Psychological Research (2000) 64: 136–148

ORIGINAL ARTICLE

Peter Kovesi

Phase congruency: A low-level image invariant

Received: 20 November 1998 / Accepted: 29 September 1999

Abstract Phase congruency is a low-level invariant property of image features. Interest in low-level image invariants has been limited. This is surprising, considering the fundamental importance of being able to obtain reliable results from low-level image operations in order to successfully perform any higher level operations. However, an impediment to the use of phase congruency to detect features has been its sensitivity to noise. This paper extends the theory behind the calculation of phase congruency in a number of ways. An effective method of noise compensation is presented that only assumes that the noise power spectrum is approximately constant. Problems with the localization of features are addressed by introducing a new, more sensitive measure of phase congruency. The existing theory that has been developed for 1D signals is extended to allow the calculation of phase congruency in 2D images. Finally, it is argued that *high-pass* filtering should be used to obtain image information at different scales. With this approach, the choice of scale only affects the relative significance of features without degrading their localization.

Introduction

The ability to construct measures of image features that remain constant over wide ranges of viewing conditions is an important goal of any visual system, whether it is a biological system or a computer based system. Such invariant quantities provide powerful tools for the analysis of images, allowing image processing algorithms to work more reliably and over wider classes of images. The work presented in this paper concentrates on invariant feature detection in low-level or early vision. While the work described in this paper is biologically inspired, its intended application is in computer vision. However, it is hoped that the issues involved in producing an effective implementation of an invariant low-level feature detector will be of interest to those working in human vision.

Some effort has been devoted to investigating invariant measures of higher level structures in images; for example, Hu (1962) developed a series of invariant moments for recognizing binary objects. More recently, there has been considerable interest in geometric invariance, the study of geometric properties of objects that remain invariant to imaging transformations. A collection of papers in this area can be found in the book by Mundy and Zisserman (1992). However, little attention has been paid to the invariant quantities that might exist in low-level or early vision for tasks such as feature detection or feature matching. Some limited exceptions to this include the work of Koenderink and van Doorn (1975), who recognized the importance of differential invariants associated with motion fields, and Florack, ter Harr Romeny and Koenderink (1992), who propose differential invariants for characterizing a number of image contour properties.

The human visual system is able to reliably identify the significance of image features under widely varying conditions. If the illumination of a scene is changed by several orders of magnitude, our interpretation of it will be largely unchanged. Similarly, our interpretation is not greatly affected by changes in spatial magnification, though not to the same degree of tolerance we have for illumination changes. Thus, in the detection of low-level image features, the main form of invariance that is required is invariance to image illumination and contrast, and to a lesser extent, image magnification. Ultimately, one has to make a decision as to whether a feature is significant or not – the thresholding problem. If one has an invariant measure of the significance of features, the thresholding problem is greatly eased.

P. Kovesi
Department of Computer Science,
The University of Western Australia,
Nedlands, W.A. 6907
e-mail: pk@cs.uwa.edu.au

The thresholding problem has plagued computer vision. Gradient based edge detection methods such as those developed by Sobel (Pringle, 1969), Marr and Hildreth (1980), Canny (1983, 1986), and others are sensitive to variations in image illumination, blurring, and magnification. The image gradient values that correspond to significant edges are usually determined empirically. Efforts to determine threshold values automatically have been very limited and not very successful (Canny, 1983; Fleck, 1992a, 1992b; Kundu and Pal, 1986). Achieving invariance with feature detectors developed in the spatial domain of an image is difficult because it is hard to avoid using dimensional measures such as intensity gradients, contrast levels, or equivalent quantities.

In the search for low-level invariant quantities in images, the approach taken here is to make use of data from representations of the image in the *frequency domain*. A model of feature perception called the Local Energy Model has been developed by Morrone, Ross, Burr, and Owens (1986) and Morrone and Owens (1987). This model postulates that features are perceived at points in an image where the Fourier components are maximally in phase. Other work on this model of feature perception can be found in Morrone and Burr (1988), Owens, Venkatesh, and Ross (1989), Venkatesh and Owens (1990), Kovesi (1991, 1993, 1996a), Owens (1994), Morrone, Navangione, and Burr (1995), and Robbins and Owens (1997). A wide range of feature types give rise to points of high phase congruency. These include step edges, line and roof edges, and Mach bands. Morrone and Burr (1988) show that this model successfully explains a number of psychophysical effects in human feature perception.

Almost all work done so far has concentrated on finding points of maximal phase congruency by looking for maxima in local energy. Local energy is a quantity that is proportional to phase congruency and is readily calculated. However, local energy is a dimensional quantity that depends on local contrast. One is unable to specify beforehand what level of local energy corresponds to a significant feature – the thresholding problem again.

This paper concentrates on the issues in calculating phase congruency itself – a dimensionless measure. Values of phase congruency vary from a maximum of 1, indicating a very significant feature, down to 0, indicating no significance. This allows one to specify a threshold to pick out features *before* an image is seen. However, so far, phase congruency has not been used successfully for feature detection for the following reasons:

1. Being a normalized quantity, phase congruency is highly sensitive to noise.
2. The calculation of phase congruency is ill-conditioned if all the frequency components of the signal are very small, or if there is only one (or nearly only one) frequency component present in the signal.
3. The existing measure of phase congruency does not provide good localization of features.

The greatest problem is the sensitivity of phase congruency to noise. Any system that uses normalization to provide invariance to contrast and illumination must inevitably suffer from sensitivity to noise. This indeed must be of concern to biological visual systems.

The paper is organized as follows: After introducing the relationship between local energy and phase congruency it is shown how phase congruency can be calculated from a bank of spatial filters in quadrature. The primary problem of sensitivity to noise is then addressed. Problems in the localization of features are considered next. It is shown that it is important to consider the spread of frequencies present at a point of phase congruency. In addition a new, more sensitive, measure of phase congruency is presented. This is followed by a section covering the issues involved in extending this theory to 2D images. Finally, the issue of analysis at different scales is considered, and it is argued that *high-pass* filtering should be used to obtain image information at different scales instead of the more usually applied low-pass filtering.

Local energy and phase congruency

The local energy model postulates that features are perceived at points of maximum phase congruency in an image. For example, when one looks at the Fourier series that makes up a square wave, all the Fourier components are sine waves that are exactly in phase at the point of the step. Similarly, one finds that phase congruency is a maximum at the peaks of a triangular wave.

Congruency of phase at *any angle* produces a clearly perceived feature. Figure 2 shows a grating constructed from the series

$$s(x) = \sum_{n=0}^{\infty} \frac{1}{(2n+1)} \sin[(2n+1)x + \phi] \qquad (1)$$

where ϕ, the offset at which congruence of phase occurs, is varied from 0 to $\pi/2$.

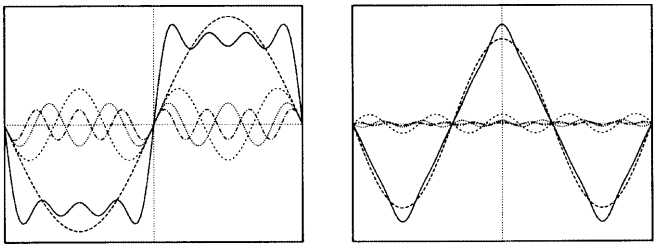

Fig. 1 Construction of square and triangular waveforms from their Fourier series. In both diagrams the first few terms of the respective Fourier series are plotted with *broken lines*, the sum of these terms is the *solid line*. Notice how the Fourier components are all in phase at the point of the step in the square wave, and at the peaks and troughs of the triangular wave

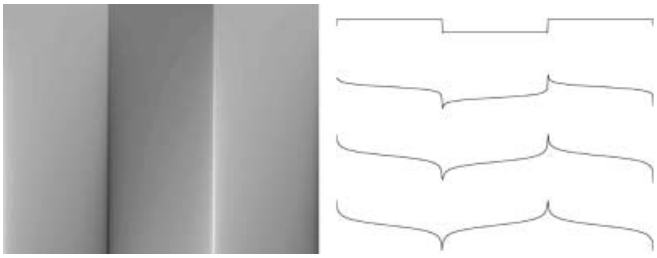

Fig. 2 Interpolation of a step feature to a line feature by continuously varying the angle of congruence of phase from 0 at the top to $\pi/2$ at the bottom. Profiles of this grating corresponding to congruence of phase at 0, $\pi/6$, $\pi/3$ and $\pi/2$ are shown on the right

Morrone and Owens (1987) define the phase congruency function in terms of the Fourier series expansion of a signal at some location x as

$$PC(x) = max_{\bar{\phi}(x)\in[0,2\pi]} \frac{\sum_n A_n \cos(\phi_n(x) - \bar{\phi}(x))}{\sum_n A_n} \quad (2)$$

where A_n represents the amplitude of the n^{th} Fourier component, and $\phi_n(x)$ represents the *local* phase of the Fourier component at position x. The value of $\bar{\phi}(x)$ that maximizes this equation is the amplitude weighted mean local phase angle of all the Fourier terms at the point being considered.

As it stands, phase congruency is a rather awkward quantity to calculate. As an alternative to this, Venkatesh and Owens (1989) show that points of maximum phase congruency can be calculated equivalently by searching for peaks in the local energy function. The local energy function is defined for a one-dimensional luminance profile, $I(x)$, as

$$E(x) = \sqrt{F^2(x) + H^2(x)} \quad (3)$$

where $F(x)$ is the signal $I(x)$ with its DC component removed, and $H(x)$ is the Hilbert transform of $F(x)$ [a 90-degree phase shift of $F(x)$]. Typically, approximations to the components $F(x)$ and $H(x)$ are obtained by convolving the signal with a quadrature pair of filters. Venkatesh and Owens show that energy is equal to phase congruency scaled by the sum of the Fourier amplitudes, that is

$$E(x) = PC(x) \sum_n A_n \quad (4)$$

Thus, the local energy function is directly proportional to the phase congruency function, so peaks in local energy will correspond to peaks in phase congruency.

The calculation of energy from spatial filters in quadrature pairs has been central to many models of human visual perception, for example, those proposed by Heeger (1987, 1988, 1992) and Adelson and Bergen (1985). Other researchers who have studied the use of local energy for feature detection are Perona and Malik (1990), Freeman (1992), and Ronse (1993, 1997). Rosenthaler, Heitger, Kubler, and Heydt (1992) make a comprehensive study of the behaviour of local energy at 2D image feature points. Wang and Jenkin (1992) use

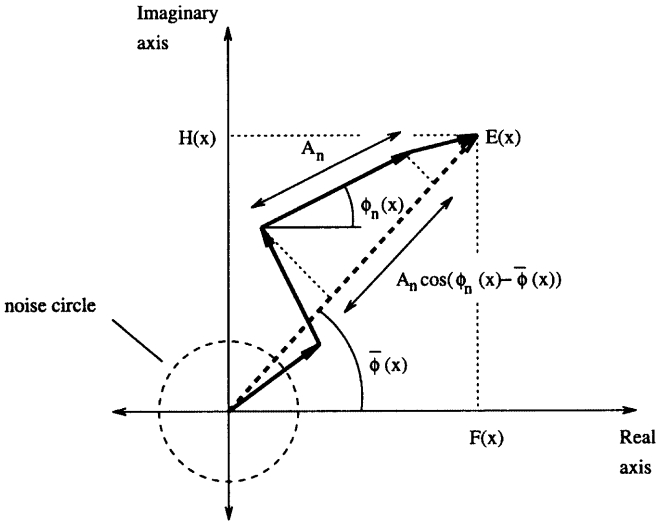

Fig. 3 Polar diagram showing the Fourier components at a location in the signal plotted head to tail. This arrangement illustrates the construction of energy, the sum of the Fourier amplitudes, and phase congruency from the Fourier components of a signal. The noise circle represents the level of $E(x)$ one can expect just from the noise in the signal

complex Gabor filters to detect edges and bars in images. They recognize that step edges and bars have specific local phase properties which can be detected using filters in quadrature; however, they do not connect the significance of high local energy with the concept of phase congruency. It should also be noted that Grossman (1998) recognized that wavelets could be used for the detection of discontinuities. He recognized the fact that discontinuities have no intrinsic scale, and this is reflected in the wavelet coefficient values. However, here too, the connection with the concept of phase congruency was not made.

The relationship between phase congruency, local energy, and the sum of the Fourier amplitudes can be seen geometrically in Fig. 3. The local Fourier components are plotted as complex vectors adding head to tail. The sum of these components projected onto the real axis represent $F(x)$, the original signal with DC component removed, and the projection onto the imaginary axis represents $H(x)$, the Hilbert transform. The magnitude of the vector from the origin to the end point is the total energy, $E(x)$. One can see that $E(x)$ is equal to $\sum_n A_n \cos(\phi_n(x) - \bar{\phi}(x))$. Phase congruency is the ratio of $E(x)$ to the overall path length taken by the local Fourier components in reaching the end point. Thus, one can clearly see that the degree of phase congruency is independent of the overall magnitude of the signal. This provides invariance to variations in image illumination and/or contrast.

Referring to Fig. 2, one can see the following problems in the calculation of phase congruency:

1. The calculation of phase congruency becomes ill-conditioned if all the Fourier amplitudes are very small ($\sum_n A_n \simeq E(x) \simeq 0$).

2. If the value of $E(x)$ falls within the "circle of noise," values of phase congruency lose all significance.
3. If there is only one (or nearly only one) frequency component present in the signal phase, congruency degenerates to a value of one ($\sum_n A_n = E(x)$).
4. The definition of phase congruency as provided by Equation 2 does not provide good localization, as this function only varies with the *cosine* of a phase deviation, rather than say, phase deviation itself.

The problem of phase congruency becoming ill-conditioned if all the Fourier amplitudes are very small can be addressed by adding a small positive constant, ε, to the denominator of the expression for phase congruency. Thus,

$$PC(x) = \frac{E(x)}{\sum_n A_n(x) + \varepsilon} \tag{5}$$

where $E(x) = \sqrt{F(x)^2 + H(x)^2}$. The appropriate value of ε depends on the precision with which we are able to extract frequency information from our signal; it does not depend on the signal itself.

The other problems in the calculation of phase congruency, particularly its sensitivity to noise, are addressed in subsequent sections. However, first it is explained how phase congruency can be calculated from the outputs of filters in quadrature.

Calculating phase congruency via filters in quadrature

In a biological system, localized frequency information is available through banks of filters in quadrature tuned to different spatial frequencies. Logarithmic Gabor functions, as suggested by Field (1987), are a convenient filter model to use. These are filters having a Gaussian transfer function when viewed on the logarithmic frequency scale. Log Gabor filters allow arbitrarily large bandwidth filters to be constructed while still maintaining a zero DC component in the even-symmetric filter. (A zero DC value cannot be maintained in Gabor functions for bandwidths over 1 octave.) On the linear frequency scale, the log Gabor function has a transfer function of the form

$$\mathcal{G}(\omega) = \mathrm{e}^{\frac{-(\log(\omega/\omega_o))^2}{2(\log(\kappa/\omega_o))^2}} \tag{6}$$

where ω_o is the filter's centre frequency. To obtain constant shape ratio filters[1] the term κ/ω_0 must also be held constant for varying ω_0. For example, a κ/ω_0 value of .75 will result in a filter bandwidth of approximately one octave, and a value of .55 will result in a two-octave bandwidth. It is of interest to note that the spatial extent of log Gabor filters appears to be minimized when they are constructed with a bandwidth of approximately two octaves (Kovesi 1996a).

[1] That is, filters that are all geometric scaling of some reference filter.

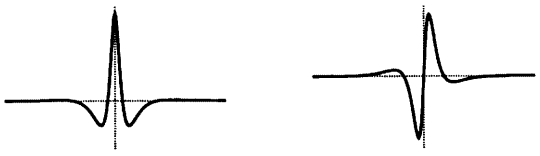

Fig. 4 Even and odd log Gabor wavelets, each having a bandwidth of two octaves

The design of the filter bank needs to be such that the transfer function of each filter overlaps sufficiently with its neighbours so that the sum of all the transfer functions forms a relatively uniform coverage of the spectrum. We wish to retain a broad range of frequencies in our signal because phase congruency is only of interest if it occurs over a wide range of frequencies.

If we let I denote the signal and M_n^e and M_n^o denote the even-symmetric (cosine) and odd-symmetric (sine) filters at a scale n, we can think of the responses of each quadrature pair of filters as forming a response vector:

$$[e_n(x), o_n(x)] = [I(x) * M_n^e, I(x) * M_n^o] \tag{7}$$

The amplitude of the transform at a given filter scale is given by

$$A_n(x) = \sqrt{e_n(x)^2 + o_n(x)^2} \tag{8}$$

and the phase is given by

$$\phi_n(x) = atan2(o_n(x), e_n(x)) \tag{9}$$

At each point x in a signal, we will have an array of these response vectors, one vector for each scale of filter[2]. These response vectors form the basis of our localized representation of the signal, and they can be used in exactly the same way as Fourier components can be used to calculate phase congruency.

An estimate of $F(x)$ can be formed by summing the even filter convolutions. Similarly, $H(x)$ can be estimated from the odd filter convolutions:

$$F(x) \simeq \sum_n e_n(x) \tag{10}$$

$$H(x) \simeq \sum_n o_n(x) \quad \text{and} \tag{11}$$

$$\sum_n A_n(x) \simeq \sum_n \sqrt{e_n(x)^2 + o_n(x)^2} \tag{12}$$

With these three components we are able to calculate phase congruency.

Noise

The crucial problem with phase congruency is its response to noise. Figure 5 illustrates the phase congruency of a step function with and without noise. In the

[2] Note that from now on, n will be used to refer to filter scale. (Previously, n denoted frequency in the Fourier series of a signal.)

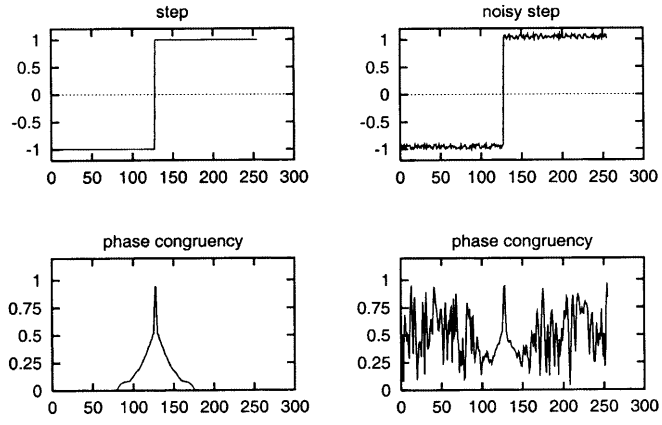

Fig. 5 Phase congruency of a step function with and without noise

vicinity of the step, phase congruency is only high at the point of the step. However, away from the step, the fluctuations due to noise are considered to be significant relative to the surrounding signal (which is noise). This will occur no matter how small the noise level is. This is the price one pays for using a normalized measure such as phase congruency.

The human visual system can adjust to illumination variations of several orders of magnitude. Clearly, some kind of adaption and normalization must take place, and somehow this is done without inducing great sensitivity to noise. Any adaptive system that uses some kind of normalization of the signal must be capable of identifying, and compensating for, noise in the signal.

To compensate for the effect of noise in the calculation of phase congruency, we need to identify the level of $E(x)$ that arises naturally from noise in the image (the "circle of noise" in Fig. 3) Once this is determined, phase congruency can then be calculated with respect to the amount that $E(x)$ exceeds this noise level. It is possible to estimate the level of $E(x)$ induced by noise if we make the following three assumptions: Image noise is additive, the noise power spectrum is constant, and features, such as edges, only occur at isolated locations in an image. While these assumptions may be considered simplistic, given the limited and sometimes conflicting data on the nature of noise in real images – see, for example, Fleck (1992b) and McIvor (1990) – it can be argued that one has little basis on which to build a more formal model.

In the following discussion we shall use the following expression for energy:

$$E = \sqrt{\left(\sum_n e_n\right)^2 + \left(\sum_n o_n\right)^2} \tag{13}$$

where e_n and o_n are the outputs of the even and odd symmetric filters at scale n. Energy is the magnitude of a vector sum (Fig. 3). If our noise is Gaussian with random phase, each vector in this sum will be made up of two independent normally distributed variables. Thus, the distribution of the position of each vector will be a 2D Gaussian centred on the origin.

The distribution of the sum of these vectors is obtained by successively convolving the position distributions for the noise vectors at each scale. As these are all 2D Gaussians, the final distribution of the end position of the energy vector will also be Gaussian. However, what we are interested in is the distribution of the *magnitude* of the energy vector. This will have a Rayleigh distribution (Weisstein, 1998) of the form:

$$R(x) = \frac{x}{\sigma_G^2} e^{\frac{-x^2}{2\sigma_G^2}} \tag{14}$$

where σ_G^2 is the variance of the Gaussian distribution describing the position of the total energy vector. The mean of the Rayleigh distribution is:

$$\mu_R = \sigma_G \sqrt{\frac{\pi}{2}} \tag{15}$$

and its variance is given by:

$$\sigma_R^2 = \frac{4-\pi}{2}\sigma_G^2 \tag{16}$$

If one can determine an expected value of energy due to noise, we can use this as an estimate of the mean of the energy's Rayleigh distribution and hence determine its variance. A noise threshold can then be set in terms of a specified number of standard deviations above the mean.

Rather than construct the expected value of E, it is more convenient to estimate E^2. Note that while E will have a Rayleigh distribution, E^2 will have a χ^2 distribution with two degrees of freedom. The expected value of E^2 will correspond to the 2nd moment of the Rayleigh distribution with respect to the origin, this is:

$$\mathbb{E}(E^2) = 2\sigma_G^2 \tag{17}$$

where $\mathbb{E}$ denotes the expected value. The expected value for E^2 in terms of our filter responses is

$$\begin{aligned}
\mathbb{E}(E^2) &= \mathbb{E}\left(\left(\sum_n e_n\right)^2 + \left(\sum_n o_n\right)^2\right) \\
&= \mathbb{E}\left(\left(\sum_n e_n\right)^2\right) + \mathbb{E}\left(\left(\sum_n o_n\right)^2\right) \\
&\quad + \mathbb{E}\left(2\sum_{i<j}(e_i e_j + o_i o_j)\right) \\
&= 2\mathbb{E}\left(\left(\sum_n e_n\right)^2\right) + 4\mathbb{E}\left(\sum_{i<j}(e_i e_j)\right)
\end{aligned} \tag{18}$$

this last step being possible because the distributions of e_n and o_n are identical, but independent.

Given that e_n is obtained by convolving the noise signal g with a filter M_n, and denoting the Fourier transform $\mathcal{F}(f) = \hat{f}$ we obtain:

$$\begin{aligned}\mathbb{E}(E^2) &= 2\mathbb{E}\left(\sum_n (M_n * g)^2\right) + 4\mathbb{E}\left(\sum_{i<j} (M_i * g)\cdot(M_j * g)\right)\\ &= 2\mathbb{E}\left(\sum_n (\hat{M}_n \cdot \hat{g})^2\right)\\ &\quad + 4\mathbb{E}\left(\sum_{i<j} \mathscr{F}^{-1}(\hat{M}_i \cdot \hat{g}) * (\hat{M}_j \cdot \hat{g})\right)\\ &= 2|\hat{g}|^2\mathbb{E}\left(\sum_n \hat{M}_n^2\right) + 4\mathbb{E}\left(\sum_{i<j} \mathscr{F}^{-1}(|\hat{g}|^2 \cdot (\hat{M}_i * \hat{M}_j)\right)\\ &= 2|\hat{g}|^2\mathbb{E}\left(\sum_n M_n^2\right) + 4|\hat{g}|^2\mathbb{E}\left(\sum_{i<j} (M_i \cdot M_j)\right) \qquad (19)\end{aligned}$$

Note that we are assuming that g has zero mean and $|\hat{g}|$ is constant. The components of Equation 19 involving the filters M_n can be evaluated numerically, but what we do not know is the amplitude of the noise spectrum, $|\hat{g}|$. However, we can estimate $|\hat{g}|$ from the response of the smallest scale filter pair in the wavelet bank as follows.

The smallest scale filter has the largest bandwidth, and as such will give the strongest noise response. Only at feature points will the response differ from the background noise response, but the regions where it will be responding to features will be small due to the small spatial extent of the filter. Thus, the distribution of the squared amplitude response from the smallest scale filter pair across the whole image will be primarily the noise distribution, a scaled 2 DOF χ^2 distribution, with some contamination as a result of the response of the filters to feature points in the image.

We can obtain a robust estimate of the mean of the squared amplitude response of the smallest scale filter via the median response. The median of a 2 DOF χ^2 distribution is the value x such that:

$$\int_0^x \frac{1}{2} e^{\frac{-x}{2}} = \frac{1}{2} \qquad (20)$$

$$\Rightarrow \text{median} = -2\ln(1/2)\ . \qquad (21)$$

Noting that the mean of a 2 DOF χ^2 distribution is 2, we obtain:

$$\mathbb{E}(A_N^2) = \frac{-\text{median}(A_N^2)}{\ln(1/2)} \qquad (22)$$

where N is the index of the smallest scale filter. This allows us to form the estimate:

$$|\hat{g}|^2 \simeq \frac{\mathbb{E}(A_N^2)}{\mathbb{E}(\hat{M}_N^2)} \qquad (23)$$

This can be substituted back into Equation 19 to obtain a value for $\mathbb{E}(E^2)$, then, using Equations 17, 16, and 15, we can obtain the mean μ_R and variance σ_R^2 of the Rayleigh distribution describing the noise energy response.

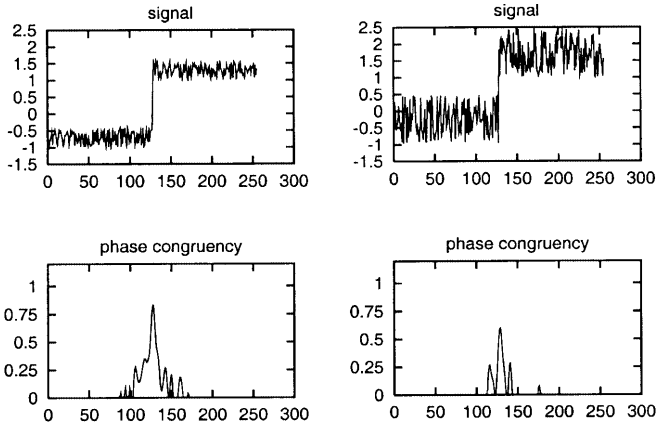

Fig. 6 Noise compensated phase congruency of two step profiles

The radius, T of the noise circle (shown in Fig. 3) is taken to be the mean noise response plus some multiple, k, of σ_R

$$T = \mu_R + k\sigma_R \qquad (24)$$

where typically k is in the range 2 to 3. If we subtract this estimated noise effect from the local energy before normalizing it by the sum of the wavelet response amplitudes, we will eliminate spurious responses to noise. Thus, we modify the expression for phase congruency to the following:

$$PC(x) = \frac{\lfloor E(x) - T \rfloor}{\sum_n A_n(x) + \varepsilon} \qquad (25)$$

where $\lfloor\ \rfloor$ denotes that the enclosed quantity is equal to itself when its value is positive, and zero otherwise. This approach to noise compensation has parallels to Donoho's techniques for de-noising via soft thresholding (Donoho, 1992).

The phase congruency of a legitimate feature will be reduced according to the magnitude of the noise's local energy relative to the feature. Thus, we end up with a measure of confidence that the feature is significant relative to the level of noise. Figure 6 shows the results of processing two noisy step profiles. In both cases, a k value of 3 was used to estimate the maximum influence of noise no local energy.

The importance of frequency spread

Clearly, a point of phase congruency is only significant if it occurs over a wide range of frequencies. In the degenerate case where there is only one frequency component (a pure sine wave), phase congruency will be 1 everywhere. A more common situation is where a feature has undergone Gaussian smoothing. The smoothing reduces the high frequency components in the signal and reduces the frequency spread.

Thus, as a measure of feature significance, phase congruency should be weighted by some measure of the

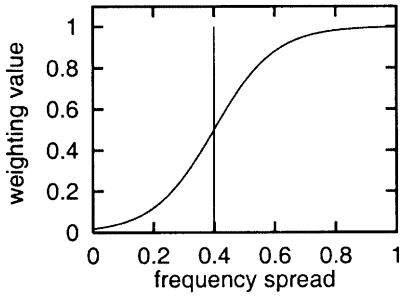

Fig. 7 Frequency spread weighting function with a cut-off value of 0.4 and γ value of 10

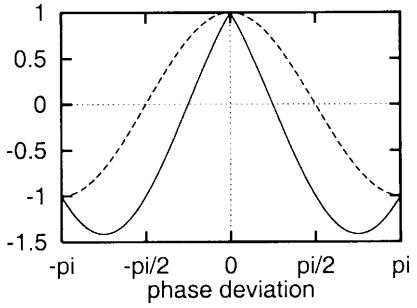

Fig. 8 Comparison between $\cos(x)$ (*dotted line*) and $\cos(x) - |\sin(x)|$ (*solid line*)

spread of frequencies present. What, then, is a significant distribution of frequencies? The amplitude spectrum of a square wave is of the form $1/\omega$. With a set of geometrically scaled filters, the bandwidth of each filter is proportional to its centre frequency. On a spectrum of this kind, the net result is that filter responses are uniform, independent of filter centre frequency. Field (1987) points out that in many cases, images of natural scenes have overall spectral distributions that fall of inversely proportional to frequency, and for this reason he also advocates the use of geometrically scaled filter banks. Under these conditions, filters at all scales will, on average, be responding with equal magnitudes. This is likely to maximize the precision of any computation (numerical or neural) that we make with the filter outputs.

However, if one considers a delta function (corresponding to a line feature), which has a uniform amplitude spectrum, the distribution of filter responses will be strongly skewed to the high frequency end. On the other hand, for a triangular waveform, where the amplitude spectrum falls off at $1/\omega^2$, the distribution of the filter amplitude responses will be strongly skewed to the low frequency end.

Thus, the difficulty we face here is that there is no one ideal distribution of filter responses. All we can say is that the distribution of filter responses should not be too narrow in some general sense. We can also say that a uniform distribution is of particular significance, as step discontinuities are common in images.

Accordingly, we can construct a weighting function that devalues phase congruency at locations where the spread of filter responses is narrow. A measure of filter response spread can be generated by taking the sum of the amplitudes of the responses and dividing by the highest individual response to obtain some notional "width" of the distribution. If this is then normalized by the number of scales being used, we obtain a fractional measure of spread that varies between 0 and 1. This spread is given by:

$$s(x) = \frac{1}{N}\left(\frac{\sum_n A_n(x)}{A_{max}(x) + \varepsilon}\right) \tag{26}$$

where N is the total number of scales being considered, $A_{max}(x)$ is the amplitude of the filter pair having maximum response at x, and ε is used to avoid division by zero and to discount the result should both $\sum A_n(x)$ and $A_{max}(x)$ be very small.

A phase congruency weighting function can then be constructed by applying a sigmoid function to the filter response spread value, namely:

$$W(x) = \frac{1}{1 + e^{\gamma(c - s(x))}} \tag{27}$$

where c is the "cut-off" value of filter response spread, below which phase congruency values become penalized, and γ is a gain factor that controls the sharpness of the cut-off. Note the sigmoid function has been merely chosen for its simplicity and ease of manipulation. Thus,

$$PC(x) = \frac{W(x)\lfloor E(x) - T \rfloor}{\sum_n A_n(x) + \varepsilon} \quad . \tag{28}$$

Weighting by frequency spread, as well as reducing spurious responses where the frequency spread is low, has the additional benefit of sharpening the localization of features, especially those that have been smoothed.

A new measure of phase congruency

Even with a addition of frequency spread weighting, one finds that the localization of phase congruency remains poor on blurred features. The reason for this is evident when one studies the expression for energy. Energy is proportional to the cosine of the deviation of phase angle, $\phi_n(x)$ from the overall mean phase angle, $\bar{\phi}(x)$. While the cosine function is maximized when $\phi_n(x) = \bar{\phi}(x)$, it requires a significant difference between $\phi_n(x)$ and $\bar{\phi}(x)$ before its value falls appreciably.

If one incorporates the sine of the phase difference, in addition to the cosine, we can construct a more sensitive phase deviation measure:

$$\Delta\Phi(x) = \cos(\phi_n(x) - \bar{\phi}(x)) - |\sin(\phi_n(x) - \bar{\phi}(x))| \tag{29}$$

Figure 8 plots this function along with the cosine function for comparison. The function falls very nearly linearly as phase deviation moves from 0 to $\pm\pi/2$. Thus, a near direct measure of phase deviation is obtained.

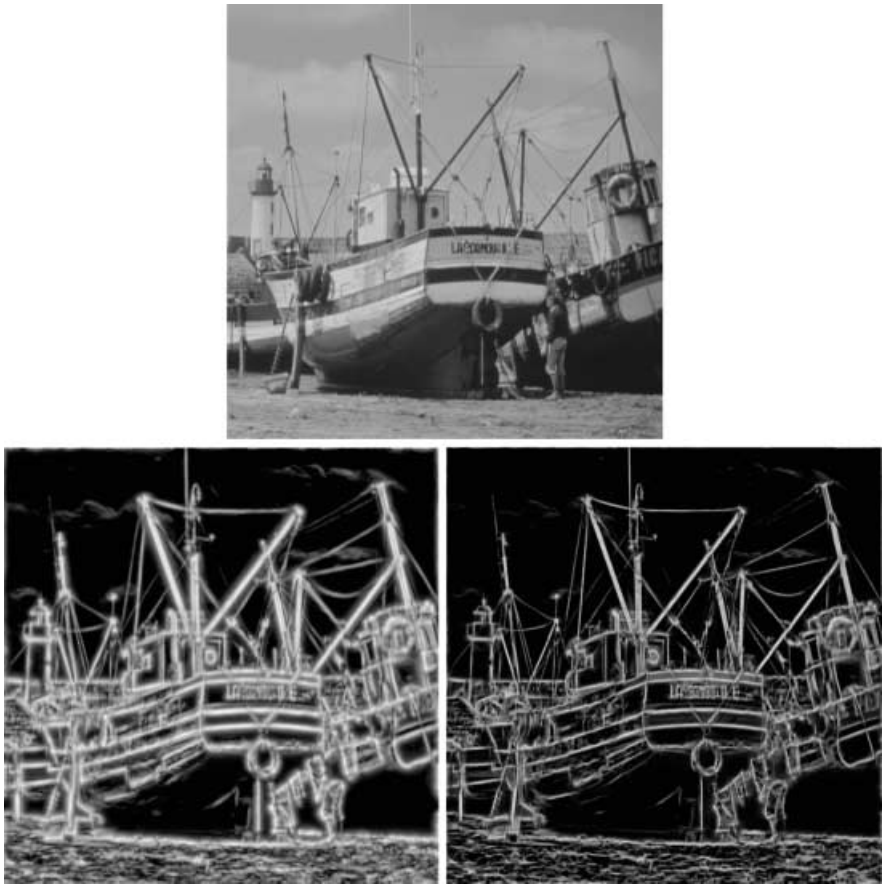

Fig. 9 Raw phase congruency calculated on the Boat image comparing the localization achieved with phase congruency measures PC_1 (*left*) and PC_2 (*right*)

Using this new measure of phase deviation, $\Delta\Phi(x)$, a new measure of phase congruency can be defined as:

$$PC_2(x) = \frac{\sum_n W(x)\lfloor A_n(x)\Delta\Phi_n(x) - T\rfloor}{\sum_n A_n(x) + \varepsilon} \tag{30}$$

where, as before, ε is a small constant to avoid division by zero and T is the estimated noise influence[3]. Note that this expression for phase congruency is called $PC_2(x)$ to distinguish it from the previous definition of phase congruency, which will now be referred to as $PC_1(x)$.

The calculation of $PC_2(x)$ requires the calculation of the cosine and sine of $(\phi_n(x) - \bar{\phi}(x))$. The unit vector representing the direction of the weighted mean phase angle, $\bar{\phi}(x)$ is given by:

$$(\bar{\phi}_e(x), \bar{\phi}_o(x)) = \frac{1}{\sqrt{(F(x)^2 + H(x)^2)}}(F(x), H(x)) \tag{31}$$

Now, using the magnitude of dot and cross products between the filter response vectors, one can calculate the new weighted phase deviation measure directly from the filter outputs:

$$\begin{aligned} A_n(x)(\cos(\phi_n(x) - \bar{\phi}(x)) - |\sin(\phi_n(x) - \bar{\phi}(x))|) \\ = (e_n(x) \cdot \bar{\phi}_e(x) + o_n(x) \cdot \bar{\phi}_o(x)) \\ - |e_n(x) \cdot \bar{\phi}_o(x) - o_n(x) \cdot \bar{\phi}_e(x)| \end{aligned} \tag{32}$$

Figure 9 provides an example of how the PC_2 measure produces a more localized response to features and allows better detection of the detail in images. Other measures similar to the PC_2 measure could be devised. However, the advantage of the expression derived above is that one obtains an approximation to the absolute phase deviation directly from the filter outputs, without having to employ inverse trigonometric functions.

Extension to two dimensions

So far, our discussion has been limited to signals in one dimension. Calculation of phase congruency requires the formation of a 90° phase shift of the signal, which we have done using odd-symmetric filters. As one cannot construct rotationally symmetric odd-symmetric filters, one is forced to analyze a two-dimensional image by applying our one-dimensional analysis over several orientations and combining the results in some way. There are two issues to be resolved: The shape of the filters in

[3] This equation for phase congruency does not lend itself readily to the noise analysis described in section 4 (Noise). In practice, it is found that the analysis used in that section approximately applies to the PC_2 measure, but typically the noise effect is overestimated. This can be compensated for by rescaling the value for T calculated in Equation 24 by a factor of 0.5 to 0.7.

two dimensions, and the way in which the results from each orientation are combined.

2D filter design

A logical way to construct 2D filters in the frequency domain is to use polar-separable 2D Gaussians. In the radial direction, along the frequency axis, the filters are log Gaussians with geometrically increasing centre frequencies and bandwidths. In the angular direction, the filters have Gaussian cross-sections. Thus, the cross-section of the transfer function in the angular direction is:

$$G(\theta) = \mathrm{e}^{-\frac{(\theta-\theta_0)^2}{2\sigma_\theta^2}} \tag{33}$$

where θ_0 is the orientation angle of the filter, and σ_θ is the standard deviation of the Gaussian spreading function in the angular direction. This is set to be some fixed ratio of the orientation spacing between the filters to ensure even coverage of the spectrum in all orientations. The final arrangement of filters results in a "rosette" of overlapping polar-separable 2D Gaussians in the frequency plane.

There is a good reason for choosing a Gaussian spreading function in the angular direction. We are interested in the phase information in the image, and the important thing to ensure is that convolution with the spreading function does not corrupt this phase data. Any funtion convolved with a Gaussian undergoes amplitude modulation of its components, but phase is unaffected. If, on the other hand, we were to, say, use a rectangular spreading function, some phase angles in the signal would be reversed because the transfer function (as sine function) has negative values. Phase congruency at features would then be corrupted.

Combining data over several orientations

The approach that has been adopted is as follows: At each location in the image we calculate energy, $E(x)$ in each orientation, compensate for the influence of noise, apply the weighting for frequency spread, and then form the sum over all orientations. This sum of energies is then normalized by dividing by the sum over all orientations and scales of the amplitudes of the individual filter responses at that location in the image. This produces the following equation for 2D phase congruency:

$$PC_2(x) = \frac{\sum_o \sum_n W_o(x) \lfloor A_{no}(x)\Delta\Phi_{no}(x) - T_o \rfloor}{\sum_o \sum_n A_{no}(x) + \varepsilon} \tag{34}$$

where o denotes the index over orientations. It is also important that the normalization of energy to form phase congruency is done *after* summing energies over all orientations. This ensures that the result from each orientation contributes to the overall normalised result in proportion to its energy. Notice in the equation above that noise compensation is performed in each orientation idependently. This has been found to give significantly better results. The perceived noise content as deduced from the average power response of the smallest scale filter pair can vary significantly with orientation. This is due to the correlation in noise along scan lines that can occur in the digitization of an image.

Scale via high-pass filtering

The traditional approach to analyzing an image at different scales is to consider various low-pass or band-passed versions of the image. A major problem with this approach to multi-scale analysis is that the number of features present in an image, and their locations, vary with the scale used. It seems very unsatisfactory for the location of a feature to depend on the scale at which it is analyzed. *Feature locations should not be a function of the scale of analysis, only the relative significance of features should change*.

Much of the thinking about scale has been strongly influenced by the problems in applying differential operators to images and the need to suppress noise. The use of phase congruency to measure feature significance allows one to consider an alternative interpretation of feature scale. Phase congruency at some point in a signal depends on how the feature is built up from the local frequency components. Depending on the size of the analysis window, features some distance from the point of interest may contribute to the local frequency components considered to be present. Thus, features are not considered in isolation, but in context with their surrounding features.

Therefore, as far as phase congruency, the natural scale parameter to vary is the size of the window in the image over which we perform the local frequency analysis. In the context of our use of filters to calculate phase congruency, the scale of analysis is specified by the spatial extent of the largest filter in the filter bank. With this approach, we are using *high-pass* filtering to specify the analysis scale. We cut out low frequency components (those having wavelengths larger than the window size), while leaving the high frequency components intact.

If we use a small analysis window, each feature will be treated with a great degree of independence from other features in the image. We will only be comparing each feature to a small number of other features that are nearby, and hence each feature is likely to be perceived more important locally. At the largest scale (window size equal to image size), each feature is considered in relation to all other features, and we obtain a sense of global significance for each feature.

In summary, it is proposed that multi-scale analysis be done by considering phase congruency of differing high-passed versions of an image. The high-pass images are constructed from the sum of band-passed images, with the sum ranging from the highest frequency band down to some cut-off frequency. With this approach, no matter what scale we consider, all features are localized

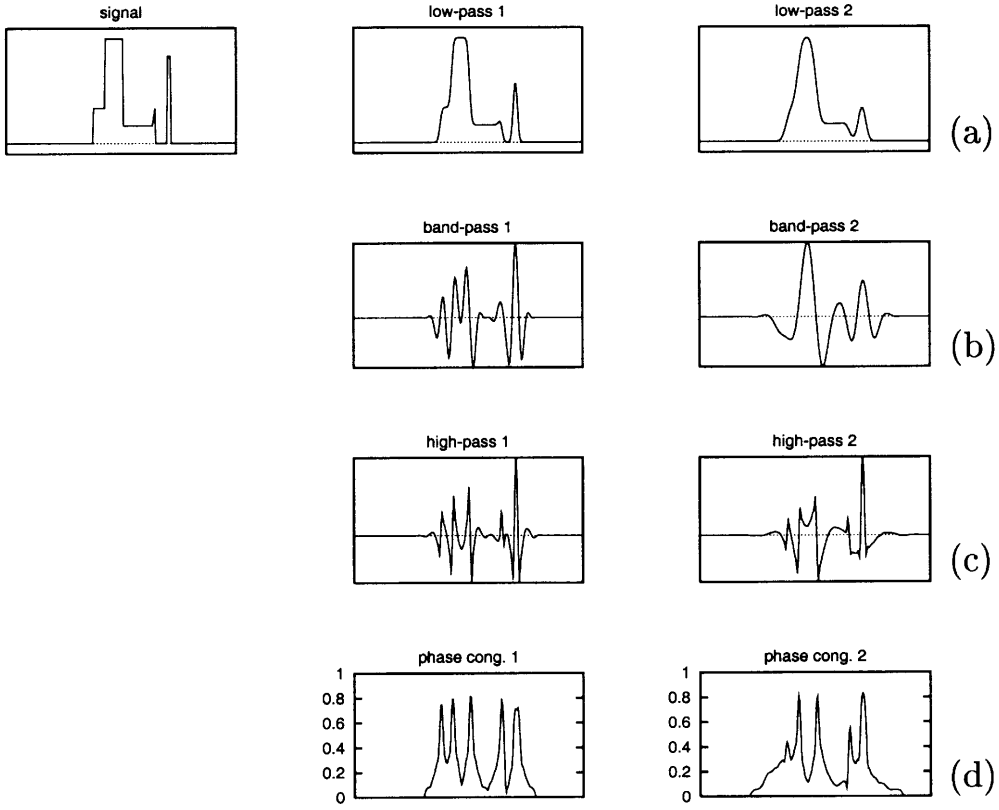

Fig. 10 Analysis at different scales. **a** A 1D signal and two different low-pass versions of the signal. **b** Two different band-pass versions of the signal. **c** Two different high-pass versions of the signal. **d** Phase congruency at both scales of high-pass filtering shown in **c**. Note that the number and location of features as measured by phase congruency remains constant, and only their relative significance varies with scale. Under low-pass and band-pass filtering, the number and locations of features varies

precisely and in a stable manner. There is no "drift" of features that occurs with low-pass filtering. All that changes with analysis at different scales is the *relative* significance of features (Fig. 10).

Experimental results

The performance of the phase congruency detector is demonstrated on two test images and on two real images on the following pages. For comparison, the output of the Canny detector is also presented. The implementation of the Canny detector used here follows the modifications suggested by Fleck (1992a). The raw, gradient magnitude image is displayed so that comparison can be made without having to consider any artifacts that may be introduced by non-maximal suppression and thresholding processes. The purpose of providing this comparison is to illustrate some of the qualitative differences in performance between the two detectors. Quantitative comparisons are difficult, because the design objectives of the two detectors are completely different. One of them is seeking to localize step edges, and the other is seeking to identify points of phase congruency.

The main qualitative difference between the two detectors is the wide range of response values from the Canny operator relative to the phase congruency detector. This illustrates the difficulty of threshold selection for the Canny operator. The other obvious difference is that the Canny detector produces twin responses on features that have congruency of phase at angles other than 0 degrees (that is, features other than step edges), whereas the phase congruency detector produces a single uniform response no matter what the angle of congruence is (Fig. 11). This illustrates the difficulty in specifying quantitative performance criteria for feature detectors that are relevant to human visual perception. The Canny operator was developed to optimally detect step edges in the presence of noise. Clearly, this performance criterion has nothing to do with human visual perception.

It should be emphasized that all the results presented on the following pages were obtained by applying the same parameter and threshold values to every image. The uniformity of the results demonstrates the invariant properties of phase congruency.

The raw phase congruency images were obtained by applying Equation 34 to the images with following parameters: Local frequency information was obtained using two octave bandwidth filters over four scales and six orientations. The wavelength of the smallest scale filters was 3 pixels, the scaling between successive filters was 2. The filters were constructed directly in the frequency domain as polar separable functions – a logarithmic Gaussian function in the radial direction and a Gaussian in the angular direction. In the angular direction, the ratio between the angular spacing of the filters and angular standard deviation of the Gaussians was 1.2. This results in a coverage of the spectrum that varies by less than 1%. A noise compensation k value of 2.0 was used. The frequency spread weighting function cut-off fraction c, was set at 0.4, and the gain parameter γ, was set at 10. The

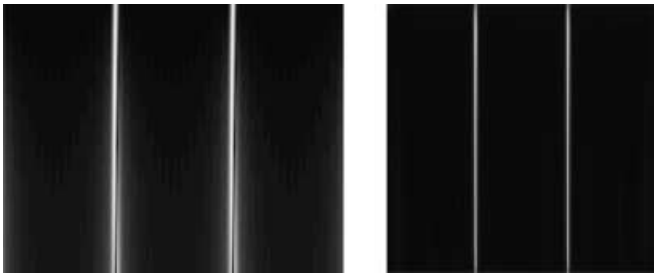

Fig. 11 Output of the Canny detector (*left*) and phase congruency operator (*right*) on the grating shown in Fig. 2. The grating interpolates a step profile at the top (congruence of phase at 0 degrees) to a line feature at the bottom (congruence of phase at 90 degrees). Where the congruence of phase occurs at angles away from 0 degrees, the Canny detector sees two features rather than one

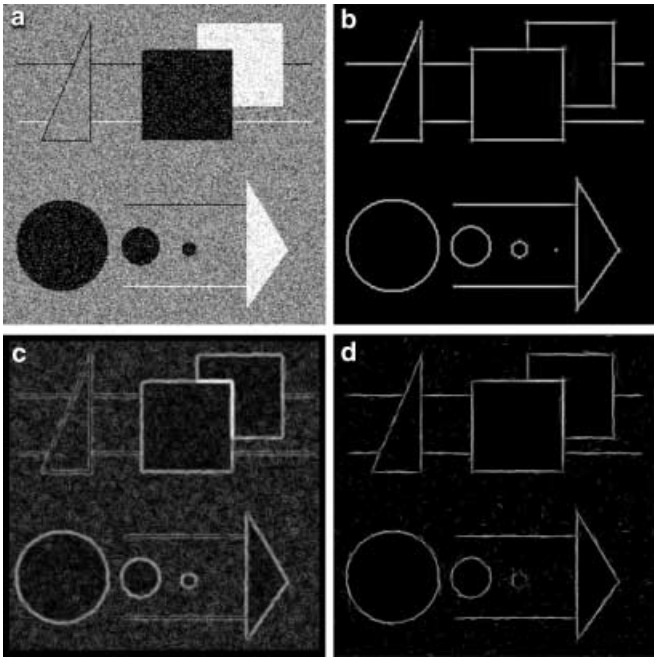

Fig. 12 Illustration of noise compensation. **a** Test image with additive Gaussian noise. **b** Raw phase congruency on noise-free image. **c** Canny edge strength on noisy image. **d** Raw phase congruency on noisy image. In this example, additive Gaussian noise with a standard deviation of 40 grey levels has been applied to a 256 grey level test image. The noise has been successfully ignored in the smooth regions of the image, though the "confidence" of the measured phase congruency at features has been reduced. Notice how phase congruency marks the line features with a single response, not two, as the Canny operator does, and that the magnitude of the phase congruency response is largely independent of local contrast

value of ε, the small constant used to prevent division by zero in the case where local energy in the image becomes very small, was set at 0.01. None of these parameter values are particularly critical. The phase congruency feature maps were obtained by performing non-maximal suppression on the raw phase congruency images followed by hysteresis thresholding with upper and lower hysteresis threshold values fixed at phase congruency values of 0.3 and 0.15.

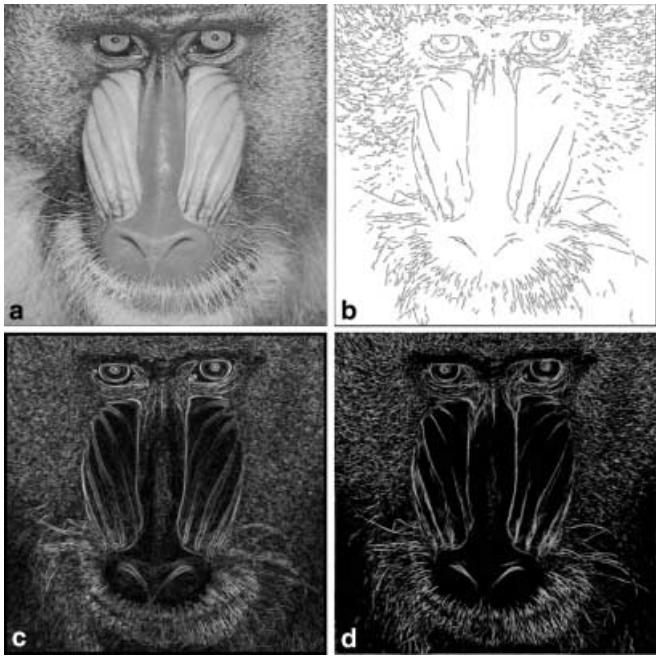

Fig. 13 Mandrill image. **a** Original image. **b** Phase congruency feature map. **c** Canny edge strength image. **d** Raw phase congruency image. This image is largely made up of line features, and this highlights the difference between phase congruence and first derivative edge operators. The Canny detector marks edges *around* all the hairs, while phase congruency marks the hairs directly as line features

Fig. 14 Goldhill image. **a** Original image. **b** Phase congruency feature map. **c** Canny edge strength image. **d** Raw phase congruency image. This image illustrates the ability of the PC_2 measure to pick out fine features. The window panes and roof tiles in the nearer houses are clearly marked

MATLAB code for performing the calculation of phase congruency, non-maximal suppression, and hysteresis thresholding is provided by Kovesi (1996b) for those wishing to replicate the results presented here.

Conclusion

We have argued for the importance of invariant low-level feature detectors in vision. These are building blocks of higher-level processes. Invariance in the outputs of low-level operators ensures reliability of these subseqent higher-level operations. Phase congruency provides a low-level invariant measure of the significance of features, but its sensitivity to noise highlights the fundamental problem of noise that any visual system must overcome if it is to use any form of normalization to achieve invariance to contrast.

This paper extends the theory behind the calculation of phase congruency in a number of ways. An effective method of noise compensation is presented that only assumes that the noise power spectrum is approximately constant. Problems with the localization of features are addressed by introducing a new, more sensitive measure of phase congruency. It is shown that weighting phase congruency by some measure of the spread of the frequencies present at a feature is also important. It is observed that when geometrically scaled filters are used, a uniform distribution of responses is a particularly significant one. This distribution matches typical spectral statistics of images and corresponds to the distribution that arises at step discontinuities. It has also been shown how phase congruency can be calculated via log Gabor filters, and the particular issues in calculating phase congruency in 2D images are covered.

Another contribution of this work is to offer a new approach to the concept of scale in image analysis. The natural scale parameter to vary in the calculation of phase congruency is the size of the analysis window over which to calculate local frequency information. Thus, under these conditions, scale is varied using high-pass filtering rather than low-pass or band-pass filtering. The significant advantage of this approach is that feature locations remain constant over scale, and only their significance relative to each other varies.

Invariant measures are important in low-level vision. Phase information is a good basis on which to construct such measures.

Acknowledgements The author would like to acknowledge many useful discussions with John Ross, James Trevelyan, Robyn Owens, Ben Robbins, Chris Pudney, Adrian Baddeley, and Concetta Morrone.

References

Adelson, E. H., and Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*, 284–299.

Canny, J. F. (1983). *Finding edges and lines in images* (TR-720). Master's thesis, Masschusettes Institute of Technology, Cambridge, MA: AI Lab.

Canny, J. F. (1986). A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, *8*, 112–131.

Donoho, D. L. (1992). *De-noising by soft thresholding* (Technical Report 409). Stanford University, CA: Department of Statistics.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of The Optical Society of America A*, *4*, 2379–2394.

Fleck, M. M. (1992a). Multiple widths yield reliable finite differences. *IEEE T-PAMI*, *14*, 412–429.

Fleck, M. M. (1992b). Some defects in finite-difference edge finders. *IEEE T-PAMI*, *14*, 337–345.

Florack, L. M. J., ter Harr Romeny, B. M., Koenderink, J. J., and Viergever, M. A. (1992). Scale and the differential structure of images. *Image and Vision Computing*, *10*, 376–388.

Freeman, W. T. (1992). *Steerable filters and local analysis of image structure*. PhD thesis, MIT Media Lab. TR-190.

Grossman, A. (1988). Wavelet transforms and edge detection. In Albeverio, S., Blanchard, P., Hazewinkel, M., and Streit, L. (Eds.), *Stochastic processes in physics and engineering*, (pp. 149–157). D. Reidel Publishing Company.

Heeger, D. J. (1987). Optical flow from spatiotemporal filters. *Proceedings, 1st International Conference on Computer Vision*, 181–190. London.

Heeger, D. J. (1988). Optical flow using spatiotemporal filters. *International Journal of Computer Vision*, *1,* 279–302.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience, 9*, 181–197.

Hu, M. K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions in Information Theory*, *IT-8*, 179–187.

Koenderink, J. J., and van Doorn, A. J. (1975). Invariant properties of the motion parallax field due to the movement of rigid bodies relative to an observer. *Optica ACTA*, *22*, 773–791.

Kovesi, P. D. (1991). A dimensionless measure of edge significance. *The Australian Pattern Recognition Society*, *Conference on Digital Image Computing: Techniques and Applications*, 281–288. Melbourne.

Kovesi, P. D. (1993). A dimensionless measure of edge significance from phase congruency calculated via wavelets. *First New Zealand Conference on Image and Vision Computing*, 87–94. Auckland.

Kovesi, P. D. (1996a). *Invariant measures of image features from phase information*. PhD thesis, The University of Western Australia.

Kovesi, P. D. (1996b). MATLAB code for calculating phase congruency and phase symmetry/asymmetry. Available: http://www.cs.uwa.edu.au/~pk/Research/research.html.

Kundu, M. K., and Pal, S. K. (1986). Thresholding for edge detection using human psychovisual phenomena. *Pattern Recognition Letters*, *4*, 433–411.

Marr, D., and Hildreth, E. C. (1980). Theory of edge detection. *Proceedings of the Royal Society*, *London B*, *207*, 187–217.

McIvor, A. M. (1990). A test of camera noise models. *British Machine Vision Conference*, 355–359. Oxford.

Morrone, M. C., and Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society*, *London B*, *235*, 221–245.

Morrone, M. C., Navangione, A., and Burr, D. (1995). An adaptive approach to scale selection for line and edge detection. *Pattern Recognition Letters*, *16*, 667–677.

Morrone, M. C., and Owens, R. A. (1987). Feature detection from local energy. *Pattern Recognition Letters*, *6*, 303–313.

Morrone, M. C., Ross, J. R., Burr, D. C., and Owens, R. A. (1986). Mach bands are phase dependent. *Nature*, *324(6094)*, 250–253.

Mundy, J. L., and Zisserman, A. (Eds.). (1992). *Geometric invariance in computer vision*. Artificial Intelligence Series. Cambridge, MA: MIT Press

Owens, R. A. (1994). Feature-free images. *Pattern Recognition Letters*, *15*, 35–44.

Owens, R. A., Venkatesh, S., and Ross, J. (1989). Edge detection is a projection. *Pattern Recognition Letters*, *9*, 223–244.

Perona, P., and Malik, J. (1990). Detecting and localizing edges and composed of steps, peaks and roofs. *Proceedings of 3rd International Conference on Computer Vision*, 52–57. Osaka.

Pringle, K. K. (1969). Visual perception by a computer. In Grasselli, A. (Ed.), *Automatic interpretation and classification of images,* 277–284. Academic Press, New York.

Robbins, B., and Owens, R. (1997). 2D feature detection via local energy. *Image and Vision Computing*, *15*, 353–368.

Ronse, C. (1993). On idempotence and related requirements in edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*, 484–491.

Ronse, C. (1997). The phase congruence model for edge detection in multi-dimensional pictures (Technical report 97/16). Département d'Informatique, Université Louis Pasteur, Strasbourg, France.

Rosenthaler, L., Heitger, F., Kubler, O., and Heydt, R. V. (1992). Detection of general edges and keypoints. In *ECCV92*, *Springer-Verlag lecture notes in computer science*, *588,* 78–86. Santa Margherita Ligure, Italy: Springer-Verlag.

Venkatesh, S., and Owens, R. A. (1989). An energy feature detection scheme. *The International Conference on Image Processing*, 553–557. Singapore.

Venkatesh, S., and Owens, R. (1990). On the classification of image features. *Pattern Recognition Letters*, *11*, 339–349.

Wang, Z., and Jenkin, M. (1992). Using complex Gabor filters to detect and localize edges and bars. In Archibald, C. and Petriu, E. (Eds.), *Advances in machine vision: Strategies and applications*, *world scientific series in computer science*, *32,* 151–170. Singapore: World Scientific Press.

Weisstein, E. W. (1998). *The CRC concise encyclopedia of mathematics*. CRC Press.

ANNOUNCEMENT

The Department of Psychology at Humboldt-University, Berlin, Germany, will celebrate its 100th anniversary on 11 and 12 January 2001. Over the years, the department has been home to many well-known and influential psychologists, including Ebbinghaus, Köhler, Stumpf, Gottschaldt, Wertheimer, Lewin, and Klix.

Part of the 100th anniversary celebration will be a scientific program with internationally known psychologists from diverse areas of psychology. Speakers include P.B. Baltes, M. Denis, M. Heidelberger, W. Kintsch, F. Klix, J. Kruschke, D. Magnusson, R. Näätänen, L. Squire, B. Velichkovsky, and K. Zimmer.

The Department of Psychology invites all interested scientists to attend the 2-day celebration. Detailed information is available under: http://www.psychologie.hu-berlin.de/ or can be obtained from Peter A. Frensch (e-mail: psy.direktor@psychologie.hu-berlin.de; Tel.: 49 + 30-20246622).