Published in: Proceedings of 3rd IEEE International Conference on Image Processing.
Vol. III, pp. 185-188. Lausanne, Switzerland. 16-19 September 1996.

A ROTATION-INVARIANT PATTERN SIGNATURE

Eero P. Simoncelli

GRASP Laboratory, rm. 335C
Computer and Information Science Dept.
University of Pennsylvania
Philadelphia, PA 19104-6228

We propose a "signature" for rotation-invariant representation of local image structure. The signature is a complex-valued vector constructed analytically from the projections of the image onto a set of oriented basis kernels. The components of the signature form an overcomplete set of algebraic invariants, but are chosen to avoid instabilities associated with previously developed algebraic invariants. We demonstrate the use of this signature for representing and classifying junctions in grayscale imagery.

Local image symmetry provides important cues for visual interpretation. In particular, the local arrangement of oriented contours is a powerful source of information in applications ranging from optical character recognition, to texture-based segmentation, to occlusion boundaries detection. It is typically the *relative* orientation of such contours that carries the important information: the absolute orientation is often irrelevant. It is thus of interest to develop stable, unique, rotation-invariant representations of such structures.

Many authors begin by projecting the image structure onto a local rotation-invariant basis (e.g., [6, 3, 7, 4, 10, 2, 5, 8, 11]). Examples of such decompositions are various types of local moment, derivative operators (which are moments in the Fourier domain), or angular harmonics. These decompositions are closely related, often differing only by a linear transformation.

Consider the problem of matching an observed local image intensity pattern against a set of candidate patterns. A brute-force solution, in which one rotates the image pattern through a set of discretized orientations searching for an optimal match is inelegant, inefficient, and highly susceptible to local minima. A number of authors have taken the approach of first estimating a "dominant" orientation from the projection onto low-order basis functions (e.g., the gradient), and using this estimate to align the two patterns for comparison (e.g., [7, 14, 15, 16, 5]). This type of approach, while efficient, becomes unstable for patterns lacking a strongly dominant orientation.

More generally, one can use the theory of algebraic invariants to construct rotation-invariant representations of image content [1, 12, 13, 14]. The theory allows one to construct a complete set of such invariants. But the set is non-unique, and depends on the initial choice of basis. Many such invariants are highly noise-sensitive, and thus unsuitable for applications. In the present paper, we propose a simple, stable, unique, rotation-invariant signature, consisting of a set of invariants of the angular Fourier decomposition.

1. RELATIVE PHASE

Consider a local decomposition of image structure via projections onto a set of angular Fourier basis kernels:

$$f_n = \int dr \int d\theta \; I(r,\theta) g(r) e^{-in\theta}, \quad 1 \le n \le N,$$

where $g(r)$ is an arbitrary integrable radial function, and $I(r,\theta)$ is a polar parameterization of the image about an (arbitrary) origin point. The case $N = 1$ corresponds to a gradient operation, and the magnitude $|f_1|$ provides a natural rotation-invariant quantity. But complex local structures cannot be represented with a single harmonic; thus we seek a rotation-invariant signature based on projections onto multiple harmonics.

The effect of rotating the image by angle α on each angular Fourier component is well known. Each component is phase shifted by an amount depending on its harmonic number:

$$R_\alpha(f_n) = e^{in\alpha} f_n. \tag{1}$$

The magnitude of each component is a rotation-invariant quantity. But the set of magnitudes is incomplete: patterns with vastly different spatial structure can have identical angular Fourier magnitudes.

Intuitively, one senses that the missing invariant quantities are the *relative orientations* of the Fourier components. It is not straightforward to encode these by comparing component phases, since the phase of the

Research partially supported by an NSF CAREER grant to EPS.

nth harmonic component has an n-fold redundancy. For example, if the phase of the third harmonic term is ϕ_3, the absolute orientation of this sinusoidal basis function is either ϕ_3, $\phi_3 + 2\pi/3$ or $\phi_3 + 4\pi/3$. Thus, a direct attempt to encode relative orientation by phase comparisons will fail.

One simple approach, closely related to that described in [14], involves choosing the phase of the first Fourier component as a pattern-dependent angular origin, and encoding all other phases relative to this one. These relative-phase invariants may be expressed as:

$$p_n \equiv \angle\{f_n f_1^{*n}\}, \qquad \forall n > 1, \tag{2}$$

where f_1^{*n} indicates the complex conjugate of the nth power of f_1, and the function $\angle\{\}$ is the branch of the complex phase in the interval $(-\pi, \pi]$.

Rotation-invariance in this situation is easily verified: rotation by an amount α produces a relative phase for the nth component as follows

$$\angle\{R_\alpha(f_n)R_\alpha(f_1^{*n})\} = \angle\left\{f_n e^{in\alpha} f_1^{*n} e^{-in\alpha}\right\} = p_n.$$

The components of a signature vector that is both rotation-invariant and contrast-invariant may be constructed by combining the relative-phase invariants p_n with the set of magnitudes, normalized by the first harmonic magnitude:

$$s_n = \frac{|f_n|}{|f_1|} e^{ip_n}, \qquad 1 < n \leq N.$$

This (complex) signature vector may be used to represent and classify patterns with significant first-harmonic content. But, as mentioned in the introduction, the calculation is singular for patterns with $|f_1| = 0$. Even without the magnitude normalization, the relative phase encoding is unstable when the first harmonic is small.

The set of relative phases and magnitudes discussed above is algebraically complete: all other invariants based on the same Fourier expansion may be expressed as functions of this set. But given their instability in the presence of a small first harmonic component, we choose to construct an overcomplete set of invariants as described in the next section.

2. A STABLE OVERCOMPLETE SET OF ROTATION INVARIANTS

Consider the following quantity, which is a natural extension of the phase invariants of the previous section:

$$p_{nm} \equiv \angle\{f_n^m f_m^{*n}\}, \qquad n < m.$$

This quantity corresponds to the relative phase of the nth and mth Fourier components. It is rotation-invariant, as can be verified by substitution of equation (1). The subset for which $m = 1$ correspond to the phase invariants given in equation (2).

When n and m have a common factor of, say, k, the product $f_n^m f_m^{*n}$ will have a k-fold rotational symmetry. This means that the invariant cannot distinguish certain phase combinations. For example, let the second and fourth-order components be of opposite phase: $f_2 = 1$ and $f_4 = e^{i\pi}$. Then $p_{24} = \angle\left\{e^{i2\pi}\right\} = 0$. But the value of p_{24} is also zero when the two components are *aligned* in phase: for example, $f_2 = 1$ and $f_4 = 1$. This multiplicity is due to the common sub-periodicity in these two Fourier harmonics, and may be eliminated by using a modified relative-phase invariant:

$$\phi_{nm} \equiv \angle\left\{f_n^{(l(n,m)/n)} f_m^{*(l(n,m)/m)}\right\}, \qquad n < m, \tag{3}$$

where $l(n, m)$ is the "least common multiple" function.

In addition to the phase invariants, we have a set of magnitude invariants. The two can be unified in a set of complex signature components, each with the same intensity dependence, as follows:

$$s_{nm} = \sqrt{|f_n f_m|} e^{i(\phi_{nm})}, \qquad n \leq m. \tag{4}$$

Note that the terms for which n and m are equal correspond to the Fourier component magnitudes. If contrast-invariance is also desired, the entire vector of signature components, $\vec{s}$, may be normalized.

The advantage of this signature over, for example, the invariants presented in [14] is that it does not rely on any specific Fourier component being nonzero. The drawback is the increased dimensionality: there are $N(N+1)/2$ signature components.

3. ORIENTED ENERGY

In some contexts, one wishes to compute a signature for junction geometry that is independent of whether this geometry is formed from lines or edges [16]. A standard technique for representing such line/edge-invariant contours is to compute an oriented "energy" measure as the sum of squared responses of a quadrature pair of filters (e.g., [7, 15, 16, 5]).

As a substrate for this computation, we use a set of asymmetric steerable "wedge" filters [18]. Filters in a quadrature filter pair are typically symmetric or anti-symmetric, and their oriented energy response is thus constrained to be periodic with period π, independent of image structure. The asymmetry of the wedge filters relieves us of this constraint. An example of such filters, for $N = 8$, is shown in figure 1.

The filters are polar-separable, with a somewhat arbitrary radial portion. The angular portions consist of N even-symmetric and N odd-symmetric functions constructed from the first N terms of a Fourier series:

$$h_e(\theta - \alpha_k) \quad = \quad \sum_{n=1}^{N} w_n \cos(n(\theta - \alpha_k)),$$

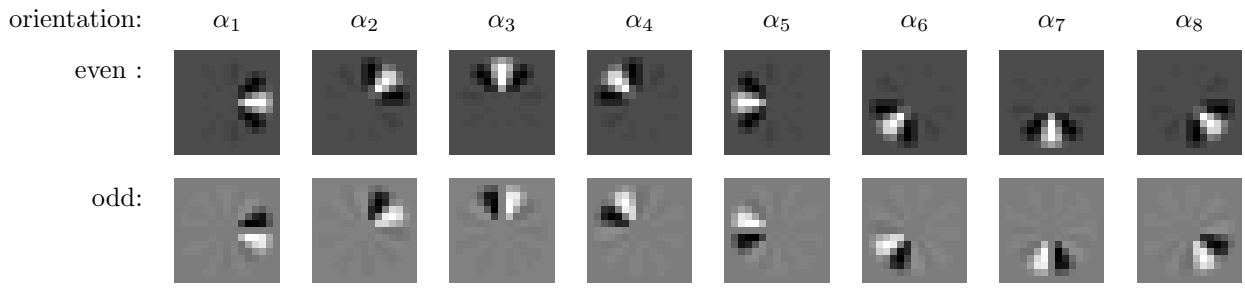

Figure 1. Example set of 15×15 steerable wedge functions for $N = 8$. From this basis set, a filter of either symmetry (even or odd) may be synthesized at *any* orientation.

$$h_o(\theta - \alpha_k) \quad = \quad \sum_{n=1}^{N} w_n \sin(n(\theta - \alpha_k)),$$

for $\alpha_k = 2\pi(k-1)/N,\ k \in [1, 2, \ldots, N]$. The weights w_n are chosen to maximize a measure of angular localization.

This set of $2N$ functions span a rotation-invariant subspace, and any rotated copy of either function may be written as a linear combination of the set:

$$h_e(\theta - \alpha) \quad = \quad \sum_{k=1}^{N} [a_k(\alpha) h_e(\theta - \alpha_k) + b_k(\alpha) h_o(\theta - \alpha_k)]$$

$$h_o(\theta - \alpha) \quad = \quad \sum_{k=1}^{N} [c_k(\alpha) h_e(\theta - \alpha_k) + d_k(\alpha) h_o(\theta - \alpha_k)]$$

The interpolation coefficients $a_k(\alpha), b_k(\alpha), c_k(\alpha)$ and $d_k(\alpha)$, are written in terms of trigonometric functions of the rotation angle, α, and are given in [18].

Naturally, the inner products (or convolutions) of these filters with an image will obey the *same* linear constraint. If the filter responses are denoted r_*, then:

$$r_e(\alpha) = \sum_{k=1}^{N} [a_k(\alpha) r_e(\alpha_k) + b_k(\alpha) r_o(\alpha_k)].$$

The energy measure is simply the sum of squares of the two responses: $E(\alpha) = r_e^2(\alpha) + r_o^2(\alpha)$. As such, it may be written as a linear combination of the set of all pairwise products of the filter responses drawn from the set $\{r_e(\alpha_k), r_o(\alpha_k) | k = 1, 2, \ldots N\}$. Thus, the Fourier expansion of the local orientation energy function may be computed directly from (quadratic combinations of) the filter outputs [19].

4. RESULTS

We calculated signature vectors for the set of prototype junctions illustrated in figure 2. We computed inner products of each of the 16 wedge kernels with each junction image, computed quadratic combinations of these responses, and computed the Fourier expansion of the local orientation energy. Using only the first 6 Fourier components of the orientation energy, we computed a vector of 21 rotation-invariant signature components, via equations (3) and (4).

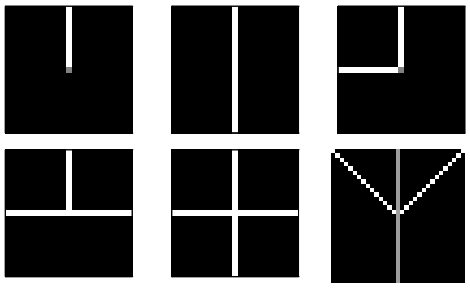

Figure 2. Set of six prototype junction images: $\frac{1}{2}$-line, line, corner, T-junction, cross, and Ψ-junction.

In order to quantify the signature stability, we measured the average change in each signature vector resulting from addition of uniform white noise to the image. Changes in the signature vectors were measured via Euclidean distance to the noise-free signature. Averages were calculated over 25 trials. Figure 3 contains plots of the signature signal-to-noise ratio (SNR) as a function of image SNR for each junction. All signatures behave stably: they are accurate at low noise levels, and degrade gracefully as the noise levels increases.

Signatures were also computed for 4 locations in a real image, shown in figure 4. These were compared with the prototype signatures using a Euclidean metric[1], and the closest prototype chosen as a label for the lo-

[1]Euclidean distance is clearly non-optimal: a proper measure of similarity should take into account the distribution of signatures of the prototype junction set, as well as the statistics of non-junction signatures. We reserve this for future work!

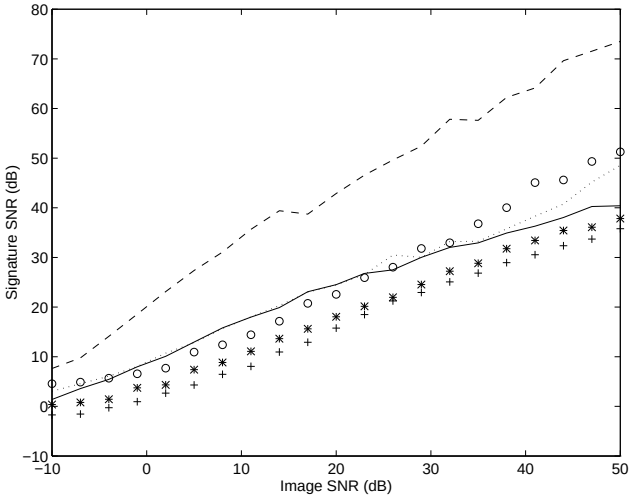

Figure 3. Signature SNR in the presence of additive white image noise. Each curve corresponds to the signature for a prototype junction: $\frac{1}{2}$-line (dashed); line (solid); corner (dots); T-junction ($\circ$); cross ($+$); Ψ-junction ($*$).

Figure 4. Signature vector $\vec{s}$ was computed at the indicated locations. Least-squares comparisons with the idealized signatures produced labels "line", "T-junction", "corner", and "line" (ordered from lowest to highest in the image).

cal image content. These preliminary tests indicate that the signature is fairly robust, although extensive testing is still necessary to quantify this.

We have described a rotation-invariant pattern signature that is unique, stable, and reasonably efficient. The signature may be built on the Fourier components up to any order, but does not rely on any of these components being nonzero.

5. REFERENCES

[1] M Hu. Visual pattern recognition by moment invariants. *IRE Trans Info Theory*, IT-8:179–187, Feb 1962.

[2] J B Martens. The Hermite transform – theory. *IEEE Trans Acoust Speech Sig Proc*, 38(9):1595–1606, 1990.

[3] Per-Erik Danielsson. Rotation-invariant linear operators with directional response. In *5th Int'l Conf Patt Rec*, Miami, Dec 1980.

[4] J J Koenderink and A J van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*, 55:367–375, 1987.

[5] W T Freeman and E H Adelson. The design and use of steerable filters. *IEEE Trans on Patt Anal Mach Intell*, 13(9):891–906, 1991.

[6] G H Granlund. In search of a general picture processing operator. *Comp Graphics Im Proc*, 8:155–173, 1978.

[7] H Knutsson and G H Granlund. Texture analysis using two-dimensional quadrature filters. *IEEE Comp Soc Workshop on Comp Arch Patt Anal Image Database Mgmt*, pages 388–397, 1983.

[8] P Perona. Steerable-scalable kernels for edge detection and junction analysis. *Image and Vision Computing*, 10(10):663–672, 1992.

[9] W T Freeman. *Steerable Filters and Local Analysis of Image Structure*. PhD thesis, MIT Media Lab, Cambridge, MA, May 1992.

[10] J Bigün. Pattern recognition by detection of local symmetries. In *Proc Patt Rec in Practice III*. Elsevier Science, May 1988.

[11] M Michaelis and G Sommer. Junction classification by multiple orientation detection. In *Proc Euro Conf Comp Vision*, 1994.

[12] S S Reddi. Radial and angular moment invariants for image identification. *IEEE Trans Patt Anal Mach Intell*, PAMI-3(2), Mar 1981.

[13] M R Teague. Image analysis via the general theory of moments. *J Opt Soc Am*, 70:920–930, Aug 1980.

[14] Y S Abu-Mostofa and D Psaltis. Image normalization by complex moments. *IEEE Trans Patt Anal Mach Intell*, PAMI-7(1), Jan 1985.

[15] M Kass and A Witkin. Analyzing oriented patterns. *Comp Vision Graphics Image Proc*, 37:362–385, 1987.

[16] P Perona and J Malik. Detecting and localizing edges composed of steps, peaks and roofs. In *Proc Int'l Conf Comp Vision*, 1990.

[17] E Simoncelli and H Farid. Steerable wedge filters. In *Proc Int'l Conf Comp Vision*, Boston, MA, June 1995.

[18] E Simoncelli and H Farid. Steerable wedge filters for local orientation analysis. *IEEE Trans Image Proc*, Aug 1996. To Appear.

[19] E P Simoncelli. *Distributed Analysis and Representation of Visual Motion*. PhD thesis, MIT Dept Elec Eng & Comp Sci, Cambridge, MA, Jan 1993.