Multiple Light Source Optical Flow

Robert J. Woodham*

Laboratory for Computational Vision
University of British Columbia
Vancouver, BC, Canada

Abstract

A new method is described to compute a dense, local representation of optical flow. The idea is to use the intensity values recorded from multiple images of moving objects acquired simultaneously under different conditions of illumination. Each image is assumed to satisfy the standard optical flow constraint equation. Multiple images give rise to multiple constraint equations. When the optical flow and the 2D motion field coincide, these multiple equations are in the same unknowns. Two light source directions are sufficient, in principle, to determine the motion field. Three (or more) light source directions overdetermine the solution, avoid local degeneracies, and help to make the computation more robust.

This paper describes the basic theory and illustrates the theory on a real image motion sequence. All computations are local, independent and relatively simple. No iteration steps are required. It is suggested that the requirement to obtain simultaneous images under different conditions of illumination be satisfied by using spectrally distinct illumination and sensing.

1. Introduction

In general, a point on a moving object induces motion at the corresponding image point. The relation between scene motion and image motion is determined entirely by the geometry of image projection. The *2D motion field* is the 2D velocity vector at each point in an image.

Object motion typically also induces change in the brightness values measured in an image. The relationship between object motion and brightness change is not purely geometric because it also depends on radiometric factors, including the illumination and the reflectance properties of the objects in view. The *optical flow* is the 2D velocity vector of the brightness values themselves.

It is not possible to recover optical flow locally due to the well-known aperture problem. Additional information is required. Horn and Schunck[1] imposed a global smoothness constraint on optical flow in their now classic treatment. Their computed optical flow can vary locally, provided the variation is smooth. In some circumstances, discontinuities in the motion field also are detected. Other approaches impose local constraint. For example, Kearney *et al.*[2] require that optical flow be constant in a selected region while both Mailloux *et al.*[3] and Verri *et al.*[8] model optical flow as locally satisfying a general affine transformation. Ohta[4] uses multispectral images to obtain additional local constraint.

This paper reconsiders the possibility of computing a dense, local representation of optical flow. The novel idea is to use the intensity values recorded from multiple images of moving objects acquired simultaneously under different conditions of illumination. Photometric stereo[10] exploits multiple conditions of illumination to determine shape from shading. Here, the same idea is applied to moving objects to determine optical flow. Each image is assumed to satisfy the standard optical flow constraint equation. Where the optical flow and the 2D motion field coincide, these multiple equations can be solved together to determine the 2D motion field. The method can be applied to every point in an image. Typically, the result is well-defined except at a sparse set of features. The method imposes no smoothness condition on optical flow and requires no *a priori* assumptions about how the image ought to be segmented.

Section 2 develops the theory of multiple light source optical flow in more detail. Section 3 discusses a particular implementation. Section 4 presents an example of the method applied to a simple toy object. Section 5 discusses the possibility of practical realization. Finally, conclusions are summarized in Section 6.

2. Background and Theory

2.1 The Optical Flow Constraint Equation

Let $E = E(x,y,t)$ denote image brightness at image point (x,y) as a function of time, t. Then, by a chain rule for differentiation, the total derivative of E with respect to time, dE/dt, can be written as

$$dE/dt = E_x u + E_y v + E_t \qquad (1)$$

where $E_x = \partial E/\partial x$, $E_y = \partial E/\partial y$ and $E_t = \partial E/\partial t$ are the partial derivatives of E with respect to x, y and t and where $u = dx/dt$ and $v = dy/dt$ determines the instantaneous flow at (x,y). If the brightness of an object point does not change as a consequence of motion, then $dE/dt = 0$. Under this assumption, the optical flow constraint equation is given by

* Fellow of the Canadian Institute for Advanced Research.

CH2934-8/90/0000/0042$01.00 © 1990 IEEE

$$E_x u + E_y v + E_t = 0 \quad (2)$$

From (1), it is clear that (2) is satisfied if and only if the total derivative dE/dt is zero. Much has been said concerning imaging situations that satisfy, approximately satisfy, or fail to satisfy the optical flow constraint equation (2). Some essential elements of the discussion are summarized in section 2.2.

For now, assume (2) holds so that its consequences can be examined in more detail. When (2) holds, optical flow and the 2D motion field coincide so that the vector $[u,v]$ is a purely geometric quantity describing the 2D motion of image point (x,y) as a function of time. The vector $[E_x,E_y,E_t]$ is a radiometric quantity describing the partial derivatives of image brightness with respect to position and time. Thus, (2) is of value because it relates something that can be measured, the vector $[E_x,E_y,E_t]$, to a geometric quantity of interest, the 2D motion field. Although (2) is a radiometric equation, it makes no assumption about the reflectance properties of the objects in view. In particular, it does not assume Lambertian (or any other) reflectance function.

Equation (2) cannot be solved locally because it is one equation in two unknowns, u and v. Given a measured $[E_x,E_y,E_t]$, equation (2) determines a line in uv–space, called the *flow constraint line.*

2.2 Geometric and Radiometric Considerations

One interpretation of equation (2) is that it requires that the vector $[u,v]$ be the only factor necessary to account for the temporal variation in image brightness. Equation (2) holds (equivalently, $dE/dt = 0$) for purely translational motion, orthographic projection and incident illumination that is uniform (i.e., does not vary) across the scene. Conversely, in any other circumstance, one can expect $dE/dt \neq 0$ so that (2) does not hold exactly.

With perspective projection or object rotation, there are geometric "foreshortening" effects that cause dE/dt to be non-zero. With rotation, surface normal vectors change relative to the direction to the light source and to the viewer. Scene radiance is a function both of the incident angle and of the view angle. Thus, scene radiance (and therefore image brightness) changes in a way that cannot be predicted without *a priori* knowledge of surface reflectance and scene illumination. Verri and Poggio[9] have quantified the expected difference between optical flow and motion for a number of special cases. Pentland[5] distinguishes brightness changes due to radiometric factors, which he calls "photometric motion", from those due to geometric factors.

Spatial or temporal variation in scene illumination will cause $dE/dt \neq 0$. Moving objects alter the scene illumination, both because they cast shadows and because they act as indirect sources of illumination via interreflection.

Finally, at locations where image brightness is discontinuous, $[E_x,E_y,E_t]$ becomes undefined. Schunck[6] argues that the optical flow constraint equation still is satisfied across discontinuities, provided the number of discontinuities is finite. Subsequently, he developed a version of his algorithm that could handle textured regions[7].

2.3 Using Multiple Light Sources

With multiple light sources, we get additional equations. For the two light source case, one obtains

$$E_{1x} u + E_{1y} v + E_{1t} = 0 \quad (3)$$

$$E_{2x} u + E_{2y} v + E_{2t} = 0 \quad (4)$$

which can be solved for $[u,v]$ as follows

$$\begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} E_{1x} & E_{1y} \\ E_{2x} & E_{2y} \end{bmatrix}^{-1} \begin{bmatrix} E_{1t} \\ E_{2t} \end{bmatrix} \quad (5)$$

provided the required matrix inverse exists.

It is useful to distinguish cases when the matrix inverse in (5) fails to exist. All gradient based methods for the determination of optical flow, including the one described here, produce no useful local information at points where $[E_x,E_y]$ is zero. It can also happen that the two brightness gradients, $[E_{1x},E_{1y}]$ and $[E_{2x},E_{2y}]$, are parallel. Typically, this happens when the brightness gradient is dominated by a local surface feature, such as a shape discontinuity or surface marking, independent of the illumination. This also can happen owing to an "accidental alignment" between surface shape and illumination direction(s). Local degeneracies can be resolved if a third light source image is provided.

For a three light source case, one obtains

$$E_{1x} u + E_{1y} v + E_{1t} = 0 \quad (6)$$

$$E_{2x} u + E_{2y} v + E_{2t} = 0 \quad (7)$$

$$E_{3x} u + E_{3y} v + E_{3t} = 0 \quad (8)$$

Equations (6-8) can be written as the matrix equation

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (9)$$

where $\mathbf{x} = [u,v]^{\mathrm{T}}$, $\mathbf{b} = -[E_{1t},E_{2t},E_{3t}]^{\mathrm{T}}$ and

$$\mathbf{A} = \begin{bmatrix} E_{1x} & E_{1y} \\ E_{2x} & E_{2y} \\ E_{3x} & E_{3y} \end{bmatrix}$$

In principle, there are many ways in which one could solve (9) for $\mathbf{x}$. The standard least squares solution, $\hat{\mathbf{x}}$, is given by

$$\hat{\mathbf{x}} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{b} \quad (10)$$

The solution is unique provided that the rank of $\mathbf{A}$ is 2. (Equation (9) can be extended, in the obvious way, to situations in which more than three light sources are used). It is important to note that the magnitude of $[E_x,E_y,E_t]$ plays the role of a "weight" that pulls the solution towards a flow constraint line for which the magnitude of $[E_x,E_y,E_t]$ is large (and consequently, away from a flow constraint line for which the magnitude of $[E_x,E_y,E_t]$ is small). This has a desirable effect in multiple light source optical flow. Locations in an image for which

brightness is nearly constant will have $[E_x, E_y, E_t]$ near zero. These locations contribute minimal information, and thus it is good that they are discounted. Because of this, points that are shadowed with respect to one of the light sources need not be considered as a special case.

2.4 Validating the Multiple Light Source Method

One would like to exploit the redundancy inherent in an overdetermined problem in order evaluate the validity of the solution method. Validating the multiple light source method requires empirical support for the assumption that the set of linear equations (6-8) is indeed defined in identical variables u and v, independent of the conditions of illumination. A good fit of this model to the measurement data would suggest that the vector $[u,v]$ is an illumination invariant measure, whatever its geometric interpretation. Loosely put, one would like to determine what component of the measurement, $\mathbf{b}$, is accounted for by $\hat{\mathbf{x}}$ under the model $\mathbf{Ax} = \mathbf{b}$. Fit to the model can be expressed quantitatively by the residual

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\hat{\mathbf{x}} \tag{11}$$

A relative error term is given by

$$\frac{|\mathbf{r}|}{|\mathbf{b}|} = \frac{|\mathbf{b} - \mathbf{A}\hat{\mathbf{x}}|}{|\mathbf{b}|} \tag{12}$$

This error term combines components due to both measurement uncertainty and to systematic modeling error. If $[E_x, E_y, E_t]$ is zero at a point in one image, equation (10) becomes equivalent to the two light source solution, given by equation (5), applied to the measurements obtained from the other two images. In this case, the relative error term will be zero since the measurements no longer are redundant.

3. Implementation

For a three light source configuration, six images are required. The first three images are obtained under different conditions of illumination at time $t = t_0$. The second three images are obtained at time $t = t_1$ respectively under the identical conditions of illumination used for the three images at $t = t_0$. It is essential that the estimated derivatives all refer to the same point in space and time. Derivatives of $E(x,y,t)$ are estimated using first differences in a 2×2×2 cube of brightness values, treating the x, y and t dimensions symmetrically. (See [1; pp 189-190] for details). Initial brightness quantization was 8 bits-per-pixel. Even if these values were noise free, this quantizes first differences into too small a set of discrete values to permit reliable estimation of derivatives. The six images are smoothed, using a 2D Gaussian filter, prior to derivative estimation. Since the 2D Gaussian is separable, filtering is implemented as the successive convolution of a 1D Gaussian filter. The filter coefficients of the 1D Gaussian were scaled to sum to 256. No bits were thrown away in the convolution so that the net effect is to interpolate 8 bits-per-pixel data to 24 bits-per-pixel smoothed data. In the example that follows, $\sigma = 1.0$.

Once the nine partial derivatives, (E_{ix}, E_{iy}, E_{it}, i = 1,2,3), are estimated, computation of optical flow proceeds point-by-point, according to equation (10). Two local checks are made to guarantee that the computation is not degenerate. First, at least two of the spatial brightness gradients, $[E_x, E_y]$, must be non-zero. Second, the rank of the matrix $\mathbf{A}$ must be two. Points that fail either of these two checks are noted and the result set as $u = 0$ and $v = 0$.

4. Example: Stay-Puft Marshmallow Man

The Stay-Puft Marshmallow Man is a commercial toy. Most of its surface has a white semi-gloss finish. The eyes and mouth are black, the collar and hat band are dark blue and the bow, in front, is red. The object was placed on a small black platform in a small "studio" constructed with matte black walls and ceiling.

Figure 1 shows the three images at $t = t_0$. Figures 1(a)-(c) have identical geometry but different illumination. Figure 1(a) has light source 1 from the upper right, Figure 1(b) has light source 2 from the upper left, and Figure 1(c) has light source 3 from almost directly behind the camera. The toy, but not the background, was then moved a small amount horizontally right to left (approximately two pixels). Three additional images, for $t = t_1$, were then obtained respectively under the same conditions of illumination as Figures 1(a)-(c). All images were of dimension 256×256.

Figure 2 illustrates the computation at a test point (row:160 column:158). Figure 2(a) marks the test point on the light source 3 image at $t = t_0$. Figure 2(b) plots the three flow constraint lines in uv–space. E1, E2 and E3 are respectively the flow constraint lines corresponding to the light source 1, 2 and 3 images for the motion from $t = t_0$ to $t = t_1$. The measurements of $[E_x, E_y, E_t]$, taken together, provide an accurate and well-conditioned estimate of u and v. (The tick marks on the axes in Figure 2(b) are spaced one unit apart so that we do indeed see that the solution is approximately $u = -2.0$ and $v = 0.0$).

Figure 3 shows the results for the entire image. Optical flow is computed at every point. To prevent clutter, Figure 3 plots results at every second point. Thus, if each estimate was exact (i.e., $u = -2.0$ and $v = 0.0$), the vectors would form a connected set of parallel lines running horizontally right to left. The estimates are good at points where the object surface is smoothly shaded. Some points on the collar fail to produce a result because they are too dark. The estimates are inaccurate at surface discontinuities and at surface reflectance boundaries because there changes in image brightness due to scene features dominate changes due to smooth shading.

Optical flow vectors also are estimated for many points on the background gray wedge and on the registration dots. Of course, there is no corresponding motion since the background was stationary. Nevertheless, there is non-zero optical flow owing to shadows and interreflection.

Figure 4 displays the relative error term computed according to equation (12). (The larger the error the darker the point). The error computation suggests that any attempt to assign a unique motion, $[u,v]$, to points on the gray wedge or on the registration dots is suspect. Similar error values occur at points on the toy's collar and, to a lesser extent, at points on the toy's body oriented towards the viewer. Points on the toy's collar are very dark with minimal brightness variation. Points on the toy's body oriented towards the viewer, while bright, also have minimal brightness variation. It is likely that in both these latter cases the local measurements are dominated by sensor noise (rather than lack of fit to the model). When the model is good and there is measurement uncertainty, the inherent redundancy in the data allows more robust estimate of the motion. Certainly, more work needs to be done to develop measures that distinguish measurement uncertainty from modeling error. Both are essential.

5. Discussion

Current experimental work requires that the objects be stationary while the multiple images are acquired. Once a set of images is obtained, one for each light source, the object is allowed to move slightly and the cycle is repeated. While this constitutes a valid demonstration of the method, it fails as a methodology for practical realization since light sources cannot easily be turned on and off rapidly enough to support the tracking of continuously moving objects.

The paradox of requiring simultaneous images of a continuously moving object under different conditions of illumination can be resolved by multiplexing the spectral dimension. Suppose three narrow-band collimated light sources, say red, green and blue, continuously illuminate a work space from three different directions. Many color cameras employ three distinct internal imaging systems, each producing a separate black and white image corresponding to a different spectral channel. Suppose these three images are acquired simultaneously and suppose there is minimal overlap in their spectral response. Then this color camera and light source configuration would support multiple light source optical flow, as described above, for objects moving continuously in the work space. Further, there is no requirement that the spectral channels chosen be in the visible portion of the spectrum. Channels in the near (i.e., reflective) infrared also are a possibility and can be chosen not to interfere with other vision algorithms working in the visible portion of the spectrum. Future work will exploit these ideas.

6. Conclusions

Multiple light source optical flow is a principled way to obtain additional local constraint. At points where the optical flow constraint equations (6-8) are satisfied exactly, three light source optical flow supports the fast, accurate and robust estimation of geometric motion. The inherent redundancy in the measurements can be exploited to validate the computation locally, including the ability to determine locations where optical flow and the 2D motion field do not coincide.

Multiple light source optical flow is complementary to techniques based on contour analysis or sparse feature matching because it works best where those techniques fail (and *vice versa*). Multiple light source optical flow works best on smoothly curved surfaces, without distinct surface markings, because local brightness then depends primarily on local shading. This makes it possible to obtain dense, non-redundant, local information by varying the direction of illumination. The method will be unreliable at surface discontinuities and surface markings because then local brightness change is dominated by scene features largely independent of the direction of illumination.

Acknowledgement

Stay-Puft Marshmallow Man is a registered trademark of Kenner Parker Toys International Inc. Uri Ascher, Rod Barman, Mike Bolotski, Alan Mackworth, Jim Little and David Lowe all provided useful comments on earlier versions of this work. Support was provided by the Natural Sciences and Engineering Research Council of Canada (NSERC) and by the Canadian Institute for Advanced Research (CIAR).

References

[1] Horn, B.K.P. & B.G. Schunck (1981), "Determining Optical Flow", *Artificial Intelligence* (17)185-203.

[2] Kearney, J.K., W.B. Thompson & D.L. Boley (1987), "Optical flow estimation: an error analysis of gradient-based methods with local optimization", *IEEE Trans. Pattern Analysis and Machine Intelligence* (9)229-244.

[3] Mailloux, G.E., F. Langlois, P.Y. Simard & M. Bertrand (1989), "Restoration of the velocity field of the heart from two-dimensional echocardiograms", *IEEE Trans. Medical Imaging* (8)143-153.

[4] Ohta, N. (1989), "Optical flow detection by color images", *Proc. IEEE Int. Conf. Image Processing* pp 801-805, Pan Pacific, Singapore.

[5] Pentland, A.P. (1989), "Photometric motion", Vision Sciences TR-120, MIT Media Lab, MIT, Cambridge, MA.

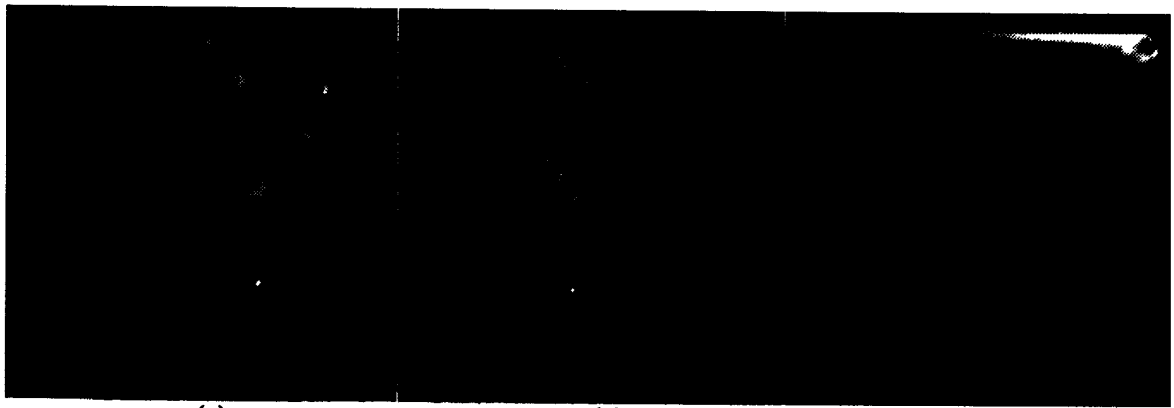
[6] Schunck, B.G. (1984), "The motion constraint equation for optical flow", *Proc. 7th Int. Conf. Pattern Recognition*, pp 20-22, Montreal, PQ.

[7] Schunck, B.G. (1988), "Image flow: fundamentals and algorithms", in *Motion Understanding: Robot and Human Vision*, W.N. Martin & J.K. Aggarwal (eds.), pp 23-80, Kluwer Academic Publishers, Boston, MA.

[8] Verri, A., F. Girosi & V. Torre (1990), "Differential techniques for optical flow", *J. Optical Society of America A*, (7)912-922.

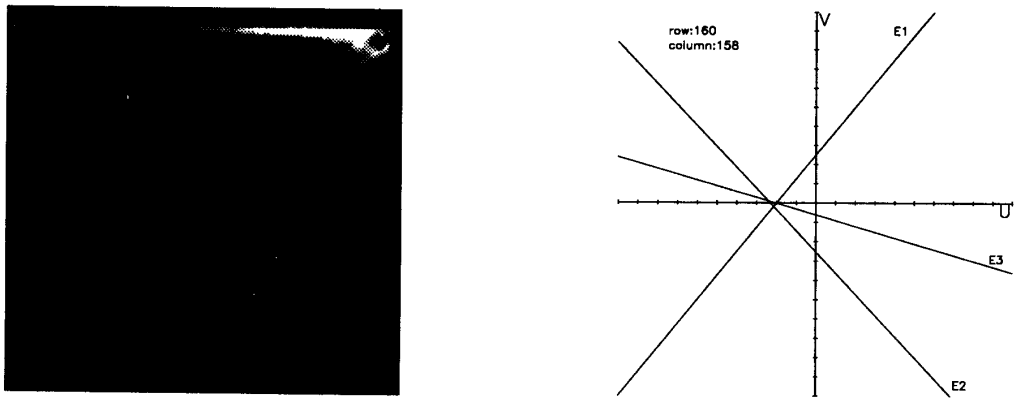
[9] Verri, A. & T. Poggio (1989), "Motion field and optical flow: qualitative properties", *IEEE Trans. Pattern Analysis and Machine Intelligence* (11)490-498.

[10]Woodham, R.J. (1980), "Photometric method for determining surface orientation from multiple images", *Optical Engineering* (19)139-144.



(a) (b) (c)

Figure 1. Three images of the Stay-Puft Marshmallow Man. Figures (a), (b) and (c) show images of the toy under three different conditions of illumination at $t = t_0$. Motion from $t = t_0$ to $t = t_1$ was a small translation from right to left. The object moved, but not the background.



(a) (b)

Figure 2. Multiple light source optical flow computation at one point on the Stay-Puft Marshmallow Man example. Figure (a) marks the point. Figure (b) shows the three flow constraint lines.

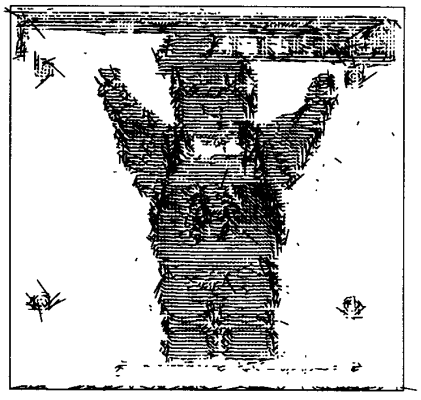


Figure 3. Optical flow computed for the Stay-Puft Marshmallow Man.

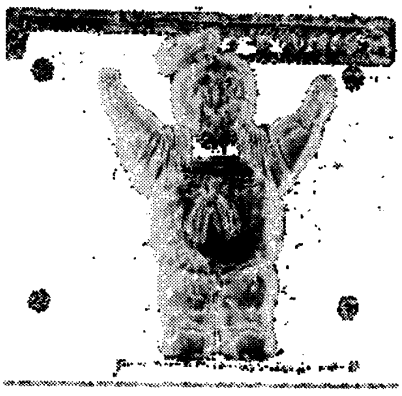


Figure 4. Error result for the Stay-Puft Marshmallow Man example.