MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19. 1998, Makuhari, Chiba, Japan

12—1 Pointwise "Flow" from a Variational Principle

Harit P Trivedi
Defence Evaluation & Research Agency, Chobham Lane, Chertsey, Surrey KT16 0EE, UK
e-mail:hptrivedi@dera.gov.uk

Abstract

A variational principle which (1) accommodates a deforming signal without an explicit model, and (2) provides a confidence measure for its result, is suggested for signal flow calculations.

Motivation

It is well known that a translating pattern can be represented as a function $I(\mathbf{x}-\mathbf{v}t)$. Elementary differentiation then yields the identity $(\nabla + \mathbf{v}.\ \partial/\partial t)I = 0$, well-known in the vision community as the brightness constraint equation[2]. For a 1-dimensional domain $\mathbf{x}$, the equation provides a unique solution for the (scalar) velocity v at every point, but no means of verifying to what extent the actual behaviour of the pattern is described by $I(\mathbf{x}-\mathbf{v}t)$. For higher dimensional domain $\mathbf{x}$, even this limited achievement vanishes and we have to consider second derivatives. It can be shown, again with little effort, that $(\mathbf{v}, 1)$ forms the null space of the Hessian $\mathbf{H}$ of I ($H_{ij} \equiv \partial^2 I/\partial x_i \partial x_j$; x_i, x_j range over $\mathbf{x}, t$). This provides both $\mathbf{v}$, and via the smallest eigenvalue λ, a cross-check on the validity of the form $I(\mathbf{x}-\mathbf{v}t)$. In practice, non-zero λ could arise from noise in the data, deformation of the pattern (i.e., deviation from the presumed functional form), or both. If, as in most visual image sequences, the latter possibility cannot be ruled out, there is no way of uncoupling the two effects in the absence of knowledge of the specific form of the deformation. Modelling the deformation introduces extra parameters and the attendant risk of over/under-modelling.

A variational principle, that of minimum curvature on the surface $I(\mathbf{x},t)$, offers a way out. It leads to the generalised eigenvalue problem $\mathbf{H}\psi = \lambda\mathbf{G}\psi$, where $\mathbf{G} \equiv \mathrm{I} + (\nabla I)(\nabla I)^T$ (i.e., $G_{ij} = \delta_{ij} + I_i I_j$). Eigenvectors ψ specify the arc direction and λ^2 the curvature value of the stationary solutions, known as geodesics[8,9]. This does away with the assumption of the special functional form $I(\mathbf{x}-\mathbf{v}t)$. It can be shown, however, if $I(\mathbf{x},t)$ is of the form $I(\mathbf{x}-\mathbf{v}t)$, that the problem possesses a null eigenvalue and that $\psi=(\mathbf{v}, 1)$ is the eigenvector belonging to it. This variational principle thus subsumes translation as a special case and is therefore a possibly useful generalisation of "flow" calculation. The formalism accommodates any deformation of the signal pattern automatically.

Background

It is often desirable to be able to compute optic flow at every point in an image sequence [1], rather than at just some isolated feature points (corners and edges) over whose locations and density the user has very little control. Some applications (for example the detailed 3D structure of the viewed surfaces) even require the dense flow information. Many methods already exist in the literature, some are differential [2-5] while others involve Fourier [6] and other [7] transforms. We propose a differential scheme invoking a variational principle, that of minimum curvature [8, 9], suggested above.

This suggestion is motivated by the desire to free the interpretation of the temporal behaviour from an explicit model fitting paradigm (such as a translating pattern) in favour of a variational principle *which at the same time automatically reproduces the model results when appropriate.* An advantage of employing a variational principle is that it circumvents over/under/mis-modelling of the situation.

Algorithm

The problem of finding the arc direction with minimum curvature comes down to solving the generalised eigenvalue/vector problem: $\mathbf{H}\psi = \lambda\mathbf{G}\psi$, where $\mathbf{H} \equiv (\nabla\nabla I)$ (i.e., $H_{ij} = I_{ij}$), and $\mathbf{G} \equiv \mathrm{I} + (\nabla I)(\nabla I)^T$ (i.e., $G_{ij} = \delta_{ij} + I_i I_j$) are the hessian and the metric matrices respectively, and $\psi =$

$(\mathbf{v}, 1)^T$. The solution with the smallest $|\lambda|$ is sought. Collapsed signal domain dimensionality (e.g. a long 1D edge) is detected and the null component is prevented from contaminating the solution. The algorithm is structured thus:

- If $|\partial/\partial t\, |(\nabla I)|^2\, | < \varepsilon_{abs}$, or $< \varepsilon_{rel}\, |(\nabla I)|^2$, implies no motion, hence $\mathbf{v} = 0$.
- Else,
- Rotate $\mathbf{H}$ and ∇I in the *spatial* subdomain so as to diagonalise $\mathbf{H}$ in *that* subdomain
  - Every small eigenvalue characterises a collapsed domain dimension
- If a small eigenvalue (relatively or absolutely) - e.g. H_{xx} small,
  - set the corresponding row and column as well as the component of ∇I to 0 - i.e., $H_{xi} = H_{ix} = 0$ and $I_x = 0$. This makes the null component ineffectual.
  - *and* the eigenvalue itself (the diagonal) to a *large* number - i.e., H_{xx} = LARGE which prevents the redundant direction from being selected below.
  - Hessian now has reduced dimensions, augmented with a null row and column (except for the diagonal) corresponding to the dummy dimension
- Construct $\mathbf{G}$ (from ∇I) and solve $\mathbf{H}.(\mathbf{v}, 1)^T = \lambda \mathbf{G}\, .(\mathbf{v}, 1)^T$ in the full domain. If $|\lambda_1| \le |\lambda_2| \le |\lambda_3|$, then $|\lambda_1|$ determines the minimum curvature and the corresponding eigenvector $(\mathbf{v}, 1)$ normally yields the velocity $\mathbf{v}$:
  - with great ($\cong$100%) confidence if $|\lambda_1|/|\lambda_3| << 1$ *and* $|\lambda_1|/|\lambda_2| << 1$
  - medium ($\cong$50%) confidence if $|\lambda_1|/|\lambda_3| << 1$ *but* $\lambda_1 \cong -\lambda_2$, implying two equally likely solutions (recalling that only $|\lambda|$ matters)
  - and no ($\cong$0%) confidence if $\lambda_1 \cong \lambda_2$, implying infinitely many equally likely solutions on account of degeneracy
  - the confidence refers to the *uniqueness* of $\mathbf{v}$ (or the eigenvector), *not* its numerical value.

  (The numerical uncertainty depends on data measurement and discretisation errors, derivative estimation errors and the consequent errors in eigenvector estimation.)

- Occasionally, the eigenvector belonging to λ_1 has a negligible time component, giving ridiculously high velocity. While mathematically correct, such a solution is rejected on physical grounds in favour of λ_2 and *its* eigenvector. The confidence level is reassessed by comparing λ_2 now with λ_3.

Results

It is instructive to consider a signal $I(\mathbf{x},t)$ in 1-dimensional domain $\mathbf{x}$ because it can be readily visualised. Suppose that $I(\mathbf{x},t)$ traces out an ellipsoid: $(I)^2 + (x/15)^2 + (t/28)^2 = 1$, $I \ge 0$ (to keep it single-valued). This is a deforming pattern without translation. Figure 1 shows the signal at 6 evenly spaced time frames *t1-t6* and the trajectory vector $(\delta x, \delta I)$ at several points along the curve at time frame *t5* ($t = 15$) for *(a)* the *brightness conservation, (b)* the *hessian,* and *(c)* the *geodesic* flows; δI in *(b)* and *(c)* being accurate through second order in δx, δt. The flow converging towards the centre is the most intuitive. Figure 2 displays the same information for the ellipsoidal surface rotated by *0.5 rad* in the *x-t* plane. The signal at any time *t* represents an oblique slice of the ellipsoid, with its centre now translating. This illustrates a deforming pattern with translation. The centripetal flow appears to accord best with intuition.

Good results (using the error criterion of Barron *et al* [1]) are obtained with real as well as synthetic imagery before post-processing.

The geodesic optic flow appears to be insensitive to constant spatiotemporal intensity gradients across image sequences. This is illustrated with an example (figure 3) in which the brightness of the translating image rises by 1 grey level every frame. This insensitivity provides some protection against small changes in scene illumination, e.g. displacement and/or dimming/brightening of the light source(s). *Changing intensity* (s.t. $\delta^2 I$ is smallest) is associated with the geodesic flow. In contrast, the brightness constancy governed flow *preserves* the intensity along the flow path, and so must, in general, effect *a different path.*

The geodesic principle and algorithm clearly apply to signals in domains of *any* dimensionality (not just 2 as in images).

Conclusions

Once the correspondence between the geodesics and the optic flow is acknowledged, it becomes possible to pass from framewise knowledge of the flow field *to tracking individual motion trajectories* over several frames by integrating the corresponding Euler-Lagrange equation. These trajectories couple to each other via the surface terrain and its derivatives. Equally, a dense set of geodesics could be propagated in time to predict the as yet unseen terrain or to fill in missing parts thereof. Since the trajectories mutually interact they evolve in a self-consistent manner. This should make the predicted images nontrivial and interesting.

Acknowledgement

This article is published with the permission of the controller of Her Britannic Majesty's Staionery Office.

References

[1] J L Barron, D J Fleet, S S Beauchemin, and T A Burkitt, "Performance of Optical Flow Techniques", in t*he Proceedings of CVPR-1992*, pp. 236-242.

[2] B K P Horn and B G Schunk, "Determining optical flow", *Artificial Intelligence*, Vol **17**, 1981, pp. 185-204.

[3] B D Lucas and T Kanade, "An iterative image-registration technique with an application to stereo vision", in *Image Understanding Workshop, DARPA, 1981*, pp. 121-130.

[4] J Weber and J Malik, "Robust Computation of Optical Flow in a Multi-Scale Differential Framework", *Int. J. Comp. Vis.*, Vol. **14**, 1995, pp. 67-81.

[5] S Uras, F Girosi, A Verri, and V Torre, "A Computational Approach to Motion Perception", *Biological Cybernetics*, Vol **60**, 1988, pp 79-97.

[6] D J Fleet and A D Jepson, "Computation of component image velocity from local phase information", *Int. J. Comp. Vis.*, Vol. **5**, 1990, pp. 77-104.

[7] Y Xiong and S A Shafer, "Moment and Hypergeometric Filters for High Precision Computation of Focus, Stereo and Optic Flow", *Int. J. Comp. Vis.*, Vol. **22**, 1997, pp 25-59.

[8] H Hertz, *The Principles of Mechanics*, Macmillan & Co, London, 1899.

[9] H Goldstein, *Classical Mechanics*, Addison-Wesley, Reading (Ma), 1964, p.234

Figures

FIGURE 1: (left) signal values at 6 evenly spaced time frames *t1-t6* for an ellipsoidal signal surface whose axes are aligned with the coordinate axes, and, (right) the flow vector $(dx,dI)/dt$ at several x values along the signal curve at time *t5*.

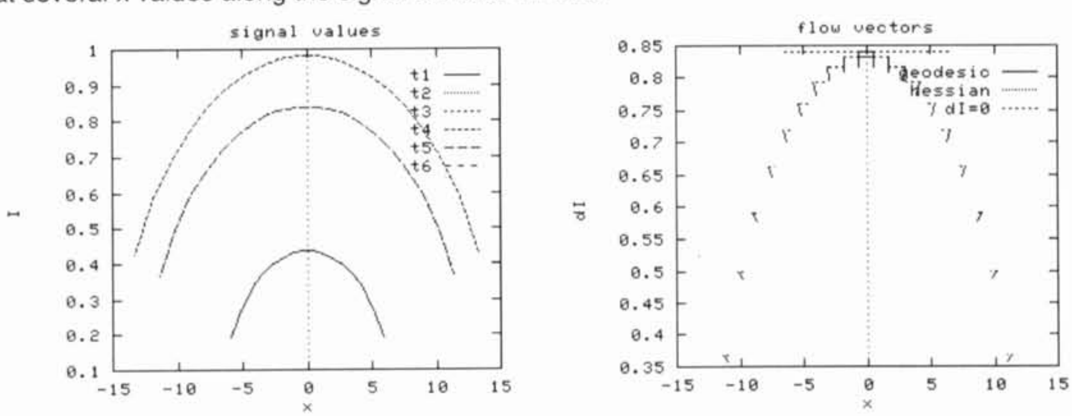

FIGURE 2: (left) signal values at 6 evenly spaced time frames *t1-t6* for the ellipsoidal signal surface of fig.(1) but with axes rotated by *0.5 rad* in the x-t plane, and, (right) the flow vector (*dx,dI*)/*dt* at several *x* values along the signal curve at time *t5*.

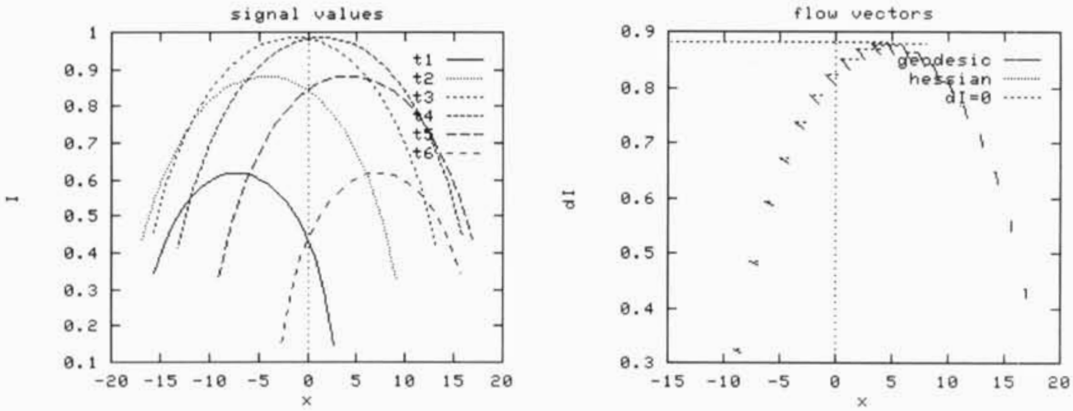

FIGURE 3 *Top row*: Uniformly translating and brightening pattern (left). Reconstructed (right) image with geodesics. *Middle row*: Flow vectors from Horn & Schunk's algorithm (courtesy Barron et al [1]) without (left) and with (right) brightening (note the patterned disturbance). *Bottom row*: Flow vectors from geodesics without (left) and with (right) brightening (unchanged, apart from mainly aliasing related glitches in both cases).