

Improving Phase-Congruency Based Feature Detection through Automatic Scale-Selection

Veit U.B. Schenk^{1,2} and Michael Brady²

¹ Laboratory of Computational Engineering, Helsinki University of Technology, P.O. Box 9203, FIN-02015 HUT, Finland, vubs@robots.ox.ac.uk

² Robotics Research Group, Dept. of Engineering Science, Ewert House, Oxford, OX2 7SG UK

Abstract. In this paper we present a novel method for computing phase-congruency by automatically selecting the range of scales over which a locally one-dimensional feature exists. Our method is based on the use of local energy computed in a multi-resolution steerable filter framework. We observe the behaviour of phase over scale to determine both the type of the underlying features and the optimal range of scales over which they exist. This additional information can be used to provide a more complete description of image-features which can be utilized in a variety of applications that require high-quality low-level descriptors. We apply our algorithm to both synthetic and real images.

Keywords: Phase-congruency, local energy feature-detection, scale-detection, steerable filters

1 Introduction

Phase congruency [7] is a very appealing concept for general feature detection because it permits feature detection independent of the actual feature type, i.e. rather than being optimized to detect edges or ridges or valleys it can be used to detect almost any type of feature in a unified framework. The underlying principle is that phase is constant or *congruent* over all scales at the location of what the human visual system would perceive as a locally one-dimensional feature, such as an edge, a ridge or a valley (as opposed to two-dimensional features, e.g. junctions) [6]. An advantage of phase-congruency is that the type of feature can be classified using the phase-value at which congruency occurs. Furthermore, the degree to which phases are congruent can readily be computed as a ratio of some ‘ideal’ value and the actual phase-values leading to a measure that is contrast and brightness independent. There are interesting parallels between the concept of phase-congruency and Lindeberg’s [5] concept of the *scale-space edge*, which is defined as a connected set of points in scale-space: features are found where a certain measure persists over a consecutive range of scales. The method presented in this paper identifies locations in image signals at which phase takes on a small set of fixed values over a range of subsequent scales.

The use of phase is particularly appealing for a number of reasons: it has been demonstrated experimentally [8] that most of the information in a signal is stored in the phase, rather than the amplitude: the phase effectively encodes the ‘location’ at which

the individual sinusoid contribute to the overall signal. Furthermore, phase is stable not only under translation but also geometric deformations and contrast variations [1].

One of the great advantages of the use of phase is that the exact position of a feature can be determined easily to sub-pixel accuracy [3] without the need for explicit sub-pixel feature detection.

Morrone and Owens [7] show that for a one-dimensional signal $I(x)$ which has the short-term Fourier Transform expansion:

$$I(x) = \sum_{n>0} A_n \cos(n\omega x + \phi_n) = \sum_{n>0} A_n \cos(\phi_n(x)) \quad (1)$$

where A_n, ϕ_n are respectively the n^{th} components of amplitude and phase at position x , all phase-components $\phi_n(x)$ are (near) identical at the location x of a feature.

Phase congruency can be shown to be directly related to local energy [10] which in the one-dimensional case is defined as

$$LE_{1D} = \sqrt{I^2 + H^2} \quad (2)$$

where I is the input signal and H its Hilbert transform.

In practice, phase-congruency is obtained by computing local energy at a number of scales and integrating the resulting coefficients appropriately. Kovessi [4] presents the first such computationally efficient implementation of phase-congruency. He defines a phase congruency measure $PC(x)$ at signal location x as:

$$PC(x) = \max_{\phi} \frac{\sum_{n>0} A_n \cos(\phi_n(x) - \bar{\phi}(x))}{\sum_{n>0} A_n} = \frac{LE}{\sum_{n>0} A_n} \quad (3)$$

The value of $\bar{\phi}(x)$ that maximizes (3) is the amplitude weighted mean local phase-angle. It can be shown that the numerator LE in the above expression is the local energy of the signal [10]; PC is consequently the (local) maximum of the amplitude-weighted sum of the phases (computed over a range of scales) normalized by the amplitude sum. Kovessi extends the concept to two dimensional signals and addresses issues such as noise or advanced phase-congruency based operators for symmetry and asymmetry. Although PC gives good results for a wide range of synthetic and natural images and produces feature maps that compare favourably with e.g. a Canny edge detector [4], there are a number of drawbacks which are related to one major oversight: the fact that local energy is proportional to phase-congruency (see (3)) is only valid for isolated features and is no longer valid as soon as the spatial extent of the filter used overlaps neighbouring features in the input signal. This is illustrated in Fig. 1 which shows in (A) a signal consisting of a square wave with decreasing frequency (from left to right). Subfigure (B) is the energy amplitude scaleogram (computed using (2) over a range of centre frequencies $\lambda = 0.1 : 0.65$). The corresponding phaseogram is shown in Subfig. (C). A number of observations are immediately obvious:

- The range of scales over which edges have significant energy is much larger than the rather small range of scales over which ridges and valleys respond.

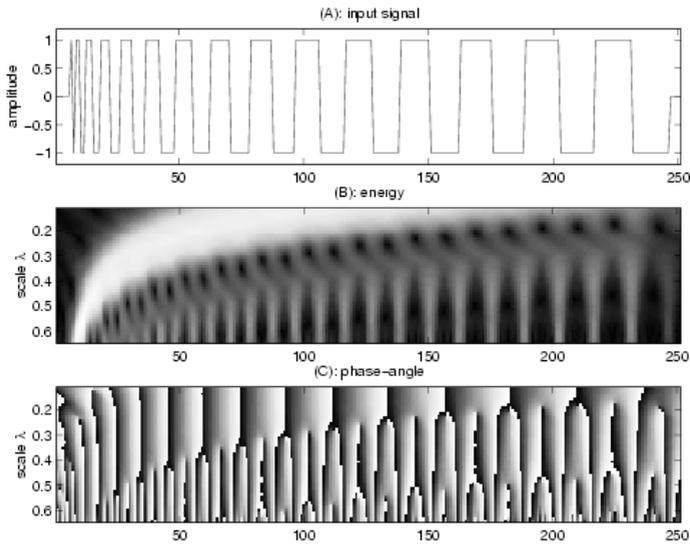


Fig. 1. Locality of energy response: the energy distribution over scale depends not only on the underlying feature, but also its neighbourhood. See text for details.

- There is interference as filters overlap neighbouring features leading to the cancellation of responses which in (B) correspond to the ‘black holes’ along the centre of the scaleogram (or alternatively, the regions in (C) where two columns of constant phase merge). In these regions phase is unstable and must not be used for further processing [3].
- A second form of interference occurs towards the low-frequencies (small λ): responses due to combinations (rather than isolated) of neighbouring features add up to high energy amplitudes as shown by the broad high intensity curve in the scaleogram. The response is high even in between what we would regard as a feature and therefore must be excluded from further processing.

Since PC is computed over *all* scales, rather than the appropriate range of scales and – as illustrated above – strong features (edges) dominate neighbourhoods, PC responds predominantly to edges and isolated very thin lines and produces spurious responses in areas with high energy which is due to interference of low-frequency responses. This is illustrated in Fig. 2 and Fig. 3. The left of Fig. 2 shows a synthetic image with a variety of structures including ridges, valleys and edges over a wide range of orientations, scales and with different amounts of smoothing.

The raw PC output is shown on the right in Fig. 2. There are high responses near most of the edges apart from the rather smooth edges in the lower left quadrant. The responses at the edges are also significantly stronger than those corresponding to the ridges and valleys in between these edges. As a result, the post-processed version shown in Fig. 3 marks as interesting features only edges and is rather unsuccessful at identifying the smoothed features in the lower left quadrant. The method proposed in this paper

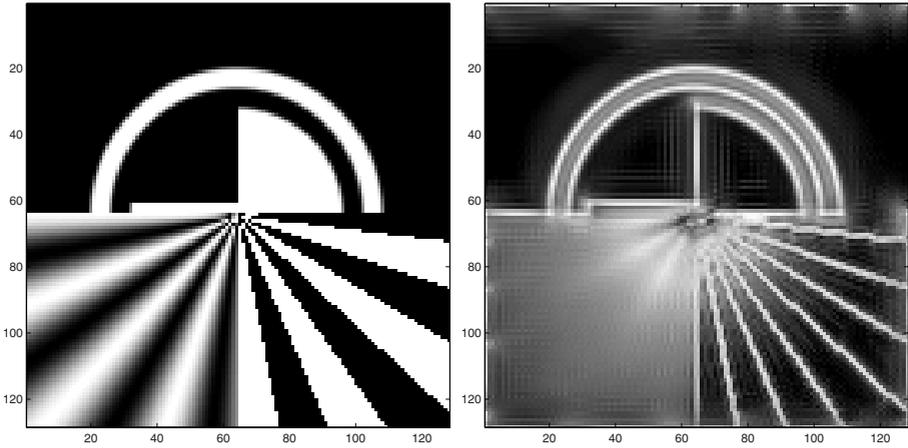


Fig. 2. Left: Synthetic signal with variety of locally one-dimensional features (edges, ridges, valleys). Right: corresponding *PC* map. Note the strong responses at edges, almost uniform response in bottom-left quadrant and weak responses at ridges/valleys.

identifies a wider range of geometric features than just edges and provides a better response to heavily smoothed features, as in the lower left quadrant of the image, thus taking advantage of local energy's inherent ability to respond to *any* type of locally one-dimensional feature equally well.

2 Methodology

Our approach is based on the use of local energy for feature detection. At each location in the image, we compute an energy response, which can be used to obtain an energy-amplitude and a phase-angle ϕ . Since we are interested in the energy response and phase-angle at each image feature independent of its orientation, we interpolate the exact response at the orientation θ_l of each image-point through the use of steerable filters [2]. We define θ_l to be *perpendicular* to the locally one-dimensional feature, i.e. along the orientation of maximum variation. Only regions of the image with high local energy will be selected for further processing. Since the post-processing stages involve tracking phase ϕ over multiple scales, the signal is decomposed at multiple resolutions.

2.1 Decomposition and Orientation Computation

Local energy is computed using quadrature pairs of odd-symmetric and even-symmetric bandpass filters. We decompose the image using a steerable filter bank at $N_o = 4$ odd-symmetric and $N_e = 5$ even-symmetric orientations θ_o and θ_e distributed evenly over the half-circle of orientations $0.. \pi$ (i.e. $\theta_o = \frac{\pi}{4}[0, 1, 2, 3]$ and $\theta_e = \frac{\pi}{5}[0, \dots, 4]$) and $S = 15$ closely spaced scales (spacing of 0.2 octaves for filters with 1-octave bandwidth) in order to approximate a continuous decomposition (in scale). In the following, where

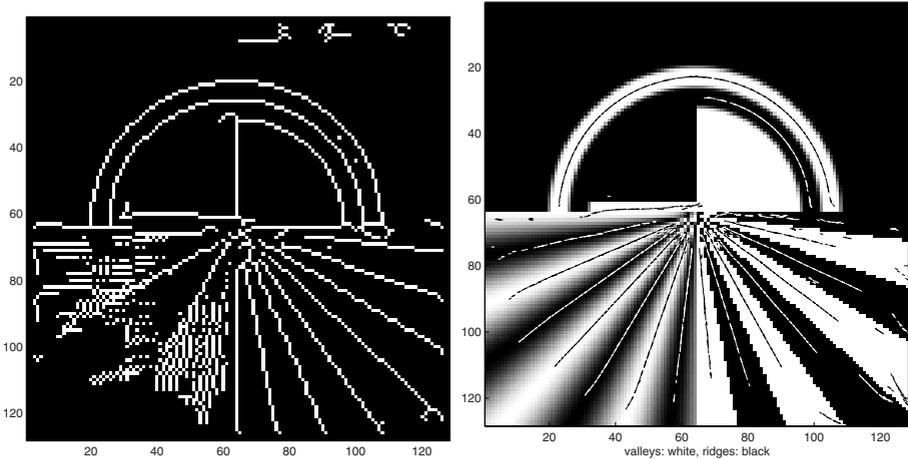


Fig. 3. Left: *PC* map after postprocessing (non-maximum suppression and hysteresis thresholding). Note the noisy response in bottom left quadrant. Furthermore, only edges are detected. Right: *LE*: Ridge (black) and valley (white) feature points overlaid onto original image. See text for more details

appropriate, the scale index s is omitted to improve legibility. The polar-separable filters are constructed in the Fourier domain with a cosine raised to the power 2 on a log-scale as the radial component and a quadrature pair of filters for the angular component which has a \cos^3 and $|\cos^3|$ cross-section for the odd/even-symmetric filter-parts respectively [9].

The image is convolved with the resulting filterbank by multiplying the Fourier Transform of the image with the individual filters and inverse Fourier transforming. The resulting subbands are kept at full resolution, i.e. no pyramid scheme is used, which facilitates the tracking of phase through scale. The 2D extension of (2) is

$$LE_n = \sqrt{O_n^2 + E_n^2} \quad (4)$$

where O_n, E_n are the n^{th} odd/even-symmetric responses for each of the N_o orientations. Since $N_o \neq N_e$ and the odd-symmetric and even-symmetric filters consequently are not aligned, they need to be aligned prior to computing *LE*. This is achieved by steering the even-symmetric filters to the orientations θ_o .

Orientation Computation. At each point and for every scale local orientation θ_l is found as

$$\theta_l = \frac{1}{2} \arctan(\text{imag}(\bar{V})/\text{real}(\bar{V})) \quad (5)$$

where $\bar{V} = \sum_{n=0}^{N_o-1} \exp(i2n\pi/N_o)|LE_n|$ is the amplitude weighted sum of oriented unit length vectors aligned with θ_o .

At each point and scale S , the N_o subband coefficients are steered to θ_l . The result of this operation is a cuboid of S steered energy response-maps.

2.2 Tracking Phase over Scale

The basic idea for determining the correct range of scales over which to compute phase-congruency is to track the precise, i.e. sub-pixel contours of phase over scale and to label a location in the image as ‘interesting’ if phase is congruent over a minimum number of scales. The candidate points are obtained using the Matlab function `contourc` on each response map of the cuboid using the appropriate phase-angles for positive/negative edges and ridges and valleys.

In order to avoid having to exhaustively search the entire image, we use hysteresis thresholding of local energy amplitude to reduce the number of candidate points.

For each of the (equally spaced along a phase-contour) candidate points p , we iterate through all scales λ . If p has correspondences with the same phase-value in a minimum number l_{min} of subsequent scales, then p is marked as a feature point for that particular phase-value. The start and end-scales λ_b and λ_e of where phase is congruent are recorded. All points in the chain from λ_b to λ_e are removed from the set of candidate points, i.e. they will not be considered at subsequent stages. Correspondence is established if the contours at subsequent scales are no further apart (measured as the distance between two parallel contours) than a specified fraction d_{max} of pixels. The subpixel location of p is recorded as the mean of all points in the chain.

Through extensive experimentation on a large set of images we found that the values $d_{max}=0.3$ pixels and $l_{min}=4$ scales give excellent results for a wide range of images. The choice of d_{max} is not critical, values between $0.2 \leq d_{max} \leq 0.5$ give acceptable results. The choice of l_{min} directly affects

- 1) the sensitivity to the degree of smoothing a feature has undergone: strong smoothing leads to responses with a small frequency spread, i.e. there are only significant responses over a (very) small range of scales. Values of $l_{min} \gg 4$ therefore eliminate strongly smoothed features and favour sharp transitions. Since edges exist over a wider range of scales than lines, too large a value for l_{min} would therefore favour edges over lines.
- 2) How features in closely packed neighbourhoods are treated: as the filter-size increases, the filters quickly start to overlap multiple features, leading to interference in responses. As a result, stability in phase over scale is only guaranteed over a small range of scales. If l_{min} is too large, regions with densely packed features are eliminated. Note that $l_{min}=4$ is a value we obtained empirically for a decomposition over $S=15$ scales with a corresponding step-size of 0.2 octaves. $l_{min}=4$ therefore implies that a feature needs to exist over just under one octave.

The method is stable with respect to the thresholds, small changes lead to gradual changes in the output, rather than catastrophic failure.

3 Results

The methods described in the previous sections were applied to a number of synthetic and real images with varying noise and contrast levels and different densities of features. The results of LE are displayed as follows: since the location of the feature-point is computed at subpixel accuracy, the resulting points are displayed as short lines perpendicular to θ_l , i.e. parallel to the orientation of the locally one-dimensional feature.

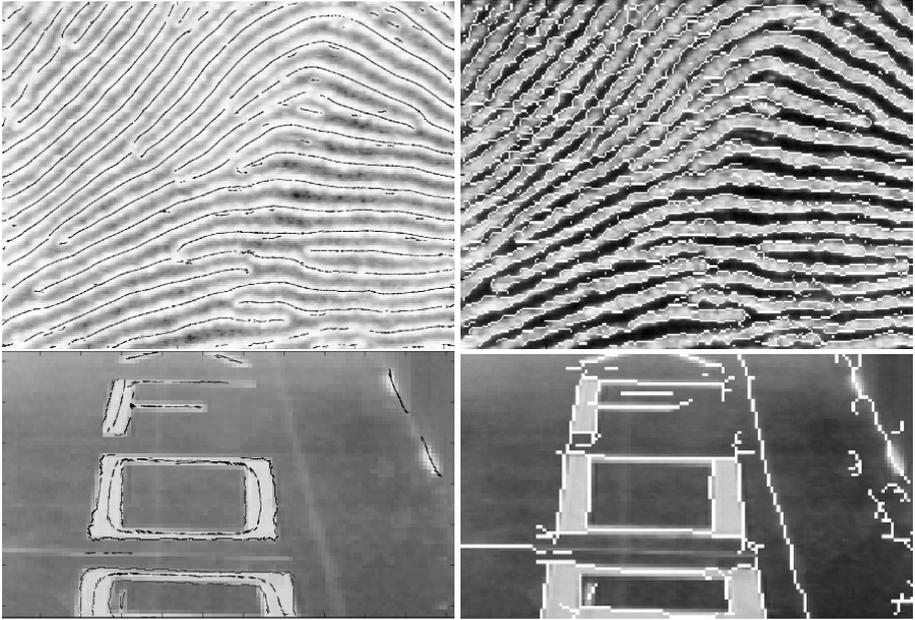


Fig. 4. Top: Fingerprint. Bottom: Writing on floor. See text for details

The right of Fig. 3 shows the ridges and valleys found in the synthetic signal of Fig. 2 overlaid in black and white respectively. For clarity the edges were not marked. Note that we have found and marked a range of features which were completely missed by *PC* as shown on the left in Fig. 3. In particular, we have found the relevant ridges and valleys independently of their actual width not only in the regions with sharp transitions, but also in the smoothed region of the image. It is clear that broad ridges and valleys towards the edge of the image are not marked, this is due to the fact that the filters used were not large enough to respond to features of this size. As expected, the edges of these features are detected however. This example also illustrates one of the problems of the method: very fine lines or edges in neighbourhoods with closely spaced features are not detected reliably as shown e.g. in the centre of the image. At these locations, phase is only valid over a very limited range of scales, due to interference from neighbouring features. This is however a property of phase-congruency per se, rather than the fault of *LE*.

In the two images at the top of Fig. 4 the ridges found in an image of a fingerprint are highlighted. The left image shows the output of the *LE* method (for clarity, the edges and valleys were not marked). Despite the noise-levels and variation in width, the exact centres of the ridges have been marked correctly. This means, not only can we detect the ridges, but we can also describe it in terms of width, orientation etc. The right image shows the output of *PC*. As with the other examples, the edges dominate most other features. Additionally, because the features are relatively closely spaced, there is significant interference between neighbouring features, leading to a high level of spurious, noisy responses.

The two images at the bottom of Fig. 4 shows the image of a piece of writing on a floor. Note how the width of the letters gets narrower towards the top of the image (rear of the scene). On the left we see the result of *LE*. Despite this considerable variation in width of the lines, all letters were detected correctly. Note however, that the faint lines on the floor have not been marked. This is due to the use of the global hysteresis thresholds $[0.15, 0.30]$ of the maximum energy amplitude. Since local energy is computed at a local level as the name implies, a future improvement would be to use hysteresis thresholding based on a local rather than a global threshold. The image on the right shows the output of hysteresis thresholded *PC*. Once again, where lines are broad enough to have distinct edges, the edges dominate the broad lines.

4 Discussion and Conclusions

We have introduced a novel method *LE* for computing phase-congruency over the *appropriate* range of scales, rather than averaging over *all* scales. The features obtained using *LE* are a great improvement over existing phase-congruency implementations and a further step towards a truly general feature detector.

The correct range of scales can be used to further describe a feature and the position of a feature is computed to subpixel precision. We have demonstrated the usefulness of the new method using both synthetic as well as real images.

References

1. D. J. Fleet and A. D. Jepson. Stability of phase information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):1253–1268, 1993.
2. William T. Freeman and Edward H. Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891–906, September 1991.
3. A.D. Jepson and D.J. Fleet. Phase singularities in scale-space. *Image and Vision Computing*, 9(5):338–343, 1991.
4. Peter Kovési. Image features from phase congruency. *Videre: A Journal of Computer Vision Research*, 1(3), 1999.
5. Tony Lindeberg. Edge detection and ridge detection with automatic scale selection. *International Journal of Computer Vision*, 30(2):117–154, 1998.
6. M. Morrone and D. Burr. Feature detection in human vision: a phase dependent energy model. *Proc. Royal Soc. London Bulletin*, pages 221–245, 1988.
7. M. C. Morrone and R. A. Owens. Feature detection from local energy. *Pattern Recognition Letters*, 6:303–313, 1987.
8. A.V. Oppenheim and J.S. Lim. The importance of phase in signals. *Proc. of IEEE 69*, pages 529–541, 1981.
9. Veit U B Schenk. *Visual Identification of Fine Surface Incisions*. PhD thesis, University of Oxford, Dept. of Engineering Science, March 2001.
10. S. Venkatesh and R.A. Owens. On the classification of image features. *Pattern Recognition Letters*, 11:339–349, 1990.