CENTRE-FREQUENCY ADAPTIVE IIR TEMPORAL FILTERS FOR PHASE-BASED IMAGE VELOCITY ESTIMATION

C.W.G. Clifford [1]   K. Langley[1]   D.J. Fleet[2]

[1]University College London, U.K.   [2]Queen's University, Kingston, Ontario, Canada.

This paper proposes an application of adaptive IIR filters to the problem of image velocity estimation. A phase-based motion algorithm (Fleet and Jepson [4], Langley and Fleet [9]) is employed to measure velocity locally within an image sequence from the outputs of a set of complex space-time separable band-pass filters. The filters' temporal tunings are adaptively modified on the basis of measured velocity to optimise the representation of image motion. In computer simulations the scheme is shown to provide accurate estimates of velocity even at high levels of image noise.

INTRODUCTION

The method we use for measuring image velocity is based on the phase information in the output of band-pass filters [4]. In a comparative study of the quantiative accuracy of several optical flow techniques, Barron et al [1] found that phase-based methods produced the most reliable estimates of image velocity. However, a common criticism of the many techniques that have been developed for computing image motion, and one which applies particularly to phase-based approaches, is the computational expense and the amount of memory necessary to store the required number of frames and the different filters' outputs. The phase-based approach of Fleet and Jepson [4], for example, involved 22 filters at each spatial scale, each of which was an FIR filter with a temporal extent of 15 frames or more.

In an attempt to create motion algorithms that run at or near frame-rates it is necessary to address the issue of efficiency. In general, the most common way of reducing storage requirements has been to restrict the duration of temporal support to a small number of frames (e.g. 2 or 3). An alternative is to introduce recursive temporal filtering, allowing performance to be maintained while improving markedly on the efficiency of implementation ([9], Fleet and Langley [6]). Here we describe the use of a class of adaptive IIR temporal filters within a phase-based framework.

Recursive centre-frequency adaptive band-pass filters were originally designed for the tracking and enhancement of one-dimensional signals in biomedicine and communications (Kumar and Pal [8]). Kumar and Pal's filters adapt according to a gradient algorithm operating on the basis of power output, while ours use measurements derived from phase. Phase information is reasonably stable with respect to small signal perturbations (Fleet and Jepson [5]), and it is therefore hoped that phase-based adaptation may offer faster convergence and wide applicability in signal processing.

The purpose of introducing adaptation into image velocity estimation is twofold: to reduce computation and storage requirements by reducing the number of independent temporal filters required; and to increase accuracy by maximising filtered image power. Assuming a model of constant image translation, $I(\mathbf{x}, t) = I_s(\mathbf{x} - \mathbf{v}t)$, image power will be constrained to lie on a plane through the origin in frequency space:

$$\hat{I}(\mathbf{k}, \omega) \ = \ \hat{I}_s(\mathbf{k})\delta(\mathbf{v} \cdot \mathbf{k} - \omega) \ , \tag{1}$$

where $\mathbf{k} = (k_1, k_2)$ and ω denote spatial and temporal frequencies. Therefore, if $I_s(\mathbf{x})$ is spatially band-pass, then the spatiotemporal signal $I(\mathbf{x}, t)$ is temporally band-pass, with its temporal frequencies constrained by its spatial frequency content and the image velocity. We consider this situation very suitable for the use of centre-frequency adaptive temporal filters. Using a model of local image translation, and filters of fixed spatial tuning, we feedback measured velocity to adjust the peak temporal frequency tuning of the IIR filters to match the temporal characteristics of the spatially filtered image signal (see Fig.1).

MEASURING VELOCITY

The phase-based approach to velocity measurement operates on the outputs of complex band-pass filters and their derivatives, so causal band-pass temporal filters and corresponding derivative filters are required. The phase derivatives of the filter outputs are combined over a small spatiotemporal region to arrive at an estimate of local velocity and a measure of confidence in that estimate. To avoid explicit calculation and differentation of phase, and the attendant problem of phase unwrapping, we use the following identity to extract phase derivatives ([4], Franks [7]):

$$\phi_t(\mathbf{x}, t) \ = \ \frac{\mathcal{I}m[R^*(\mathbf{x}, t) \ R_t(\mathbf{x}, t)]}{|R(\mathbf{x}, t)|^2} \ , \tag{2}$$

where $R(\mathbf{x}, t)$ is the complex-valued filter output, $\phi(\mathbf{x}, t)$ is its phase, and R_t is the partial temporal derivative of R. We compute velocity using a weighted least squares fit of local first-order constraints on phase
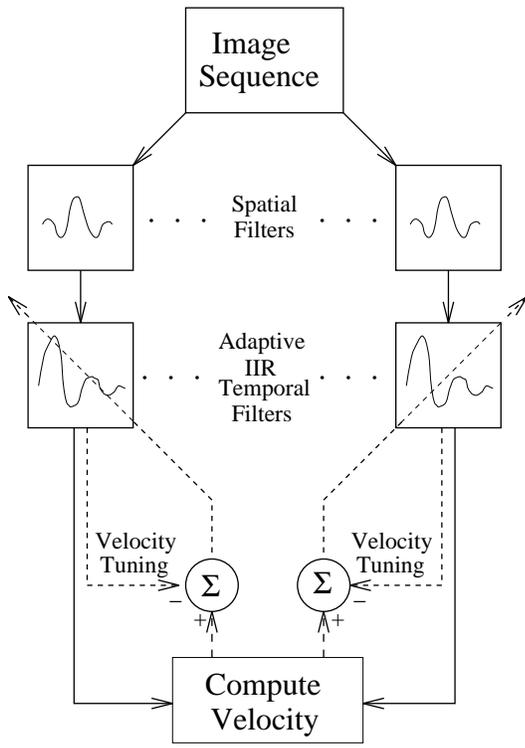
Figure 1: *Schematic diagram of the method of image velocity measurement described in this paper. Velocity measurements computed from the outputs of space-time separable band-pass filters are fed back to adapt the filters' temporal tunings so that they match the space-time characteristics of the image signal.*

to a model of constant velocity, $\mathbf{v}(\mathbf{x}, t)$, in each small spatiotemporal neighbourhood, N, by minimizing:

$$\sum_{(\mathbf{x}, t) \in N} W(\mathbf{x}, t)[|\nabla \phi(\mathbf{x}, t) \cdot \mathbf{v}(\mathbf{x}, t) + \phi_t(\mathbf{x}, t)|]^2 , \quad (3)$$

where $W(\mathbf{x}, t)$ is a window that gives more weight to constraints near the centre of the neighbourhood. The minimization of (3) leads to a linear system of the form $\mathbf{W}\,\mathbf{\Phi_X}\mathbf{v} = \mathbf{W}\,\mathbf{\Phi}_t$, the solution of which is given by:

$$\mathbf{v} = (\mathbf{\Phi_X^T}\,\mathbf{W}\,\mathbf{\Phi_X})^{-1}\mathbf{\Phi_X^T}\,\mathbf{W}\,\mathbf{\Phi}_t , \quad (4)$$

where $\mathbf{v}$ is the image velocity vector, $\mathbf{\Phi_X}$ and $\mathbf{\Phi}_t$ are the nx2 matrix and nx1 vector of spatial and temporal phase derivatives, and $\mathbf{W}$ is the diagonal weight matrix. Unreliable estimates of velocity may be identified using the eigenvalues of $\mathbf{\Phi_X^T}\,\mathbf{W}\,\mathbf{\Phi_X}$ [1]. Here we use the magnitude of the smallest eigenvalue as a measure of confidence in the associated velocity estimate.

IIR TEMPORAL FILTERS

To implement the above phase-based scheme we use space-time separable filters, allowing us to consider temporal filtering in isolation. We design the filters in the continuous-time domain and transform them to obtain a discrete-time transfer function. The class of filters we use are derived from the truncated exponential, a causal low-pass filter with a nearly linear phase spectrum [6]. We modulate with a complex sinusoid to obtain band-pass filters, and cascade to localise in the frequency domain. As the number of cascades increases, the resulting filter approaches a Gabor function, widely used in motion analysis. The impulse response after n cascades $A_n(t)$, is:

$$A_n(t) = \begin{cases} \frac{t^{n-1}b^n}{(n-1)!}\exp[-bt + j\omega_0 t] & , \quad t \geq 0 \\ 0 & , \quad t < 0 \end{cases} \quad (5)$$

where ω_0 is the temporal frequency tuning, and the temporal support of the filter and its bandwidth are determined by a combination of b^{-1} (the exponential time constant when $n = 1$) and n.

Digital Filter Design

To realise a discrete IIR implementation of the continuous filter in (5) we map the Laplace transform (the s-domain) representation of the filter onto the discrete-time z-domain (Bozic [2]). The Laplace transform of $A_n(t)$ is given by:

$$\mathcal{L}[A_n(t)] = \left[\frac{b}{s + b - j\omega_0}\right]^n , \quad (6)$$

and that of the derivative filter follows from the derivative theorem for Laplace transforms (Bracewell [3]): $\mathcal{L}[dA_n(t)/dt] = s\mathcal{L}[A_n(t)]$. Here we consider the discrete implementation of a 3rd-order filter and its derivative according to the bilinear transform. Although the impulse invariant method is somewhat simpler, it is susceptible to severe aliasing problems. The bilinear transform effectively warps frequency space to avoid aliasing [2], and it is then a simple matter to unwarp the phase derivatives in converting back to the analogue domain. The z-transforms of the band-pass filter *and* its derivative can both be expressed in the following form:

$$H_3(z) = q^3 \left(\frac{m_0 + m_1 z^{-1} + m_2 z^{-2} + m_3 z^{-3}}{1 + 3rz^{-1} + 3r^2 z^{-2} + r^3 z^{-3}}\right) , \quad (7)$$

where $q = b/(b - j\omega_0 + 2)$, $r = (b - j\omega_0 - 2)/(b - j\omega_0 + 2)$, and the values of the m_i are given in Table 1.

Filter	m_0	m_1	m_2	m_3
band-pass	1	3	3	1
derivative	2	2	-2	-2

Table 1: *Parameters for the discrete-time transfer function of the 3rd-order filter and its derivative according to (7).*

It should be noted that the transfer functions of the band-pass filter and its derivative share the same de-
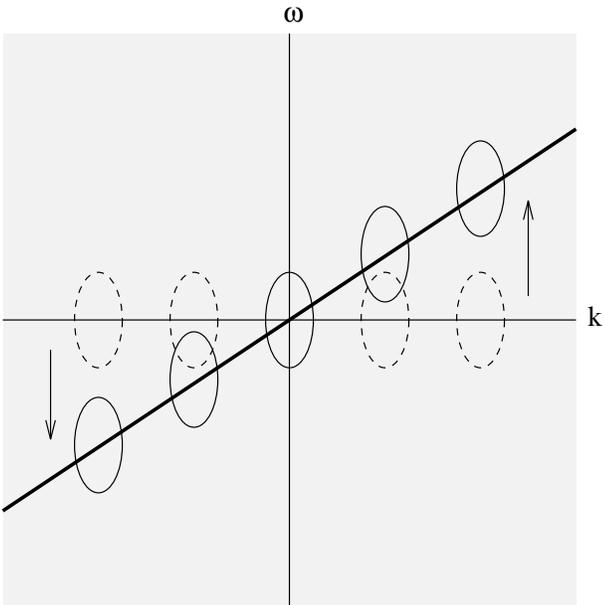
Figure 2: *1-D schematic diagram of a set of filters adapting their temporal tunings to lie on the line (in 2-D: plane) in frequency space corresponding to the velocity of the local image signal.*

nominator, allowing a further reduction in computational storage. Discrete implementations follow from the z-transform. We realise the common feedback path as the cascade of three first-order sections to ensure stability and to reduce coefficient sensitivity compared to the direct form (Tam et al [11], Nayeri and Jenkins [10]).

TEMPORAL ADAPTATION

The goal of adaptation as discussed here is to minimise the error in image velocity measured from a small number of spatiotemporal filters. Assuming the least-squares solution to image velocity given in (4), we define the state of a set of filter kernels $\mathbf{v}_f$ by their local velocity selectivity in a similar manner:

$$\mathbf{v}_f = (\mathbf{K}^T \, \mathbf{W} \, \mathbf{K})^{-1} \mathbf{K}^T \, \mathbf{W} \, \mathbf{\Omega} = \mathbf{C}\mathbf{\Omega} \, , \qquad (8)$$

where $\mathbf{K}, \mathbf{\Omega}$ refer to the spatial and temporal frequency tunings of the filter kernels. In this formulation the matrix $\mathbf{C}$ is a constant because only the temporal parameters are adapted.

The adaptive scheme is posed in terms of a single-flow model of local image translation. In frequency space, all the power in a translating image lies on a plane through the origin (1). Here, filters adapt their temporal centre-frequency tunings towards the plane of power in frequency space corresponding to the measured velocity of the local image signal, thereby tending to maximise the signal-to-noise ratio of their outputs (see Fig.2). The centre frequency tuning of the nth-

order cascaded kernel $A_n(t)$ (5) is simply the temporal modulation frequency, ω_0. Thus, to lie on the plane of power of the image signal, the centre-frequency temporal tuning of any filter tuned to a spatial frequency of $\mathbf{k}_0$ must satisfy:

$$\omega_0 \; = \; \mathbf{v} \cdot \mathbf{k}_0 \; . \qquad (9)$$

By adjusting the parameter ω_0, and hence the weights q and r in the IIR implementation (7), we can control the peak temporal frequency tuning of the filter. Here the value of ω_0 is under adaptive control in an effort to minimise the distance of the peak tuning of the filter from the plane of power of the local image signal.

Dynamics of Adaptation

At each time step, measured velocity is fed back to adapt the filters. The feedback mechanism that we have chosen to implement is a vector formulation of the LMS algorithm of Widrow and Stearns [12]. Treating the locally measured velocity as the desired response vector, the temporal state of the filter kernels is adapted on the basis of the difference equation:

$$\epsilon = \mathbf{v} - \mathbf{v_f} = \mathbf{v} - \mathbf{C}\mathbf{\Omega} \, , \qquad (10)$$

where ϵ is a vector error signal. Following [12], we take the instantaneous value of the squared error, $|\epsilon|^2$, and minimise it with respect to $\mathbf{\Omega}$ to arrive at a gradient descent update equation:

$$\mathbf{\Omega}_{t+1}^T = \mathbf{\Omega}_t^T - \frac{\eta}{2}\nabla_{\mathbf{\Omega}}|\epsilon_t|^2 \, , \qquad (11)$$

$$\mathbf{\Omega}_{t+1}^T = \mathbf{\Omega}_t^T + \eta\epsilon_t^T \mathbf{C} \, , \qquad (12)$$

where η controls the rate of adaptation. Without noise, when $\eta = 1.0$, the adaptive scheme convergences in one iteration. In the simulations described below, using noisy input images, η is set at 0.1 to ensure convergence.

EXPERIMENTAL RESULTS

We illustrate the performance of the adaptive scheme described above with results from two image sequences, one synthetic and one real. The synthetic diverging tree sequence [1, 6] has a known 2-D motion field, allowing a quantiative comparison of the performance of the adaptive and non-adaptive versions of the algorithm at various levels of additive white spatiotemporal noise. Results are presented graphically in the form of needle diagrams and confidence maps, and numerically as error statistics for the synthetic images.

Implementation

The spatial filters used in our implementation are complex band-pass Gabor filters and their derivatives. The filters are tuned to each of 6 spatial orientations, with centre-frequency spatial tunings of 0.2 cycles per pixel

and envelope standard deviations of 2.5 pixels. Each spatial filter feeds into a fixed low-pass temporal filter (as described in (5) but with $\omega_0 = 0$), and a pair of band-pass temporal filters of equal and opposite peak tuning. Under the adaptive scheme, the centre-frequency tunings of the band-pass temporal filters are adapted from an initial value of zero. In the non-adaptive scheme, used for comparison, the temporal tunings of the band-pass filters are set to 0.2 cycles per frame, equivalent to a velocity tuning of 1 pixel per frame in the preferred direction. The temporal filter time constant (5), b^{-1}, is set at 1.25 frames, giving an implicit delay of 3 frames in measurements of velocity. Following [6], the spatiotemporal window $W(\mathbf{x}, t)$ in (3) is a Gaussian in space (with a standard deviation of 1.2 pixels) and an exponential in time (with a time constant of 3.33 frames). Under the adaptive scheme η (12) is set at 0.1.

Error Measures

Following [1], we view velocity as spatiotemporal orientation and measure error as an angle in space-time. If velocity, $\mathbf{v} = (v_x, v_y)^T$, is represented as a unit vector, $\vec{\mathbf{v}}$, in space-time:

$$\vec{\mathbf{v}} = \frac{1}{\sqrt{v_x^2 + v_y^2 + 1}}(v_x, v_y, 1)^T , \qquad (13)$$

then the error, ψ_E, between the correct velocity, $\vec{\mathbf{v}}_c$, and an estimate, $\vec{\mathbf{v}}_e$, is given by:

$$\psi_E = \arccos(\vec{\mathbf{v}}_c \cdot \vec{\mathbf{v}}_e) . \qquad (14)$$

For results on synthetic image sequences we quote a mean error and a density. The error is calculated as the mean angular error of all points satisfying an arbitrary confidence threshold. The density gives the percentage of points satisfying that threshold. The confidence measure we use is the smallest eigenvalue of the spatial covariance matrix (3), which depends on the magnitudes of the spatial gradients and the range of their orientations [1].

Figure 3: **Left:** *Frame 20 of the diverging tree sequence.* **Right:** *Confidence map obtained under the adaptive scheme with no added noise.*
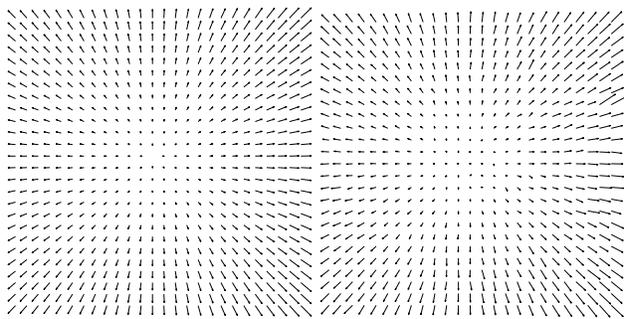
Figure 4: **Left:** *Correct motion field for frame 20 of the diverging tree sequence.* **Right:** *Measured motion field under the adaptive scheme with no added noise.*

Noise	Fixed filters		Adaptive filters	
(%)	error	density	error	density
0	2.04^o	43.6%	2.44^o	44.7%
5	2.33^o	41.1%	2.71^o	42.1%
10	3.19^o	39.1%	3.10^o	40.0%
15	4.96^o	37.9%	3.99^o	38.7%
20	7.71^o	37.4%	5.91^o	38.1%
25	11.31^o	37.6%	9.22^o	38.3%

Table 2: *Mean errors obtained from frame 37 of the diverging tree sequence at various levels of added noise using fixed and adaptive filters.*

Results

During the diverging tree sequence the camera moves along its line of sight, the focus of expansion being the centre of the image. Speeds vary from 1.4 pixels per frame on the left of the image to 2.0 on the right. Figure 3 (left) shows a frame from the sequence, and Figure 4 (left) the true flow field. The measured motion field under the adaptive scheme is shown in Figure 4 (right) to correspond closely to the true flow field. Encouragingly, the most noticable deviation from the veridical field occurs in a region where confidence estimates are low. Confidence estimates are shown in Figure 3 (right), with regions of high confidence being represented by bright points in the image.

To generate image noise we linearly scale the noiseless sequence, $I_0(\mathbf{x}, t)$, and then add the random variable, $n(\mathbf{x}, t)$, drawn from a distribution uniform over the range of values of the noiseless sequence. We create the noisy image, $I_n(\mathbf{x}, t)$, from the equation:

$$I_n(\mathbf{x}, t) = (1 - \alpha)I_0(\mathbf{x}, t) + \alpha n(\mathbf{x}, t) , \qquad (15)$$

where 100α is the percentage noise level. Numerical results from the fixed and adaptive schemes show that at noise levels exceeding 5% the adaptive scheme yields lower mean errors at a higher density of measurement (Table 2). The performance of the adaptive scheme ap-

proaches that of its non-adaptive counterpart at low-levels of noise, whilst showing an advantage which increases with noise level for 10% noise and above. We would expect to be able to improve the performance of the fixed scheme by incorporating additional temporal filters, but even under the proposed IIR implementation this would involve a considerable increase in computational storage.

The rotating Rubik cube sequence [1] is a real image sequence in which a Rubik's cube is rotating anti-clockwise on a turntable. Figure 5 shows a frame from the sequence and Figure 6 the flow field measured under the adaptive scheme after confidence thresholding.

Figure 5: *Frame from the rotating Rubik cube sequence.*
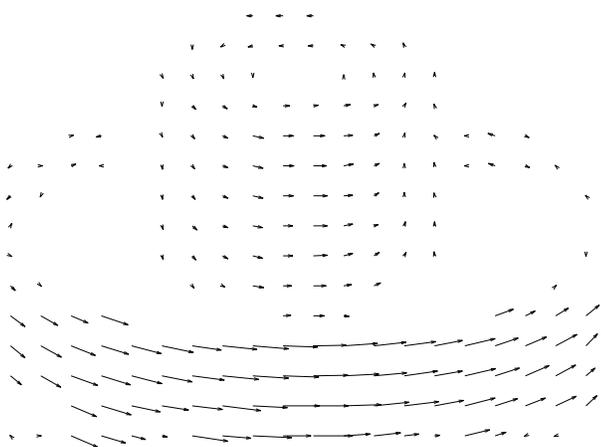
Figure 6: *Measured flow field from the rotating Rubik cube sequence, after thresholding, under the adaptive scheme.*

CONCLUSION

We have described an algorithm for computing image motion from the phase output of centre-frequency

adaptive IIR filters. The algorithm may find application in biomedical imaging, for example in profusion or cardiac studies where image noise levels are potentially high. The main limitation of the technique is that some error is necessarily introduced into the velocity estimates by the process of adaptation itself, although for noisy image sequences this factor is outweighed by the improved signal-to-noise response of the adaptive filters. Indeed, as hardware developments allow faster frame-rates for real-time implementation of motion analysis algorithms so the proportion of time during which significant adaptation occurs will decrease, and the benefit of superior image signal representation will outweigh the cost of adaptation even at low levels of noise.

REFERENCES

[1] Barron.J.L, Fleet.D.J. and Beauchemin.S.S., 1994, "Performance of optical flow techniques", *Int. J. Comp. Vision*, **12**, 43–77.

[2] Bozic.S.M., 1986, "Digital and Kalman filtering", Arnold.

[3] Bracewell.R.N., 1986, "The Fourier transform and its applications", McGraw-Hill, London.

[4] Fleet.D.J. and Jepson.A.D., 1990, "Computation of component velocity from local phase information", *Int. J. Comp. Vision*, **i5**, 77–101.

[5] Fleet.D.J. and Jepson.A.D., 1993, "Stability of phase information", *IEEE PAMI*, **15**, 1253–1268.

[6] Fleet.D.J. and Langley.K., 1995, "Recursive filters for optical flow", *IEEE PAMI*, **17**, 61–67.

[7] Franks.L., 1969, "Signal analysis", Prentice-Hall, NJ.

[8] Kumar.R.V.R. and Pal.R.N., 1985, "A gradient algorithm for centre-frequency adaptive recursive bandpass filters", *Proc. IEEE*, **73**, 371–372.

[9] Langley.K. and Fleet.D.J., 1992, "Recursive filters for phase-based optic flow", *Proc. 9th Israeli Symp. on AI and Comp. Vision*, 255–264.

[10] Nayeri.M. and Jenkins.W.K., 1989, "Alternate realizations to adaptive IIR filters and properties of their performance surfaces", *IEEE Trans. Circuits Syst.*, **36**, 485–496.

[11] Tam.Y.H., Ching.P.C., and Chan.Y.T., 1987, "Adaptive recursive filters in cascade form", *IEE Proc. F*, **134**, 245–252.

[12] Widrow.B. and Stearns.S.D., 1985, "Adaptive signal processing", Prentice-Hall, 1985.