

STABLE ESTIMATION OF IMAGE ORIENTATION

Leif Haglund

Department of Electrical Engineering
Linköping University
S-581 83 Linköping
Sweden
leif@isy.liu.se

David J. Fleet

Department of Computing Science
Queen's University
Kingston, Ontario
Canada, K7L 3N6
fleet@qucis.queensu.ca

ABSTRACT

We examine the performance of phase-based and energy-based techniques for estimating image orientation, emphasizing measurement stability under deformations that commonly occur in natural images. We also present a new energy-based technique for estimating multiple orientations in an neighbourhood.

1. INTRODUCTION

Image orientation is central to many tasks in machine vision and image processing, including the extraction of edges, lines and curves [5, 9, 11, 16], and image enhancement [12, 14]. It is also central to the measurement of normal velocity constraints for optical flow (e.g. [2]) and stereo disparity estimation (e.g. [15]). More recent uses of image orientation concern the determination of surface slant from texture [10, 18], from binocular viewing [13, 20], or from a moving camera.

Many methods for orientation estimation implicitly assume a single orientation in each local neighbourhood. But many of the tasks above exploit multiple orientation estimates. The accuracy of orientation measurement is also a primary concern – e.g., Jones and Malik [13] showed that under typical stereo viewing geometries, orientation differences between binocular views of slanted 3d surfaces are usually less than 5 degrees. Orientation measurements must be accurate and stable with respect to deformations between left and right views to be useful in this context.

We examine the accuracy with which one can expect to measure image orientation using energy-based and phase-based methods, and propose a new energy-based method for reliably extracting multiple orientations from textured image regions. Rather than evaluate the performance of such methods with only synthetic inputs with known orientations, such as lines and edges, we concentrate on the stability of measurements under deformations of images like those that typically occur between different views of 3d scenes. This is clearly an important criterion if one wants the measurements to depend primarily on the image struc-

ture, independent of small changes of viewing geometry and the filters used.

Toward this end we develop a framework based on steerable (adaptive) filters in which one can steer orientation-tuned quadrature-pair filter responses, their amplitude and phase, and their amplitude and phase derivatives with respect to orientation and spatial position. In this way one can extract rich descriptions of local image structure appropriate for energy and phase-based methods.

Interestingly, our initial results based on instantaneous frequency (following [7]) proved somewhat unstable relative to the accuracy required for slant estimation, although multiple local orientations were extracted. Estimates from a simple energy-based approach (following [9, 14]) also proved inappropriate as the density of estimates was low since just one orientation was extracted in each neighbourhood.

We have found a significant improvement with a new method that extracts multiple local orientations based on maxima of energy as a function of orientation. Central to the method is the set of steerable filters mentioned above, the extraction of local amplitude maxima via zero-crossings in the amplitude derivative, a correspondence technique that matches orientations in left and right views, and stability measures that identify unreliable estimates.

We show that this method produces superior results compared to other methods in terms of measurement accuracy, stability with respect to noise and deformation, and the number of independent measurements, often with two or three orientations in a single neighbourhood. The confidence measures are particularly important to this performance, as they are in most methods, e.g. for measuring optical flow [2].

2. COMPUTATIONAL FRAMEWORK

We first outline a class of steerable filters that are suitable for both energy and phase-based methods.

2.1. Adaptive Quadrature Filters

Following Andersson [1], we consider polar-separable amplitude spectra of the form

$$Q(\omega, \phi; \psi) = H(\omega)S(\phi; \psi) \quad (1)$$

where (ω, ϕ) denotes polar coordinates in frequency space, $H(\omega)$ is a bandpass function of radial frequency, and $S(\phi; \psi)$

This paper was presented at the IEEE International Conference on Image Processing, Austin, November 1994, Vol. III, pp. 68-72. The work was financially supported in part by ITRC (Ontario Centre of Excellence) and NSERC Canada. Leif Haglund was also supported by TFR, Sweden.

denotes an angular spectrum, as a function of ϕ , that is tuned (rotated) to ψ . The ability to interpolate such filters over a continuous range of orientations follows from the trigonometric identity

$$\cos(n(\phi - \psi)) = \cos(n\phi) \cos(n\psi) + \sin(n\phi) \sin(n\psi)$$

That is, for any n one can compute $\cos(n(\phi - \psi))$, at any ψ , using a weighted sum of two basis functions, namely, $\cos(n\phi)$ and $\sin(n\phi)$. It is therefore convenient to express the angular spectra of orientation-tuned filters as weighted sums of cosinusoidal functions, i.e.,

$$Q(\omega, \phi; \psi) = H(\omega) \sum_{n=0}^N a_n \cos(n(\phi - \psi)) \quad (2)$$

The basis set contains $2N + 1$ real-valued functions of the form $H(\omega) \cos(n\phi)$ and $H(\omega) \sin(n\phi)$, that are weighted with the coefficients a_n .

To design quadrature-pair filters, the weights a_n are chosen to minimize the energy in the spectrum (2) in the rear half-plane, i.e., for $|\phi - \psi| > 0.5\pi$ [1]. That is, we minimize

$$\int_0^{2\pi} W(\phi) |Q(\omega, \phi; \psi)|^2 d\phi = \int_0^{2\pi} W(\phi) \left| \sum_{n=0}^N a_n \cos(n(\phi)) \right|^2 d\phi \quad (3)$$

where the weighting function $W(\phi)$ is

$$W(\phi) = \begin{cases} 1 & \text{if } |\phi - \psi| > \pi/2 \\ 0 & \text{if } |\phi - \psi| \leq \pi/2 \end{cases} \quad (4)$$

Other weighting functions are clearly possible. For example, in [1] a smooth weighting function of the form $W(\phi) = \sin^{2r}(0.5(\phi))$ is examined from which the resulting filter closely approximates a quadrature filter, and by changing r one can obtain different angular bandwidths. The cost of narrower angular bandwidth is a small amount of ringing in the rear half plane. As the number of basis filters increases, the energy in the rear half-plane and the orientation bandwidth of the filters both decrease; e.g., with $N = 6$ the angular bandwidth at half height can be $50^\circ - 65^\circ$, depending on weighting function. For the band-pass function $H(\omega)$ we currently use lognormal functions (e.g. [14]).

$$H(\omega) = \exp\{-4 \ln^2(\omega/\omega_0)/(B^2 \ln(2))\} \quad (5)$$

In the experiments we set the centre frequency to $\omega_0 = \pi/(2\sqrt{2})$, where the Nyquist frequency is π and the relative bandwidth $B = \sqrt{2}$ octaves. We then take an inverse 2D-DFT of $H(\omega) \cos(n\phi)$ and $H(\omega) \sin(n\phi)$ to obtain the discrete basis kernels $b_n(\mathbf{x})$, where $\mathbf{x} = (x, y)$. The orientation tuned filters $q(\mathbf{x}; \psi)$ are then constructed as weighted sums of the $b_n(\mathbf{x})$ according to (2).

The oriented filters $q(\mathbf{x}; \psi)$ and their responses $R(\mathbf{x}; \psi)$ are complex-valued. The real part is constructed from even-order basis functions (i.e. $n = 0, 2, 4, 6$) and the imaginary part from odd-order basis functions (i.e. $n = 1, 3, 5$). The response can also be expressed in terms of amplitude $\rho(\mathbf{x}; \psi)$ and phase $\theta(\mathbf{x}; \psi)$, as in

$$R(\mathbf{x}; \psi) = \rho(\mathbf{x}; \psi) e^{i\theta(\mathbf{x}; \psi)} \quad (6)$$

where $\rho(\mathbf{x}; \psi) = |R(\mathbf{x}; \psi)|$, $\theta(\mathbf{x}; \psi) = \arg[R(\mathbf{x}; \psi)]$, and $i^2 = -1$. The final concept we need is instantaneous frequency, defined as the derivative of phase with respect to spatial position, $\nabla \phi(\mathbf{x})$, where ∇ is the usual gradient operator. Instantaneous frequency has played a central role in phase-based methods for computing image velocity [7] and binocular disparity [8, 17, 19].

2.2. Steerable Derivatives

This steerable set allows one to interpolate the complex-valued response $R(\mathbf{x}; \psi)$, and its amplitude and phase with ease. Unlike previous uses of steerable filters, we also want to interpolate amplitude and phase derivatives with respect to spatial position and orientation. The amplitude derivative with respect to orientation will be used for localizing energy maxima in orientation, and as a stability measure. The phase derivative in orientation could be useful as a stability measure [1]. Derivatives of phase and amplitude with respect to spatial position are needed for phase-based methods, and for stability measures as discussed below.

The mathematical derivation of the derivatives follows from (6)

$$\frac{\partial R(\mathbf{x}; \psi)}{\partial s} = [\rho_s(\mathbf{x}; \psi) + i\rho(\mathbf{x}; \psi)\theta_s(\mathbf{x}; \psi)] e^{i\theta(\mathbf{x}; \psi)} \quad (7)$$

where s denotes one of x , y or ψ . As a consequence [4, 7]

$$\begin{aligned} \frac{\partial \theta(\mathbf{X}; \psi)}{\partial s} &= \frac{\text{Im}[R_s(\mathbf{X}; \psi)R^*(\mathbf{X}; \psi)]}{|R(\mathbf{X}; \psi)|^2} \\ \frac{\partial \rho(\mathbf{X}; \psi)}{\partial s} &= \frac{\text{Re}[R_s(\mathbf{X}; \psi)R^*(\mathbf{X}; \psi)]}{|R(\mathbf{X}; \psi)|} \end{aligned} \quad (8)$$

where the superscript $*$ denotes complex conjugate.

Using these identities, amplitude and phase derivatives with respect to orientation are easily computed (steered) using the original basis filters $b_n(\mathbf{x})$. In particular, from (2) it follows that

$$\frac{\partial Q(\omega, \phi; \psi)}{\partial \psi} = H(\omega) \sum n a_n \sin n(\phi - \psi) \quad (9)$$

Then by using the trigonometric identity

$$\sin(n(\phi - \psi)) = \sin(n\phi) \cos(n\psi) - \cos(n\phi) \sin(n\psi)$$

and (8) one can easily steer phase and amplitude derivatives with respect to orientation using the responses of $b_n(\mathbf{x})$, already available. That is, the derivative w.r.t the tuning angle of the quadrature response is given by another weighting of the responses of the same basis filters.

We also require derivatives of amplitude and phase with respect to spatial position. Although not as simple as orientation derivatives, they are easily obtained, e.g.

$$\frac{\partial}{\partial x} q(x, y; \psi) = \sum a_n \frac{\partial}{\partial x} b_n(x, y) \quad (10)$$

or equivalently, in the Fourier domain, with $\omega_x = \omega \cos(\phi)$

$$\omega_x Q(\omega, \phi; \psi) = \sum a_n \omega_x H(\omega) \cos n(\phi - \psi) \quad (11)$$

For efficiency we numerically differentiate the basis filter outputs instead of using a second set of 2d filters applied to

the original image¹. This procedure gives us full steerability for the quadrature response and its derivatives. It is now easy to form an estimate of the instantaneous frequency, $\nabla \theta$ according to (8). Note that we do not have to steer responses at neighbouring pixels to the same filter tuning to differentiate the steered filter output.

3. EXPERIMENTAL METHODS

We outline our experimental method before discussing the actual estimation techniques we have employed. In the evaluation of the estimation techniques we are interested in stability with respect to deformations that typically occur between binocular views, or between frames in an image sequence. The procedure we have been using involves

1. Create an image pair by applying a known deformation (warping) to a natural image. Cubic interpolation has been used for the warping. The deformations have been rotations ($2^\circ - 10^\circ$), anisotropic dilations ($2\% - 10\%$) and shears ($2\% - 10\%$). For example, the following deformation was used for anisotropic dilation.

$$\begin{pmatrix} 1+a & 0 \\ 0 & 1 \end{pmatrix}$$

where $a \in [0.02 \ 0.04 \ 0.06 \ 0.08 \ 0.10]$.

2. Apply the algorithm on both the original and the deformed image.
3. Calculate error statistics for the estimates, utilizing confidence measures, while compensating for the known deformation. In the results the mean orientation difference, the standard deviation of the difference, and the coverage are reported.

4. STABLE ESTIMATION OF ORIENTATION

We have tested the stability of orientation estimation for phase-based and energy-based techniques.

4.1. Phase-Based Technique

As mentioned above, although phase-based orientation measurements from band-pass filter outputs produce multiple local orientations and are reasonably stable under certain lighting and geometric deformations of the input [6, 8], they proved too unstable (with fixed filters) for computing orientational disparities. For example, Table 1 shows the mean error and standard deviation of orientation differences, as functions of rotations of the image shown in Figure 1, with a single orientation-tuned filter. The orientation estimates in this case are given by the orientation of the instantaneous frequency, $\nabla \phi(\mathbf{x})$.

4.2. Energy-Based Techniques

Energy-based methods are another technique to measure image orientation, which typically rely on the fact that a

¹The computational cost for differentiating the filter outputs, with a 8 tap filter, is $2 \cdot 8 \cdot 13 = 208$ operations per pixel, which should be compared to $15 \cdot 15 \cdot 13 \cdot 2 = 5850$ operations per pixel if the 2d filters are assumed to be 15x15 pixels in size

rotation	mean	Std dev
2°	1.21°	1.52°
4°	2.13°	3.53°
6°	3.92°	4.34°
8°	5.54°	5.61°

rotation	mean	Std dev
2°	1.21°	1.52°
4°	2.13°	3.53°
6°	3.92°	4.34°
8°	5.54°	5.61°

Table 1: *Stability of phase-based orientation from a single filter for rotations of the input. Columns show mean and standard deviations for errors in orientation differences.*



Figure 1: *The used test image.*

perfectly oriented signal will have its Fourier power concentrated on a line through the origin. By estimating the orientation of this line the image orientation is also obtained.

Estimation of a Dominant Orientation

Simple energy-based methods [14, 9] assume a single orientation in each neighbourhood. In essence these methods estimate the RMS orientation of the angular energy distribution. These techniques will yield stable results under local rotations, but not necessarily under shear or dilation if more than one local orientation is present. For these reasons we have used a stability measure based on the ratio of the *oriented energy* to the total local energy, to detect neighbourhoods with multiple local orientations, (e.g.[12]). Tables 2-4 show the results. We have also tested a gradient-based scheme [3] and found the same type of behaviour and similar estimation errors. Although accuracy is generally good, there remain two drawbacks of these techniques:

- the neighbourhoods with the most information (multiple orientations) are explicitly removed. For instance, note the coverage of estimates in Tables 2-4.
- if multiple orientations are too close, the confidence measure is weakened.

Estimation of Multiple Orientations

It is therefore natural to try to extract multiple local orientations based on energy maxima of the steerable filters [1, 13, 16]. The issues that arise include how to find the local maxima, how to do the matching between two views, and the development of stability measures to identify outliers. The method outlined below significantly increases the

rotation	coverage	mean	Std dev
2°	40%	-0.005°	0.35°
4°	40%	-0.015°	0.59°
6°	39%	-0.026°	0.80°
8°	39%	-0.033°	1.02°
10°	39%	-0.044°	1.16°

Table 2: *Stability of energy-based estimates with rotations of the input (using a confidence measure).*

dilation	coverage	mean	Std dev
2%	38%	-0.005°	0.68°
4%	38%	0.003°	0.88°
6%	38%	0.010°	1.15°
8%	37%	0.036°	1.44°
10%	37%	0.041°	1.67°

Table 3: *Stability of the energy-based orientation estimation algorithm w.r.t. anisotropic dilation.*

shear	coverage	mean	Std dev
2%	38%	-0.014°	0.49°
4%	38%	-0.002°	0.62°
6%	37%	-0.002°	0.78°
8%	37%	0.006°	0.95°
10%	37%	0.011°	1.16°

Table 4: *Stability of the energy-based orientation estimation algorithm w.r.t. shear of the input image.*

number of independent local measurements and their stability compared to the single orientation techniques.

Our current technique uses 6th-order quadrature-pair filters steered to 18 orientations at which the response R , its derivative R_ψ , and the amplitude derivative ρ_ψ are computed using (2), (9), and (8).² Energy maxima are detected from zero-crossings of ρ_ψ (using linear interpolation of ρ_ψ when changes in its sign are found between the 18 sampling points).

In order to test the stability of the measurements with respect to deformations between two views we also need a correspondence algorithm to associate the orientations in the two views. At present we only use an ordering constraint, remembering that orientation estimates are modulo π (see Fig. 2), and a simple cost function according to (12):

$$\epsilon_{AB} = \sum_i |\sin(\phi_{Ai} - \phi_{Bi})| \quad (12)$$

The match is chosen to minimize this measure.

Confidence Measures

Finally, stability constraints (confidence measures) are used to detect unreliable measurements, and therefore reduce the number of outliers. In the current implementation four different confidence measures are used.

²With $N = 6$, a minimum sampling of 13 orientations is required by the sampling theorem to encode the response as a function of orientation. We oversample slightly to permit simpler interpolation.

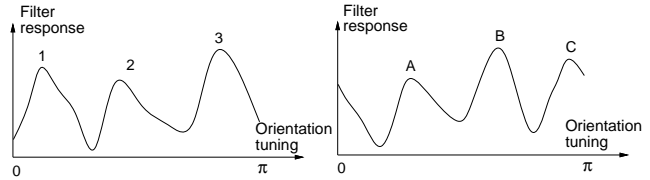


Figure 2: *Two estimation curves that should be matched, with an ordering constraint three matches are possible, (1-A, 2-B, 3-C); (1-C, 2-A, 3-B) or (1-B, 2-C, 3-A).*

SNR-threshold: For each of the two views the amplitude of the filter output has to exceed 5% of the maximum value over the whole image. The objective is to remove estimates where the signal-to-noise levels are problematic.

Sharpness of orientation maxima: From the filter design we know the shape of the quadrature response for an ideal single orientation, and especially we know how sharp the curve should be at the maxima. To avoid estimates where two or more orientations are within the angular bandwidth of the filter we calculate the quotient $\rho_{\psi\psi}/\rho$. This measure must be larger than 0.6 for the estimate to be considered.

IF versus Energy: The filtering framework allows us to obtain the image orientation from both an energy-based technique (allowing multiple orientations) and from the instantaneous frequency from a quadrature-pair filter steered to the same orientation. The orientation difference between the estimates from the two input images, taken separately for the techniques, is compared and if they behave differently this is an indication that the orientation estimate is not stable. This effectively removes many outliers in the procedure. This difference is also related to singularity neighbourhoods [6] where the quadrature response is particularly unstable.

Phase difference: The final confidence measure is based on the difference in phase between the quadrature-pair responses steered to the orientation maxima in each of the input images. A large phase difference between the responses is a strong indication of misalignment, since this is the information that disparity estimation from phase differences relies on, [8, 15, 17, 19].

It is worth noting that the absolute levels of these thresholds are not critical to the results of the algorithm.

Tables 5-7 show the errors in orientation differences based on the two maxima with the highest energy at each pixel. Note the increased coverage of measurements, and the high density of multiple measurements (indicated by the coverage of the secondary maxima), compared to previous results in Tables 2-4. The smaller errors for the first peak are especially encouraging, and as these are difference errors, the effective accuracy of the individual orientation estimates should be significantly smaller (assuming IID errors).

rotation	coverage	mean	Std dev
2°	78%	0.028°	0.46°
4°	75%	0.026°	0.54°
6°	73%	0.000°	0.57°
8°	71%	0.008°	0.59°
10°	70%	0.012°	0.60°

rotation	coverage	mean	Std dev
2°	48%	0.078°	0.69°
4°	42%	0.062°	0.81°
6°	40%	-0.003°	0.84°
8°	39%	-0.043°	0.82°
10°	38%	0.036°	0.82°

Table 5: *Stability of energy-based estimates w.r.t rotation. The upper table is for the maxima with highest energy, and the lower is the second highest.*

dilation	coverage	mean	Std dev
2%	68%	-0.004°	0.56°
4%	68%	-0.014°	0.74°
6%	66%	-0.016°	0.98°
8%	63%	0.004°	1.16°
10%	59%	0.017°	1.36°

dilation	coverage	mean	Std dev
2%	35%	0.021°	0.74°
4%	33%	0.060°	1.03°
6%	30%	0.099°	1.37°
8%	28%	0.150°	1.66°
10%	25%	0.185°	1.97°

Table 6: *Stability of the energy-based estimates w.r.t anisotropic dilation. The upper table is for the peak with the highest energy and the lower is the second highest peak.*

5. SUMMARY

This paper examines the performance of phase-based and energy-based techniques for estimating image orientation. The stability of these estimates under deformations that commonly occur with projections of 3d scenes. Jones and Malik [13] and Perona [16] have suggested that orientation measurements can be attained with an accuracy of about 1°, suggesting orientation difference errors about $\sqrt{2}$ larger. With a new energy-based framework for extracting orientations, we find that even better results are possible³, but these depend critically on suitable stability constraints.

Our next goal is to use the orientation disparity measurements for surface slant/tilt estimation, where we have promising, but still very preliminary results.

6. REFERENCES

- [1] M. T. Andersson. *Controllable Multidimensional Filters in Low Level Computer Vision*. PhD thesis, Linköping University, Sweden, 1992.

³With synthetic inputs, this method shows much better results with standard deviations of about 0.2°.

shear	coverage	mean	Std dev
2%	68%	-0.027°	0.42°
4%	67%	-0.029°	0.52°
6%	66%	-0.041°	0.67°
8%	65%	-0.050°	0.83°
10%	63%	-0.059°	0.99°

shear	coverage	mean	Std dev
2%	35%	-0.021°	0.60°
4%	34%	-0.033°	0.81°
6%	32%	-0.028°	1.07°
8%	31%	-0.031°	1.36°
10%	29%	0.001°	1.62°

Table 7: *Stability of the energy-based estimates w.r.t shear. The upper table is for the peak with the highest energy and the lower is the second highest peak.*

- [2] J. Barron, D. Fleet, and S. Beauchemin. Performance of optical flow techniques. *IJCV*, 12(1):43–77, 1994.
- [3] J. Bigün. *Local Symmetry Features in Image Processing*. PhD thesis, Linköping University, Sweden, 1988.
- [4] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal. *Proc. IEEE*, 80(4), 1992.
- [5] J. Canny. A computational approach to edge detection. *IEEE-PAMI*, 8(6):255–274, 1986.
- [6] D. J. Fleet and A. D. Jepson. Stability of phase information. *IEEE-PAMI*, 15(12):1253–1268, 1994.
- [7] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *IJCV*, 5(1):77–104, 1990.
- [8] D.J. Fleet, A. D. Jepson, and M. Jenkin. Phase-based disparity measurement. *CVGIP Im. Understanding* 53(2):198–210, 1991.
- [9] W. T. Freeman and E. H. Adelson. The design and use of steerable filters. *IEEE-PAMI*, 13(9):891–906, 1991.
- [10] J. Gårding. *Shape from surface markings*. PhD thesis, Royal Institute of Technology, Stockholm, May 1991.
- [11] G. H. Granlund. In search of a general picture processing operator. *CGIP*, 8(2):155–178, 1978.
- [12] L. Haglund. *Adaptive Multidimensional Filtering*. PhD thesis, Linköping University, Sweden, 1992.
- [13] D. Jones and J. Malik. Determining three-dimensional shape from orientation and spatial frequency disparities I – using corresponding line elements. TR UCB-CSD 91-656, Univ. California, Berkeley, 1991.
- [14] H. Knutsson. *Filtering and Reconstruction in Image Processing*. PhD thesis, Linköping Univ., Sweden, 1982.
- [15] K. Langley, T. J. Atherton, R.G. Wilson, and M. Lacombe. Vertical and horizontal disparities from phase. In *Proc. ECCV*, pp. 315–325, 1990.
- [16] P. Perona. Steerable-scalable kernels for edge detection and junction analysis. In *Proc. ECCV*, pp. 3–18, 1992.
- [17] T. D. Sanger. Stereo disparity computation using gabor filters. *Biological Cybernetics*, 59:405–418, 1988.

- [18] B. J. Super and A. C. Bovik. Shape-from-texture by wavelet-based measurement of local spectral moments. In *Proc. CVPR*, pp. 296–301, 1992.
- [19] C.-J. Westelius. Preattentive Gaze Control for Robot Vision. Lic-thesis, Linköping University, Sweden, 1992.
- [20] R. P. Wildes. Direct recovery of three-dimensional scene geometry from binocular stereo disparity. *IEEE-PAMI*, 3(8):761–774, 1991.