OBSERVER MOTION ESTIMATION AND CONTROL FROM OPTICAL FLOW

Liangyin Yu, Charles R. Dyer

Computer Sciences Department
University of Wisconsin
Madison, Wisconsin 53706 USA

ABSTRACT

The information conveyed by optical flow is analytically linked to the observer motion in this paper by decomposing the optical flow field into its vector field components. It is shown that the observer may recover his ego-motion by interpreting the decomposed optical flow field, and he may further utilize his mobility to actively control the shape of the optical flow field, which directly reflects the surface shape of the object. The information of surface geometry discontinuity can be derived more directly from optical flow field by segmenting the whole field. A new method for the segmentation is proposed here, which combines both the magnitude and phase parts of the optical flow. The integration of these two different kinds of information proves to be effective in making various surface geometry boundary explicit.

1. INTRODUCTION

Optical flow conveys information about a surface through depth cues. This information may be used to identify object boundaries (optical flow segmentation) [1, 2], to determine the observer's ego-motion [3, 4], or to recover surface geometry.

Optical flow itself provides depth cues of the scene [5], as well as surface geometry [6, 7], and its segmentation generally corresponds to object boundaries (the apparent contour). When the spatial and temporal derivatives can be computed reliably, the surface geometry can be recovered [8, 9]. However, the local measurement and computation of optical flow are generally noisy and inaccurate, as opposed to global behaviors of the field. This is especially useful for tasks that do not require full scene recovery such as navigation where the relationship between the observer and the scene geometry is made explicit [10].

In this paper, we investigate two kinds of global behaviors of optical flow: the vector field decomposition and segmentation. For the former, instead of computing the gradient, divergence and curl fields of the optical flow as proposed by Koenderink [7], we relate the eigenvalues of the component fields to observer motions directly. We show that an active observer can utilize its mobility to control the shape of the optical flow field so that both the surface geometry and the observer motion can be recovered more effectively and accurately. For the segmentation, we expose the discontinuities of both the surface shape and surface orientation by combining the segmentation of magnitude and orientation of the optical flow field. This segmentation process is conducted by first computing the directional derivative of either the magnitude or the orientation of optical flow in the direction of the field. This integration of information proves to be effective in making various surface geometry boundary explicit.

We start this paper with the framework of the decomposition of a general vector field into divergence, curl and deformation fields. Each provides different information of the object surface relative to the observer. This information can be used by the active observer to control his motion in order to complete relevant tasks (e.g., navigation or shape recovery). Following this, it is shown next how the eigenvalues of the decomposed fields can be related to the observer-controlled motion and surface geometry. Finally, the new method of optical flow segmentation is presented for an observer-controlled translation motion, and various examples are presented.

2. THEORETICAL FRAMEWORK: 2D VECTOR FIELD DECOMPOSITION

On a 2D Euclidean manifold (ξ, η) the *integral curves* of a 2D linear vector field $\mathbf{u} = (\mu, \nu)$ are the family of curves $\mathbf{q} = (\xi(s), \eta(s))$ defined by

$$\mu = \frac{\partial \xi}{\partial s} = \frac{\partial \mu}{\partial \xi}\xi + \frac{\partial \mu}{\partial \eta}\eta$$
$$\nu = \frac{\partial \eta}{\partial s} = \frac{\partial \nu}{\partial \xi}\xi + \frac{\partial \nu}{\partial \eta}\eta \quad (1)$$

where s is curve length. Let $\mathbf{P} = \begin{pmatrix} \frac{\partial \mu}{\partial \xi} & \frac{\partial \mu}{\partial \eta} \\ \frac{\partial \nu}{\partial \xi} & \frac{\partial \nu}{\partial \eta} \end{pmatrix}$. Eq. (1) can then be written in the form $\mathbf{u}^T = \mathbf{P}\mathbf{q}^T$. For a general vector field $\mathbf{u}$, Eq. (1) provides the first-order approximation to $d\mathbf{u}$ (by Taylor series) in the form $d\mathbf{u}^T = \mathbf{P}\,d\mathbf{q}^T$.

The matrix $\mathbf{P} = (p_{ij})$ can be decomposed into the sum of a symmetric matrix $\mathbf{P}^s = (p^s_{ij})$ and an antisymmetric matrix $\mathbf{P}^a = (p^a_{ij})$ according to $p^s_{ij} = (p_{ij} + p_{ji})/2$ and $p^a_{ij} = (p_{ij} - p_{ji})/2$. Since a symmetric matrix can always be diagonalized by a similar transform, $\mathbf{P}^s$ can be put into the form

$$\mathbf{P}^s = \mathbf{Q}^{-1}\begin{pmatrix}\zeta_1 & 0\\ 0 & \zeta_2\end{pmatrix}\mathbf{Q}$$

where $\zeta_1 > \zeta_2$ and $\mathbf{Q}$ is an orthogonal matrix with $|\mathbf{Q}| = 1$. It has the property $tr\ \mathbf{P}^s = \partial\mu/\partial\xi + \partial\nu/\partial\eta = \zeta_1 + \zeta_2$. Let $\mathbf{I}_2 = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}$, $\mathbf{J}_2 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}$, $\mathbf{K}_2 = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$, and, by the above property, we have

$$2\mathbf{P}^s = (\frac{\partial\mu}{\partial\xi} + \frac{\partial\nu}{\partial\eta})\mathbf{I}_2 + (\zeta_1 - \zeta_2)\mathbf{Q}^{-1}\mathbf{J}_2\mathbf{Q}.$$

A more compact form can be reached by noting that if the 2D vector field $\mathbf{u} = (\mu, \nu)$ is treated as a 3D field $(\mu, \nu, 0)$ then $\nabla\cdot\mathbf{u} = \partial\mu/\partial\xi + \partial\nu/\partial\eta$, $\nabla\times\mathbf{u} = (\partial\nu/\partial\xi - \partial\mu/\partial\eta)\hat{\mathbf{e}}_z$. If we denote the only component of the curl as $(\nabla\times\mathbf{u})_z$, the matrix $\mathbf{P}$ has the decomposed form:

$$\mathbf{P} = \frac{1}{2}\Big[(\nabla\cdot\mathbf{u})\mathbf{I}_2 + (\nabla\times\mathbf{u})_z\mathbf{K}_2 + (\zeta_1 - \zeta_2)\mathbf{Q}^{-1}\mathbf{J}_2\mathbf{Q}\Big]. \tag{2}$$

Following [6] we will refer these three decomposed components of $\mathbf{u}$ as the *divergence, curl* and *deformation* fields, respectively.

The integrated vector field is determined by three canonical subfields formulated in Eq. (2). Its properties can be investigated by examining the eigenvalues of $\mathbf{P}$. The eigenvalues are themselves functions of the image coordinates (ξ, η) defined at each point on the image plane, as is the field itself. If we insert the orthogonal matrix $\mathbf{Q}$ and let $c = (\nabla\times\mathbf{u})_z/2$, $d = (\nabla\cdot\mathbf{u})/2$, $e = (\zeta_1 - \zeta_2)/2$, $\mathbf{P}$ can be written in the form

$$\begin{aligned}\mathbf{P} &= \frac{d}{2}\mathbf{I}_2 + \frac{c}{2}\mathbf{K}_2 + \frac{e}{2}\begin{pmatrix}\cos 2\gamma & -\sin 2\gamma\\ -\sin 2\gamma & -\cos 2\gamma\end{pmatrix}\\ &= \frac{1}{2}\begin{pmatrix}d + e\cos 2\gamma & -c - e\sin 2\gamma\\ c - e\sin 2\gamma & d - e\cos 2\gamma\end{pmatrix},\end{aligned}$$

where γ is the rotation angle for $\mathbf{Q}$. Hence the characteristic equation for $\mathbf{P}$ is $\lambda^2 - 2d\lambda + (c^2 + d^2 - e^2) = 0$ and the eigenvalues are $\lambda = d \pm (e^2 - c^2)^{1/2}$. From this we can make the observation that $e^2 - c^2$ (curl and deformation) acts as an essential factor in deciding the field characteristics. One of the interesting cases is when the curl and deformation fields cancel each other so that only the divergence field shows up. This is different from vanishing curl and deformation fields, but it appears identical to the observer. The relationship between eigenvalues and observer motion will be derived in the next section.

3. OBSERVER MOTION FROM OPTICAL FLOW DECOMPOSITION

For an active observer, optical flow can be very useful if it can be used to infer the relative relationship between the observer and the surface so that further observer motion can be planned relative to the surface. In order to achieve this, we have to relate optical flow to observer motion. This was done in [7] by expressing the canonical fields (i.e., divergence, curl and deformation) in terms of the rotation Ω and translation $\mathbf{v}$ by the observer. In this section, a different formulation is developed, which relates the eigenvalues of these fields to translational observer motion. This alternative form makes the observer motion explicit in order to control the optical flow.

Let's consider an observer undergoing an instantaneous translation in a static environment. Let $\mathbf{x} = (x, y, z)$ be the coordinates in the observer frame and $\mathbf{q} = (\xi, \eta, 1)$ the projected image coordinates in the 3D Euclidean space. The relative translation velocity is given by $\mathbf{v} = -\partial\mathbf{x}/\partial t = (v_x, v_y, v_z)$. The observer-centered coordinate system is set up so that the image plane is located at $z = 1$. The object surface can then be represented as $(x, y, z(x, y))$. Since $(z_x, z_y, -1)$ is the normal vector to the tangent plane of the object surface at $\mathbf{x}$, we will use $\mathbf{n}$ to denote $(z_x, z_y, -1)$. In this set-up, we have the relationship: $\mathbf{q} = \mathbf{x}/z$, and the optical flow is given by $\mathbf{u} = \partial\mathbf{q}/\partial t$.

When a rotation $\mathbf{\Omega}$ is involved in observer motion, the transversing translation velocity perpendicular to the line of sight $\hat{\mathbf{x}} = \mathbf{x}/|\mathbf{x}|$ is given by $\tilde{\mathbf{v}}_t = \tilde{\mathbf{v}} - (\tilde{\mathbf{v}}\cdot\hat{\mathbf{x}})\hat{\mathbf{x}}$, where $\tilde{\mathbf{v}} = \mathbf{v}/z$, and the canonical fields can be expressed as (see [7]):

$$\begin{aligned}\nabla\cdot\mathbf{u} &= \mathbf{n}\cdot\tilde{\mathbf{v}}_t + 2\tilde{\mathbf{v}}\cdot\hat{\mathbf{x}}\\ (\nabla\times\mathbf{u})_z &= (\mathbf{n}\times\tilde{\mathbf{v}}_t)_z - 2\mathbf{\Omega}\cdot\hat{\mathbf{x}}\\ \zeta_1 - \zeta_2 &= |\mathbf{n} + \hat{\mathbf{e}}_z||\tilde{\mathbf{v}}_t|\end{aligned} \tag{3}$$

By definition, $\mathbf{v} = -\partial\mathbf{x}/\partial t$ for a point $\mathbf{x}$ on an object surface, so we can derive $\mathbf{u} = \partial\mathbf{q}/\partial t = (-\mathbf{v} + v_z\mathbf{q})/z = -\tilde{\mathbf{v}} + \tilde{v}_z\mathbf{q}$. By the inverse function theorem it is straightforward to show that $(\partial z/\partial\xi, \partial z/\partial\eta) = -(z_x, z_y)z/(\mathbf{n}\cdot\mathbf{q})$, where z_x and z_y are differentials of z with respect to x and y. It follows that

$$\begin{pmatrix}\frac{\partial\mu}{\partial\xi} & \frac{\partial\mu}{\partial\eta}\\ \frac{\partial\nu}{\partial\xi} & \frac{\partial\nu}{\partial\eta}\end{pmatrix} = \tilde{v}_z\mathbf{I}_2 + \frac{1}{\mathbf{n}\cdot\mathbf{q}}\begin{pmatrix}\mu z_x & \mu z_y\\ \nu z_x & \nu z_y\end{pmatrix} \tag{4}$$

where $\mathbf{I}_2$ is the 2×2 identity matrix.

We can derive the following formulas for the divergence and curl of $\mathbf{u}$:

$$\begin{aligned}\nabla\cdot\mathbf{u} &= 2\tilde{v}_z + \frac{\mathbf{n}\cdot\mathbf{u}}{\mathbf{n}\cdot\mathbf{q}} = 3\tilde{v}_z - \frac{\mathbf{n}\cdot\tilde{\mathbf{v}}}{\mathbf{n}\cdot\mathbf{q}}\\ (\nabla\times\mathbf{u})_z &= \frac{(\mathbf{n}\times\mathbf{u})_z}{\mathbf{n}\cdot\mathbf{q}} = \frac{\big[\mathbf{n}\times(\tilde{v}_z\mathbf{q} - \tilde{\mathbf{v}})\big]_z}{\mathbf{n}\cdot\mathbf{q}}.\end{aligned} \tag{5}$$

In addition, the symmetric part of $\mathbf{P}$ matrix is given by

$$\mathbf{P}^s = \tilde{v}_z \mathbf{I}_2 + \frac{1}{2\mathbf{n}\cdot\mathbf{q}} \begin{pmatrix} 2\mu z_x & \mu z_y + \nu z_x \\ \mu z_y + \nu z_x & 2\nu z_y \end{pmatrix}.$$

This symmetric matrix can be diagonalized if we choose $\tilde{\mathbf{v}}$ to be such that $\mu z_y + \nu z_x = 0$, i.e.,

$$\tilde{\mathbf{v}} \cdot (-z_y, -z_x, \xi z_y + \eta z_x) = 0. \tag{6}$$

Alternatively we can compute the rotation matrix $\mathbf{Q}$ such that $\mathbf{Q}\mathbf{P}^s\mathbf{Q}^{-1}$ is diagonalized, if the surface shape (z_x, z_y) is already known. The former corresponds to an observer-controlled motion and the latter corresponds to an "off-line" computation. If we diagonalize $\mathbf{P}^s$ by observer motion using Eq. (6), it can be shown that the deformation is

$$\zeta_1 - \zeta_2 = \frac{z_x^2 + z_y^2}{\xi z_y + \eta z_x} \frac{(\mathbf{q} \times \tilde{\mathbf{v}})_z}{(\mathbf{n}\cdot\mathbf{q})}.$$

If we diagonalize explicitly by rotation defined by $\mathbf{Q}$, the deformation is

$$\zeta_1 - \zeta_2 = \frac{(z_x^2 + z_y^2)^{1/2}}{\mathbf{n}\cdot\mathbf{q}} (\mu^2 + \nu^2)^{1/2} = \frac{(z_x^2 + z_y^2)^{1/2} |\mathbf{u}|}{\mathbf{n}\cdot\mathbf{q}} \tag{7}$$

with the rotation angle γ given by $(\mu z_y + \nu z_x)/(\mu z_x - \nu z_y)$. From Eq. (5) we can express the optical flow field $\mathbf{u}$ in terms of its curl $(\nabla \times \mathbf{u})$, divergence $(\nabla \cdot \mathbf{u})$ and surface tilt (z_x, z_y):

$$\mathbf{u} = \frac{\mathbf{n}\cdot\mathbf{q}}{z_x^2 + z_y^2} \left[(\nabla\cdot\mathbf{u} - 2\tilde{v}_z)(z_x, z_y) + \nabla \times \mathbf{u}(z_y, z_x) \right]. \tag{8}$$

The eigenvalues of the linear vector field is then given by

$$2\lambda_{1,2} = \nabla\cdot\mathbf{u} \pm \left[(\zeta_1 - \zeta_2)^2 - (\nabla \times \mathbf{u})_z^2 \right]^{1/2}.$$

From Eq. (5) and Eq. (7) it can be shown that $(e^2 - c^2)^{1/2} = d - \tilde{v}_z$. Hence the two eigenvalues are given by

$$\lambda_{1,2} = \tilde{v}_z,\ \tilde{v}_z + \frac{\mathbf{n}\cdot\mathbf{u}}{\mathbf{n}\cdot\mathbf{q}}. \tag{9}$$

From this form, we can see that the eigenvalues are always real, which is consistent with the elimination of the curl field in the first place, since it is not useful in solving for scene geometry [7]. Without a curl field, the matrix $\mathbf{P}$ is symmetric and the deformation is given by

$$|\lambda_1 - \lambda_2| = \left| \frac{\mathbf{n}\cdot\mathbf{u}}{\mathbf{n}\cdot\mathbf{q}} \right|. \tag{10}$$

If we assume that the canonical fields can be observed and computed by the observer, Eq. (10) allows us to determine the surface normal without having to compute the divergence field. In essence, the deformation field tells us the surface orientation.

4. SEGMENTATION OF THE OPTICAL FLOW FIELD

One of the purposes of optical flow segmentation is to identify discontinuity of surface geometry. The occurrence of discontinuity is due to either the presence of discontinuous contours on an object surface or to discontinuity in depth. Though there is no unique interpretation of the results from segmentation, the result does strongly constrain the problem of identifying object boundaries.

For translational motion $\mathbf{v} = (v_x, v_y, v_z)$ controlled by an observer, the optical flow field $\mathbf{u}$ is given by $\mathbf{u} = -\tilde{\mathbf{v}} + \tilde{v}_z\mathbf{q}$. To measure the smooth characteristics of the optical flow, two factors have to be considered: the orientation and the magnitude of the optical flow. Consequently, the "smoothness" of $\mathbf{u}$ can be measured by the directional derivative of the magnitude of $\mathbf{u}$ in the direction of $\mathbf{u}$:

$$\epsilon(\mathbf{u}) \stackrel{\triangle}{=} \nabla(|\mathbf{u}|) \cdot \frac{\mathbf{u}}{|\mathbf{u}|}. \tag{11}$$

Since $|\mathbf{u}| = A^{1/2}/z$, where $A = (-v_x + \xi v_z)^2 + (-v_y + \eta v_z)^2$, it follows that

$$\nabla(|\mathbf{u}|) \cdot \frac{\mathbf{u}}{|\mathbf{u}|} = -\frac{\mathbf{u}\cdot\nabla z}{z} + \frac{z v_z}{A} \mathbf{u} \cdot (\mu\nabla\xi + \nu\nabla\eta)$$

Simplify the expression and it can be shown that

$$\epsilon(\mathbf{u}) = \frac{1}{z} \left[-(\mathbf{n}\cdot\mathbf{u}) + \tilde{v}_z \left(1 - \frac{(\mathbf{q}\cdot\mathbf{u})(\mathbf{n}\cdot\mathbf{u})}{|\mathbf{u}|^2} \right) \right] \tag{12}$$

If we treat $\epsilon(\mathbf{u}(\xi, \eta)) = \epsilon(\xi, \eta)$ as a function defined on image plane, changes (i.e., discontinuities) in the structure of ϵ will correspond to image contours. The boundary where these changes occur is a consequence of changes in scene geometry, which involves surface normal $\mathbf{n}$ and depth z as reflected in Eq. (12). Alternatively, we can derive the magnitude and orientation of the optical flow individually and combine them in a separate phase. This may provide more flexibility in terms of how the information is used.

5. EXAMPLES

The first test sequence is a sequence of 20 images with both curl and divergence fields. The integral curves of the flow field with respect to time are made apparent by smoothing the local image structure.

For optical flow, a window of 5 frames was used to compute the integral curve for the flow field. These curves were computed at four scales: 1.5, 2, 3, and 4 pixels. The optical flow at each point was computed using the method in a previously published paper, and the result for 10th frame is shown in Figure 1. The components of the curl and divergence fields are clear in the spiral shape of the optical flow.

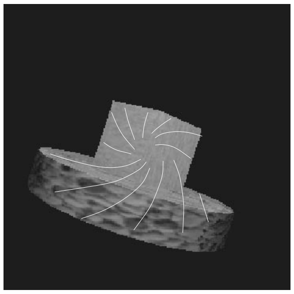

Fig. 1. The integral curve of the optical flow field.

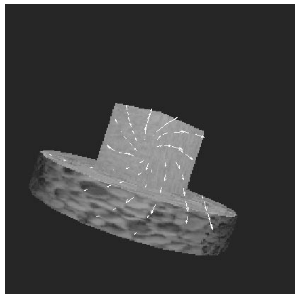

Fig. 2. The optical flow field.

The scale which was used to identify the magnitude of the optical flow is not shown in the results, but it is an indication of the nature of local texture.

Fig. 3. The magnitude of the optical flow for frame 8.

Fig. 4. The orientation of the optical flow for frame 8.

The second test sequence is the synthetic sequence of 15 frames that shows a fly-through of Yosemite valley. The segmentation method formulated in Eq. (11) was applied to the optical flow centered at the eighth frame of the sequence. The magnitude ($|\mathbf{u}|$) and the orientation ($\mathbf{u}/|\mathbf{u}|$) parts are represented by gray-level images in Figures 3 and 4, respectively. The measure of directional derivative, $\epsilon(\mathbf{u})$, is shown in Figure 5, encoded and equalized for gray-level representation. By combining the measure of segmentation and the cues provided by the magnitude and orientation of the optical flow, the boundaries of objects are shown in Figure 6.

Fig. 5. The gray-level segmentation of optical flow.

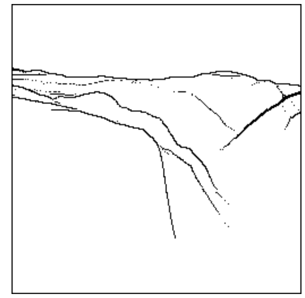

Fig. 6. The binary segmentation of optical flow.

6. SUMMARY

We present results that exploit the fact that global behavior of optical flow is a more accurate information source to use than the local measurement of the field. The major results presented in this paper are : (1) the component vector fields of optical flow can be related directly to observer motion, and this relation may be used to to recover either observer motion or scene geometry, and (2) the segmentation of optical flow can be derived by combining the global changes of magnitude and orientation of the field. This provides the observer with strong hypotheses regarding the type and location of object boundaries, The information can be effectively derived when the observer is active and can control his motion.

7. REFERENCES

[1] G. Adiv, "Determining three-dimensional motion and structure from optical flow generated by several moving objects," *IEEE Trans. Patt. Anal. Machine Intell.*, 1985.

[2] S. Srinvasan, "Extracting structure from optical flow using the fast error search technique," *Int. J. Computer Vision*, vol. 37, no. 3, pp. 203–230, 2000.

[3] A. R. Bruss and B. K. P. Horn, "Passive navigation," *Computer Vision, Graphics, and Image Processing*, vol. 21, 1983.

[4] C. Fermuller and Y. Aloimonos, "Qualitative egomotion," *Int. J. Computer Vision*, vol. 15, pp. 7–29, 1995.

[5] H. C. Longuet-Higgins and K. Prazdny, "The interpretation of a moving retinal image," *Proc. Roy. Soc. London B*, vol. 208, pp. 385–397, 1980.

[6] J. J. Koenderink and A. J. van Doorn, "Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer," *Optica Acta*, vol. 22, no. 9, pp. 773–791, 1975.

[7] J. J. Koenderink and A. J. van Doorn, "Optic flow," *Vision Research*, vol. 26, no. 1, pp. 161–180, 1986.

[8] A. M. Waxman and S. Ullman, "Surface structure and three-dimensional motion from image flow kinematics," *Int. J. Robotics Research*, vol. 4, no. 3, pp. 72–94, 1985.

[9] C. Tomasi, *Shape and Motion from Image Streams: a Factorization Method*, Ph.D. thesis, Carnegie Mellon University, 1991.

[10] R. Cipolla and A. Blake, "Motion planning using image divergence and deformation," in *Active Vision*, A. Blake and A. Yuille, Eds. MIT Press, 1992.