

Disparity from Local Weighted Phase-Correlation *

David J. Fleet

Department of Computing Science

Queen's University, Canada

fleet@qucis.queensu.ca

Abstract

Phase-based methods for extracting binocular disparity are discussed, including phase-difference methods and phase-correlation. A third method is also described that combines some of their properties, and appears consistent with recent physiological data.

1 Introduction

This paper outlines a method for extracting binocular disparity. It borrows from two existing approaches, namely, phase-difference methods [15, 5, 2, 16], and phase-correlation methods [8, 4, 13], and was designed to be consistent with recent physiological data [12].

We begin with a review of phase-difference and phase-correlation methods. They are both shown to be instances of the same basic approach, differing in the form of band-pass filters, stability constraints, and the control strategy, where phase-correlation looks more like a voting scheme than a coarse-to-fine approach. From this perspective, noting the advantages and disadvantages of both methods, we outline a new model that combines desirable properties of both, namely, the robustness and reliability of wavelet-based phase-differences, and the voting strategy of phase-correlation. The approach is also computationally simple, being composed primarily of linear operations throughout, with no explicit coarse-to-fine control strategy.

2 Phase-Difference Methods

Central to phase-based methods are the filters that decompose the images into band-pass signals. Perhaps the easiest way to extract phase is to use complex-valued quadrature-pair filters, the real and imaginary parts of which are $\pi/2$ radians out of phase, with identical amplitude spectra. Often one assumes that the real and imaginary parts of the kernel are even and odd-symmetric, but this is not strictly necessary.

Let $K_j(x)$ be the impulse response of the j^{th} filter,¹ and let the complex-valued outputs of its convolution with the left and right images, $I_l(x)$ and $I_r(x)$, be

$$\begin{aligned} O_l(x) &= \text{Re}[O_l(x)] + i \text{Im}[O_l(x)] \\ O_r(x) &= \text{Re}[O_r(x)] + i \text{Im}[O_r(x)] \end{aligned}$$

where $\text{Re}[z]$ and $\text{Im}[z]$ denote the real and imaginary parts of z . Amplitude ρ and phase ϕ constitute a polar representation of the real and imaginary parts in the complex plane,

$$O_l(x) = \rho_l(x) e^{i\phi_l(x)}, \quad O_r(x) = \rho_r(x) e^{i\phi_r(x)} \quad (1)$$

where amplitude is the magnitude of response, $\rho_l(x) = |O_l(x)|$, and the phase denotes the argument of the complex response, $\phi_l(x) = \arg[O_l(x)]$. Here, $\phi(x)$ and $\rho(x)$ are often called instantaneous phase and amplitude to emphasize their local nature. Also useful for phase-difference methods is the concept of *instantaneous frequency*, usually defined as the derivative of instantaneous phase with respect to spatial position [14]:

$$\tilde{\omega}_l(x) = \frac{d\phi_l(x)}{dx}, \quad \tilde{\omega}_r(x) = \frac{d\phi_r(x)}{dx} \quad (2)$$

This provides the frequency of the band-pass signal at each spatial position.

Phase-based matching methods define disparity as the shift necessary to align the phase values of band-pass filtered versions of the two signals in (1). To understand the reasons for the use of phase it is helpful to examine the typical behaviour of band-pass signals. Fig. 1 shows the real part of the output of a one octave Gabor filter applied to a sample of white noise, along with its amplitude and phase components. Not surprisingly, such outputs are well modelled in local regions by a sinusoidal signal with a slowly varying amplitude and a slowly varying frequency that remains close to the filter's tuning frequency [1]. Among other things, as shown in Fig. 1, this sinusoidal behaviour means that phase is predominantly linear. With a *purely* sinusoidal signal, phase will be perfectly linear.

*This paper was presented at IEEE International Conference on Systems, Man and Cybernetics, San Antonio, October 1994, pp. 48-56

¹Although the mathematical development is presented in 1d assuming epipolar lines, the implementation below is in 2d.

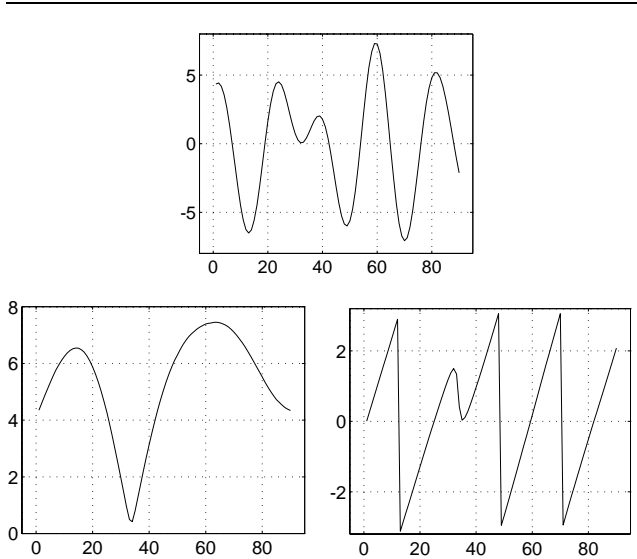


Figure 1. **Phase Linearity:** (top) Response of cosine-Gabor filter convolved with white noise. (bottom) Amplitude and phase parts of response.

The importance of phase linearity follows from the fact that displacements of a linear function are easy to estimate. For instance, with sinusoidal inputs and disparity d , the filter outputs would be

$$O_l(x) = \rho e^{i\omega_0 x} \quad , \quad O_r(x) = \rho e^{i\omega_0(x-d)} \quad . \quad (3)$$

The left and right phase functions are then $\phi_l(x) = \omega_0 x$ and $\phi_r(x) = \omega_0(x - d)$, so disparity is given by

$$d = \frac{\phi_l(x) - \phi_r(x)}{\omega_0} \quad (4)$$

Of course, phase is only uniquely defined in the interval $(-\pi, \pi]$. It is expected to wrap around from π to $-\pi$ every wavelength, and therefore we can only measure disparities to half a wavelength.

With general inputs we still expect phase to be predominantly linear [1], and hence reliable approximations to disparity are available from phase differences and the instantaneous frequency of the signal. The corresponding disparity estimator may take the form [2]

$$d = \frac{[\phi_l(x) - \phi_r(x)]_{2\pi}}{0.5(\tilde{\omega}_l(x) + \tilde{\omega}_r(x))} \quad (5)$$

where $[\psi]_{2\pi}$ denotes the principal part of ψ that lies between $-\pi$ and π . The size of the neighbourhood within which phase is monotonic determines the range of disparities that can be measured uniquely by the estimator. Filter outputs containing short wavelengths

can measure only small disparities while those with longer wavelengths can handle larger disparities.² This typically leads to some form of coarse-to-fine strategy, in which an initial guess is provided from coarser scales with which the images at the current scale are preshifted (warped) to bring them into registration within the appropriate domain of convergence (e.g. [11, 5, 7]). The wavelength at the coarsest scale should be more than twice the largest expected disparities.

There are several ways to measure phase differences. One can compute them explicitly, or one can take the complex-valued product of left output and the complex conjugate of the right [7]:

$$\begin{aligned} C(x) &= O_l(x)O_r^*(x) \\ &= \rho_l(x)\rho_r(x) [\cos \Delta\phi(x) + i \sin \Delta\phi(x)] \end{aligned} \quad (6)$$

where $\Delta\phi(x) = \phi_l(x) - \phi_r(x)$. The real and imaginary parts of $C(x)$ can be computed directly from the real-valued filter outputs as follows:

$$\rho_l \rho_r \cos \Delta\phi = \text{Re}[O_l] \text{Re}[O_r] + \text{Im}[O_l] \text{Im}[O_r] \quad (7a)$$

$$\rho_l \rho_r \sin \Delta\phi = \text{Im}[O_l] \text{Re}[O_r] - \text{Re}[O_l] \text{Im}[O_r] \quad (7b)$$

Interestingly, physiological data suggest that the terms on the right side of (7) may model the basic binocular interaction of simple cells, while their sums in (7a) and (7b) model complex-cell responses [12].

A second major reason for the success of phase-based methods is the stability of phase with respect to geometric deformations and contrast variations between left and right views [1]. Although most methods for estimating disparity are derived from a model of image translation, the importance of robustness with respect to affine deformations, like those that occur regularly with 3d surfaces, should not be overlooked.

Although usually stable, it can also be shown that phase exhibits a common form of instability where it is very sensitive to changes in spatial position and scale between the left and right signals. This instability occurs in the neighbourhoods of phase singularities, where the amplitude of the signal goes through the origin in the complex plane, and may be detected with constraints on the instantaneous frequency and the amplitude derivative of the filter output [1]. The detection and removal of measurements that occur in regions of phase instability is an important ingredient of current phase-difference methods [2, 6].

²Note that gradient-methods applied to the real-valued filter output directly (effectively linearizing the outputs) have an even narrower disparity range since the filter outputs are linear only in narrow regions about the zero-crossings.

3 Phase-Correlation

Phase-correlation methods use Fourier phase for signal registration [8, 4, 13]. The method is often derived assuming pure translation between two images:

$$I_r(x) = I_l(x - d). \quad (8)$$

The Fourier shift theorem tells us that their Fourier transforms satisfy $\hat{I}_r(\omega) = \hat{I}_l(\omega)e^{-id\omega}$. Their amplitude spectra are identical, $A_l(\omega) = A_r(\omega)$, and their phase spectra differ by $d\omega$, i.e., $\psi_r(\omega) = \psi_l(\omega) - d\omega$.

Taking the product of the left Fourier spectra and the complex conjugate of the right, and then dividing by the product of their amplitude spectra, we obtain

$$\frac{\hat{I}_l(\omega) \hat{I}_r^*(\omega)}{A_l(\omega) A_r(\omega)} = \frac{A_l(\omega) A_r(\omega) e^{i(\psi_l(\omega) - \psi_r(\omega))}}{A_l(\omega) A_r(\omega)} = e^{i\omega d} \quad (9)$$

Here, $e^{i\omega d}$ is a sinusoidal function in the frequency domain, and therefore its inverse Fourier transform is an impulse, located at the disparity d , that is, $\delta(x + d)$. Thus, in short, phase-correlation methods measure disparity by finding peaks in

$$\mathcal{F}^{-1} \left[\frac{\hat{I}_l(\omega) \hat{I}_r^*(\omega)}{A_l(\omega) A_r(\omega)} \right] \quad (10)$$

In practice, it is desirable to measure disparity locally. Accordingly, it is common to use windowed regions of the left and right images rather than the original images. The windows must be considerably larger than the expected displacements if no initial registration information is available. The large windows ensure that there is sufficient information in common in the two windows that can then be used for matching.

The reason for reviewing phase-correlation here is to first show its relationship to phase-difference methods, and to borrow some of its properties in designing a new approach. In comparing phase correlation to phase-difference methods, note first that a windowed Fourier transform is in fact a set of linear band-pass filters. To see this, consider the windowed Fourier transform of a signal $I(x)$, with the window centred at x_0 :

$$\begin{aligned} \mathcal{F}[W(x - x_0)I(x)] &= \int W(x - x_0)I(x)e^{-i\omega x} dx \\ &= e^{-i\omega x_0} R(\omega) \end{aligned} \quad (11)$$

where $R(\omega) = H(\omega) \otimes I(\omega)$ is the convolution of the input with a filter $H(\omega) = W(-\omega)e^{i\omega x_0}$. The Fourier coefficient, at frequency ω , of the windowed Fourier transform is equal to the output of the filter $H(\omega)$ tuned to

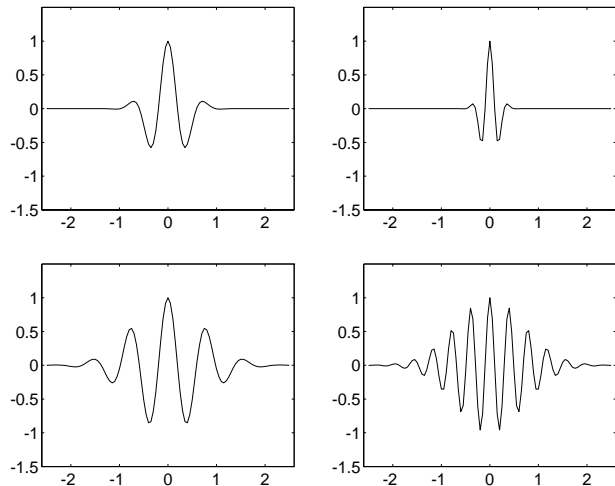


Figure 2. **Wavelets and Windowed FT:** (top) Typical kernels from a wavelet transform. (bottom) Typical kernels from a windowed Fourier transform.

the same frequency ω , at the location of the window. The filter is a sinusoidally modulated window (like a Gabor function). However, in contrast to the wavelet filters typically used in phase-difference methods, with constant octave bandwidths, the spatial extent of the implicit filters here does not depend on the frequency of the modulation as illustrated in Fig. 2, and thus they have a constant linear bandwidth [9].

By viewing the windowed Fourier transform as a set of filters, one can then see that the phase differences implicit in (10) are analogous to those provided in phase-difference methods in (6). But the product in (10) is a function of frequency; in effect, it represents a set of phase differences, one for each filter. If one were to extract these phase differences explicitly from the product in (10), divide by frequency, and obtain disparity estimates, then this would be a phase-difference method, but with a particular set of filters.

However, phase-correlation does not involve explicit phase differences, instantaneous frequencies, explicit disparity estimates, nor does it appear to involve a coarse-to-fine strategy which is common to many disparity techniques. Instead, phase-correlation uses a voting scheme to find the disparity. If one views the Fourier transform as a decomposition of a function into a sum of sinusoids, then the inverse Fourier transform amounts to the reconstruction of the function as a weighted sum of sinusoids. Here, the inverse Fourier transform is a sum of phase-shifted sinusoidal functions,

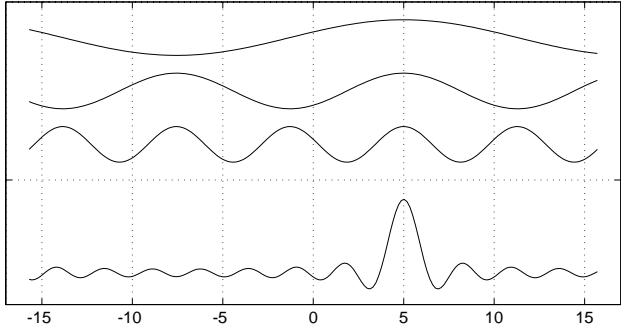


Figure 3. Voting Scheme: *Phase-shifted sinusoids in the inverse Fourier transform act as voting functions (top), where the peaks coincide in the sum to form a single peak (bottom).*

each of which has unit magnitude owing to the normalization in (10):

$$\begin{aligned} \mathcal{F}^{-1} \left[e^{i\Delta\phi(\omega)} \right] &= \int e^{i\Delta\phi(\omega)} e^{i\omega x} d\omega \\ &= \int [\cos(\omega x + \Delta\phi(\omega)) + \\ &\quad i \sin(\omega x + \Delta\phi(\omega))] d\omega \quad (12) \end{aligned}$$

One can view this as a voting scheme in which each band-pass channel votes in a sinusoidal manner, where the locations of the peaks are determined by the interocular phase differences. Ideally, as in Fig. 3, there will be a single disparity at which peaks coincide across a wide range of frequencies to form a distinct peak.

Finally, it may be interesting to note that phase-correlation also resembles a correlation technique in which the magnitude of the cross-correlation at each frequency is normalized to unity [13]. In other words, the product of the Fourier spectra in (10) is equivalent to the cross-correlation of the two inputs, but here they have been prewhitened by normalizing the amplitude spectra. This effectively enhances the influence of the higher frequencies in the cross-correlation, yielding a sharper peak in ideal cases.

4 Discussion

Phase-difference approaches have a variety of appealing properties, many of which are the result of research in recent years. Some of the main advantages result from the use of local wavelet filters and the stability constraints that significantly improve the robustness of the measurements. The coarse-to-fine control strategy, although common, is often thought to be somewhat un-

satisfactory. If the coarsest scale yields a poor estimate, or if phase at intermediate scales is unstable, then finer channels may receive a poor initial guess, in which case the rest of the process may converge to an incorrect disparity. In addition, the warping required at each level is not always convenient for parallel computation.

As a biological model there are further difficulties with phase-difference methods. Although the Ohzawa-DeAngelis-Freeman model [12] appears to compute the terms of the complex-valued product in (6) according to the receptive field combinations in (7a) and (7b), there is little or no physiological evidence for explicit representations of phase differences, instantaneous frequency, or disparity for that matter. In addition, there is growing evidence against coarse-to-fine control strategies in the psychophysical literature (eg. [10]).

Phase-correlation, by comparison, offers an interesting alternative. In particular, the voting scheme determines disparity based on the consistency of information at different scales and orientations. This does away with the coarse-to-fine strategy, and allows matches in which all band-pass channels are shifted by more than $\pi/2$ radians. Thus, this approach could in principle succeed where a coarse-to-fine approach might fail.

On the other hand, there are a variety of unappealing properties of phase-correlation. One problem arises from the windowed Fourier transform as the initial set of filters. Because the window size of the effective filters is fixed, the higher frequencies will have very narrow bandwidths compared to lower frequencies. From the results in [1] this implies that higher frequency channels will be very sensitive to even small scale changes between left and right images. Another concern with such filters is that the spatial resolution does not improve with high frequencies as it would with wavelet filters. Because the window sizes in phase-correlation methods are kept large to ensure enough structure for reliable matching, the spatial resolution of the disparity map will be coarse.

Two further problems with phase-correlation concern the voting strategy: First, all frequencies are weighted equally in (12), even though higher frequency channels are likely to be more sensitive to geometric deformations between views and to signal-to-noise problems. Unfortunately, phase-correlation methods do not exploit phase stability constraints like those used (at all scales) in phase-difference methods. A second concern with (12) is that the frequencies of the sinusoidal voting functions are determined by the tuning frequencies of the implicit filters rather than the instantaneous frequency of their responses. This implies significant er-

rors for larger disparities, and may cause the greatest problems at low frequencies, where the bandwidths are large, so that instantaneous frequency can differ significantly from the filter’s tuning frequency.

5 Local Weighted Phase-Correlation

In what follows we outline a local version of phase correlation that combines the basic robustness of phase-difference methods with the voting strategy of phase-correlation. The initial Fourier transforms are replaced with a family of quadrature-pair filters tuned in both orientation and scale with a constant octave bandwidth. The product of left and right filter outputs is used to obtain the phase differences as in phase-correlation methods. However, instead of assuming purely sinusoidal voting functions, we construct sampled versions of them using a series of preshifts of one of the two images. The voting functions are then summed across the different filters, from which the disparity measurements are extracted. The following paragraphs describe these ideas in more detail.

The initial quadrature-pair filters are currently implemented with a Gaussian pyramid, each scale of which is then decomposed using oriented quadrature-pair filters. We then have access to several oriented band-pass filter responses at each scale of the pyramid.

Instead of taking the normalized product of left and right outputs directly as in (10) or (6), here we introduce a small amount of smoothing [6], yielding

$$C_j(x, \tau) = \frac{W(x) \otimes [O_l(x) O_r^*(x+\tau)]}{\sqrt{W(x) \otimes |O_l(x)|^2} \sqrt{W(x) \otimes |O_r(x+\tau)|^2}} \quad (13)$$

where $W(x)$ is a small, localized window, τ acts as a preshift of the right filter output, and the subscript j refers to the j^{th} filter, whose output is used to compute $C_j(x, \tau)$. Peaks in the real part of $C_j(x, \tau)$ act as votes for candidate disparities τ between left and right filter outputs at location x . Rather than assume perfectly sinusoidal voting functions as in (12) and Fig. 3, preshifts τ of the right signal are introduced so that the effective voting function is sampled explicitly. Before considering the integration of functions $C_j(x, \tau)$ and the extraction of disparity, we first outline some of their properties.

The first important property of $C_j(x, \tau)$ is that its phase, like the phase of $C(x)$ in (6), is a phase difference that encodes the shift required to match the phases of the left and right band-pass signals. More precisely, one can show that the phase of $C_j(x, \tau)$ at location x_0 and preshift τ_0 , corresponds to the phase of the complex scalar z needed to minimize the squared difference

between the left signal and a phase-shifted version of the right signal.

$$\int W(x - x_0) |O_l(x) - z O_r(x + \tau_0)|^2 dx \quad (14)$$

It is interesting to contrast this with the more traditional approach that minimizes the squared difference between the left signal and a translated version of the right,

$$\int W(x - x_0) |O_l(x) - O_r(x + s)|^2 dx \quad (15)$$

the minimum of which occurs at the shift s that maximizes the normalized cross-correlation of O_l and O_r .

The second important property of $C_j(x, \tau)$ is that its magnitude (bounded between 0 and 1) provides a confidence measure for the goodness of fit between the phase-shifted left and right signals [6]. To see this, rewrite the convolution in (13) at some location as a local spatial average of vectors in the complex plane (ignoring the window weights for convenience):

$$C_j(x_0, \tau_0) \approx \frac{\sum \rho_l \rho_r e^{i\Delta\phi}}{\sqrt{\sum \rho_l^2} \sqrt{\sum \rho_r^2}} \quad (16)$$

where each vector has magnitude $\rho_l \rho_r$ and orientation $\Delta\phi$. This vector sum will be large in magnitude when there is little or no orientation variation among the vectors. It will be small when the orientations vary significantly, and therefore cancel one another when the vectors are summed. In this way the magnitude of $C_j(x, \tau)$ depends on the local consistency of the phase differences within the window, and replaces the more explicit stability constraints used in phase-difference methods. When phase in one or both views varies rapidly as a function of spatial position, so will the phase difference, thereby causing the magnitude of $C_j(x, \tau)$ to decrease.

Furthermore, notice that when all phase differences are the same, the magnitude of (16) equals the cross-correlation coefficient between the two local amplitude components of the filter output. Thus the magnitude of $C_j(x, \tau)$ depends on the cross-correlation of the amplitude components of the left and right filter outputs, and on the local stability (consistency) of the phase difference. This means that the voting functions will *not* all have unit magnitude.

A further point of interest is that $C_j(x, \tau)$ is expected to be band-pass in τ and low-pass in x . As a consequence it can be subsampled along both dimensions at a rate that depends on their linear bandwidths, which are expected to be similar. Coarse scales can be sampled sparsely in τ and in x . However, in order to integrate information across different filter outputs it will

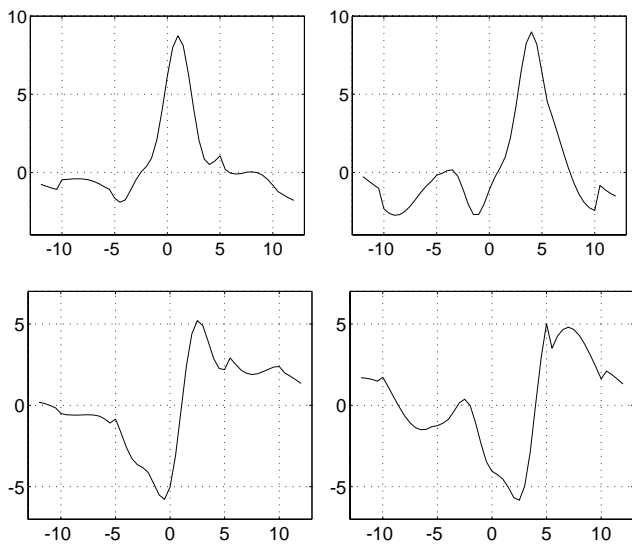


Figure 4. **Behaviour of $S(x, \tau)$ (with random-dot stereograms):** *Real (top) and imaginary (bottom) parts of $S(x, \tau)$, as functions of τ , at 2 spatial positions, with disparities of 1 pixel (left), and 4 pixels (right).*

be convenient to interpolate the $C_j(x, \tau)$ at a common set of positions x and disparities τ . Because $C_j(x, \tau)$ is band-pass in τ with frequencies close to the filter’s tuning frequency, it is easiest to demodulate the signal, then interpolate the corresponding low-pass signal, followed by modulation to undo the initial demodulation.

Finally, as mentioned above, the 2d functions $C_j(x, \tau)$ give us one voting function from each filter output. Although there are several interesting ways to use such functions to measure interocular disparity, here we consider the simplest approach, which, like phase-correlation methods, involves a simple summation:

$$S(x, \tau) = \sum_j C_j(x, \tau) \quad (17)$$

Near the true disparity we expect to find a zero in the imaginary part of $S(x, \tau)$ and a peak in its real part. Away from the true disparity, we expect the magnitude of the functions $C_j(x, \tau)$ to decrease, and we expect the phase differences to vary across the different scales and orientations of the filters, so that the net result of the sum should be relatively small. Examples of $S(x, \tau)$ from our implementation, as a function τ at two positions in a random-dot-stereogram are shown in Fig. 4.

6 Implementation Results

The results of a simple implementation are given below. At present we use 3 scales of a Gaussian pyramid, subsampled by a factor of 2 horizontally and vertically at each level. Three quadrature-pair filters are then applied at each level, tuned to orientations 0° , $+45^\circ$ and -45° , where 0 is vertical. The filters have an octave bandwidth of about 1.2 octaves, and are sampled with 4 samples per wavelength of centre frequency.

The voting functions $C_j(x, \tau)$ are computed using a Gaussian window $W(x)$ with a standard deviation of one half a wavelength of the filter’s tuning frequency. Preshifts are computed at one pixel intervals on the subsampled lattice at each scale, which also means a sampling rate of about 4 samples per wavelength of the expected modulation in τ . Remember that $C_j(x, \tau)$ will be low-pass in space and band-pass in disparity, with similar linear bandwidths in each dimension.

To compute $S(x, \tau)$, the voting functions $C_j(x, \tau)$ are interpolated back to the resolution of the original image. The spatial interpolation is done by constant interpolation (replicating pixel values), while interpolation in τ is done using demodulation, linear interpolation of the low-pass signal, followed by modulation.

Given $S(x, \tau)$, maxima (peaks) in its real part serve as crude estimates of disparity (to pixel accuracy). Sub-pixel accuracy is obtained using linear interpolation of the zero-crossing in the imaginary part of $S(x, \tau)$ that is nearest to the maxima in the real part.

Disparity estimates from two binocular pairs are shown in Fig. 5 and 6. Fig. 5 is a random-dot stereogram with 3 levels of disparity, namely, 1, 4, and 7 pixels. Fig. 6 shows one frame from the SRI Tree sequence (courtesy of SRI), and the disparities computed using frames 2 and 4 from the sequence, involving displacements up to 4 or 5 pixels.

7 Summary

We outline a new method for extracting binocular disparity that combines the robustness of wavelet-based phase-difference methods [15, 5, 2, 16], and the basic control strategy of phase-correlation methods [8, 4, 13]. It is computationally simple, being composed primarily of linear operations throughout, with no explicit coarse-to-fine control strategy.

Acknowledgements: Comments and suggestions from L. Haglund, D. Heeger, A. Jepson, and H. Wagner have helped shape these ideas. This research has been supported by grants from Queen’s University, NSERC Canada, and ITRC.

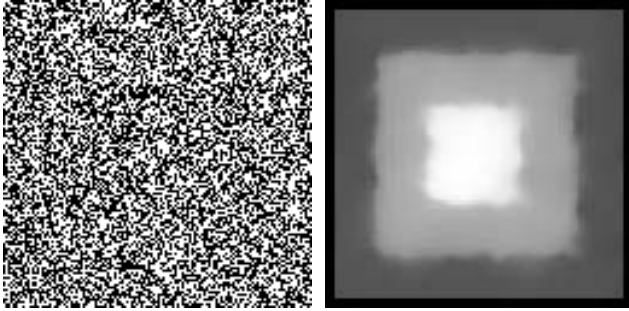


Figure 5. **Random-Dot Stereogram:** (left) Input. (right) Disparity estimates.

References

- [1] Fleet, D. and Jepson, A. (1993) Stability of phase information. *IEEE Trans. PAMI* 15:1253-1268
- [2] Fleet, D., Jepson, A. and Jenkin, M. (1991) Phase-based disparity measurement. *CVGIP: Image Understanding* 53:198-210
- [3] Gabor, D. (1946) Theory of communication. *J. IEE* 93:429-457
- [4] Girod, B. and Kuo, D. (1989) Direct estimation of displacement histograms. *Proc. OSA Meeting: Image Understanding and Machine Vision*, pp. 73-76, Cape Cod
- [5] Jenkin, M. and Jepson, A. (1988) The measurement of binocular disparity. in *Computational Processes in Human Vision*, (ed.) Z. Pylyshyn, Ablex Press, NJ
- [6] Jenkin, M. and Jepson, A. (1994) Recovering local surface structure through local phase difference measurements. *CVGIP: Image Understanding* 59:72-93
- [7] Jepson, A. and Jenkin, M. (1989) The fast computation of disparity from phase differences. *Proc. IEEE CVPR*, pp. 398-403, San Diego
- [8] Kuglin, C. and Hines, D. (1975) The phase correlation image alignment method. *Proc. IEEE Int. Conf. Cybern. Soc.*, pp. 163-165
- [9] Mallat, S. (1989) Multifrequency channel decomposition of images and wavelet models. *IEEE Trans. ASSP* 37:2091-2110
- [10] Mallot, H., Dartsch, S. and Arndt, P. (1994) Is correspondence search in human stereo vision a coarse-to-fine process? Technical Memo No. 4, Max Plank Institute for Biological Cybernetic, Tubingen, Germany
- [11] Marr, D. and Poggio, T. (1979) A computational theory of human stereo vision. *Proc. R. Soc. Lond.* B204:301-328
- [12] Ohzawa, I., DeAngelis, G., and Freeman, R. (1990) Stereoscopic depth discrimination in the visual cortex:

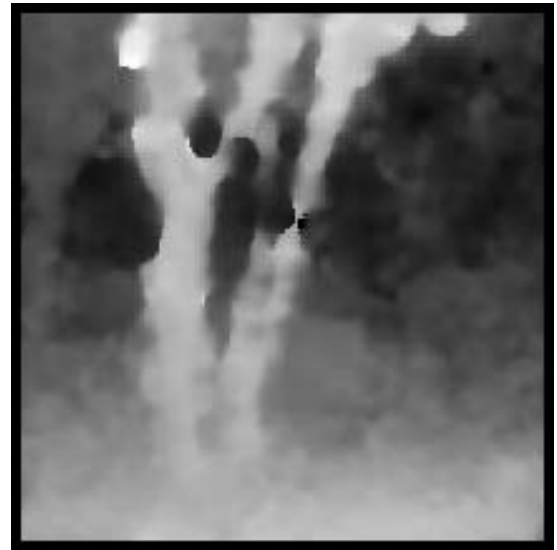


Figure 6. **SRI Tree Sequence:** (top) Input. (bottom) Disparity estimates.

Neurons ideally suited as disparity detectors. *Science* 249, pp. 1037-1041

- [13] Olson, T. and Potter, R. (1989) Real-time vergence control. *Proc. IEEE CVPR*, San Diego, pp. 404-409
- [14] Papoulis, A. (1977) *Signal Analysis*. McGraw-Hill, New York
- [15] Sanger, T. (1988) Stereo disparity computation using Gabor filters. *Biological Cybernetics* 59:405-418
- [16] Weng, J. (1994) Image matching using windowed Fourier phase. *Int. J. Comp. Vision* 11:211-236