First Order Optic Flow from Log-Polar Sampled Images

H. Tunley and D. Young

School of Cognitive and Computing Sciences, University of Sussex, Brighton, England BN1 9QH

Abstract. The first-order spatial derivatives of optic flow – dilation, shear and rotation – provide powerful information about motion and surface layout. The log-polar sampled image (LSI) is of increasing interest for active vision, and is particularly well-suited to the measurement of local first-order flow. We explain why this is, propose a simple least-squares method for measuring first-order flow in an LSI sequence, and demonstrate that the method works well when applied to real images.

1 Introduction

The log-polar sampled image has a long history in computer vision ([3] [5], [6], [8]). Tistarelli and Sandini [5] [6], for example, have demonstrated how measurements of LSI optic flow can yield variables relevant to the control of action, by relating the time to nearest approach of a point in the environment to the flow and its partial derivatives. Despite this work the LSI has received relatively little interest, due to the dominance of raster-scan imaging technology, and the requirement of an active vision system capable of controlling fixation to perform eye movements and stabilise images.

Optic flow estimation involves a trade-off between spatio-temporal localisation and accuracy of parameter estimates. In an active vision system, where computational resources can be rapidly switched to different parts of the optic array, it may be a good strategy to use a simple flow model for the current region of interest, to give a robust approximation to the data, rather than forming dense but noisy estimates over the entire image. Our approach is to assume that the optic flow field can be modelled locally as a first-order function of image coordinates, and to estimate the parameters of this function. These parameters specify the local zero-order flow, divergence, shear and rotation, which provide valuable information for the control of action [2, 4]. This model is a good approximation to the true flow for sufficiently local regions of images of smooth surfaces.

We assume that the LSI will normally be centred on a particular surface whose orientation and relative motion are of interest. Given the concentration of samples close to the origin, an unweighted estimate of a parameter will be dominated by information from a relatively small region of the image, which will usually correspond to a coherent surface. Active vision should allow rapid exploration of the visual environment using a succession of camera movements

Lecture Notes in Computer Science, Vol. 800
Jan-Olof Eklundh (Ed.)
Computer Vision – ECCV '94

(like saccades) to change the fixation. Each provides a relatively small amount of reliable information, rather than pixel-by-pixel maps.

Log-polar mapping produces a sampling density which decreases linearly with distance from the origin. Image points are indexed by a logarithmic distance from the centre, ξ, and an angle, γ The LSI is often displayed with ξ and γ as orthogonal rectilinear coordinates (usually ξ horizontal and γ vertical). This is like cutting the round LSI at the 3 o'clock position and deforming it into a rectangle by opening out the cut. An example of a conventionally-sampled image (CSI) and its centred LSI transformation are shown in Fig. 1 and Fig.2. If an object is fixated then expansion maps to a shift along the ξ axis and rotation about the origin maps to a shift along the γ axis.

Fig. 1. CSI

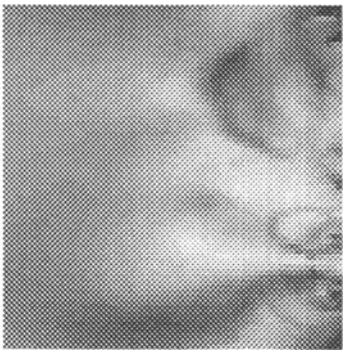

Fig. 2. LSI

2 First-Order Optic Flow: Method and Implementation

The LSI is particularly well-matched to the measurement of first-order flow, for the following reason. Consider the Taylor expansion of the optic flow $\mathbf{v}$ at the origin (ignoring second-order and higher terms):

$$\mathbf{v}(\mathbf{r}) = \mathbf{v}_0 + \mathbf{T}\mathbf{r} = \begin{bmatrix} v_{x_0} \\ v_{y_0} \end{bmatrix} + \begin{bmatrix} D + S_1 & S_2 - R \\ S_2 + R & D - S_1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{1}$$

where $\mathbf{r}$ is image position; $\mathbf{v}_0$ is the uniform, zero-order flow at the origin; $\mathbf{T}$ is the deformation rate tensor (the matrix of first derivatives of the flow at $\mathbf{r}$); D is dilation, R, rotation, and S_1 and S_2 are the components of shear. If the point at the origin is tracked ($\mathbf{v}_0 = 0$), and the observed surface is reasonably smooth (higher-order terms negligible), then in some region round the origin the first-order term will dominate to give:

$$\mathbf{v}(\mathbf{r}) \simeq \mathbf{T}\mathbf{r} \tag{2}$$

The magnitude of $\mathbf{v}$ is therefore proportional to the distance from the origin, suggesting an image sampling strategy where the spatial sample separation is also proportional to $\mathbf{r}$, as in the LSI. The image motion between successive frames will then be a constant fraction of the sample separation at every point in the image. This provides the best basis for integrating information across an image region; in a CSI, the first-order flow is likely to vary from far less than a pixel per frame close to the tracked point to many pixels per frame far from it, giving useful information in only a relatively small ring. By tracking the image feature at the LSI origin a match is maintained between between flow magnitude and sample spacing. One way to achieve this is to estimate $\mathbf{v}_0$ from the LSI data to provide feedback into a tracking system. We show below how to estimate both the zero-order term $\mathbf{v}_0$ and the components of $\mathbf{T}$ from the grey levels of an LSI sequence.

Log-polar resampling was carried out using software. Sub-pixel density central regions were processed using bilinear interpolation between the four nearest CSI neighbours, whilst sparsely-sampled peripheral regions used simple grey-level averaging within a circular neighbourhood. A switch between strategies occurred at an intermediate point.

2.1 First-Order Flow in the Log-Sampled Image

An LSI image position has the log-polar coordinates (ξ, γ):

$$\xi = \log_a \rho - p \quad \text{and} \quad \gamma = q\eta \tag{3}$$

where p and a are constants and (ρ, η) are the polar coordinates. The optic flow field in log-polar coordinates is related to the conventional form by:

$$\begin{bmatrix} \dot{\xi} \\ \dot{\gamma} \end{bmatrix} = \begin{bmatrix} \frac{\partial \xi}{\partial x} & \frac{\partial \xi}{\partial y} \\ \frac{\partial \gamma}{\partial x} & \frac{\partial \gamma}{\partial y} \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \frac{1}{\rho} \begin{bmatrix} \frac{\cos\eta}{\log a} & \frac{\sin\eta}{\log a} \\ -q\sin\eta & q\cos\eta \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} \tag{4}$$

Combining (1) and (4) gives first-order flow in (ξ, γ) coordinates:

$$\begin{bmatrix} \dot{\xi}\log_e a \\ \frac{\dot{\gamma}}{q} \end{bmatrix} = \begin{bmatrix} \frac{1}{\rho}\cos\eta & \frac{1}{\rho}\sin\eta & 1 & 0 & \cos 2\eta & \sin 2\eta \\ -\frac{1}{\rho}\sin\eta & \frac{1}{\rho}\cos\eta & 0 & 1 & -\sin 2\eta & \cos 2\eta \end{bmatrix} \begin{bmatrix} v_{x_0} \\ v_{y_0} \\ D \\ R \\ S_1 \\ S_2 \end{bmatrix} \tag{5}$$

2.2 Least-Squares Estimates of Flow Parameters

Given LSI optic flow estimates $(\dot{\xi}, \dot{\gamma})$ it is straightforward to obtain a least squares estimate of the flow parameters. Alternatively, the brightness constancy assumption, $\frac{dI}{dt} = 0$, gives a relationship that can be used to estimate the flow parameters directly from the spatial and temporal grey-level gradients.

$$\dot{I} = -\left[\frac{\partial I}{\partial \xi} \; \frac{\partial I}{\partial \gamma}\right]\begin{bmatrix}\dot{\xi}\\ \dot{\gamma}\end{bmatrix} \tag{6}$$

Substituting (5) into (6) gives:

$$\dot{I} = -\left[\frac{g_x}{\rho} \; \frac{g_y}{\rho} \; g_\xi \; g_\gamma \; g_u \; g_v\right]\begin{bmatrix}v_{x_0}\\ v_{y_0}\\ D\\ R\\ S_1\\ S_2\end{bmatrix} \tag{7}$$

$$g_\xi = \frac{1}{\log a}\frac{\partial I}{\partial \xi} \quad \text{and} \quad g_\gamma = q\frac{\partial I}{\partial \gamma}$$
$$g_x = g_\xi \cos\eta - g_\gamma \sin\eta \text{ and } g_y = g_\xi \sin\eta + g_\gamma \cos\eta$$
$$g_u = g_\xi \cos 2\eta - g_\gamma \sin 2\eta \text{ and } g_v = g_\xi \sin 2\eta + g_\gamma \cos 2\eta \tag{8}$$

Six or more measurements of $\dot{I}$, $\frac{\partial I}{\partial \xi}$ and $\frac{\partial I}{\partial \gamma}$ at different points in the LSI allow least-squares estimation of the flow parameters. The implicit noise model is simplest if we choose $q = \frac{1}{log_e a}$, which corresponds to equal sample spacing in the radial and tangential directions at each point in the LSI.

3 Results

3.1 Parameter Extraction from Affine Transformations

Simulated optic flow was generated by applying affine transformations to real images to create realistic grey-level statistics with known flow parameters. For each base image, a sequence was generated by applying the appropriate transform to successive frames (all parameters altered simultaneously). A similar transformation was also applied to a range of CSI using the same number of sample points as the given LSI for a fair comparison. Results were gained for a number of different images and deformation rates per frame and are presented graphically. The degree of deformation is represented on the x axis, whilst the y axis indicates the RMS error from the correct (i.e. input) increment over a total of 5 frames. The images contained a single object centred on the LSI.

3.2 First-Order Flow Extraction: Dilation, Rotation and Shear

The accuracy of the LSI and CSI methods when determining dilation, rotation and shear are shown in Fig. 3. The LSI accuracy is superior to the CSI, especially at faster deformation rates. The superior CSI performance for the smaller parameter deformation rates is probably due to the software oversampling at the centre of the LSI introducing errors for very small displacements – a factor which would not affect an LSI obtained using a hardware implementation.

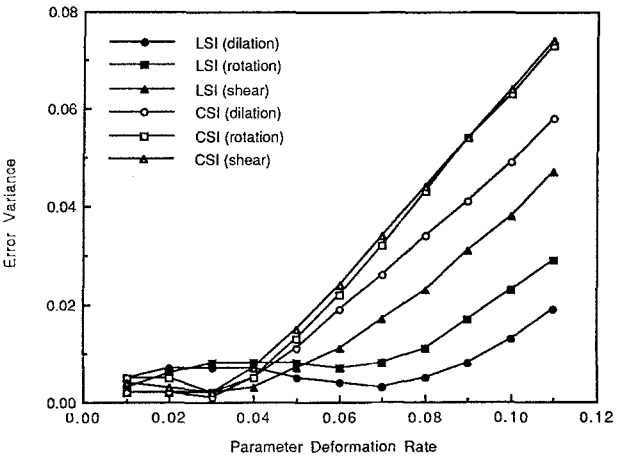

Fig. 3. Graph of Dilation, Rotation and Shear Accuracy as a Function of Deformation Rate

3.3 Determining Dilation from an Image Sequence

It is acknowledged that the use of affine transformations is not entirely realistic. In reality surfaces at different depths from the viewer, undergo differing distortions and require segmentation (e.g. [1]). However, for coherent surfaces, the simplification of regarding the scene as an approximately planar surface, without recourse to attempts at segmentation, allows reasonable accuracy to be achieved in real world situations. This is illustrated by an example sequence from a camera translating towards a distant table in front of a textured wall. The rate of dilation, D for the sequence was determined and the time to contact, t_c, (defined as $\frac{1}{D}$) plotted against frame number in Fig. 4. Despite no attempt being made to ensure that a stable point was fixated (the slight 'glitch' in the trace is a consequence of camera bounce introducing a strong translational component) there is a clear trend towards a reduction in t_c with frame progression.

4 Conclusion

We have taken advantage of the match between the spatial structure of first-order optic flow and the LSI sampling pattern to produce a robust and efficient method of estimating first-order flow parameters. This has been tested with a range of images and deformation rates using both affine transformations and a sequence from a translating camera. The next stage is to work with real image sequences containing multiple surfaces at differing depths in an attempt to assess the performance over more limited regions of the field of view. To truly exploit the nature of log-polar sampling requires an active vision system to track features at the centre of the field of view and increase the accuracy with which other parameters can be detected. The use of feedback to minimise $\mathbf{v_0}$ is thus also a priority.

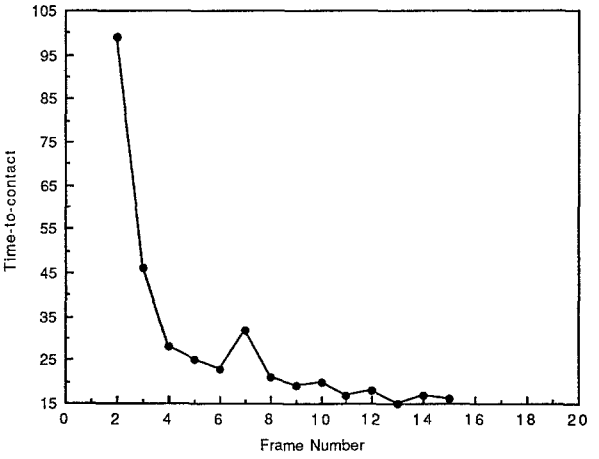

Fig. 4. Graph of Time-to-Contact as a Function of Frame Number

5 Acknowledgements

This work was supported by the UK Joint Research Council Initiative in Cognitive Science and Human Computer Interaction, and the Science and Engineering Research Council Image Interpretation Initiative.

References

1. Bouthemy, P., François, E.: Motion segmentation and qualitative dynamic scene analysis from an image sequence. Int. J. Comp. Vis. **10(2)** (1993), 157–182
2. Cipolla, R., Blake, A.: Surface orientation and time to contact from image divergence and deformation. in Proc. ECCV (1992) 187–202
3. Jain, R.C., Bartlett, S.L., O'Brien, N.: Motion stereo using ego-motion complex logarithmic mapping. IEEE Trans. PAMI **9(3)** (1987) 356–369
4. Koenderink, J.J., Van Doorn, A.J.: Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. Optica Acta **22(9)** (1975) 773–791
5. Tistarelli, M., Sandini, G.: On the advantages of polar and log-polar mapping for direct estimation of time-to-impact from optical flow. IEEE Trans. PAMI **15(4)** (1993) 401–410
6. Tistarelli, M., Sandini, G.: Dynamic aspects in active vision. CVGIP.: Image Understanding **56(1)** (1992) 108–129
7. Young, D.S.: First order optic flow. Cognitive Science Research Paper **313**, University of Sussex (1993)
8. Young, D.S.: Logarithmic sampling of images for computer vision. In Cohn, T. (Ed.) Proceedings of the 7th Conference of the Society for the Study of Artificial Intelligence and Simulation of Behaviour. Pitman, London (1989), 145–150