Confidence Testing Optical-Flow Estimates

M. Fleury, A. C. Downton and A. F. Clark

Department of Electronic Systems Engineering, University of Essex,

Wivenhoe Park, Colchester, CO4 4SQ, U.K

tel: +44 - 1206 - 872795

fax: +44 - 1206 - 872900

e-mail fleum@essex.ac.uk

November 3, 1998

Abstract

Localised estimates of the optical-flow field in a video-frame sequence are noise prone. By observing a similarity between first differential filtering used for two differing purposes, confidence testing of estimates is shown to be reliant on the curvature of an intensity surface. A successful confidence test moves towards hardware implementation.

1 Introduction

2-D optical flow leading to a motion field can be extracted from a sequence of progressive video frames by processing the image intensity over time, without matching corresponding features in the 3-D scene structure. Some potential applications, such as video-compression, surveillance, and moving target indication, require video-rate responses. Certain differential methods of calculating optical flow having the required algorithmic regularity are suggestive of a VLSI solution provided there is a means of rejecting unreliable full estimates. This letter proposes that the first fundamental form of differential geometry is the basis of a confidence test.

In the Lucas-Kanade (LK) method [1], the velocity vector for an image patch is averaged out based on an understanding that: per-pixel image brightness remains constant over a small time interval (one frame), all pixels within the patch have the same velocity vector being from the same object, and the image intensity will vary across the patch (though not as a random pattern). Averaging out, in the LK case by Linear Least-Squares Error (LLSE), may be seen as equivalent to using the aggregate intensity gradients at a single pixel site. Two ways to disturb the LK model are if the scene reflectance is not Lambertian (and contrast is variable), or if additive Gaussian noise is present. [2] proposed a Bayesian estimation model to compensate the LK method for such disturbances. In [3], the modified LK method was shown in a wide-ranging and thorough study to perform well compared to other optical flow methods when tested on a set of synthetic video sequences. However, [3] modified the Bayesian estimator, apparently on heuristic grounds though noticing a link with the 'aperture problem'. Setting aside the problem of occluded objects, there is another well-known reason to reject some estimates from a method using local measurements, i.e. if the image patch intensity surface is locally planar, or marked by a continuous straight edge. This letter explains why the Bayesian estimator will accept estimates even when the intensity structure is ambiguous, explains how the modified estimator may be working, and indicates how a confidence threshold can be set to a fixed (low) value without eigenvector calculation as is appropriate for hardware implementation.

2 Confidence Testing

The optical-flow equation (OFE) for image-intensity field I is

$$\mathbf{I_s} \cdot \mathbf{v} + I_t = 0, \tag{1}$$

with $\mathbf{I_s} = (I_x = \delta I/\delta x, I_y = \delta I/\delta y)$, and $\mathbf{v} = (u = dx/dt, v = dy/dt)$, the optical-flow vector. Perturb to form

$$\mathbf{I_s} \cdot \mathbf{v} + I_t = \mathbf{I_s} \cdot \mathbf{n_1} + \mathbf{n_2},\ \mathbf{n_i} \sim N(0, \Sigma_i), \tag{2}$$

which can be represented as the conditional probability $P(I_t|\mathbf{v},\mathbf{I_s})$. Use Bayes' rule to form

$$P(\mathbf{v}|\mathbf{I_s},I_t) = \frac{P(I_t|\mathbf{v},\mathbf{I_s}).P(\mathbf{v})}{P(I_t)}, \tag{3}$$

which with $P(\mathbf{v}) \sim N(0,\Sigma_p)$ leads to the Maximum Likelihood Estimate (MLE), $\mu_{\mathbf{v}}$, set as:

$$\Sigma_{\mathbf{v}} = \left[\frac{\mathbf{M}}{(\Sigma_1||\mathbf{I_s}||_2^2+\Sigma_2)} + \Sigma_p^{-1}\right]^{-1} \tag{4}$$

$$\mu_{\mathbf{v}} = -\Sigma_{\mathbf{v}}.\frac{\mathbf{b}}{(\Sigma_1||I_s||_2^2+\Sigma_2)}, \tag{5}$$

where $\mathbf{M} = \mathbf{I_s}\mathbf{I_s}^T$ and $\mathbf{b} = I_t\mathbf{I_s}$. The LLSE solution for (1) when summed over a local image patch is

$$\mathbf{M}.\mathbf{v} + \mathbf{b} = \mathbf{0}, \tag{6}$$

which resembles (5). Since (5) is modified by $||\mathbf{I_s}||_2^2 = \lambda_1 + \lambda_2$, which is the trace of $\mathbf{M}$ with $\{\lambda_i\}$ the eigenvalues, [2] use the trace as a confidence measure for $\mathbf{v}$. [3] threshold the eigenvalues (by either 1.0 or 5.0). If both eigenvalues are larger $\mathbf{v}$ is accepted and if only one is above the threshold an estimate in the direction of $\mathbf{I_s}$ is used.

Consider an image surface $\mathbf{S}(\mathbf{x},\mathbf{y}) = x\mathbf{i} + y\mathbf{j} + I(x,y)\mathbf{k}$ as a coordinate (Monge) patch. The first fundamental form, $\Theta_1 = d\mathbf{S}\cdot d\mathbf{S}$ with $d\mathbf{S} = \mathbf{S}_x dx + \mathbf{S}_y dy$, has covariant metric tensor

$$\mathbf{G} = \begin{bmatrix} 1+I_x^2 & I_xI_y \\ I_xI_y & 1+I_y^2 \end{bmatrix}, \tag{7}$$

which viewed as a matrix has a similar structure to $\mathbf{M}$. Setting $\mu_i = 1 - \lambda_i$ in the characteristic equation for $\mathbf{G}$, $\{\lambda_i\}$ again the principal values, results in a solution: $\mu_1 = 0$ and $\boldsymbol{\mu_1} = (-I_y, I_x)$ together with $\mu_2 = ||\nabla I||_2^2$ and $\boldsymbol{\mu_2} = (I_x, I_y)$, the $\{\boldsymbol{\mu_i}\}$ being the associated eigenvectors. Suppose an average in the form of an LLSE estimate of $\mathbf{G}$ is taken over an image patch adjacent to a step edge aligned with the y axis, then $\mathbf{G}$ is already the average tensor, I_x will be high and consequently so will μ_2. Therefore, trace($\mathbf{G}$) is inadequate as a confidence measure for optical flow despite the MLE, though a low trace value will reject planar patches. Turn to det($\mathbf{G}$) $= \mu_1\mu_2$ where now $\mathbf{G}$ is a weighted average over an image patch with a variety of per-pixel intensity gradients (so that det($\mathbf{G}$) is no longer zero).

$$\det(\mathbf{G}) = ||\mathbf{S}_x \times \mathbf{S}_y||_2^2 = (||1+I_x||_2.||1+I_y||_2)^2 \sin^2\nu, \tag{8}$$

where ν is the angle between the coordinate vectors which govern the warp imposed on the coordinate patch by the surface. In fact, $\sin \nu$ is directly related to the intensity surface normal curvature at a point by Euler's equation. Therefore, a high value of both eigenvalues is required for a successful confidence test. Notice that sometimes second, not first, differentials from the Hessian matrix are taken to find the invariant Gaussian curvature, typically to determine the type of curvature.

[4] used a similar analysis to investigate a corner detector but differed in the interpretation of aggregate $\mathbf{G}$. The intensity gradients for a video sequence also differ as (importantly) they are taken after Gaussian smoothing over a set of frames. An eigenvalue test will be more successful in the LK method because it is only necessary to establish a sound optical-flow vector and not find the type of corner. As in the corner detector, if the test utilises $\mathrm{trace}(\mathbf{G})/\det(\mathbf{G})$, which is akin to a conditioning number, then the curvature factor can be isolated. An analysis of the LLSE solution for the OFE as an inverse problem or through an inertia tensor method also will give rise this to test. Table 1, using the synthetic sequences available by anonymous ftp to csd.uwo.ca in /pub/vision/, shows the equivalence of results, which approach lower-bound error. Tuning of the test is required to set the acceptance density.

3 Conclusions

A confidence test requires an understanding of why it works. The Bayesian explanation is not fully satisfactory as it relies on *a priori* assumptions. The differential geometry explanation is helpful as it results in a test without formal calculation of eigenvalues. First differentials are sufficient to indicate curvature for this application.

Acknowledgement

This work is being carried out under EPSRC research contract GR/K40277 'Parallel software tools for embedded signal processing applications' as part of the EPSRC Portable Software Tools for Parallel Architectures directed programme.

References

[1] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In $7^{\underline{th}}$ *International Joint Conference on Artificial Intelligence*, pages 674–679, 1981.

[2] E. P. Simoncelli, E. H. Adelson, and D. J. Heeger. Probability distributions of optical flow. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 310–315, 1991.

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

[4] A. Noble. Finding corners. In $3^{\underline{rd}}$ *Alvey Vision Conference*, pages 267–274, 1987.

Sequence name	Frame no.	Trace > 2.0	λ_1, λ_2 ≥ 1.0	λ_1, λ_2 ≥ 5.0	trace/det ≤ 1.05	trace/det ≤ 0.3
yosemite	9	10.4 ± 15.8, 86.0	4.3 ± 10.1, 39.8	3.3 ± 8.6, 12.2	4.3 ± 10.2, 39.2	3.6 ± 9.2, 15.9
square1	9	20.4 ± 24.83, 41.6	0.3 ± 0.3, 7.1	0.2 ± 0.2, 4.8	0.3 ± 0.3, 7.2	0.3 ± 0.3, 5.3
tree-div	20	2.4 ± 2.5, 82.5	2.0 ± 2.0, 54.4	1.7 ± 1.5, 30.1	2.0 ± 2.0, 54.1	1.8 ± 1.8, 35.6
tree-tran	20	0.9 ± 1.2, 85.0	0.6 ± 0.7, 46.0	0.6 ± 0.6, 18.3	0.6 ± 0.7, 45.9	0.6 ± 0.6, 23.9

Table 1: Angular Error with Standard Deviation (degrees) and Density (%) for Different Confidence Tests