + (u_y + \alpha t_y) f_y + f_t = 0. \tag{5}$$

Thus, this model produces a set of planes in the frequency domain. This set of planes forms a *bowtie* (see Fig. 1) and follows the two following propositions:

Proposition 1: The planes of the bowtie intersect at a common line, called the *axis of the bowtie*, that passes through the origin.

Proposition 2: The axis of the bowtie is in direction $(-t_y, t_x, u_x t_y - u_y t_x)$[1]. By normalizing (t_x, t_y), this direction becomes $(-t_y, t_x, |U|sin(\phi))$, where ϕ is the angle between vectors (t_x, t_y) and (u_x, u_y).

Applied to the egomotion problem, it was noted in [5] that Eq. (4) corresponds to

$$(v_x, v_y) = (-\omega_y + \alpha t_x, \omega_x + \alpha t_y) \tag{6}$$

In other words, camera rotation generates a constant velocity component $(-\omega_y, \omega_x)$ for small field of views ($\pm 20°$) and lateral translation (t_x, t_y) generates a velocity inversely proportional to depth. Components ω_z and t_z were assumed to be 0. Therefore, this model cannot recover arbitrary camera motions.

The main contribution of this paper is to show that complementary camera motions, i.e. following components ω_z and t_z (all other components assumed to be 0), produce non-lateral optical snow sequences that can be rectified to be analyzed by existing optical snow methods. We will describe how to find components ω_z and t_z.

3 Motion Field

The motion field equation contains 7 variables which are the depth P_z at each pixel, the translation vector (t_x, t_y, t_z) and the rotation vector $(\omega_x, \omega_y, \omega_z)$.

More precisely, the velocity field for a pixel (x,y), as defined in [6], is:

$$(v_x, v_y)^T = \begin{pmatrix} p_x p_y \omega_x - (1 + p_x^2)\omega_y + p_y \omega_z + \frac{p_x t_z}{P_z} - \frac{t_x}{P_z} \\ (1 + p_y^2)\omega_x - p_x p_y \omega_y - p_x \omega_z + \frac{p_y t_z}{P_z} - \frac{t_y}{P_z} \end{pmatrix}$$

By assuming that only P_z, t_z and ω_z are non-zero, we are left with:

$$(v_x, v_y)^T = \begin{pmatrix} p_y \omega_z + \frac{p_x t_z}{P_z} \\ -p_x \omega_z + \frac{p_y t_z}{P_z} \end{pmatrix}$$

Note that rotation component ω_z is perpendicular to translation component t_z for each pixel (x,y), and that ω_z is independent of depth. From this observation, rectification is performed on image sequences to obtain optical snow motion $(0, \omega_z) + \alpha(t_z, 0)$.

3.1 Motion Field Rectification

The rectification is a polar transformation around the FOE [8] which is the image center in our case. First, we rectify the motion field induced by pure ω_z rotation. Consider a point p located in the image at (1,0) on the camera projection plane (see Fig. 2-a). The path followed through time by p for a pure ω_z movement is a circle of radius 1, and its speed is exactly equal to $|\omega_z|$. The rectification is done by "unfolding" images such that this circle becomes straight and vertical, as shown in Fig. 2-b. For a field of view of 90° and images of size $N \times N$, the vertical length of the rectified image is $2\pi \frac{N}{2} = \pi N$ pixels. Note that flow lines of a pure forward motion become horizontal. The velocities in the rectified motion field correspond to $(v'_x, v'_y)^T = (\alpha t_z \sqrt{p_x^2 + p_y^2}, \omega_z)^T$.

Standard optical snow produces velocites that only depend on depth. However, the velocities in rectified non-lateral motion sequences vary according to depth and to image position. Therefore, horizontal lines in the rectified sequence must be resampled according to factor $\frac{1}{\sqrt{p_x^2 + p_y^2}}$, where ($p_x$,$p_y$) is the position of a pixel on the camera projection plane in the original sequence. In theory, we get infinite sampling at center (0,0). In practice, we do not rectify near the image center as illustrated in Fig. 2-c. Notice that image corners are cut to remove empty spaces in the rectified sequence.

[1] For simplicity, all equations in this paper assume image sequences of equal dimension in space and time.

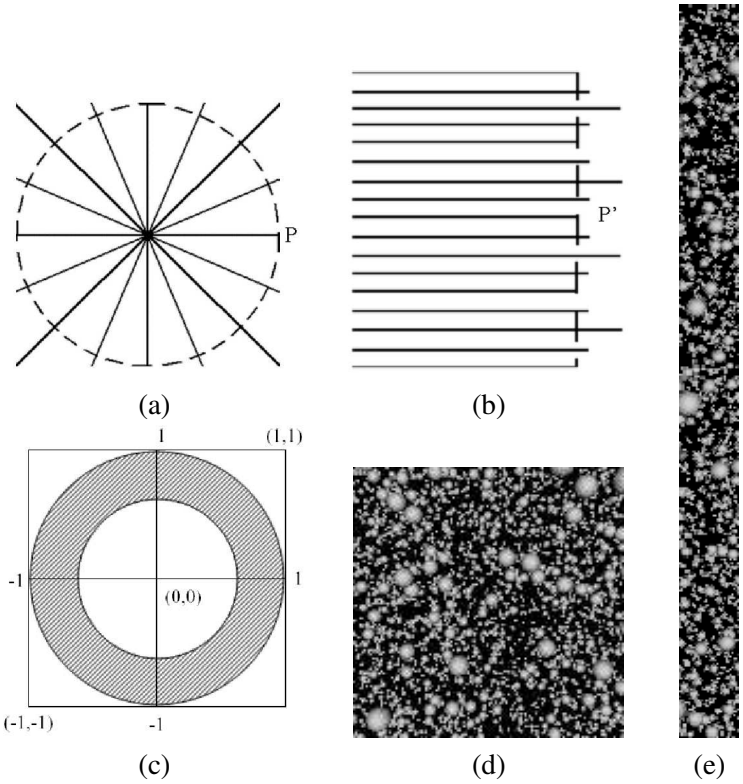


Figure 2: (a) Original motion field (b) Rectified motion field (c) Rectified region of the original sequence (d) Original frame (e) Rectified frame

3.2 Finding rotation ω_z

The bowtie axis can be computed from the rectified sequence using Principal Components Analysis as described in [2]. For standard optical snow, the angle ϕ in the bowtie axis equation (see Proposition 2) is unknown. For non-lateral optical snow, however, we know that $\phi = 90°$. Hence, the bowtie axis equation becomes $(0, 1, \omega_z)$. Thus, the third component of the bowtie axis gives us ω_z directly with speed given in pixels/frame in the rectified sequence. The rotation given in degrees, for a field of view of $90°$, is then $360\frac{\omega_z}{\pi N}$.

The bowtie axis can also be found using the best fit plane [2]. Let (n_x, n_y, n_z) be the normal of the best fit plane π. Since t_z is the only component affected by depth, the bowtie axis is the line on the best fit plane in direction $(0, 1, -\frac{n_y}{n_z})$.

3.3 Finding translation t_z

From Eq. 5, and since t_z generates only horizontal velocities and ω_z only vertical velocities, planes forming the bowtie have equation $(\alpha t_z, \omega_z, 1)$. The best fit plane π is defined as $(\frac{n_x}{n_z}, \frac{n_y}{n_z}, 1) = (\overline{\alpha} t_z, \omega_z, 1)$, where $\overline{\alpha}$ is the weighted "average" of the slopes of the motion planes that compose the bowtie. The weights depend on depth distribution in the scene as well as image contrast contributed by each object. Since this information is unknown, we can only compute t_z up to a scale factor $\overline{\alpha}$.

4 Experimental results

To evaluate our method, we rendered several synthetic image sequences of scenes containing lambertian spheres (see Fig. 3-a). Image motion was generated by moving a camera ($90°$ field of view) through the scene with various t_z and ω_z parameters.

Table 1 shows results for various rectified sequences. The last two rows correspond to real image sequences, respectively the lab sequence (see Fig. 3-b) and the plants sequence (see Fig. 3-c). The ω_z component is found almost exactly. The running time is about 1.6 seconds for 128x128x32 sequences on a 1.3GHz AMD Athlon machine.

(a) (b) (c)

Figure 3: (a) Synthetic scenes were constituted of small balls at different depths (64 frames of size 128x128 pixels). (b) Lab sequence taken from a camera rotating around the z-axis (40 frames of 128x128 pixels). (c) Plants sequence taken from a camera making a forward motion (32 frames of 128x128 pixels).

Table 1 : Results for synthetic and real scenes

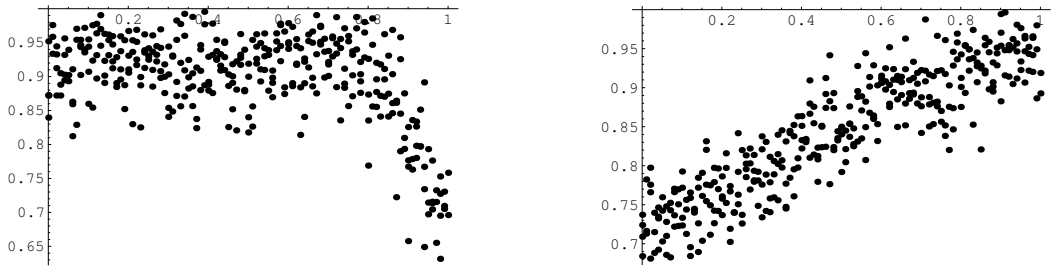
True Rotation (degrees/frame)	**True Translation** (pixel/frame)	**Rotation Found**	**Translation Found** (up to a scale factor)
0.00	1.00	0.00	0.15
1.80	0.00	1.78	0.00
1.80	1.00	1.81	0.22
≈ 0.50	≈ 0.00	0.35	-0.05
≈ 0.00	≈ 0.80	0.00	0.31

4.1 Comparing the eigenvalues

When analyzing image sequences globally, we would like to estimate how well the motion fits the optical snow model, i.e. if a bowtie is present. For instance, an unrectifed forward motion or a rectified lateral motion do not produce a bowtie signature. For such cases, the motion field features velocities oriented in all directions. Detecting these situations would be a great benefit.

In [2], depth range is evaluated by comparing the two largest eigenvalues of the Principal Components Analysis method. We can adapt this measure to detect the presence of a bowtie. The ratio of the first two eigenvalues λ_2 and λ_1 has a maximum of 1 and should decrease as we get closer to a bowtie signature. In fact, the presence of a bowtie creates a high power concentration along its axis which increases the first eigenvalue. The absence of a bowtie is caracterized by a uniform power distribution which makes λ_1 and λ_2 almost equal. Fig. 4-a shows a plot of $\frac{\lambda_2}{\lambda_1}$ as a function of $\frac{|t_z|}{|T|}$ for rectified sequences. As expected, the ratio falls off as the bowtie takes shape in the frequency domain. Fig. 4-b shows $\frac{\lambda_2}{\lambda_1}$ as a function of $\frac{|t_z|}{|T|}$ for non-rectified sequences. As expected, it increases linearly up to 1 as bowtie signature disappears. The sequences used for these graphs have random translation vectors of unit length.

Notice that the curve in Fig. 4-a decreases non-linearly while the other one is linear. This might be caused by the subsampling during rectification which accentuates the effect of any lateral motion. This is still under investigation. These curves, if modelled correctly, would allow merging non-lateral and lateral motions analysis into a general egomotion estimation algorithm.



(a) (b)

Figure 4: (a) $\frac{\lambda_2}{\lambda_1}$ as a function of $\frac{|t_z|}{|T|}$ for rectified sequences. It starts at 1 and decreases as a bowtie signature takes shape. (b) $\frac{\lambda_2}{\lambda_1}$ as a function of $\frac{|t_z|}{|T|}$ for non-rectified sequences. It increases linearly up to 1 as bowtie signature disappears.

5 Conclusion

This paper presented a method to analyze purely non-lateral optical snow by introducing a rectification process. Results show its accuracy and efficiency. We hope to solve the general egomotion problem in a global approach using this scheme.

References

[1] J.L. Barron, D.J. Fleet, and S.S. Beauchemin. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43–77, February 1994.

[2] V. C.-Couture, S. Roy, M. S. Langer, and R. Mann. Principal components analysis of optical snow. In *British Machine Vision Conference*, September 2004.

[3] B. Horn and B. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[4] M. S. Langer and R. Mann. Dimensional analysis of image motion. In *IEEE International Conference on Computer Vision*, pages 155–162, 2001.

[5] M.S. Langer and R. Mann. Optical snow. *International Journal of Computer Vision*, 55(1):55–71, 2003.

[6] H.C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, B-208:385–397, 1980.

[7] R. Mann and M. S. Langer. Estimating camera motion through a 3d cluttered scene. In *Canadian Conference on Computer and Robot Vision*, London, Canada, May (to appear) 2004.

[8] M. Pollefeys, R. Koch, and L. Van Gool. A simple and efficient rectification method for general motion. In *International Conference on Computer Vision*, pages 496–501, Corfy, Greece, 1999.

[9] J. H. Rieger and D. T. Lawton. Processing differential image motion. *J. Opt. Soc. Am. A*, 2:254–260, 1985.

[10] W.H. Warren and D.J. Hannon. Eye movements and optical flow. *Journal of the Optical Society of America A*, 7(1):160–169, 1990.

[11] A. Watson and A. Ahumada. Model of human visual-motion sensing. *Journal of the Optical Society of America*, 2(2):322–342, 1985.