Stability of Phase Information[1]

David J. Fleet
Department of Computing and Information Science,
Queen's University,
Kingston, Canada K7L 3N6

Allan D. Jepson
Department of Computer Science,
University of Toronto,
Toronto, Canada M5S 1A4

Abstract

This paper concerns the robustness of local phase information for measuring image velocity and binocular disparity. It addresses the dependence of phase behaviour on the initial filters as well as the image variations that exist between different views of a 3d scene. We are particularly interested in the stability of phase with respect to geometric deformations, and its linearity as a function of spatial position. These properties are important to the use of phase information, and are shown to depend on the form of the filters as well as their frequency bandwidths. Phase instabilities are also discussed using the model of phase singularities described by Jepson and Fleet [14]. In addition to phase-based methods, these results are directly relevant to differential optical flow methods and zero-crossing tracking.

1 Introduction

An important class of image matching techniques has emerged based on phase information; that is, the phase behaviour in band-pass filtered versions of different views of a 3-d scene [3, 7, 8, 11, 13, 15, 17, 18, 25, 27, 30, 32, 33]. These include phase-difference and phase-correlation techniques for discrete two-view matching, and the use of the phase gradient for the measurement of image orientation and optical flow.

Numerous desirable properties of these techniques have been reported: For two-view matching, the disparity estimates are obtained with sub-pixel accuracy, without requiring explicit sub-pixel signal reconstruction or sub-pixel feature detection and localization. Matching can exploit all phase values (not just zeros), and therefore extensive use is made of the available signal so that a dense set of estimates is often extracted. Furthermore, because phase is amplitude invariant, the measurements are robust with respect to smooth shading and lighting variations. With temporal sequences of images, the phase difference can be replaced by a temporal phase derivative, thereby producing more accurate

[1]Published in *IEEE Trans. PAMI*, 15(12): 1253–1268, 1993

measurements. These computations are straightforward and local in space-time, yielding efficient implementations on both serial and parallel machines. There is the additional advantage in space-time that different filters can be used initially to decompose the local image structure according to velocity and scale. Then, the phase behaviour in each velocity-tuned channel can be used independently to make multiple measurements of speed and orientation within a single image neighbourhood. This is useful in the case of a single, densely textured surface where there exist several oriented image structures with different contrasts, or multiple velocities due to transparency, specular reflection, shadows, or occlusion [6, 7, 19]. In a recent comparison of several different optical flow techniques, phase-based approaches often produced the most accurate results [1].

But despite these advantages, we lack a satisfying understanding of phase-based techniques and the reasons for their success. The usual justification for phase-based approaches consists of the Fourier shift theorem, an assumption that the phase output of band-pass filters is linear as a function of spatial position, and a model of image translation between different views. But because of the local spatiotemporal support of the filters used in practice, the Fourier shift theorem does not strictly apply. For example, when viewed through a window, a signal and a translated version of it , $W(x)s(x)$ and $W(x)s(x-d)$, are not simply phase-shifted versions of one another with identical amplitude spectra. For similar reasons, phase is not a linear function of spatial position (and time) for almost all inputs, even in the case of pure translation. The quasi-linearity of phase often reported in the literature depends on the form of the input and the filters used, and has not been addressed in detail. Also unaddressed is the extent to which these techniques produce accurate measurements when there are deviations from image translation. Fleet et al. [7, 8] suggested that phase has the important property of being stable with respect to small geometric deformations of the input that occur with perspective projections of 3-d scenes. They showed that amplitude is sensitive to geometric deformation, but they provided no concrete justification for the stability of phase.

This paper addresses several issues concerning phase-based matching, the behaviour of phase information, and its dependence on the band-pass filters. It presents justification for the claims of phase stability with respect to geometric deformations, and of phase linearity. By the stability of some image property, we mean that a small deformation of the input signal causes a similar deformation in the image property, so that the behaviour of that property reflects the structure of the input that we wish to measure. Here, we concentrate on small affine deformations like those that occur between left and right stereo views of a slanted surface. For example, scale variations between left and right binocular views of a smooth surface are often as large as 20% [24]. Although phase deformations do not exactly match input deformations, they are usually close enough to provide reasonable measurements for many vision applications.

We address these issues in the restricted case of 1-d signals, where the relevant deformations are translations and dilations. Using a scale-space framework, we simulate changes in the scale of the input by changing the tuning of a band-pass filter. In this context our concerns include the extent to which phase is stable under small scale perturbations of the input, and the extent to which phase is generally linear through space. These properties are shown to depend on the form of the filters used and their frequency bandwidths. Situations in which phase is clearly unstable and leads to inaccurate matching are also discussed, as are methods for their detection. Although we deal here with one dimension, the stability analysis extends to affine deformations in multiple dimensions.

We begin in Section 2 with a brief review of phase-based matching methods and several comments on the dependence of phase on the initial filters. Section 3 outlines the scale-space framework used in the theoretical development that follows in Sections 4 and 5, and Section 6 discusses phase instabilities in the neighbourhoods of phase singularities. Although the majority of the theoretical results assume white noise as input, Section 7 briefly discusses the expected differences that arise with natural images. Finally, Sections 8 and 9 draw conclusions and outline some topics that require further research.

2 Phase Information from Band-Pass Filters

Phase, as a function of space and/or time, is defined here as the complex argument of a complex-valued band-pass signal. The band-pass signal is typically generated by a linear filter with a complex-valued impulse response (or kernel), the real and imaginery components of which are usually even and odd symmetric. Gabor functions (sinusoidally-modulated Gaussian windows [10]) are perhaps the most commonly used filter [7, 8, 15, 18, 27, 30, 33]. Other choices of filter include: sinusoidally-modulated square-wave (constant) windows [30]; filters with nonlinear phase that cycles between $-\pi$ and π only once within the support width of the kernel [31, 32, 33]; quadrature-pair steerable pyramids constructed with symmetric kernels and approximations to their Hilbert tranforms [9, 28]; log-normal filters [5, 31]; and derivative of Gaussian filters, and real and imaginary parts of which are the first and second directional derivatives of a Gaussian envelope [31]. Phase can also be defined from a single real-valued band-pass signal by creating a complex signal, the real and imaginary parts of which are the original band-pass signal and its Hilbert transform. In all these cases, the expected behaviour of phase depends on the filter.

Given the initial filters, phase-based optical flow techniques define image velocity in terms of the instantaneous velocity of level (constant) phase contours (given by the spatiotemporal phase gradient). Phase-based matching methods, based on two images, define disparity as the shift necessary to align the phase values of band-pass filtered versions of the two signals [13, 27, 30]. In this case, a disparity predictor can be constructed using phase differences between the two views: For example, let $R_l(x)$ and $R_r(x)$ denote the output of band-pass filters applied to left and right binocular signals (along epipolar lines). Their respective phase components are then written as $\phi_l(x) = \arg[R_l(x)]$ and $\phi_r(x) = \arg[R_r(x)]$, and binocular disparity is measured (predicted) as

$$d(x) \;=\; \frac{[\phi_l(x) - \phi_r(x)]_{2\pi}}{k(x)} \tag{1}$$

where $[\phi]_{2\pi}$ denotes the principal part of ϕ, that is, $[\phi]_{2\pi} \in (-\pi,\, \pi]$, and $k(x)$ denotes some measure of the underlying frequency of the signal in the neighbourhood of x. The frequency $k(x)$ may be approximated by the frequency to which the filter is tuned, or measured using the average phase derivative (i.e., $(\phi_l'(x) + \phi_r'(x))/2$) from the left and right filter outputs [8, 18]. The phase derivative is often referred to as the instantaneous frequency of a band-pass signal [2, 26]. Note that linear phase, and therefore constant instantaneous frequency, implies a sinusoidal signal.

One reason for the success of the predictor in (1) is the fact that phase is often nearly linear over relatively large spatial extents, with instantaneous frequencies close to the filter tuning. In other words, the filter output is nearly sinusoidal, so that phase and instantaneous frequency at a single location provide a good model for the local phase structure of the filter response, in terms of a complex signal winding sinusoidally (with some amplitude variation) about the origin of the complex plane. When the left and right signals are shifted versions of one another, and phase is precisely linear over a distance larger than the shift, then the predictor produces an exact measurement. That is, displacements of a linear function $f(x)$ are given by differences in function value divided by the derivative of the function: $d = [f(x) - f(x-d)]/f'(x)$.

In more general cases the predictor produces estimates of disparity, and may be used iteratively to converge to accurate matches of the two local phase signals [8]. The size of the neighbourhood within which phase is monotonic (and therefore unique) determines the range of disparities that may be handled correctly by the predictor, and therefore its domain of convergence. Of course, for a sinusoidal signal the predictor can handle disparities up to half the wavelength of the signal. Because the domain of convergence depends on the wavelength of the band-pass signal, and hence on the scale of the filter, small kernels tuned to high frequencies have small domains of convergence. This necessitates some form of control strategy, such as coarse-to-fine propagation of disparity estimates. A detailed analysis of the domain of convergence, and control strategies are beyond the scope of this paper.

Within the domain of convergence, the accuracy of the predictor, and hence the speed with which it converges, depends on the linearity of phase. As discussed in [8], predictor errors are a function of the magnitude of disparity and higher-order derivatives of phase. The accuracy of the final measurement, to which the predictor has converged, will also depend critically on phase stability (discussed below in detail).

Phase linearity and monotonicity are functions of both the input signal and the form of the filters. With respect to the filters, it is generally accepted that the measurement of binocular disparity and optical flow should require only local support, so that the filters should be local in space-time as well as band-pass. For example, although filter kernels of the form $\exp[ik_0x]$ (i.e. Fourier transforms) produce signals with linear phase, they are not local. It is also important to consider the correlation between real and imaginary components of the kernel, as well as differences in their amplitude spectra. To the extent that they are correlated the phase signal will be nonlinear because the filter response will form an elliptical path, elongated along orientations of $\pm\pi/4$ in the complex plane. If they have different amplitude spectra, then the real and imaginary parts will typically contain different amounts of power, causing elliptical paths in the complex plane elongated along the real or imaginary axes. Finally, it is natural to ensure that the filters have no *dc* sensitivity, for this will often mean that the complex filter response will not wind about the origin, thereby also introducing nonlinearities, and smaller domains of convergence [30, 31]. Fortunately, we can deal with these general concerns relatively easily by employing complex-valued filters with local support in space-time, the imaginary parts of which are Hilbert transforms of the real parts (i.e. quadrature-pair filters). The real and imaginary parts are then uncorrelated with similar amplitude spectra and no dc sensitivity.

Gabors functions have Gaussian amplitude spectra with infinite extent in the frequency domain, and therefore some amount of dc sensitivity. Because of the substantial power in natural signals at low frequencies this dc sensitivity often introduces a positive bias in the real part of the response. This problem can be avoided for the most part with the use of small bandwidths (usually less than one octave), for which the real and imaginary parts of Gabor functions are close approximations to quadrature pairs. It is also possible to modify the real (symmetric) part of the Gabor kernel by subtracting its dc sensitivity from it (e.g., [7]). Log-normal kernels have Gaussian amplitude spectra in log-frequency space, and are therefore quadrature-pair filters. Modulated square-wave filters [30] and derivative-based real and imaginary parts may have problems because their real and imaginary parts do not have the same amplitude spectra. In both cases, the odd-symmetric components of the kernels are more sensitive to low frequencies than the even-symmetric component. This introduces a bias towards phase values of $\pm\pi/2$.

3 Scale-Space Framework

To address questions of phase stability and phase linearity, we restrict our attention to band-pass filters with complex-valued kernels $K(x, \lambda)$, the real and imaginary parts of which form quadrature pairs (i.e. they are Hilbert transforms of one another [26]). Let $\lambda > 0$ denote a scale parameter that determines the frequency pass-band to which the filter is tuned. Also, let the kernels be normalized such that

$$\| K(x, \lambda) \| \;=\; 1 \,, \tag{2}$$

where $\| K(x, \lambda) \|^2 \equiv \langle K(x, \lambda), K(x, \lambda) \rangle$, which is defined by

$$\langle f(\mathbf{x}), g(\mathbf{x}) \rangle \;=\; \int_{-\infty}^{\infty} f(\mathbf{x})^* \, g(\mathbf{x}) \, d\mathbf{x} \,, \tag{3}$$

where f^* denotes the complex conjugate of f. For convenience, we also assume translational invariance and self-similarity across scale (i.e., wavelets [20]), so that $K(x, \lambda)$ satisfies

$$K(x, \lambda) \;=\; \frac{1}{\sqrt{\lambda}} \, K(x/\lambda, 1) \,. \tag{4}$$

Wavelets are convenient since, because of their self-similarity, their octave bandwidth remains constant, independent of the scale to which they are tuned. Our results also extend to filters other than wavelets such as windowed Fourier transforms for which the spatial extent of the effective kernels is independent of λ.

The convolution of $K(x, \lambda)$ with an input signal $I(x)$ is typically written as

$$S(x, \lambda) \;=\; K(x, \lambda) \,*\, I(x) \;. \tag{5}$$

Because $K(x; \lambda)$ is complex-valued, the response $S(x, \lambda) = \mathrm{Re}[S(x, \lambda)] + i\,\mathrm{Im}[S(x, \lambda)]$ is also complex-valued, and can be expressed using amplitude and phase as in $S(x, \lambda) = \rho(x, \lambda)\, e^{i\,\phi(x, \lambda)}$, where

$$\rho(x, \lambda) \;=\; |S(x, \lambda)| \;=\; \sqrt{\mathrm{Re}[S(x, \lambda)]^2 + \mathrm{Im}[S(x, \lambda)]^2} \;, \tag{6a}$$

$$\phi(x, \lambda) \;=\; \arg[S(x, \lambda)] \;=\; \mathrm{Im}[\log S(x, \lambda)] \;\in\; (-\pi, \pi] \;. \tag{6b}$$

The first main concern of this paper is the expected stability of phase under small scale perturbations of the input. If phase is not stable under scale variations between different views, then the phase-based measurements of velocity or binocular disparity will not be reliable. To examine this, we simulate changes in the scale of the input by changing the scale tuning of the filter; if one signal is a translation and dilation of another, as in

$$I_0(a(x)) \;=\; I_1(x) \;, \tag{7}$$

where $a(x) = a_0 + a_1 x$, then, because of the filters' self-similarity, the responses $S_0(x, \lambda_0)$ and $S_1(x, \lambda_1)$ will satisfy

$$\sqrt{a_1}\; S_0(a(x), \lambda_0) \;=\; S_1(x, \lambda_1) \;, \qquad \lambda_1 = \lambda_0 \,/\, a_1 \;. \tag{8}$$

That is, if filters tuned to λ_0 and λ_1 were applied to $I_0(x)$ and $I_1(x)$, then the structure extracted from each view would be similar (up to a scalar multiple) and the filter outputs would be related by precisely the same deformation $a(x)$. Of course, in practice we apply the *same* filters to $I_0(x)$ and $I_1(x)$ because the scale factor a_1 that relates the two views is unknown. For phase-matching to yield accurate estimates of $a(x)$, the phase of the filter output should be insensitive to small scale variations of the input.

The second concern of this paper is the extent to which phase is linear with respect to spatial position. Linearity affects the ease with which the phase signal can be differentiated in order to estimate the instantaneous frequency of the filter response. Linearity also affects the speed and accuracy of disparity measurement based on phase-difference disparity predictors, as well as the typical extent of the domain of convergence; if the phase signal is exactly linear, then the disparity can be computed in just one step, without requiring iterative refinement [8, 18, 27, 30], and the domain of convergence is $\pm\pi$. In practice, because of deviations from linearity and monotonicity, the reliable domain of convergence is usually less than $\pm\pi$.

For illustration, let $K(x,\,\lambda)$ be a Gabor kernel[2] [10], $Gabor(x,\,\sigma(\lambda), k(\lambda))$, where

$$Gabor(x;\,\sigma,\,k) \;=\; G(x;\sigma)\, e^{ixk} \tag{9}$$

$$G(x;\sigma) \;=\; \frac{1}{(\sqrt{\pi}\sigma)^{1/2}}\; e^{-x^2/2\sigma^2}\;. \tag{10}$$

The peak tuning frequency of the Gabor filter is given by

$$k(\lambda) \;=\; \frac{2\pi}{\lambda}\;. \tag{11a}$$

Let the extent of the Gaussian envelope be measured at one standard deviation, and let the bandwidth β be close to one octave. Then, the standard deviation of the amplitude spectra $\sigma_k = \sigma^{-1}$ satisfies $\beta = \log_2[(k(\lambda)+\sigma_k)/(k(\lambda)-\sigma_k)]$. From this, it is straightforward to show that the radius of spatial support is

$$\sigma(\lambda) \;=\; \frac{1}{k(\lambda)}\left(\frac{2^\beta+1}{2^\beta-1}\right)\;. \tag{11b}$$

Figure 1A shows a signal composed of a sample of white Gaussian noise concatenated with a sample of red Gaussian noise.[3] The two middle images show the amplitude and phase components of the scale-space Gabor response, $\rho(x,\,\lambda)$ and $\phi(x,\,\lambda)$; spatial position is shown on the horizontal axis, and log scale is shown on the vertical axis (over two octaves). The bottom images show their level contours. In the context of the scale-space expansion, an image property is said to be *stable* for image matching where its level contours are vertical. Figure 1 shows that $\rho(x,\,\lambda)$ depends significantly on scale as its level contours are not generally vertical. By contrast, the phase structure is generally stable, except for several isolated regions to be discussed below. Gradient-based techniques applied to low-pass or band-pass filtered images produce inaccurate velocity estimates, in part, because they implicitly require that *both* amplitude and phase behaviour be stable with respect to scale perturbations.

The response $S(x,\,\lambda)$ defined in (5) is referred to as a scale-space expansion of $I(x)$. It is similar to band-pass expansions defined by the Laplacian of a Gaussian ($\nabla^2 G$), but it is expressed in terms of complex-valued filters (cf. [16, 20, 34, 35]). Interestingly, zero-crossings of filters derived from directional derivatives of Gaussian envelopes output are equivalent to crossings of constant phase of complex-valued band-pass filters, the imaginary parts of which are Hilbert transforms of the corresponding Gaussian derivatives. Here we use the scale-space framework to investigate the effects of small perturbations of input scale on image properties that might be used for matching. We are not proposing a new multi-scale representation, nor are we interpreting phase behaviour in terms of specific image features such as edges.

[2] As mentioned above, although the real and imaginary parts of Gabor kernels do not have identical amplitude spectra, they are a good appoximation to a quadrature pair for small bandwidths (e.g. less than one octave, measured at one standard deviation of the Gaussian spectrum).

[3] That is, a sample of white noise smoothed with an exponential kernel $\exp(-|x|/5)$.

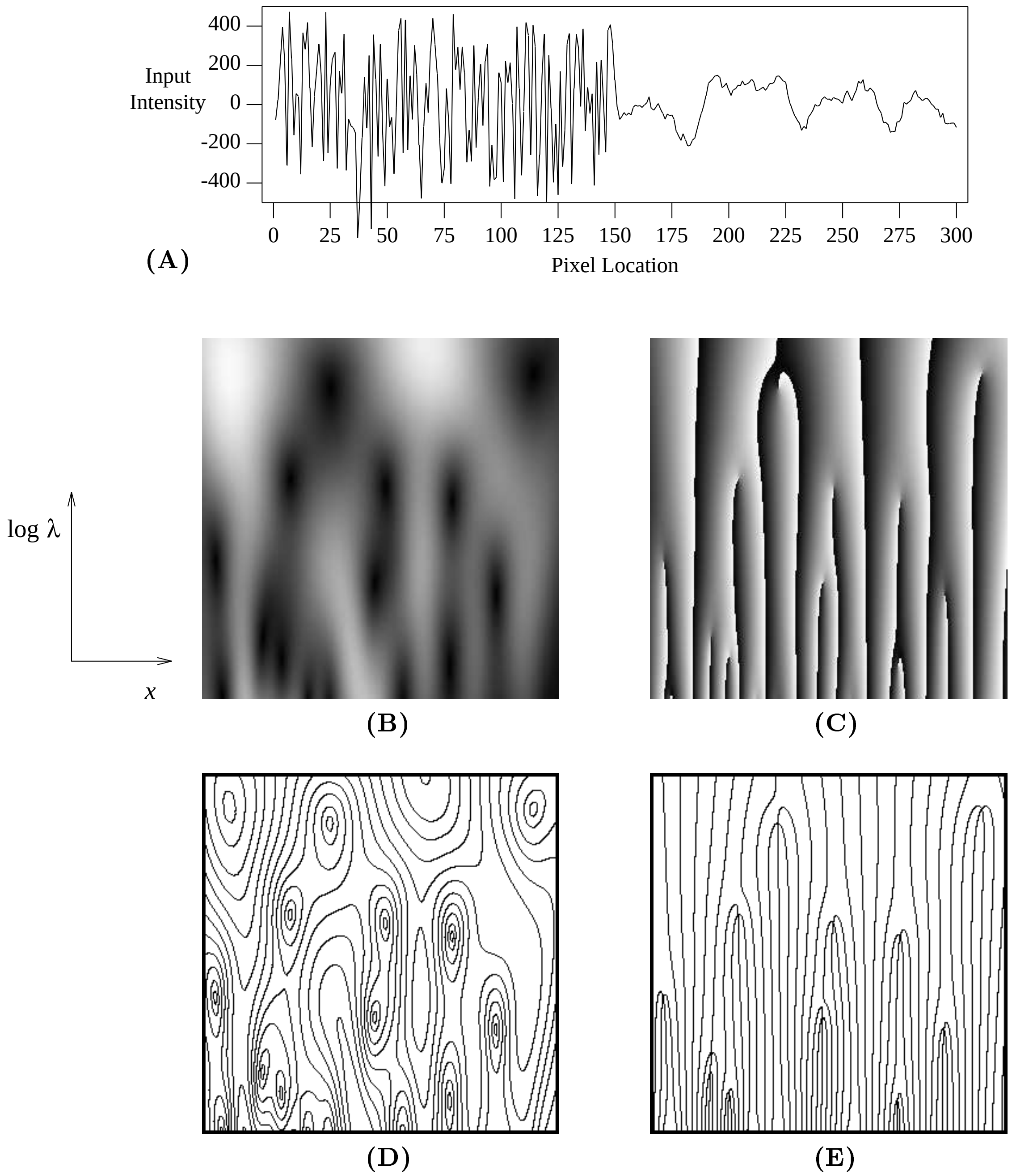


Figure 1. **Gabor Scale-Space Expansion:** *The input signal* **(A)** *consists of a sample of white Gaussian noise (left side), and a sample of red Gaussian noise (right side). The remaining panels show the amplitude and phase components of* $S(x, \lambda)$ *generated by a Gabor filter with a bandwidth of* $\beta = 0.8$ *octaves, with* $16 \le \lambda \le 64$ *pixels. The horizontal and vertical axes represent spatial position and log scale.* **(B)** *Amplitude response* $\rho(\mathbf{x}, t)$*;* **(C)** *Phase response* $\phi(\mathbf{x}, t)$*;* **(D)** *Level contours of* $\rho(\mathbf{x}, t)$*;* **(E)** *Level contours of* $\phi(\mathbf{x}, t)$*.*

4 Kernel Decomposition

The stability and linearity of phase can be examined in terms of the differences in phase between an arbitrary scale-space location and other points in its neighbourhood. Towards this end, let $S_j \equiv S(x_j, \lambda_j)$ denote the filter response at scale-space position $\mathbf{p}_j = (x_j, \lambda_j)$, and for convenience, let S_j be expressed using inner products instead of convolution:

$$S_j \quad = \quad < K_j^*(x), I(x) > \quad , \tag{12}$$

where $K_j(x) \equiv K(x_j - x, \lambda_j)$. Phase differences in the neighbourhood of an arbitrary point $\mathbf{p}_0$, as a function of relative scale-space position of neighbouring points $\mathbf{p}_1$, with $\mathbf{p}_1 - \mathbf{p}_0 = (\Delta x, \Delta\lambda)$, can be written as

$$\Delta\phi(\mathbf{p}_1, \mathbf{p}_0) \quad = \quad \arg[S_1] - \arg[S_0] \ . \tag{13}$$

Phase is perfectly stable when $\Delta\phi$ is constant with respect to changes in scale $\Delta\lambda$, and it is linear with respect to spatial position when $\Delta\phi$ is a linear function of Δx.

To model the behaviour of $S(x, \lambda)$ and $\Delta\phi$ in the neighbourhood of $\mathbf{p}_0$ we write the scale-space response at $\mathbf{p}_1$ in terms of S_0 and a residual term $R(\mathbf{p}_1, \mathbf{p}_0)$ that goes to zero as $\| \mathbf{p}_1 - \mathbf{p}_0 \| \to 0$, that is,

$$S_1 \quad = \quad z(\mathbf{p}_1, \mathbf{p}_0)\, S_0 \ + \ R(\mathbf{p}_1, \mathbf{p}_0) \ . \tag{14}$$

Equation (14) is easily derived if the effective kernel at $\mathbf{p}_1$, that is $K_1(x)$, is written as the sum of two orthogonal terms, one which is a scalar multiple of $K_0(x)$, and the other orthogonal to $K_0(x)$:

$$K_1(x) \quad = \quad z(\mathbf{p}_1, \mathbf{p}_0)\, K_0(x) \ + \ H(x; \mathbf{p}_1, \mathbf{p}_0) \ , \tag{15}$$

where the complex scalar $z(\mathbf{p}_1; \mathbf{p}_0)$ is given by

$$z(\mathbf{p}; \mathbf{p}_0) \quad = \quad < K_0(x), K_1(x) > \ , \tag{16}$$

and the residual kernel $H(x; \mathbf{p}_1, \mathbf{p}_0)$ is given by

$$H(x; \mathbf{p}_1, \mathbf{p}_0) \quad = \quad K_1(x) \ - \ z(\mathbf{p}_1, \mathbf{p}_0)\, K_0(x) \ . \tag{17}$$

Equation (14) follows from (15) with $R(\mathbf{p}_1, \mathbf{p}_0) = \ < H^*(x; \mathbf{p}_1, \mathbf{p}_0), I(x) >$. The scalar z reflects the cross-correlation of the kernels $K_0(x)$ and $K_1(x)$. The behaviour of $R(\mathbf{p}_1, \mathbf{p}_0)$ is related to the signal structure to which $K_1(x)$ responds but $K_0(x)$ does not.

For notational convenience below, let $z_1 \equiv z(\mathbf{p}_1, \mathbf{p}_0)$, $H_1(x) \equiv H(x; \mathbf{p}_1, \mathbf{p}_0)$, and $R_1 \equiv R(\mathbf{p}_1, \mathbf{p}_0)$. Remember that z_1, H_1, and R_1 are functions of scale-space position $\mathbf{p}_1$ in relation to $\mathbf{p}_0$.

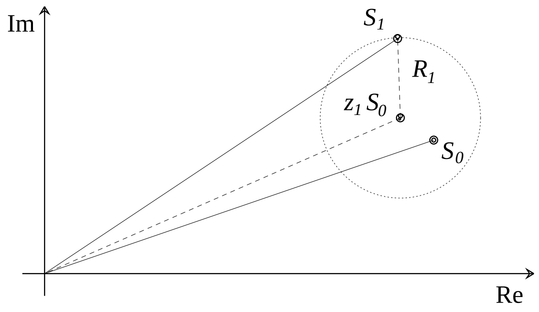


Figure 2. **Sources of Phase Variation:** *This shows the formation of S_1 in terms of S_0, the complex scalar $z_1 \equiv z(\mathbf{p}; \mathbf{p}_0)$, and the additive residual $R_1 \equiv R(\mathbf{p}_1, \mathbf{p}_0)$.*

Equation (14), depicted in Figure 2, shows that the phase of S_1 can differ from the phase of S_0 because of the additive phase shift due to z_1, and the phase shift caused by the additive residual term R_1. Phase will be *stable* under small scale perturbations whenever both the phase variation due to z_1 as a function of scale and $|R_1|/|z_1 S_0|$ are reasonably small. If $|R_1|/|z_1 S_0|$ is large then phase remains stable only when R_1 is in phase with S_0, that is, if $\arg[R_1] \approx \arg[S_0]$. Otherwise, phase may vary wildly as a function of either spatial position or small scale perturbations.

5 Phase Stability Given White Gaussian Noise

To characterize the stability of phase behaviour through scale-space we first examine the response of $K(x, \lambda)$ and its phase behaviour to stationary, white Gaussian noise. Using the kernel decomposition (15), we derive approximations to the mean phase difference $\mathrm{E}[\Delta\phi]$, and the variation about the mean $\mathrm{E}[\,|\Delta\phi - \mathrm{E}[\Delta\phi]|\,]$, where $\mathrm{E}[\cdot]$ denotes mathematical expectation. The mean provides a prediction for the phase behaviour, and the expected variation about the mean amounts to our confidence in the prediction. These approximations can be shown to depend only on the cross-correlation z_1 (16), and are outlined below; they are derived in further detail in Appendix A.

Given white Gaussian noise, the two signals R_1 and $z_1 S_0$ are independent (because the kernels $H_1(x)$ and $z_1 K_0(x)$ are orthogonal), and the phase of S_0 is uniformly distributed over $(-\pi, \pi]$. If we also assume that $\arg[R_1]$ and $\arg[z_1 S_0]$ are uncorrelated and that $\arg[R_1] - \arg[z_1 S_0]$ is uniform, then the residual signal R_1 does not affect the mean phase difference. Therefore we approximate the mean $\mathrm{E}[\Delta\phi]$ by

$$\mu(z_1) \quad = \quad \arg[z_1] \, , \tag{18}$$

where z_1 is a function of scale-space position. Then, from (13) and (18), the component of $\Delta\phi$ about the approximate mean is given by (cf. Figure 2)

$$\Delta\phi - \mu(z_1) = \arg[z_1 S_0 + R_1] - \arg[z_1 S_0] . \tag{19}$$

The expected magnitude of $\Delta\phi - \mu(z_1)$ measures of the spread of the distribution of $\Delta\phi$ about the mean;[4] it is a function of the magnitude of $z_1 S_0$ and the magnitude of the component of R_1 perpendicular to the direction of $z_1 S_0$ in the complex plane (see Figure 2). With the assumptions[5] that $|R_1| < |z_1 S_0|$ and that $\arg[R_1]$ is uniformly distributed, it is shown in Appendix A that an approximate bound, $b(z_1)$, on $\mathrm{E}[\,|\Delta\phi - \mu(z_1)|\,]$ is given by

$$b(z_1) = \frac{\sqrt{1 - |z_1|^2}}{|z_1|} . \tag{20}$$

It is tightest for small variations about the mean, that is, small values of $\Delta\phi - \mu(z_1)$.

5.1 Gabor Kernels

For illustrative purposes we apply these results to Gabor filters. Although they only approximate quadrature-pair filters for small bandwidths, they admit simple analytic derivation for z_1 which is the basis for the stability measures. For many other filters, it is more convenient to derive z_1 numerically from discrete kernels.

For $K(x, \lambda) = Gabor(x;\, \sigma(\lambda),\, k(\lambda))$, z_1 is given by (see Appendix C.1)

$$z_1 = \sqrt{2\pi}\; G(\Delta x;\, \bar{\sigma})\; G(\Delta k;\, \frac{\bar{\sigma}}{\sigma_0 \sigma_1})\; e^{i\,\Delta x\,[k_0 + (\Delta k\, k_0^2\,/\,\bar{k}^2)]} \tag{21}$$

where $k_j = k(\lambda_j)$ defines the filter tunings (11a), $\sigma_j = \sigma(\lambda_j)$ defines the support widths (11b), $\Delta k = k_1 - k_0$, $\bar{k} = \sqrt{k_0^2 + k_1^2}$, and $\bar{\sigma} = \sqrt{\sigma_0^2 + \sigma_1^2}$.

From (21), the approximate mean phase difference, that is $\arg[z_1]$, is given by

$$\mu(z_1) = \Delta x \left(k_0 + \frac{\Delta k k_0^2}{\bar{k}^2} \right) . \tag{22}$$

The expected magnitude of $\Delta\phi$ about the mean (20) can also be determined from (21) straightforwardly. In particular, from (20) we expect $b(z_1)$ to behave linearly in the neighbourhood of $\mathbf{p}_0$, because for sufficiently small $\|\,(\Delta x,\, \Delta\lambda)\,\|$ it can be shown from (21) that $|z_1| = 1 + O(\|\,(\Delta x,\, \Delta\lambda)\,\|^2)$.

Figure 3 illustrates this behaviour in the restricted cases of pure dilation and pure translation between points $\mathbf{p}_1$ and $\mathbf{p}_0$. Figure 3A shows $\mu(z_1)$ with error bars (the expected deviation about the

[4] The expected value of the $|\Delta\phi|$ is one possible measure of the spread of the probability density function. Compared to the standard deviation it is less sensitive to outliers [12], and in this case, it yields analytic results while the second moment does not (see Appendix A).

[5] The assumption that $|R_1| < |z_1 S_0|$ means that $\mathbf{p}_0$ is not in the immediate neighbourhood of a singular point, where $|S_0|$ is very small. Singularity neighbourhoods are discussed below in Section 6.

mean, $b(z_1)$) as a function of scale for a vertical slice through a Gabor scale-space expansion with bandwidth $\beta = 1$ octave. Figure 3B shows $\mu(z_1)$ and $b(z_1)$ as a function of spatial position for a horizontal slice through the same scale-space expansion. This shows clearly that the mean is constant with respect to scale changes and linear as a function of spatial position. The behaviour of $b(z_1)$, the variation about the mean, shows that the mean provides a very good model for the expected phase behaviour for small translations and dilations. For larger translations and dilations we expect a larger variation about the mean. This has a direct impact on disparity measurement when dilation occurs between two views, suggesting that errors will increase with the amount of dilation due to spatial drift of the phase contours.

In the case of translation, the increased variation about the mean is caused by two factors, namely, phase nonlinearities, and the distribution of the instantaneous frequencies in the filter output, which is a function of the bandwidth of the filter.[6] When the tuning frequency of the filter is used to approximate the instantaneous frequency in the denomenator of the disparity predictor (1) as in [27, 13, 15], then $b(z_1)$ provides a direct measure of the expected predictor errors. Otherwise, if the instantaneous frequency is measured explicitly as in [8, 18], then $b(z_1)$ can be used for an upper bound on expected predictor errors. The variation about the mean $b(z_1)$ can also be used indirectly as an indication of the range of disparities that a band-pass channel will reliably measure before aliasing occurs (when phase wraps between π to $-\pi$). This is important to consider when developing a control strategy to handle relatively large disparities with predictors at several scales and spatial positions.

A more complete illustration of the expected phase behaviour in the scale-space neighbourhood of $\mathbf{p}_0$ is given in Figure 4. Figure 4A shows level contours of $\mu(z_1)$ for a Gabor filter. The point $\mathbf{p}_0 = (0, \lambda_0)$ lies at the centre, and represents a generic location away from phase singularities (discussed below). Log scale and spatial position in the neighbourhood of $\mathbf{p}_0$ are shown on vertical and horizontal axes. These contours illustrate the expected mean phase behaviour, which has the desired properties of stability through scale and linearity through space. The contours are essentially vertical near $\mathbf{p}_0$, and for fixed Δk, the mean phase behaviour $\mu(z_1)$ is a linear function of Δx. It is also interesting to note that the mean phase behaviour does not depend significantly on the bandwidth of the filter; for Gabor filters it is evident from (21) that $\mu(z_1)$ is independent of β.

Figures 4B and 4C show level contours of $b(z_1)$ for Gabor kernels with bandwidths β of 0.8 and 1.0 octaves. In both cases, $b(z_1)$ is monotonically decreasing as one approaches $\mathbf{p}_0$ in the centre. As $\| \mathbf{p}_1 - \mathbf{p}_0 \|$ decreases so does the relative magnitude of R_1, and therefore so does the expected fluctuation in the phase of S_1 about the phase of $z_1 S_0$. By design, $b(z_1)$ is a measure of the distribution of phase differences about the mean. The contours in Figure 4 show that the expected deviation from $\mu(z_1)$ is small in the vicinity of $\mathbf{p}_0$. Since $\mu(z_1)$ has the properties we desire (stability through scale and linearity through space near $\mathbf{p}_0$), we can also view $b(z_1)$ as a direct measure of phase stability and phase linearity.

[6] The expected distribution of instantaneous frequencies for a given filter is discussed in [6]. A significant proportion of phase variability is due to the variation in instantaneous frequency in addition to phase nonlinearities. However, the separation of these two causes is beyond the scope of this paper, as it is not the principal issue.

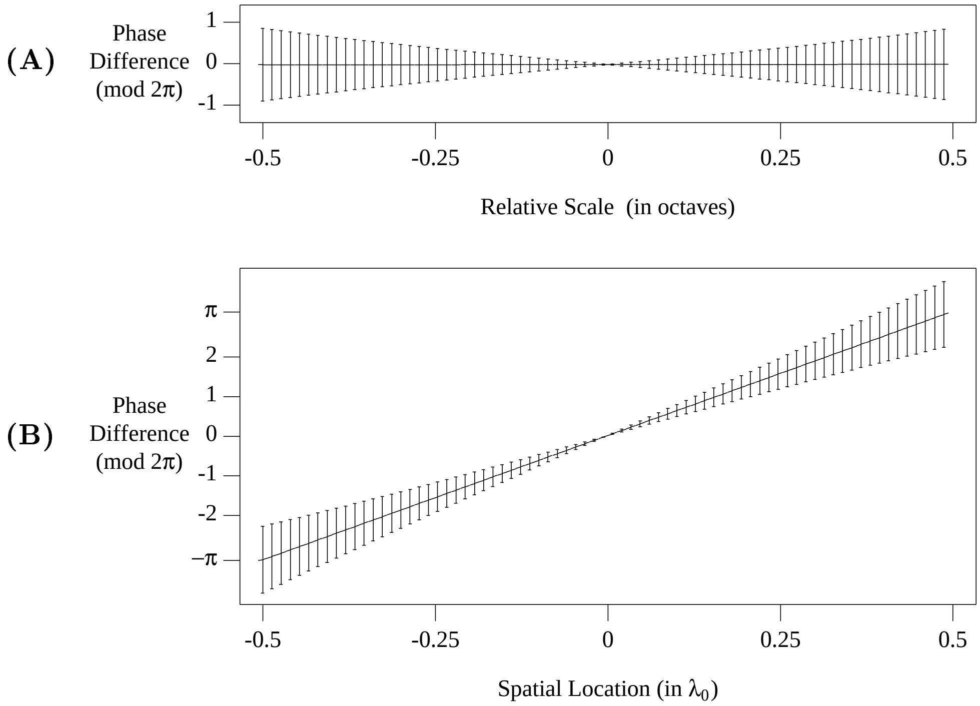
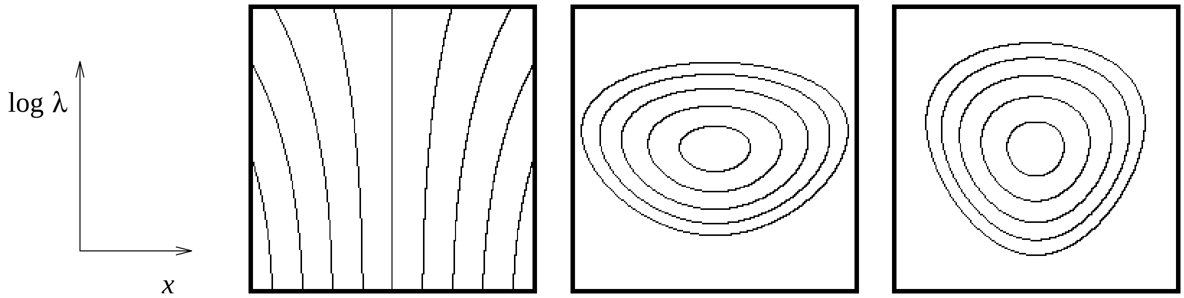


Figure 3. **Approximate Mean and Variation About the Mean of $\Delta\phi$ for 1-D Slices of Scale-Space:** *The expected behaviour of $\Delta\phi$ is shown for vertical and horizontal slices through the middle of the scale-space in Figure 4 for Gabor filters with $\beta = 1$. Both plots show the approximate mean $\mu(z_1)$ and the variation about the mean $b(z_1)$ (as error bars).* **(A)** *Phase behaviour as a function of $\Delta\lambda$ while $\Delta x = 0$.* **(B)** *Phase behaviour as a function of Δx for $\Delta\lambda = 0$.*



(A) **(B)** **(C)**

Figure 4. **Phase Stability with Gabor Kernels:** *Scale-space phase behaviour near* $\mathbf{p}_0 = (0,\, \lambda_0)$ *is shown with log scale on the vertical axis over two octaves with* $\lambda_1 \in (\frac{1}{2}\lambda_0,\, 2\lambda_0)$*, and spatial position on the horizontal axis,* $x_1 \in (-\lambda_0,\, \lambda_0)$*. The point* $\mathbf{p}_0$ *is in the centres of the figures.* **(A)** *Level contours* $\mu(z_1) = n\pi/2$*, for* $n = -4, \ldots, 4$*, as a function of scale-space position. These contours are independent of the bandwidth.* **(B)** *and* **(C)** *Level contours of* $b(z_1)$ *are shown for* $\beta = 0.8$ *and* 1.0*. Each case shows the contours* $b(z_1) = 0.3n$ *where* $n = 1, \ldots, 5$*; the innermost contours correspond to* $b(z_1) = 0.3$*.*

It is also evident from Figures 4B and 4C that $b(z_1)$ depends on the bandwidth of the filter. This can be explained from the dependence of σ_0 and σ_1 in (21) on bandwidth as given by (11b). As the bandwidth increases the amplitude spectra of filters tuned to nearby scales will overlap to a greater extent, and phase is therefore stable for larger scale perturbations of the input. On the other hand, an increase in bandwidth implies a decrease in the spatial extent of the kernels, and therefore a decrease in the spatial extent over which phase is generally linear. The smallest contours in Figures 4B and 4C encircle $\mathbf{p}_0 = (0,\, \lambda_0)$ and correspond to $b(z_1) = 0.3$ which amounts to a phase difference of about $\pm$ 5% of a wavelength. For $\beta = 1$ this contour encloses approximately $\pm$ 20% of an octave vertically and $\pm$ 20% of a wavelength spatially. In this case, relative scale changes of 10% and 20% are typically accompanied by phase shifts of less than 3.5% and 6.6% of a wavelength respectively.

5.2 Verification of Approximations

There are other ways to illustrate the behaviour of phase differences as a function of scale-space position. Although they do not yield as much explanatory insight they help to validate the approximations discussed above.

First, following Davenport and Root [4] it can be shown[7] that the probability density function for $\Delta\phi(\mathbf{p}_0, \mathbf{p}_1)$, at scale-space positions $\mathbf{p}_0$ and $\mathbf{p}_1$, for a quadrature-pair kernel and white Gaussian noise

[7] See [6] for details.

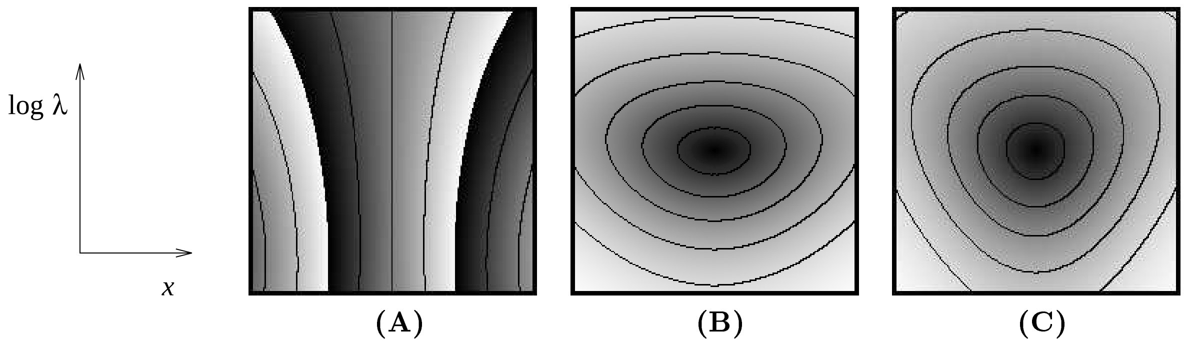


Figure 5. **$\mathbf{E}[\Delta\phi]$ and $\mathbf{E}[\,|\Delta\phi - \mathbf{E}[\Delta\phi]|\,]$ for White Noise Inputs:** *Scale-space phase behaviour based on (23) for Gabor kernels. As above, $\mathbf{p}_0 = (0, \lambda_0)$ is the centre, with $-\lambda_0 \le x_1 \le \lambda_0$ and $\frac{1}{2}\lambda_0 \le \lambda_1 \le 2\lambda_0$. The same level contours as in Figure 4 are superimposed for comparison.* **(A)** *$E[\Delta\phi]$ for $\beta = 0.8$.* **(B)** *and* **(C)** *$E[\,|\Delta\phi - E[\Delta\phi]|\,]$ for $\beta = 0.8$ and $\beta = 1.0$.*

input is

$$PDF(\Delta\phi) \;=\; \frac{2\pi\, A\,[(1-B^2)^{1/2} \,+\, B(\pi - \arccos B)]}{(1-B^2)^{3/2}} \;, \tag{23}$$

for $-\pi \le \Delta\phi < \pi$, where A and B are given by

$$A \;=\; \frac{\left(\frac{1}{4} - c_1^2 - c_2^2\right)}{\pi^2} \;, \qquad B \;=\; 2(c_1 \cos\Delta\phi + c_2 \sin\Delta\phi) \;,$$

$c_1 = \,<\!\mathrm{Re}[K_0(x)],\, \mathrm{Re}[K_1(x)]\!>$, and $c_2 = \,<\!\mathrm{Re}[K_0(x)],\, \mathrm{Im}[K_1(x)]\!>$. Given the density function, we can use numerical integration to find its mean behaviour and the expected variation about the mean. Figures 5A and 5B show the behaviour of the mean, and the absolute variation about the mean, of $\Delta\phi$ as functions of scale-space position for Gabor kernels with a bandwidth of 0.8 octaves. Figure 5C shows the expected variation of $\Delta\phi$ about the mean for Gabor filters of 1.0 octave. The mean behaviour in this case is not shown as it is almost identical to Figure 5A. In all three cases level contours have been superimposed to better illustrate the behaviour. Intensity in Figure 5A reflects values between $-\pi$ and π. Values in the other two range from 0 in the centre to $\pi/2$ at the edges where the distribution of $\Delta\phi$ becomes close to uniform. Comparing Figures 4 and 5, notice that the bound $b(z_1)$ in (20) is tightest for smaller values of $\Delta\phi$ (see Appendix A); the two smallest contours in Figures 4 and 5 are extremely close.

It is also instructive to compare the phase behaviour predicted by $\mu(z_1)$ and $b(z_1)$ with actual statistics of phase differences gathered from scale-space expansions of different input signals. Figure 6A shows behaviour of $\Delta\phi$ predicted by $\mu(z_1)$ and $b(z_1)$ for one octave Gabor filters as in Figure 3. Figures 6B and 6C show statistical estimates of $\mathrm{E}[\Delta\phi]$ and $\mathrm{E}[|\Delta\phi - \mathrm{E}[\Delta\phi]|]$ measured from scale-

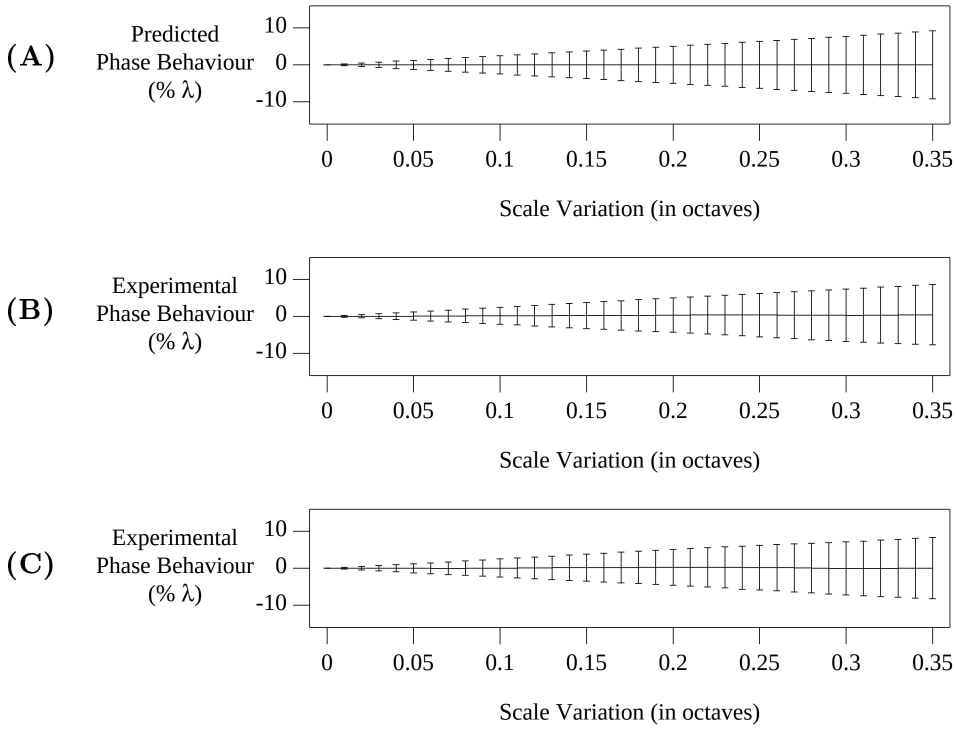


Figure 6. **Predicted Versus Actual Phase Behaviour:** **(A)** *Predicted behaviour of* $\Delta\phi$ *based on* $\mu(z_1)$ *and* $b(z_1)$ *for a Gabor filter with* $\beta = 1$. **(B)** *and* **(C)** *Statistical estimates of* $E[\Delta\phi]$ *and* $E[\,|\Delta\phi - E[\Delta\phi]|\,]$ *that were extracted from Gabor scale-space expansions of white noise (middle) and scanlines from natural images (bottom).*

space expansions of white noise and of scanlines of natural images. In both cases the observed phase behaviour is in very close agreement with the predicted behaviour. The statistical estimates of $\mathrm{E}[|\Delta\phi - \mathrm{E}[\Delta\phi]|]$ in these and other cases were typically only 1–5% below the bound $b(z_1)$ over the scales shown.

5.3 Dependence on Filter

These quantitative measures of expected phase behaviour can be used to predict the performance of phase matching as a function of the deformation between the input signals. The same measures can also be used to compare the performance of different filters; for although Gabor filters (8) have been popular [27, 7, 8, 15, 18], several alternatives have been suggested in the context of phase information. Weng [30] used a self-similar family of kernels derived from a square-wave (or constant) window:

$$K(x,\,\lambda) \;=\; \begin{cases} \frac{1}{\sqrt{\lambda}}\, e^{ik(\lambda)x} & \text{if } |x| \le \lambda/2 \\ 0 & \text{otherwise .} \end{cases} \tag{24}$$

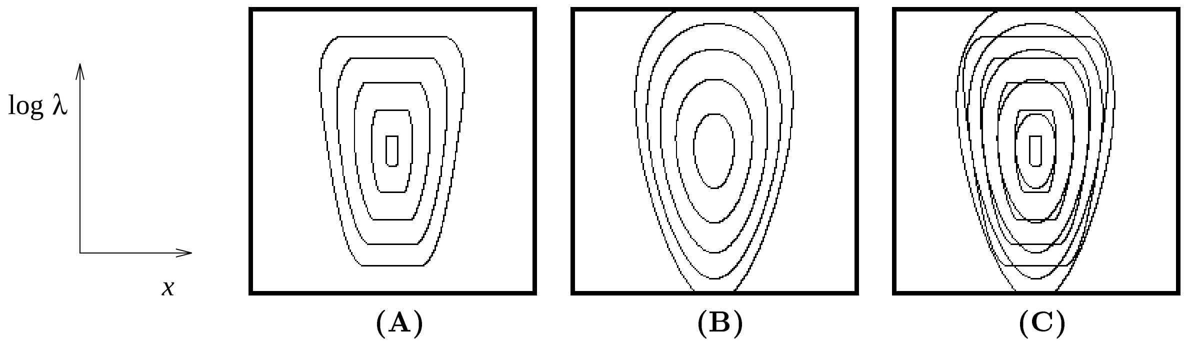


Figure 7. **Modulated Square-Wave Windows:** *Scale-space phase behaviour is shown for modulated square-wave windows (24) in comparison to Gabor filters.* **(A)** *Level contours* $b(z_1) = 0.3n$ *for* $n = 1 \ldots, 5$ *for the square-wave kernel.* **(B)** *Similar level contours of* $b(z_1)$ *for a Gabor kernel with* $\beta = 1.5$. **(C)** *The superposition of level contours from square-wave and Gabor kernels.*

Other alternatives, common to phase-correlation techniques, are families of filters with fixed spatial extents but tuned to different scales, also called windowed Fourier transforms [11, 17, 20]. This section addresses several issues in addition to those in Section 2 that are related to the choice of filter.

The effect of bandwidth on phase stability was illustrated in Figure 4. This dependence can be explained from the form of z_1 (16) and Parseval's theorem:

$$z_1 \quad = \quad < K_0(x),\, K_1(x) > \quad = \quad (2\pi)^{-1} < \hat{K}_0(k),\, \hat{K}_1(k) > \ , \tag{25}$$

where $\hat{K}(k)$ denotes the Fourier transform of $K(x)$. As the bandwidth increases the extent of the amplitude spectrum increases, and so does the range of scales $\Delta\lambda$ over which $\hat{K}_0(k)$ and $\hat{K}_1(k)$ may remain highly correlated. Similar arguments hold in space with respect to the extent of spatial support and phase linearity. For wavelets, the stability of phase with respect to input scale changes is constant across filters tuned to different scales, but the spatial extent over which phase is expected to be nearly linear decreases for higher scales as a function of the support width (and therefore the wavelength) of the filters. For windowed Fourier transforms, the support width is constant for different scales, and hence so is the expected extent of linearity, but the stability of phase with respect to input scale perturbations decreases at higher scales as the octave bandwidth decreases.

But does a simultaneous increase in the extent of the kernels' support in space and its amplitude spectrum produce better scale stability or more extensive linearity through space? For example, compared to the Gaussian which *minimizes* a measure of simultaneous extent (the uncertainty relation), the modulated square-wave kernel (24) has relatively large simultaneous extents in space and frequency domain. Its amplitude spectrum is particularly broad, and therefore we might expect better stability than with comparable Gabor filters (e.g., with $\beta \approx 1.5$).

The expected phase behaviour for the modulated square-wave filter (24) is shown in Figure 7. Its mean phase behaviour, given by $\arg[z_1]$, is very much like that exhibited by Gabor filters in Figure 4A and is not shown. Figure 7A shows the scale-space behaviour of $b(z_1)$ for the modulated square-wave kernel (24), for which z_1 is given by (see Appendix C.2)

$$z_1 \quad = \quad i\, e^{i\Delta x k_1} \left(e^{i\Delta k\, a_1} \; - \; e^{i\Delta k\, a_2} \right) \; (\Delta k \sqrt{\lambda_0 \lambda_1})^{-1} \; , \tag{26}$$

where $\Delta\lambda$ and Δk are defined above,[8] and

$$a_1 \;=\; \frac{-\lambda_0}{2} + \max\left[0,\; -\frac{\Delta\lambda}{2} - \Delta x\right] , \quad a_2 \;=\; \frac{\lambda_0}{2} + \min\left[0,\; \frac{\Delta\lambda}{2} - \Delta x\right] .$$

For comparison, Figure 7B shows level contours of $b(z_1)$ for a Gabor kernel with a bandwidth of $\beta = 1.5$, and Figure 7C shows the superposition of the level contours from Figures 7A and 7B. These figures show that the distribution of $\Delta\phi$ about the mean for the modulated suare-wave kernel is two to three times larger near $\mathbf{p}_0$, which suggests poorer stability *and* poorer linearity. Note that the innermost contour of the Gabor filter clearly encloses the innermost contour of the square-wave filter. This Gabor filter handles scale perturbations of 10% with an expected phase drift of up to $\pm$ 3.6% of a wavelength, while a perturbation of 10% for the modulated square-wave kernels gives $b(z_1) = 0.47$, which amounts to a phase difference of about $\pm 7.5\%$.

The poorer phase stability exhibited by the modulated square-wave kernel implies a wider distribution of measurement errors. Because of the phase drift due to scale changes, even with perfect phase matching, the measurements of velocity and disparity will not reflect the projected motion field and the projected disparity field as reliably. The poorer phase linearity affects the accuracy and speed of the disparity predictor, requiring more iterations to match the phase values between views. Moreover, we find that the larger variance also causes a reduction in the range of disparities that can be measured reliably from the predictor. The poorer linearity exhibited in Figure 7 also contradicts a claim in [30] that modulated square-wave filters produce more nearly linear phase behaviour.

Wider amplitude spectra do not necessarily ensure greater phase stability. Phase stability is the result of correlation between kernels at different scales. The shapes of both the amplitude and phase spectra will therefore play significant roles. The square-wave amplitude spectra is wide, but with considerable ringing so that $|z_1|$ falls off quickly with small scale changes.

Another issue concerning the choice of filter is the ease with which phase behaviour can be accurately extracted from a subsampled encoding of the filter output. It is natural that the outputs of different band-pass filters be quantized and subsampled to avoid an explosion in the number of bits needed to represent filtered versions of the input. However, because of the aliasing inherent in subsampled encodings, care must be taken in subsequent numerical interpolation/differentiation. For example, we found that, because of the broad amplitude spectrum of the modulated square-wave and its sensitivity to low frequencies, sampling rates had to be at least twice as high as those with Gabors (with comparable bandwidths, $\beta = 1.5$) to obtain reasonable numerical differentiation. If these issues

[8] As $\Delta k \rightarrow 0$, this expression for z_1 converges to $\exp[i\Delta x k_0]\,(1 - \Delta x/\lambda_0)$.

are not considered carefully, they can easily cause greater problems in phase-based matching than differences in stability or linearity between kernels. To alleviate some of these problems Weng [30] presmoothed the input signal with a Gaussian.

5.4 Amplitude Stability

Although our main concern is phase behaviour, it is also of interest to consider the expected scale-space behaviour of amplitude. Towards this end, using the same arguments as above for white noise inputs, it is shown in Appendix B that the expected (mean) amplitude variation as a function of scale-space position is constant, independent of the direction of $\mathbf{p}_1 - \mathbf{p}_0$. The expected absolute magnitude of amplitude differences, like phase variations about the mean, will depend on the relative magnitudes of $z_1 S_0$ and R_1. We expect the size of amplitude variations to increase for greater differences in scale-space distance.

This implies that amplitude often varies slowly through scale-space. However, while level phase contours exhibit predominantly vertical structure, level amplitude contours will occur at all orientations, and are therefore not consistently stable with respect to dilations between inputs. This variability is evident in Figure 1D compared to Figure 1E.

5.5 Multiple Dimensions

Finally, although beyond the scope of the current work, it is important to note that this basic framework can be extended to consider the stability of multidimensional filters with respect to other types of geometric deformation. In particular, we are interested in the phase behaviour of 2-d oriented filters with respect to small amounts of shear and rotation as well as scale changes. This analysis can be done, as above, using the cross-correlation between a generic kernel and a series of deformations of it. In this way, quantitative approximations can be found to predict the expected degree of phase drift under different geometric deformations of the input.

6 Singularity Neighbourhoods

The above analysis gives quantitative bounds on the expected stability of phase through scale and its linearity through space. But from Figure 1 it is clear that phase stability is not uniform throughout scale-space; some regions exhibit much greater instability in that the phase contours are nearly horizontal and not vertical as desired. Jepson and Fleet [14] explained that this phase instability occurs in the neighbourhoods of phase singularities (locations in space-time where the filter output passes through the origin in the complex plane). In terms of (14), S_0 is zero at a singularity, and the response S_1 in singularity neighbourhood is dominated by the residual term R_1. Zeros of $S(x, \lambda)$ appear as black spots in Figure 1B.

From the analysis described in [6, 14] the singularity neighbourhoods and the nature of phase instability can be characterized in terms of properties of the complex logarithm of the filter response

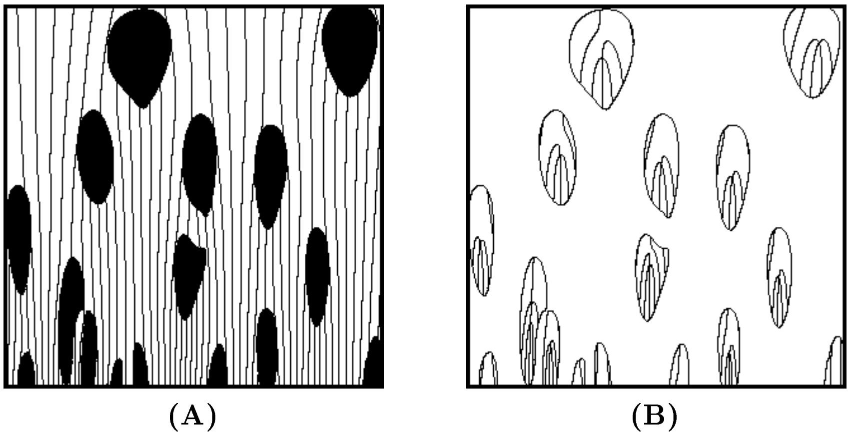


Figure 8. **Detection of Singularity Neighbourhoods:** **(A)** *Phase contours that survive the constraints.* **(B)** *Phase contours in regions removed by the constraints. The stability constraint in (28) was used with* $\tau = 1.25$.

$\log S(x, \lambda) = \log \rho(x, \lambda) + i\phi(x, \lambda)$, and its x-derivative:

$$\frac{\partial}{\partial x} \log S(x, \lambda) \;=\; \frac{[S^*(x, \lambda)\; S_x(x, \lambda)\,]}{|S(x, \lambda)|^2} \;=\; \frac{\rho_x(x, \lambda)}{\rho(x, \lambda)} + i\,\phi_x(x, \lambda)\;. \tag{27}$$

The imaginary part $\phi_x(x, \lambda)$ gives the local (instantaneous) frequency of the response [2, 26], that is, the instantaneous rate of modulation of the complex signal. The real part $\rho_x(x, \lambda) \,/\, \rho(x, \lambda)$ is the relative amplitude derivative. It can be shown that the instantaneous frequency of the filter output $\phi_x(x, \lambda)$ is expected to be within the passband of the filter, and the amplitude derivative $\rho_x(x, \lambda) \,/\, \rho(x, \lambda)$ is expected to be near zero [6]. The behaviour exhibited in singularity neighbourhoods is, however, quite different. As $|S(x, \lambda)|$ decreases to zero, $|\log S(x, \lambda)|$ increases without bound, as do the magnitudes of $\phi_x(x, \lambda)$ and/or $\rho_x(x, \lambda) \,/\, \rho(x, \lambda)$ [6]. This leads to a simple method for detecting singularity neighbourhoods so that unreliable measurements of binocular disparity and image velocity can be detected. Jepson and Fleet [14] used two constraints, one on the local frequency of response, and one on the magnitude of the amplitude derivative. Here we combine them into one:

$$\frac{1}{\sigma_k(\lambda)} \left| \frac{\partial}{\partial x} \log S(x, \lambda) \;-\; ik(\lambda) \right| \;\leq\; \tau\;, \tag{28}$$

where $k(\lambda)$ and $\sigma_k(\lambda) = 1/\sigma(\lambda)$ are given in (11a) and (11b). As τ decreases, this constraint detects larger neighbourhoods about the singular points. In effect, the left-hand side of (28) reflects the inverse scale-space distance to a singular point.

Figure 8 shows the application of (28) to the scale-space in Figure 1. Figure 8A shows the phase contours that survive the constraint, which are predominantly stable; the large black regions are the singularity neighbourhoods detected by (28). Figure 8B shows those removed by the constraint, which amounts to about 20% of the entire scale-space area. With respect to the quantitative approximations to phase behaviour presented in Section 4, we reported that statistics of mean phase differences and the absolute variation about the mean agreed closely with the bounds. When the phase behaviour in singularity neighbourhoods is ignored, so that the statistics are gathered only from outside of such singularity neighbourhoods, we find that the magnitude of the variation of $\Delta\phi$ about the mean is generally less than half of that predicted by the bound in (20). This detection of unstable regions is essential to the reliable performance of phase-based matching techniques, and it can be used to improve the performance of zero-crossing and phase-correlation techniques [11, 17, 21, 22, 29].

7 Natural Images

Unlike white noise, the Fourier harmonics of natural images are often correlated across scales, and their amplitude spectra typically decay something like $1/k$ [5]. Both of these facts affect our results concerning phase stability, the accuracy of phase matching, and instability due to singularity neighbourhoods.

First, because of amplitude spectra decay with spatial frequency, the filter responses will be biased to lower frequencies (as compared to white noise). As a result, care is required to ensure that the filter outputs do not contain too much power at low frequencies. Otherwise, there may be *a)* more distortion due to aliasing in a subsampled representation of the response, and *b)* larger singularity neighbourhoods, and hence a sparser set of reliable measurements. These problems are evident when comparing the modulated square-wave filters with Gabor filters (or similar bandwidths) because the former have greater sensitivity to low frequencies. Second, without the assumption of white noise we should expect $z_1\, S_0$ and R_1 in (14) to be correlated. This will lead to improved phase stability when R_1 and S_0 remain in phase, and poorer stability when they become systematically out of phase.

Although we lack a sufficient model of natural images, in terms of local structure, to provide a detailed treatment of phase stability on general images, several observations are readily available: For example, with many textured image regions the phase structure appears much like that in Figure 1. We find that the noise-based analysis provides a good model of the expected phase behaviour for complex structures that regularly occur in natural images.

Moreover, it appears that phase is even more stable in the neighbourhoods of salient image features, such as those that occur in man-made environments. To see this, note that the output of a filter in a small region can be viewed as a weighted sum of harmonics. In the vicinity of localized image features such as edges, bars and ramps, we expect greater phase stability because the phases of the input harmonics (unlike the white noise) are already coincident. This is clear from their Fourier transforms. Therefore changing scales slightly, or adding new harmonics at the high or low ends of the passband will not change the phase of the output significantly.

It is also worth noting that this phase coincidence at the feature locations coincides with local maxima of the amplitude response [23]. When different harmonics are in phase their amplitudes combine additively. When out of phase they cancel. Therefore it can be argued that neighbourhoods of local amplitude maxima correspond to regions in which phase is maximally stable (as long as the signal-to-noise-ratio is sufficiently high). This is independent of the absolute phase at which the different harmonics coincide,[9] and is significant for stable phase-based matching.

As one moves away from salient features, such as edges, the different harmonics may become increasingly out of phase, the responses from different features may interfere, and the amplitude of response decreases. This yields two main types of instability: *1)* where interference from nearby features causes the total response to disappear at isolated points; and *2)* where large regions have very small amplitude and are dominated by noise. The first case amounts to a phase singularity and is detectable using the stability constraint (28). In the second case, the phase behaviour in different views may be dominated by uncorrelated noise, but will *not* necessarily violate the stability constraint. For these situations a signal-to-noise constraint is necessary.

To illustrate these points, Figure 9 shows the Gabor scale-space expansion of a signal containing several step edges (Figure 9A). There are two bright bars, 30 and 25 pixels wide, 45 pixels apart. The scale-space plots were generated by Gabor filters with $\beta = 1$, spanning 2 octaves ($12 \leq \lambda \leq 48$). Figures 9A and 9B show the amplitude response and its level contours superimposed on the input signal (replicated through scale) to show the relationship between amplitude variation and the edges. Figures 9C and 9D show the scale-space phase response and its level contours, and Figures 9E and 9F show the contours that remain after the detection of singularity neighbourhoods using (28) with $\tau = 1.25$, and the phase contours in the neighbourhoods detected by (28). As expected, phase is stable near the edge locations where the local Fourier harmonics are in phase, and $\arg[R_1] \approx \arg[S_0]$ over a wide range of scales. The similarity of the interference patterns between the different edges to the singularity neighbourhoods shown in Figures 1 and 8 is clear, and these regions are detected using the stability constraint. Also detected are the regions relatively far from the edges where the amplitude and phase responses of the filter both go to zero.

However, as explained above, regions in which the filter response decreases close to zero are also very sensitive to noise. These regions can become difficult to match since uncorrelated noise between two views can dominate the response. To illustrate this, we generated a different version of the scale-space in which uncorrelated noise was added to the input independently before computing each scale (to simulate uncorrelated noise added to different views). The response to the independent noise patterns satisfied the stability constraint much of the time, but the phase structure was unstable (uncorrelated) between scales. Figure 9H shows the regions detected by the stability constraint in this case. As discussed, the regions of low amplitude in the original are now dominated by the response to the noise and are no longer detected by (28). Another constraint on the signal-to-noise ratio of the filter output appears necessary. For example, Figure 9I shows the regions in which the amplitude of the filter output is 5% or less of the maximum amplitude at that scale. This constraint in conjunction

[9] Morrone and Burr [23] argue that psychophysical salience (of spatial features) correlates well with phase coincidence only for certain absolute values of phase, namely, integer multiples of $\pi/2$, which are perceived as edges and bars of different polarities.

with those discussed earlier are sufficient to obtain phase stability comparable to that in the noiseless case (Figure 9J).

This also demonstrates the fact that an amplitude constraint alone does not serve our purposes, for the constraints on instantaneous frequency and amplitude derivative detect different regions. Although the regions detected by a single threshold applied directly to $\rho(x,\ \lambda)$ will eventually include the regions detected by (28) if the threshold is large enough, they will also enclose regions of stable phase behaviour. A single amplitude constraint will remove more of the signal than necessary if relied on to detect all the instabilities.

8 Further Research

Finally, it is clear that not all variations between two views of a scene are accounted for by affine geometric deformation and uncorrelated Gaussian noise. Other factors include contrast variations due to shading and shadows, specular anomalies, effects of anisotropic reflectance, and occlusion boundaries [6]. Although the behaviour of phase in these cases is not examined here, below we outline some of the main problems that require investigation.

Many contrast variations are smooth compared to surface texture. Thus, we expect the phase information from filters tuned to higher frequency to be insensitive to the contrast variations. In other cases, the contrast becomes a source of significant amplitude modulation in the input, which manifests itself in the phase behaviour of the filter output. Thus, some illumination changes will perturbation the disparity measurements in a non-trivial way.

Another difficulty for phase-based techniques, as well as most other matching techniques, is that the local intensity structure in one view of a 3d scene may be very different from that in another view of the same scene. Two obvious examples are occlusions, in which a surface may be visible in only one of the two views, and specular phenomena, in which case highlights may be visible from only one of two views. In both of these cases we should not expect the two views to be highly correlated in the usual sense; the difference between views is not well modelled by uncorrelated Gaussian noise. Moreover, we do not expect these cases to be detected by our stability constraint.

Another topic that is left for further research concerns measures of confidence for phase-based disparity measures. In our work we have found several ways of detecting unreliable measurements. One is the stability measure, and another is some form of SNR constraint. A further constraint that is useful in detecting poor matches, such as those due to occlusion is a correlation measure between matched regions. The use of these measures, possibly combining them into one measure has not been addressed in sufficient detail.

9 Discussion and Summary

This paper examines the robustness of phase-based techniques for measuring image velocity and binocular disparity [7, 8, 11, 13, 15, 17, 18, 25, 27, 30, 32, 33]. Our primary concerns are the effects of

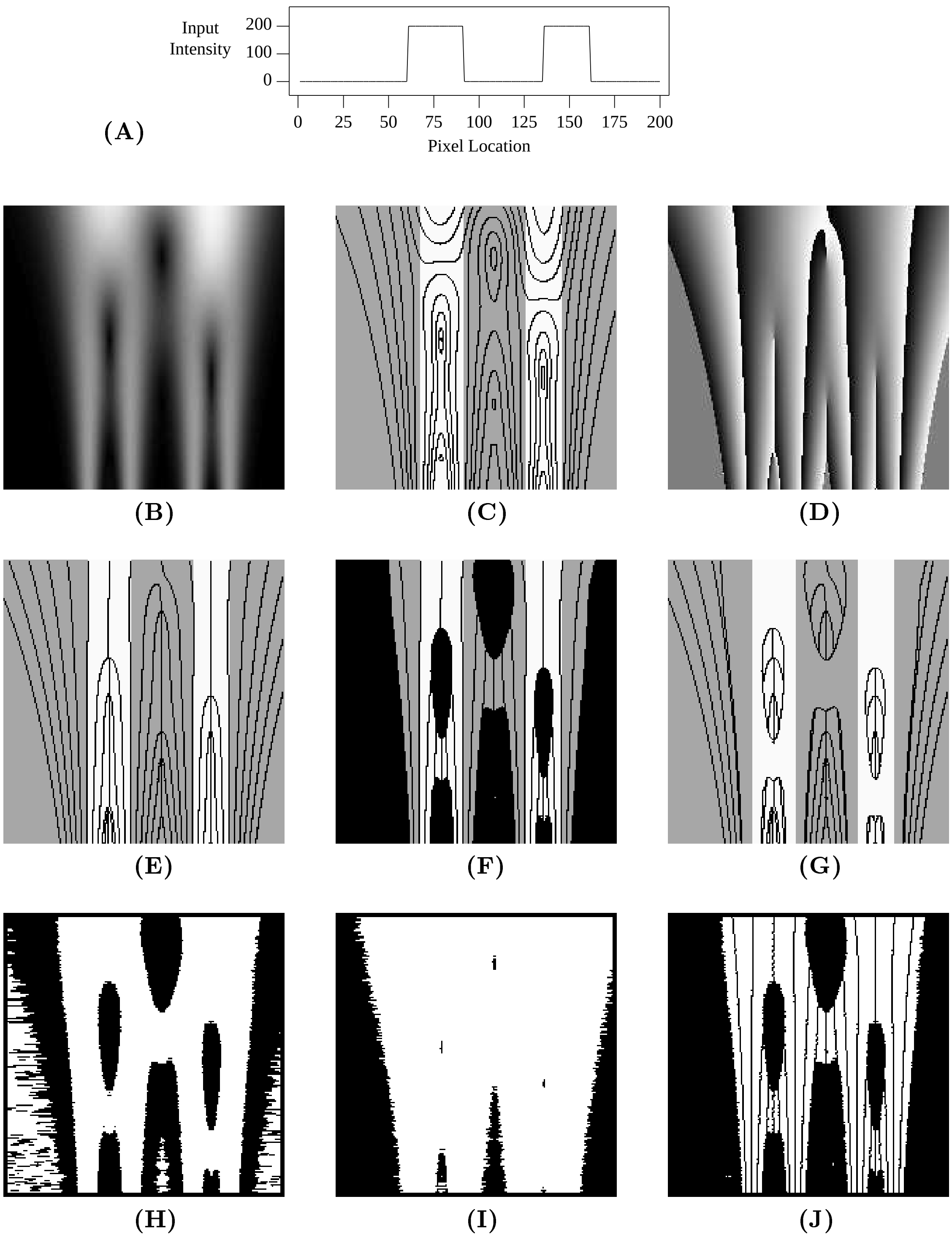


Figure 9. **Gabor Scale-Space Expansion With Step-Edge Input:** *The input signal* **(A)** *consists of two bars. Vertical and horizontal axes of the Gabor scale-space represent log scale and spatial position. Level contour have been superimposed on the input to show the relative location of the edges.* **(B)** *and* **(C)** $\rho(x, \lambda)$ *and its level contours;* **(D)** *and* **(E)** $\phi(x, \lambda)$ *and its level contours;* **(F)** *and* **(G)** *Level phase contours that survive the stability constraint (28), and those detected by it;* **(H)** *Regions of the "noisy-scale-space" that were detected by (28);* **(I)** *Regions detected by a simple amplitude constraint;* **(J)** *The level phase contours that survive the union of constraints in* **(H)** *and* **(I)**.

the filters and the stability of phase with respect to typical image deformations that occur between different views of 3-d scenes. Using a scale-space framework it was shown that phase is generally stable with respect to small scale perturbations of the input, and quasi-linear as a function of spatial position. Quantitative measures of the expected phase stability and phase linearity were derived for this purpose. From this it was shown that both phase stability and linearity depend on the form of the filters and their frequency bandwidths. For a given filter type, as the bandwidth increases, the extent of the phase stability increases, while the spatial extent over which phase is expected to be linear decreases. In the context of disparity measurement, the bandwidth of the filters should therefore depend, in part, on the expected magnitude of deformation between left and right views, since the potential accuracy of phase-based matching depends directly on phase stability.

One of the main causes of instability is the occurrence of phase singularities, the neighbourhoods of which exhibit phase behaviour that is extremely sensitive to input scale perturbations, small changes in spatial position, and small amounts of noise. Phase behaviour in these neighbourhoods is a source of significant measurement error for phase-difference and phase-gradient techniques, as well as gradient-based techniques, zero-crossing techniques, and phase-correlation techniques. Fortunately, singularity neighbourhoods can be detected automatically using a simple constraint (28) on the filter output. This stability constraint is an essential component of phase-based methods. A second constraint is also needed to ensure a reasonable signal-to-noise ratio.

This basic approach can also be used to examine the stability of multi-dimensional filters to other types of geometric deformation, such as the stability of 2-d oriented filters with respect to local affine deformation (rotation, shear, and dilation). As explained here, we may consider the behaviour of phase information using the cross-correlation of deformed filter kernels z_1 as a function of rotation and shear in addition to the case of dilation on which we concentrated in this paper. In this way, quantitative approximations can be found to predict the expected degree of phase drift under different geometric deformations of the input.

Acknowledgements

We are grateful to Michael Langer for useful comments on earlier drafts of this work. This research has been supported in part by the Natural Sciences and Engineering Research Council of Canada, and the Ontario Government under the ITRC centres.

References

[1] Barron, J.L., Fleet, D.J., Beauchemin, S., and Burkitt, T. (1992) Performance of optical flow techniques. *Proc. IEEE CVPR*, Champaign, pp. 236-242 (also see Technical Report RPL-TR-9107, Department of Computing Science, Queen's University)

[2] Boashash, B. (1992) Estimating and interpreting the instantaneous frequency of a signal. *Proc. IEEE* 80, pp. 520-568

[3] Burt, P.J., Bergen, J.R., Hinhorani, R., Kolczynski, R., Lee, W.A., Leung, A., Lubin, J., and Shvaytser, H. (1989) Object tracking with a moving camera. *Proc. IEEE Workshop on Visual Motion*, Irvine, pp. 2-12.

[4] Davenport, W.B. and Root, W. (1958) *Introduction to the Theory of Random Signals and Noise*, McGraw-Hill, New York

[5] Field, D.J. (1987) Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A* 4, pp. 2379-2393

[6] Fleet, D.J. (1992) *Measurement of Image Velocity*, Kluwer Academic Publishers, Norwell MA

[7] Fleet, D.J. and Jepson, A.D. (1990) Computation of component image velocity from local phase information. *Int. J. Computer Vision* 5, pp. 77-104

[8] Fleet, D.J., Jepson, A.D. and Jenkin, M. (1991) Phase-based disparity measurement. *CVGIP: Image Understanding* 53, pp. 198-210

[9] Freeman, W.T. and Adelson, E.H. (1991) The design and use of steerable filters. *IEEE Trans. PAMI* 13, pp. 891-906

[10] Gabor, D. (1946) Theory of communication. *J. IEE* 93, pp. 429-457

[11] Girod, B. and Kuo, D. (1989) Direct estimation of displacement histograms. *Proc. OSA Topcial Meeting on Image Understanding and Machine Vision*, pp. 73-76, Cape Cod

[12] Huber, P.J. (1981) *Robust Statistics*, John Wiley & Sons, New York

[13] Jenkin, M., and Jepson, A.D. (1988) The measurement of binocular disparity. **in** *Computational Proceses in Human Vision*, (ed.) Z. Pylyshyn, Ablex Press, New Jersey

[14] Jepson, A.D. and Fleet, D.J. (1991) Phase singularities in scale-space. *Image and Vision Computing* 9, pp. 338-343

[15] Jepson, A.D. and Jenkin, M. (1989) Fast computation of disparity from phase differences. *Proc. IEEE CVPR*, San Diego, pp. 398-403

[16] Koenderink, J.J. (1984) The structure of images. *Biological Cybernetics* 50, pp. 363-370

[17] Kuglin, C. and Hines, D. (1975) The phase correlation image alignment method. *Proc. IEEE Int. Conf. Cybern. Society*, pp. 163-165

[18] Langley, K., Atherton, T.J., Wilson, R.G., and Larcombe, M.H.E. (1990) Vertical and horizontal disparities from phase. *Proc. 1st ECCV*, Antibes, Springer-Verlag, pp. 315-325

[19] Langley, K., Fleet, D.J., and Atherton, T. (1992) Multiple motion from instantaneous frequency. *Proc. IEEE CVPR*, Champaign, pp. 846-849

[20] Mallat, S.G. (1989) Multifrequency channel decomposition of images and wavelet models. *IEEE Trans. ASSP* 37, pp. 2091-2110

[21] Marr, D. and Poggio, T. (1979) A computational theory of human stereo vision. *Proc. R. Soc. Lond.* B204, pp. 301-328

[22] Mayhew, J. and Frisby, J. (1981) Computational studies toward a theory of human stereopsis. *Artificial Intelligence* 17, pp 349-385

[23] Morrone, M.C. and Burr, D.C. (1988) Feature detection in human vision: a phase-dependent energy model. *Proc. R. Soc. Lond.* B 235, pp. 221-245

[24] Ogle, K.N. (1956) *Research in Binocular Vision*, W.B. Saunders Co., Philadelphia

[25] Olson, T. and Potter, R. (1989) Real-time vergence control. *Proc. IEEE CVPR*, San Diego, pp. 404-409

[26] Papoulis, A. (1965) *Probability, Random Variables, and Stochastic Processes*, McGraw-Hill, New York

[27] Sanger, T. (1988) Stereo disparity computation using Gabor filters. *Biological Cybernetics* 59, pp. 405-418

[28] Simoncelli, E.P., Freeman, W.T., Adelson, E.H. and Heeger, D.J. (1992) Shiftable multiscale transforms. *IEEE Trans. Info. Theory* 38, pp. 587-607

[29] Waxman, A.M., Wu, J., and Bergholm, F. (1988) Convected activation profiles: Receptive fields for real-time measurement of short-range visual motion. *Proc. IEEE CVPR*, Ann Arbor, pp. 717-723

[30] Weng, J. (1990) A theory of image matching. *Proc. 3rd ICCV*, Osaka, pp. 200-209

[31] Westelius, C-J (1992) Preattentive gaze control for robot vision. Thesis: LIU-TEK-LIC-1992:14, Dept. of Electrical Engineering, Linkoping University, Sweden

[32] Wiklund, J., Westelius, C-J, and Knuttson, H. (1992) Hierarchical phase-based disparity estimation. Technical Report: LiTH-ISY-I-1327, Dept. of Electrical Engineering, Linkoping University, Sweden

[33] Wilson, R. and Knuttson, H. (1989) A multiresolution stereopsis algorithm based on the Gabor representation. *Proc. IEE Intern. Conf. Im. Proc. and Applic.*, Warwick, U.K., pp. 19-22

[34] Witkin, A.P. (1983) Scale-space filtering. *Proc. 8th IJCAI*, Karlsruhe, pp. 1019-1022

[35] Yuille, A.L. and Poggio, T.A. (1986) Scaling theorems for zero-crossings. *IEEE Trans. PAMI* 8, pp. 15-25

A Approximations to Expected Phase Behaviour

This appendix provides greater detail about the approximations used in Sections 4 and 5 to illustrate and predict scale-space phase behaviour. In particular, it derives approximations to $E[\Delta\phi]$ and $E[\,|\Delta\phi \,-\, E[\Delta\phi]\,|\,]$ where, as defined in Section 4, $\Delta\phi(\mathbf{p}_1,\,\mathbf{p}_0)$ is the phase difference between two points in scale-space. We are interested in the behaviour of $\Delta\phi(\mathbf{p}_1,\,\mathbf{p}_0)$ as a function of points $\mathbf{p}_1$ in the neighbourhood of an arbitrary point $\mathbf{p}_0$ (away from the immediate neighbourhoods of singular points).

Following the notation in Section 4, let $K_j(x)$ and S_j be the effective kernel and its response at scale-space location $\mathbf{p}_j$. We assume that $K_j(x)$ is a quadrature-pair kernel. Then, using the decompositions in (14) and (15) we can write the response at S_1 as

$$S_1 \;=\; z_1\, S_0 \;+\; R_1\;. \tag{29}$$

where $z_1(\mathbf{p}_1,\,\mathbf{p}_0) = \,<K_0(x),\,K_1(x)>$ and R_1 is the response to the residual kernel $H_1(x)$ defined by (17). If we suppose a random process for the input, then, (29) specifies how the random variable S_1 is derived from the known quantity z_1, and the random variables S_0 and R_1. In this way we can also relate the phase of S_1 to the phase of S_0.

For mean-zero Gaussian white-noise input, the response S_0 is mean-zero Gaussian with variance $\sigma_0^2 \,=\, \|\, K_0 \,\|^2 \,=\, 1$,[10] and assuming that K_0 is a quadrature filter the real and imaginary parts of S_0 have a joint Gaussian density, with zero mean and a isotropic covariance $\sigma^2 = 1/2$ [Papoulis, 1965]. The phase of S_0 is uniformly distributed over $(-\pi,\,\pi]$ and independent of the amplitude. Based on similar arguments, $z_1\,S_0$ is also mean-zero Gaussian, but its variance is $|z_1|^2$, and that of its real and imaginary parts is $|z_1|^2/2$ and isotropic. In polar coordinates, $\rho_0 \,\equiv\, |z_1\,S_0|$ has a Rayleigh density with mean $\mathrm{E}[\rho_0] \,=\, \sqrt{\pi}\,|z_1|/2$, and second moment $\mathrm{E}[\rho_0^2] \,=\, |z_1|^2$, where $\mathrm{E}[\cdot]$ denotes mathematical expectation. Its phase angle $\psi_0 \equiv \arg[z_1\,S_0]$ is independent of ρ_0 and has a uniform density function [Papoulis, 1965].

The real and imaginary parts of the residual process R_1 are Gaussian. But their density is not isotropic with uniform phase because the real and imaginary parts of the kernel $H_1(x)$ are not orthogonal. To see this, note that

$$<\mathrm{Re}[H_1(x)],\,\mathrm{Im}[H_1(x)]> \;\;=\;\; -\,\mathrm{Im}[\,<K_1^*(x)\,,\;z_1 K_0(x)>\,]\;, \tag{30}$$

which is generally non-zero; (30) follows from the fact that $<\mathrm{Re}[z_1 K_0],\,\mathrm{Im}[z_1 K_0]>$ and $<\mathrm{Re}[K_1],\,\mathrm{Im}[K_1]>$ are both 0. As an approximation we assume that $H_1(x)$ is a quadrature filter so that its output is mean-zero with variance $\sigma_1^2 \,=\, \|\, H_1(x)\,\|^2$. In other words, we assume that $\rho_1 \equiv |R_1|$ has a Rayleigh density with mean $\mathrm{E}[\rho_1] \,=\, \sqrt{\pi}\,\|\, H_1(x)\,\|/2$, and second moment $\mathrm{E}[\rho_1^2] \,=\, \|\, H_1(x)\,\|^2$. Its phase angle $\psi_1 \equiv \arg[R_1]$ is independent of ρ_1 and has a uniform density function over $(-\pi,\,\pi]$. These expressions

[10] Given stationary white noise the spectral density of the output equals the power spectrum of the filter.

can be simplified further because $\| H_1(x) \|$ can be shown to reduce to $\sqrt{1-|z_1|^2}$. This is easily derived from $H_1(x) = K_1(x) - z_1 K_0(x)$ as follows:

$$
\begin{aligned}
\| H_1(x) \|^2 &= \; <K_1(x) - z_1 K_0(x),\, K_1(x) - z_1 K_0(x)> \\
&= \| K_1(x) \|^2 + \| z_1 K_0(x) \|^2 - (<z_1 K_0(x),\, K_1(x)> + <K_1(x),\, z_1 K_0(x)>) \\
&= 1 + |z_1|^2 - (<z_1 K_0(x),\, K_1(x)> + <z_1 K_0(x),\, K_1(x)>^*) \\
&= 1 + |z_1|^2 - 2\,\mathrm{Re}[z_1^* <K_0(x),\, K_1(x)>] \\
&= 1 - |z_1|^2 \,.
\end{aligned}
\tag{31}
$$

Finally, because the kernels $z_1 K_0(x)$ and $H_1(x)$ are orthogonal (by construction) the two signals $z_1\, S_0$ and R_1 are uncorrelated; and because the input is Gaussian they are statistically independent. With the assumption that R_1 is isotropic in its real and imaginary parts, it has no influence on the mean phase difference $\Delta\phi = \arg[S_1] - \arg[S_0]$. Thus, we approximate $\mathrm{E}[\Delta\phi]$ by $\mu(z_1) = \arg[z_1]$. By the same argument, the variation of $\Delta\phi$ about the mean is determined by the phase difference between $z_1\, S_0$ and $S_1 = z_1\, S_0 + R_1$, that is

$$
\begin{aligned}
\Delta\phi - \mu(z_1) &= (\arg[S_1] - \arg[S_0]) - \arg[z_1] \\
&= \arg[z_1\, S_0 + R_1] - \arg[z_1\, S_0] \,.
\end{aligned}
\tag{32}
$$

When $|R_1| < |z_1\, S_0|$ (i.e. with $\mathbf{p}_0$ not in the immediate neighbourhood of a phase singularity), the magnitude of $\Delta\phi - \mu(z_1)$ is given by the magnitude of the arctangent of the component of R_1 that is perpendicular to the complex direction of $z_1 S_0$, divided by the magnitude of $z_1 S_0$ (cf. Figure 2); that is,

$$
|\Delta\phi - \mu(z_1)| \;\equiv\; \left| \arctan\left(\frac{d_1}{\rho_0}\right) \right| \;\leq\; \frac{|d_1|}{\rho_0} \,, \tag{33}
$$

where $\rho_0 = |z_1\, S_0|$, and

$$
d_1 = \frac{\mathrm{Im}[(z_1\, S_0)^* R_1]}{|z_1\, S_0|} = \mathrm{Im}[e^{-i\psi_0}\, R_1] = \rho_1\, \mathrm{Im}[e^{i(\psi_1 - \psi_0)}] \,, \tag{34}
$$

where $\psi_0 = \arg[z_1\, S_0]$ and $R_1 = \rho_1\, e^{i\psi_1}$. Viewed as vectors in the complex plane, d_1 is the length of the projection of R_1 onto the unit vector normal to $z_1 S_0$. Finally, we can now formulate a bound on $\mathrm{E}[\,|\Delta\phi - \mu(z_1)|\,]$ as follows

$$
\mathrm{E}[\,|\Delta\phi - \mu(z_1)|\,] \;\leq\; \mathrm{E}\left[\frac{|d_1|}{\rho_0}\right] \;=\; \mathrm{E}[\rho_0^{-1}]\; \mathrm{E}[\rho_1]\; \mathrm{E}[\,|\sin(\psi_1 - \psi_0)|\,] \,. \tag{35}
$$

This follows from the independence of the four random quantities ρ_0, ρ_1, ψ_0, and ψ_1. From above, we know that $\mathrm{E}[\rho_1] = \sqrt{\pi}\, \| H_1(x) \| / 2 == \sqrt{\pi(1 - |z_1|^2)}/2$. Moreover, given the Rayleigh densities of ρ_0 and ρ_1, and the uniform densities of ψ_0 and ψ_1, it can be shown that $\mathrm{E}[\rho_0^{-1}] = \sqrt{\pi}\,/\,|z_1|$ and

$\mathrm{E}[\,|\sin(\psi_1 - \psi_0)|\,] \;=\; 2/\pi$. Therefore, the bound reduces to

$$\mathrm{E}[\,|\Delta\phi - \mu(z_1)|\,] \;\;\leq\;\; \frac{\sqrt{1 \;-\; |z_1|^2}}{|z_1|} \;. \tag{36}$$

It is tightest for small values of $\Delta\phi - \mu(z_1)$ because of the bound used in (33).

B Approximations to Expected Amplitude Behaviour

Using the same arguments as Appendix A, it can be shown that the expected (mean) amplitude variation as a function of scale-space position is constant, independent of the direction of $\mathbf{p}_1 - \mathbf{p}_0$. To see this, we first rewrite (29) in terms of the amplitude and phase components of $z_1 S_0$ and R_1

$$S_1 \;=\; |z_1|\rho_0 e^{i\psi_0} \;+\; \rho_1 e^{i\psi_1} \;, \tag{37}$$

where $\psi_0 \equiv \arg[z_1 S_0]$, $\rho_0 \equiv |S_0|$, $\psi_1 \equiv \arg[R_1]$, and $\rho_1 \equiv |R_1|$. With straightforward algebraic manipulation, the squared magnitude of S_1 can be written as

$$|S_1|^2 \;=\; |z_1|^2\,\rho_0^2 \;+\; \rho_1^2 \;+\; 2|z_1|\,\rho_1\rho_0\,\cos(\psi_1 - \psi_0) \;. \tag{38}$$

As above, (38) expresses the magnitude of S_1 as a function of known quantities and random variables. Its expected value is

$$E[\,|S_1|^2\,] \;=\; |z_1|^2\,E[\,\rho_0^2\,] \;+\; E[\,\rho_1^2\,] \;+\; 2|z_1|\,E[\,\rho_1\rho_0\,\cos(\psi_1 - \psi_0)] \;. \tag{39}$$

Using the same facts and assumptions as in Appendix A (i.e., the assumptions of white Gaussian input, the independence of variables ψ_0, ψ_1, ρ_0, and ρ_1, and the uniformity of $\psi_1 - \psi_0$) equation (39) becomes

$$E[\,|S_1|^2\,] \;=\; |z_1|^2 \;+\; 1 - |z_1|^2 \;+\; 2|z_1|^2 E[\,\rho_0\,\rho_1\,]\,E\,[\cos(\psi_1 - \psi_0)] \;, \tag{40}$$

From the uniformity of $\psi_1 - \psi_0$ the last term becomes zero, and the mean there reduces to

$$E[\,|S_1|^2\,] \;=\; 1 \;. \tag{41}$$

This shows that the scale-space dependence expressed in (29) and (38) does not affect the expected magnitude of S_1.

These results, of course, depend on the input distribution. With the same filters and input signals with greater power at low frequencies, we might expect a bias toward horizontal contours of constant amplitude. In the vicinity of localized image structure (e.g., such as edges) we expected a bias toward vertical amplitude contours (e.g. see Section 7).

C Derivations of z_1

Section 4 outlines the decomposition of a kernel K_1 into two components: one in the span of K_0, and the other orthogonal to K_0. The essential ingredient of this decomposition is given by z_1, the inner product of the generic kernel K_0 with deformations of itself given by K_1. This appendix derives expressions for z_1 for the cases of Gabor kernels (8) and modulated square-wave kernels (23).

C.1 Gabor Kernels

We first outline the decomposition of Gabor kernels where $G_j(x) = Gabor(x_j - x;\, \sigma_j,\, k_j)$ denotes a Gabor function at location $\mathbf{p}_j = (x_j,\, \lambda_j)^T$, with $\sigma_j = \sigma(\lambda_j)$ and $k_j = k(\lambda_j)$ as in (11a) and (11b). The decomposition of $G_1(x)$ hinges on the derivation of z_1 given by

$$z_1 \;=\; <G_0(x),\, G_1(x)> \;. \tag{42}$$

Let $\Delta x = x_1 - x_0$, and let $\Delta k = k_1 - k_0$. Without loss of generality in what follows we assume that $x_0 = 0$.

Using Parseval's theorem, where $\hat{G}(k)$ denotes the Fourier transform of $G(x)$ and the Gabor kernels have constant octave bandwidth, we can solve for z_1 as follows:

$$\begin{aligned} <G_0(x),\, G_1(x)> &= (2\pi)^{-1} <\hat{G}_0(k),\, \hat{G}_1(k)> \\ &= \frac{\sqrt{\pi\sigma_0\sigma_1}}{\pi} \int e^{-(k-k_0)^2\,\sigma_0^2/2}\; e^{-(k-k_1)^2\,\sigma_1^2/2}\; e^{i\Delta x\, k}\; dk \end{aligned} \tag{43}$$

With a change of variables ($\omega = k - k_0$), and straightforward algebraic manipulation (complete the square) this becomes

$$\frac{\sqrt{\pi\sigma_0\sigma_1}}{\pi}\, e^{i\Delta x\, k_0}\; e^{-\Delta k^2\,\sigma_1^2\;[1-(\sigma_1^2/\bar{\sigma}^2)]\,/2} \int e^{-\bar{\sigma}^2\,[\omega-(\Delta k\,\sigma_1^2/\bar{\sigma}^2)]^2\,/2}\; e^{i\Delta x\,\omega}\; d\omega \;, \tag{44}$$

where $\bar{\sigma}^2 = \sigma_0^2 + \sigma_1^2$. In (44) the integrand is a Gabor function and therefore easily integrated. After solving the integral, and a small amount of algebraic manipulation (using $\sigma_1^2\,\bar{k}^2 = \bar{\sigma}^2\,k_0^2$, where $\bar{k}^2 = k_0^2 + k_1^2$), one can show that

$$\begin{aligned} <G_0(x),\, G_1(x)> &= \frac{\sqrt{2\sigma_0\sigma_1}}{\bar{\sigma}}\; e^{-\Delta k^2\,\sigma_0^2\,\sigma_1^2\,/\,2\bar{\sigma}^2}\; e^{-\Delta x^2\,/\,2\bar{\sigma}^2}\; e^{i\Delta x\,[k_0+(\Delta k\,k_0^2\,/\,\bar{k}^2)]} \\ &= \sqrt{2\pi}\; G(\Delta x;\, \bar{\sigma})\; G(\Delta k;\, \bar{\sigma}/\sigma_0\sigma_1)\; e^{i\,\Delta x\,[k_0+(\Delta k\,k_0^2\,/\,\bar{k}^2)]} \end{aligned} \tag{45}$$

where $G(x;\, \sigma)$ is the scaled Gaussian in (10).

In the case of Gabor Kernels for which the support width σ is constant (i.e. independent of λ so that *linear* bandwidth is constant), z_1 is given by

$$z_1 \;=\; \frac{4\pi}{\sigma}\, e^{i\,\Delta x\,[k_0+(\Delta k/2)]}\; e^{-\Delta k^2\,\sigma^2/4}\; e^{-\Delta x/4\sigma^2}\;. \tag{46}$$

C.2 Modulated Square-Wave Kernels

Here we show the derivation of $z_1 \;=\; <K_0(x),\, K_1(x)>$ for modulated square-wave windows of the form described by Weng (1990). The kernel is given by

$$K(x,\,\lambda) \;=\; e^{ik(\lambda)x}\; sq(x,\,\lambda) \tag{47}$$

where $k(\lambda) \;=\; 2\pi/\lambda$ is the peak tuning frequency, and

$$sq(x,\,\lambda) \;=\; \begin{cases} \frac{1}{\sqrt{\lambda}} & \text{if } \;|x| \le \lambda/2 \\ 0 & \text{otherwise .} \end{cases} \tag{48}$$

Then, as above, with $\lambda_1 = \lambda_0 + \Delta\lambda\,,\;\; k_j \;=\; k(\lambda_j)\,$, and $\;\Delta k = k_1 - k_0$,

$$\begin{aligned} z_1 \;&=\; \int_{-\infty}^{\infty} e^{-ik_0x}\; e^{i(k_0+\Delta k)\,(x+\Delta x)}\; sq(x,\,\lambda)\; sq(x+\Delta x,\,\lambda_1)\; dx \\ &=\; \frac{1}{\sqrt{\lambda_0\lambda_1}}\; e^{i\Delta x k_1} \int_a^b e^{i\Delta k\, x}\; dx \\ &=\; i\, e^{i\Delta x k_1}\; \left(e^{i\Delta k\, a} \;-\; e^{i\Delta k\, b}\right)\; \left(\Delta k\,\sqrt{\lambda_0\lambda_1}\right)^{-1} \end{aligned} \tag{49}$$

where $b \;=\; \frac{\lambda_0}{2} + \min\left[0,\, \frac{\Delta\lambda}{2} - \Delta x\right]$, and $a \;=\; \frac{-\lambda_0}{2} + \max\left[0,\, -\frac{\Delta\lambda}{2} - \Delta x\right]$. In the limit as $\Delta k \to 0$, z_1 converges to $e^{i\Delta x k_0}\,(b-a)\,\lambda_0^{-1}$.