MULTISCALE ORIENTATION ESTIMATION OF PERCEPTUAL BOUNDARIES

Jeffrey Ng and Anil A. Bharath

Department of Bioengineering, Imperial College London
United Kingdom SW7 2AZ
{jeffrey.ng,a.bharath}@imperial.ac.uk

ABSTRACT

The dominant orientation at any point P in an image is the direction from P in which there is the least gray-level variance. It is often defined by using gradient estimates, but may be extended to employ neighbourhood operators that provide some degree of phase invariance. However, the approach to estimating the dominant orientation at P depends on the scale or size of the (usually) non-trivial neighbourhood being considered. Multiscale PCA-based orientation estimation techniques involve computationally-heavy solving of eigensystems at each scale and location. In contrast, we propose two methods which use a measure of anisotropy to select or weight orientations respectively at different scales in order to provide a single estimate of orientation at any given point. This is believed to be closer to human perception of contour direction. In this paper, results are presented for two simple orientation estimation techniques against comparisons with multiscale PCA estimation for human perceptual boundaries.

1. INTRODUCTION

Estimates of local orientation in images play a very important role in many computer vision tasks such as feature detection, image denoising, contour extraction, image segmentation and even high-level tasks such as object recognition. Local orientation may be estimated from the vector fields obtained by applying differential operators, edge detectors [1] and steerable filters [2] to the luminance channel of a colour image.

Orientation estimation is particularly important for the extraction of contour or region boundaries for image segmentation. Although in signal processing, the dominant orientation is usually dependent on the size or scale of a given neighbourhood, human subjects perceive contour boundaries and their orientations at the appropriate scale for segmentation. This indicates that a scale selection or a scale weighting process is involved in the human perception of orientation, especially for boundaries. Perona [3] applied anisotropic diffusion to orientation maps. Feng and Milanfar [4] used PCA to obtain the dominant orientation at each scale and weighted these with an eigenvalue-based measure of orientation dominance. This yields a single orientation estimate for each location. Solving eigensystems of gradient neighbourhoods for each location and scale adds a computational overhead and the particular measure of orientation dominance is limited by the evaluation of gradient energy in only two orthogonal orientations, i.e. the method assumes a maximum of two orthogonal structures at each location.

In real imagery, local neighbourhoods are not restricted to edges or ridges but may also contain composite structures such as corners and junctions. Aach et al. [5] extended the PCA-based single-scale orientation estimation techniques by decomposing the eigenvectors of an extended gradient-based tensor into the orientations of the component structures, allowing the estimation of multiple orientations at each location. We adopt a different approach by increasing the orientation-selectivity of our filters such that the interference from structures at different orientations is minimised, thereby avoiding the need to solve eigensystems at each scale.

In Section 2, we describe a steerable pyramid for decomposing an image into scale and orientation channels. In Section 3, we propose the novel use of an anisotropy measure for two techniques to select or weight orientations at different scales. An adaptive angle-wrapping step that enables accurate weighted averaging of orientations is given towards the end of this section. We provide experimental results of our two different orientation estimation techniques in Section 4 and compare them to Multiscale Principal Component Analysis technique for the orientation estimation technique of Feng and Milanfar [4]. We conclude with a summary in Section 5.

2. ORIENTATION ESTIMATION FROM A STEERABLE PYRAMID

Most orientation estimation techniques are based on the analysis of gray-level intensity gradients, which can be performed

This work was partly funded by the UK Research Council Grant GR/R87642/02.

by differential operators or spatial filtering. Differential operators and orthogonal spatial filters typically estimate gradients in only two orientations. Filter banks employing filters at more than two orientations possess higher orientation selectivity and can better resolve the orientation of structures with reduced interference from structures at other orientations. The orientation selectivity and scale selectivity characteristics of Gabor filters cannot be as easily tuned as other Fourier-domain polar separable filters. We use the complex steerable pyramid of Bharath and Ng [6] to obtain energy responses at multiple orientations and scales that are quasi-invariant to the phase of image structures, i.e. edges, ridges and anything in-between.

The filters are specified in the Fourier domain by polar separable functions, $G_\alpha^\theta(\omega,\phi)$, which allow the angular (orientation) selectivity, $G^\theta(\phi)$, to be tuned independently of the radial (scale) selectivity, $G_\alpha(\omega)$, where α and θ are the scale and orientation of the desired filter respectively

$$G_\alpha^\theta(\omega,\phi) = G_\alpha(\omega)G^\theta(\phi) \tag{1}$$

The angular selectivity of the filters is defined with third power cosine functions that provide a flat power (sum of squares) response across all orientations and higher selectivity compared to $x-y$ orientated differential operators and spatial filters

$$G^\theta(\phi) = cos^3(\phi-\theta)rect_\pi(\phi-\theta) \tag{2}$$

where $rect_\pi(\phi)$ is the unit rectangular function

$$rect_\pi(\phi) = \begin{cases} 1 & \text{, if } |\phi| \leq \frac{\pi}{2} \\ 0 & \text{, otherwise} \end{cases} \tag{3}$$

The radial selectivity of the filters control the behaviour, particularly the stability, of their responses across scales and an appropriate function is important for comparing filter responses across scales. We use Erlang functions [7] of order $n=7$ and scale $\alpha=0.5$

$$G_\alpha(\omega) = \left(\frac{\alpha e}{n}\right)^n \omega^n e^{-\alpha\omega} \tag{4}$$

The filter kernels in the spatial domain, $g_\alpha^\theta(\omega,\phi)$, are obtained by the inverse discrete Fourier transform of $G_\alpha^\theta(\omega,\phi)$. We use the same Butterworth lowpass filters as Bharath and Ng [6] prior to decimating the image by half in each axis and using the result as input to the next level of the pyramid.

3. ANISOTROPY AND MULTISCALE ORIENTATION DOMINANCE

The orientation estimates obtained from differential operators and spatial filters are dependent on the scale of the chosen neighbourhood. A scale selection process is required to obtain the orientation estimate that is adapted to the *a priori* unknown scale of a structure at a particular location. In order to detect the "intrinsic" scale of structures, Lindeberg [8] searched for extrema in normalised filter responses across scales. Such a criteria based on absolute filter responses is unsuitable for orientation estimation as it does not consider the dominance of the estimated orientation with respect to the other orientations. Feng and Milanfar [4] applied PCA to the components of the gradient field in a local neighbourhood of an image and used the difference in the two obtained eigenvalues, normalised by the sum of the eigenvalues, as a dimensionless measure of the dominance of the estimated orientation at each scale. The eigenvalues provide a measure of the signal energy along and perpendicular to the estimated orientation. The technique operates on the assumption that a maximum of two orientated structures may simultaneously exist in a local neighbourhood and that the two structures are perpendicular to each other, which may not always be the case.

In neuroscience, the orientation selective neurons of mammalian primary visual cortex may be characterised by measuring the output of the neuron during multiple presentations of oriented stimuli at many orientations. Ringach [9] proposed a measure of orientation dominance that weights vectors in each of the presented orientation by the neuronal response, performs a vector addition to find the overall preferred orientation of the cell and normalises the result by the sum of all the responses so that the result is independent of the overall sensitivity of the neuron and always lie between zero and one. Given the energy response $|f_k^{(\ell)}(m,n)|, \quad k=0..K-1$ of the K complex filters at location (m,n) and level ℓ of the complex steerable pyramid, we use the orientation dominance measure

$$\mathbf{O}^{(\ell)}(m,n) = \left| \frac{\sum_{k=0}^{K-1} |f_k^{(\ell)}(m,n)| e^{j2\phi_k}}{p + \left(\sum_{k=0}^{K-1} |f_k^{(\ell)}(m,n)|^2\right)^{\frac{1}{2}}} \right| \tag{5}$$

where we added the conditioning constant, $p = 1.25\%$ of the maximum gray-level intensity value in the image.

The multiscale estimated orientation at a particular image location can be chosen as the estimated orientation at the scale with the highest orientation dominance. However, the number of scales or levels in the pyramid is usually limited by computational power and the sampling of the scale parameter is usually coarse. A better multiscale estimation technique may be to weight the estimated orientations at each scale with their respective orientation dominance. However, the weighting of orientations require special care in the pre-processing of the orientations in addition to angle-wrapping.

Orientation is usually represented by either a vector v or an angle θ to a reference axis such as the horizontal. Both representations do not capture the property of orientation to flow in either direction to the vector, i.e. v and $-v$, or to rotations of the angle by π, i.e. $\theta + n\pi$ where $n \in \mathbb{Z}$. Following the example in Fig. 1, given two orientations θ and β which have been wrapped to the range $[-\pi, \pi]$, computing an average or a vector addition would yield an incorrect result, ω, that is close to perpendicular to the desired result ϕ. Rather than angle β, angle α should be averaged with angle θ. Moreover, averaging orientation or adding vectors that are more than $\frac{\pi}{2}$ radians apart may yield an incorrect result because of the discontinuity at either extremes of the range which are supposed to wrap around.

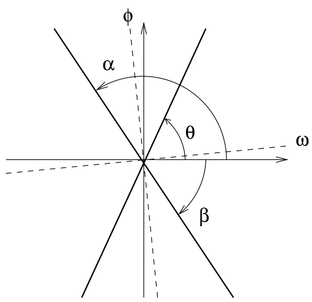


Fig. 1. Averaging or weighting two orientations.

We propose the following algorithm for adaptively wrapping the orientations prior to averaging or vector addition to address the problem. Given a set of m orientations $\theta_1, \theta_2, \theta_3, \ldots, \theta_m$,

1. Choose an arbitrary orientation, e.g. θ_1, as the reference orientation θ_r.

2. Compute the set of the number of rotations by π, $n_1, n_2, n_3, \ldots, n_m$ such that no orientation is greater than π apart from the reference orientation θ_r, i.e. $abs(\theta_i + n_i\pi - \theta_r) <= \frac{\pi}{2}$ for all $1 \leq i \leq m$.

3. If any of the compensated orientation $\theta_i + n_i\pi$ is exactly π apart from the reference orientation θ_r, choose another reference orientation and go to Step 2.

4. Otherwise, perform simple or weighted averaging or vector addition with the compensated orientations.

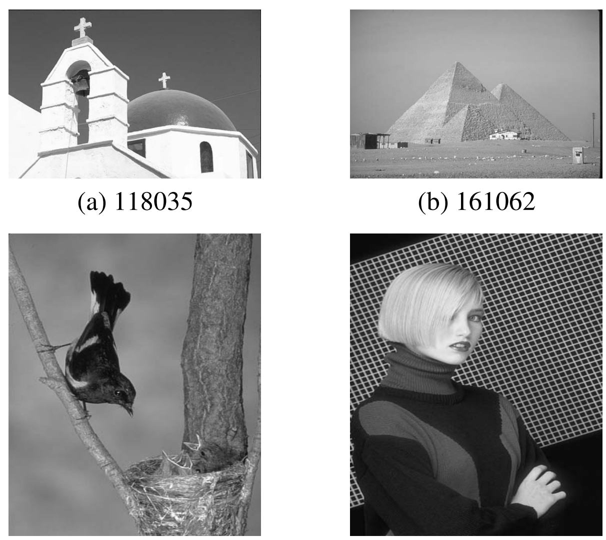
For each location in the image, we weight the adaptively wrapped estimated orientations, $\theta_e^{(\ell)}(m,n)$, at each of the L scales of the complex steerable pyramid with their measure of orientation dominance (Eqn. (5)),

$$WMScaleOri(m,n) = \frac{\sum_{\ell=1}^{L} \mathbf{O}^{(\ell)}(m,n)\theta_e^{(\ell)}(m,n)}{\sum_{\ell=1}^{L} \mathbf{O}^{(\ell)}(m,n)} \tag{6}$$

4. EXPERIMENTS

We applied both techniques of estimating multiscale orientation, comprising of scale selection and scale weighting by orientation dominance respectively, to real images where the complexity of structures would tax their performance. We used four orientations and four scales for the complex steerable pyramid and compared the results of the two techniques with our implementation of Feng and Milanfar's multiscale PCA technique [4].

The Berkeley Segmentation Dataset [10] contains human hand-segmentations of a subset of the Corel image database containing natural images. We converted the region information of the dataset into contours and estimated the orientation along each contour with Savitsky-Golay [11] filters of width 5. For each location in the image, the ground-truth orientations are adaptively wrapped and averaged across the available contour orientations (obtained from multiple human subjects) at that location. We selected images, shown in Fig. 2, where the contours could be obtained from the gray-scale edges instead of higher-order features such as texture boundaries.



(a) 118035 (b) 161062 (c) 163014 (d) 198023

Fig. 2. Test images from the Corel image database with hand segmentation ground-truths from the Berkeley Segmentation Dataset [10].

The multiscale orientation estimation results are given both in terms of mean error and root mean square error in Table 1. The techniques of scale selection and scale weighting from the processed features obtained by a complex steerable pyramid perform better than Feng and Milanfar's technique of applying PCA to gradient fields obtained by differential operators as used in their paper. We can also observe that the scale weighting technique generally per-

Image ID	Multiscale PCA	Scale Selection	Scale Weighting
118035	17.84° (27.68°)	15.53° (27.24°)	15.24° (24.48°)
161062	23.99° (33.75°)	22.04° (32.95°)	21.73° (32.00°)
163014	26.46° (36.39°)	14.17° (23.04°)	14.67° (23.29°)
198023	25.52° (35.94°)	17.91° (28.32°)	18.33° (27.43°)

Table 1. Average orientation estimation error (root mean square error in brackets) in degrees compared to human ground truth from the Berkeley Natural Image Database. Each is computed over thousands of pixels.

forms better than scale selection except in the average error in the third and fourth images. The third image contain more complex textures than the others. The r.m.s. error in the fourth image is better for scale weighting indicating a smaller frequency of extreme errors.

5. CONCLUSION

Orientation estimation by approaches based on gradient information are dependent on the scale or size of the image neighbourhoods being considered, e.g. filter pyramids provide different orientation estimates at each level for real images. A single scale-independent estimate of orientation is often desirable for tasks such as contour extraction and image segmentation, similar to how humans perceive the orientation of contours. Perona [3] applied anisotropic diffusion and Feng and Milanfar [4] used orientation dominance measures obtained from a PCA analysis of differential gradients in each neighbourhood in an image to weight orientation estimates obtained at each level from the PCA analysis. We propose the use of a complex steerable pyramid that provide orientation channels with higher orientation selectivity and a simpler measure of orientation dominance that can be applied directly to the output of the complex steerable pyramid. Furthermore, the proposed measure of orientation dominance does not make any assumption about the number of structures in the neighbourhood, nor their orthogonality as in the case of [4]. We also proposed an algorithm for adaptively wrapping orientations to prevent the wraparound of orientations from inducing errors in orientation averaging, weighting and vector addition.

In the future, we plan to use a complex steerable pyramid with more orientation channels and develop other measures of orientation dominance that can better deal with complex image structures such as corners and junctions. We also plan to explore the possibility of reformulating the multiscale orientation estimation and orientation dominance measures under a more formal framework of scale steering.

6. REFERENCES

[1] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, 1986.

[2] William T. Freeman and Edward H. Adelson, “The design and use of steerable filters,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 13, no. 9, pp. 891–906, 1991.

[3] P. Perona, “Orientation diffusions,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 457–467, 1998.

[4] X.G. Feng and P. Milanfar, “Multiscale principal components analysis for image local estimation,” in *The 36th Conference on Signals, Systems and Computers*, Pacific Grove, USA, November 2002, vol. 1, pp. 478–482.

[5] T. Aach, I. Stuke, C. Mota, and E. Barth, “Estimation of multiple local orientations in image signals,” in *IEEE International Conference on Acoustics, Speech and Signal Processing*, Montreal, Canada, May 2004.

[6] A.A. Bharath and J. Ng, “A steerable complex wavelet construction and its application to image denoising,” *IEEE Transactions on Image Processing*, To appear.

[7] Athanassios Papoulis, *Probability, Random Variables and Stochastic Processes*, McGraw–Hill, New York, 1984.

[8] Tony Lindeberg, “Principles for automatic scale selection,” in *Handbook on Computer Vision and Applications*, vol. 2, pp. 239–274. Academic Press, 1999.

[9] D. L. Ringach, R. M. Shapley, and M. J. Hawken, “Orientation selectivity in macaque V1: Diversity and laminar dependence,” *Journal of Neuroscience*, vol. 22, no. 13, pp. 5639–5651, July 2002.

[10] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proc. 8th Int’l Conf. Computer Vision*, Vancouver, Canada, July 2001, vol. 2, pp. 416–423.

[11] A. Savitsky and M.J.E. Golay, “Smoothing and differentiation of data by simplified least squares procedures,” *Analytical Chemistry*, vol. 36, no. 8, pp. 1627–1639, 1964.