Estimating mean disparity of stereo images using shift-trials of phase differences

Li-Dong Cai John Mayhew

Artificial Intelligence Vision Research Unit
University of Sheffield, England

Abstract

Disparity can be estimated using the phase differences between a pair of stereo images after Gabor filtering. In this paper a simple method is proposed for target detection, which uses shift-trials of phase differences to detect the mean disparity over the sampling window up to the precision of one pixel (the sampling interval unit). The process is posed as a problem of minimising the phase differences in terms of their matching residue's norm through direct searching. It involves no computation of the phase derivatives and is therefore robust to noise and phase singularity. Good results are obtained at an acceptable computational cost.

1 Introduction

The estimation of image disparity is a fundamental problem of early biological and computational visual processing. There exists an abundance of literature and methods on this issue. Recently, attention has focused on a new "correspondenceless" technique where disparity is expressed in terms of phase differences in the the left and right views after a local, bandpass filtering. This approach requires neither explicit signal reconstruction or feature detection and localisation. The results may be directly or iteratively used for further depth extraction, texture analysis and motion analysis *etc*[6, 7].

A kind of widely used methods is to estimate disparity from the local phase differences, as used by Fleet *et al.* and Sanger[2, 4, 8]. It can be shown that these methods are related to the Newton iterations[1], so the problem of phase singularities remains. To reduce the influence of singular points, edge information can be utilised as proposed in [9], but the problem of how to overcome the phase singularity itself has not been addressed. We propose here a simple method to deal with the singularity problem in purpose of target detection, which uses shift-trials of phase differences over the sampling window to detect the mean disparity up to the precision of one pixel, *i.e.*, the unit of sampling interval. The process is posed as a minimisation of the phase differences in terms of the matching residue's norm through direct searching. It involves no computation of the phase derivatives and is therefore robust to noise and phase singularity. Good results are obtained at an acceptable computational cost.

BMVC 1992 doi:10.5244/C.6.25

2 Minimising phase difference by shift-trials

Let $\tilde{L}(x)$ and $\tilde{R}(x)$ be a pair of stereo images of a 1-D signal with a relative shift s:

$$\tilde{L}(x) = \tilde{R}(x+s) \tag{1}$$

This relationship is maintained after a convolution with a function G, such as Gabor filtering [3], over the interval $(-\infty, \infty)$, since

$$\begin{aligned} L(x) &\stackrel{\text{def}}{=} \tilde{L}(x) * G(x) = \tilde{R}(x+s) * G(x) = \int_{-\infty}^{\infty} \tilde{R}(t+s)G(t-x)dt \\ &= \int_{-\infty}^{\infty} \tilde{R}(t)G(t-s-x)dt = \tilde{R}(x+s) * G(x+s) \stackrel{\text{def}}{=} R(x+s) \end{aligned} \tag{2}$$

The Fourier components of frequency ω of both images can be presented in the complex forms:

$$L_\omega(x) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \tilde{L}(x)e^{i\omega x}dx = Me^{i\omega x} \stackrel{\text{def}}{=} Me^{i\Phi_l(x)} \tag{3}$$

$$R_\omega(x) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} \tilde{R}(x)e^{i\omega x}dx = Me^{i\omega(x-s)} \stackrel{\text{def}}{=} Me^{i\Phi_r(x)} \tag{4}$$

where the amplitude M and phase Φ of a complex number $z = a + ib$ is defined as:

$$M \stackrel{\text{def}}{=} \sqrt{a^2+b^2} \tag{5}$$

$$\Phi \stackrel{\text{def}}{=} tan^{-1}\frac{b}{a} \tag{6}$$

In practice, given the (complex) sampling window width $T > 0$, the sampling interval $\Delta = 1$, then the tuning frequency f is limited by the Nyquist frequency $f_c = \frac{1}{2\Delta} = \frac{1}{2}$ to avoid the aliasing effect. That is

$$0 < \mid f \mid \leq f_c \tag{7}$$

Accordingly, the tuning wavelength λ is limited within the interval:

$$2 \leq \lambda \leq T \tag{8}$$

Note that any phase $\Phi(x)$ will be modulated into the interval $(-\pi, \pi]$. All detected shifts or disparities s^* will be modulated into the interval $(-\frac{\lambda}{2}, \frac{\lambda}{2}]$ accordingly. That is, any shift s and $\hat{s}$ will be seen as identical under the modulation of λ if they satisfy

$$\lambda - s = \hat{s} \tag{9}$$

$$0 \leq s \leq \lambda\ , \quad -\frac{\lambda}{2} < \hat{s} \leq \frac{\lambda}{2} \tag{10}$$

This indicates that the range of disparities that can be detected by the phase differencing methods is limited at each tuning wavelength level.

From Eq. (3) and (4), the shift s can be obtained either directly from the phase difference of $\Phi_l(x)$ and $\Phi_r(x)$:

$$s = \frac{\Phi_l(x) - \Phi_r(x)}{\omega} \tag{11}$$

this is the method used in [8], or through Newton iterations of the phase difference converging to zero, such as

$$F(x,s) \stackrel{\text{def}}{=} \Phi_l(x+\frac{s}{2}) - \Phi_r(x-\frac{s}{2}) \;\rightarrow\; 0 \tag{12}$$

this is the method used in [2]. Obviously, both methods produce results using local calculations at individual positions and large perturbations may appear at some positions in the presence of singularity of the phase function.

2.1 Problem posed

Disparity estimation is a task dependent problem. For target detection of a moving vehicle, a small sampling window is allowed and the 2-D disparity estimation can be reduced to a 1-D estimation at some scanlines of the image. Meanwhile, disparity estimation can change from local calculations at individual points on $[0,T]$ to global calculations over $[0,T]$ as a whole. These lead to the following shift-trials method, where disparity estimation is posed as a problem of minimising the norm of phase differences over $[0,T]$.

Given a pair of phase functions $\Phi_l(x)$ and $\Phi_r(x)$ over $[0,T]$, let the difference of phases $\Phi_l(x)$ and $\Phi_r(x+s)$ be

$$\Psi(x,s) \stackrel{\text{def}}{=} \Phi_l(x) - \Phi_r(x+s) \tag{13}$$

$\Psi(x,s)$ can be seen as the residue function over $[0,T]$ in the course of $\Phi_l(x)$ being approximated or matched by a family of functions $\Phi_r(x+s)$ which are generated from $\Phi_r(x)$ with (circular) shift trials s over $[0,\lambda]$.

To judge the goodness of approximation, define a norm of the residue $\Psi(x,s)$ as:

$$\nu(s) = \|\Psi(x,s)\| \stackrel{\text{def}}{=} \int_0^T |\Psi(x,s)|dx \tag{14}$$

It is a non-negative functional, whose graph is a continuous curve over $[0,\lambda)$:

$$0 \leq \|\Psi(x,s)\| \qquad s \in [0,\lambda) \tag{15}$$

When the norm $\|\Psi(x,s)\|$ tends to zero, $\Phi_r(x+s)$ converges in *mean* to $\Phi_l(x)$ over $[0,T]$. So the minimum value of $\|\Psi(x,s)\|$ gives the desired shift s^* where the shifted right phase gives the best approximation of the left phase over $[0,T]$, and the resulting shift s^* is thus called the *mean* disparity (to distinguish from the average of disparity values). Hence, by varying s over $[0,\lambda)$ the shift-trials process is virtually a direct search for the minimum phase difference in terms of the norm $\|\Psi(x,s)\|$.

2.2 Norm and singularity

The shift-trials method for target detection has some advantages. First, the disparity detected by the shift-trials method is a global quantity resulting from a function-to-function approximation over the whole sampling window $[0,T]$, rather than a set of local quantities resulting from a pointwise approximation at individual positions $s \in [0,T]$ as in [2, 8].

Second, unlike methods using Newton iterations, the shift-trials method involves no phase derivatives in the process. It is thus free from the pre-conditions of the Newton iteration:

i) A good initial guess $s^{(0)}$ of the real disparity s^*.

ii) An analytic (complex) function $\Phi(x)$.

Notice that it is unlikely for phase differencing methods to satisfy the pre-condition ii), since the phase function $\Phi(x)$ is not itself analytic everywhere, always leading to large perturbations occurring at some positions due to singularities. In fact, from the ill-defined phase function Φ in Eq. (6), we have[1]:

$$\delta\Phi = \frac{a\delta b - b\delta a}{a^2 + b^2} \tag{16}$$

Hence, when both a and b are small, for example,

$$|a| = |b| < \frac{1}{K}, \quad K > 10 \tag{17}$$

$$sgn(a\delta b) = sgn(-b\delta a) \tag{18}$$

there must be

$$\frac{|a|}{a^2 + b^2} = \frac{|b|}{a^2 + b^2} > \frac{K}{2} > 5 \tag{19}$$

$$\Phi = tan^{-1}\frac{b}{a} = \frac{\pi}{4} \tag{20}$$

therefore

$$|\delta\Phi| = \frac{|a|}{a^2 + b^2}|\delta b| + \frac{|b|}{a^2 + b^2}|\delta a| > 5(|\delta a| + |\delta b|) \tag{21}$$

$$|\frac{\delta\Phi}{\Phi}| = \frac{4}{\pi}\frac{|ab|}{a^2 + b^2}(|\frac{\delta a}{a}| + |\frac{\delta b}{b}|) > \frac{2}{\pi}(|\frac{\delta a}{a}| + |\frac{\delta b}{b}|) \tag{22}$$

That is, when both a and b are close to zero, any errors δa and δb contained in a and b could be amplified as a fairly large error $\delta\Phi$ contained in Φ, leading to an unstable computation of the phase function. Such severe errors in $\Phi(x)$ result in unreliable disparities.

In contrast, from the definition of the norm of $\Psi(x, s)$ in Eq. (12), it can be seen that singularities will be tolerated in shift-trials as it has no contribution to the integration in the continuous case and only an insignificant contribution in most of the discrete cases.

Hence, it can be expected that the shift-trials method will be more robust to noise and singularities than the Newton iteration methods or other direct searching methods based on local calculations, such as the Golden section method.

2.3 Global minimum and periodic property

As the Fourier components at a specific wavelength λ, both Gabor filtered left and right images are functions with a period λ. So are the left and right phases:

$$\Phi_p(x + \lambda) = \Phi_p(x) \qquad p = l, r \tag{23}$$

Therefore, the norm function of the phase difference $\|\Psi(x, s)\|$ will be a function with the same period λ with respect to shift s, because

$$\|\Psi(x, s + \lambda)\| = \|\Phi_l(x) - \Phi_r(x + s + \lambda)\| = \|\Phi_l(x) - \Phi_r(x + s)\| = \|\Psi(x, s)\| \tag{24}$$

Let the disparity and the shift trial are $s^*, s \in [0, \lambda)$ respectively and suppose the singularities and noise can be ignored. Then from Eq. (3), (4) and (13), we have

$$\Psi(x, s) = \Phi_l(x) - \Phi_r(x + s) = \omega x - \omega(x - s^* + s) = \omega(s^* - s) \tag{25}$$

So, $\Psi(x, s)$ is a linear, monotonic function over $[0, \lambda)$ in the absence of modulation of phase functions. After being modulated into $(-\pi, \pi]$, $\Psi(x, s)$ becomes discontinuous when it jumps between $-\pi$ and π. However, $|\Psi(x, s)|$ always remains as a piecewise linear, continuous function over $[0, \lambda)$, which has the unique minimum value 0 at the position s^* no matter s^* is zero or not.

Therefore, from the definition in Eq. (14), we have

$$\|\Psi(x, s)\| = \int_0^T |\omega(s^* - s)| dx = |\omega(s^* - s)|T \tag{26}$$

This shows that with respect to s, the norm $\|\Psi(x, s)\|$ has similar behaviour to $|\Psi(x, s)|$ over $[0, \lambda)$.

Hence, the minimal point s^* is the global minimum point of $\|\Psi(x, s)\|$ over $[0, \lambda)$. Once a minimal point s has been obtained, it must be the desired minimal point s^* of the the norm function.

Note that the singularity of phase functions has insignificant contribution to the norm $\|\Psi(x, s)\|$ and noise has the same effect on each (circular) shift trial $s \in [0, \lambda)$. Therefore it can be expected that the above conclusion can be maintained in practice.

2.4 Computational cost

The computational cost of the shift-trials method relates to the sampling window width T and the disparity s.

Without loss of generality, suppose the shift trials are made over $[0, \lambda)$. When the unknown disparity is zero, there must be $\|\Psi(x, s)\| > \|\Psi(x, 0)\|$ $\forall s \in (0, \lambda)$. So, the searching has to run over the whole interval $[0, \lambda)$ as illustrated in Fig. 1.

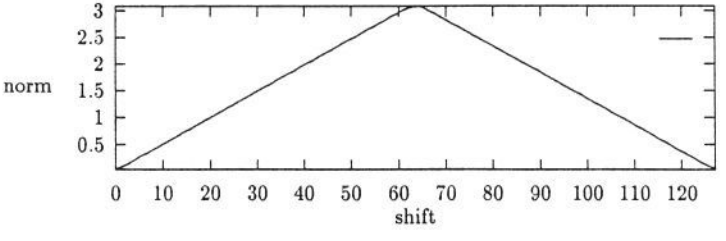

Figure 1: The searching runs over the whole shift range when the minimum point of the norm curve $s^* = 0$ is unknown (wavelength $\lambda = 128$).

On the contrary, when the unknown disparity is non-zero the desired shift s^* must be positive, therefore once a minimal point $s > 0$ has been obtained, further searching becomes redundant. So, when the disparity (*i.e.* the shift s^*) is small, the

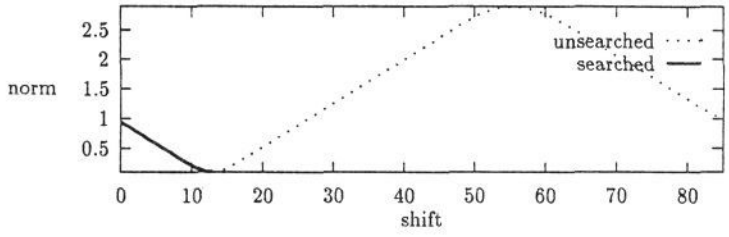

Figure 2: A further searching is unnecessary once a minimal point of the norm curve $s^* = 13$ has been attended.

searching will end at an early stage with a low computational cost as illustrated in Fig. 2.

But worse cases will occur when the disparity $s^* \in [\frac{\lambda}{2}, \lambda)$, as illustrated in Fig 3, where the searching has to finish at a late stage with a higher computational cost.

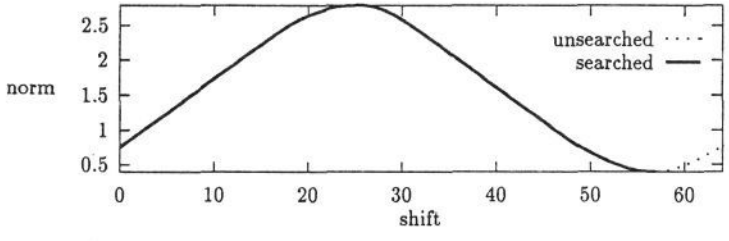

Figure 3: A worse case: more computation is needed in searching for the unmodulated minimum shift point $s^* = 57$ when wavelength $\lambda = 64$.

In this case, according to the periodic property in Eq. (24), $\|\Psi(x, s)\|$ must first increase, then decrease to the minimum point. Hence, a comparison of $\|\Psi(x, s)\|$ at $s = 0$ and $s = 1$ is enough to indicate the tendency of $\|\Psi(x, s)\|$. By letting $\hat{s} = \lambda - s^*$, we get the following results from Eq. (24):

$$\|\Psi(x, s^*)\| = \|\Psi(x, \lambda - \hat{s})\| = \|\Psi(x, -\hat{s})\| \tag{27}$$

where $\hat{s} \in (0, \frac{\lambda}{2}]$ is right the modulated value of s^* (cf. Eq. (10)).

This means that a search along the inverse direction of shifting will lead to the same modulated minimum point $\hat{s} = \lambda - s^*$, but at a reduced computational cost, as illustrated in Fig. 4.

The above treatment also shows that the searching no longer needs to run over the whole range $[0, \lambda)$ as before, a half range run is sufficient. Hence, the actual computational cost in all cases will be proportional to the disparity s^* itself as well as the sampling window width T. The average complexity of the shift trials method is therefore proportional to the wavelength $\frac{\lambda}{4}$ at which it works.

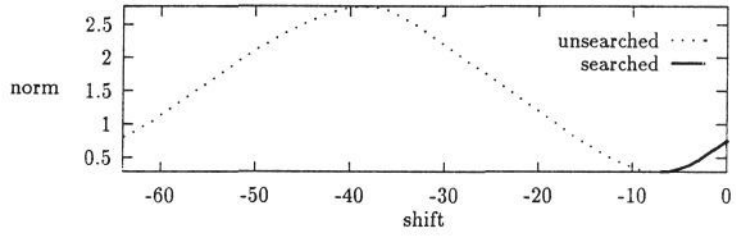

Figure 4: Computational cost is reduced by searching along the inverse direction of shifting for the modulated shift point $\hat{s} = -7$ when $\lambda = 64$.

In this experiment, to reduce the effects of noise, the shift-trial process is applied around the peak frequency of the power spectrum rather than across the whole frequency range $(0, f_c]$. In many cases the frequency range bounded by the left and right image's peak frequencies is a narrow interval, even a single frequency bin. This too reduces significantly the total computational cost at different tuning wavelength levels.

3 Experimental results

The shift-trial method was tested with synthetic and real data. Results of three sets of data are shown in this section. One set of data is a 1-D harmonic signal composed of three frequencies:

$$y = 3\cos\frac{2\pi x}{13} + 4\sin\frac{2\pi x}{13} + 2\cos\frac{2\pi x}{55} + 5\sin\frac{2\pi x}{55} + \cos\frac{2\pi x}{64} + 7\sin\frac{2\pi x}{64} \qquad (28)$$

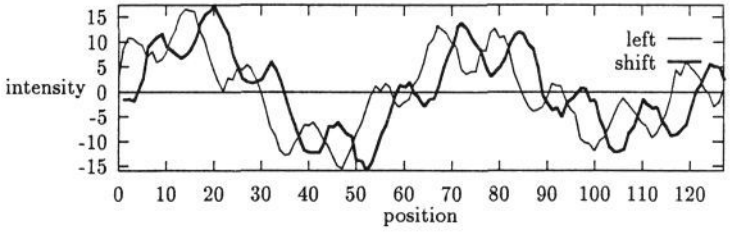

Figure 5: Left and right images of a 1-D harmonic signal.

As shown in Fig. 5 the left image is the signal and the right image its shifted version (circular shift = 5), both images are corrupted by white noise $\sigma(2,\ 0.7)$.

The shift-trials were made at the tuning wavelength $\lambda = 64$ of the Gabor filtering. The result is illustrated in Fig. 6 along with result yielded by Sanger's method as a comparison.

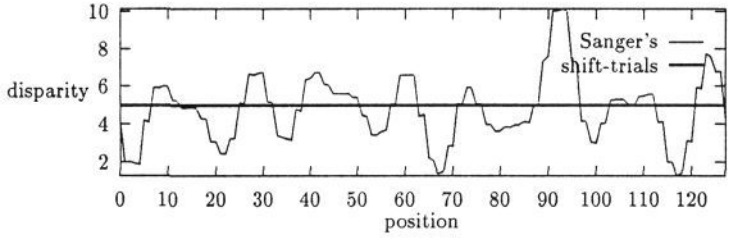

Figure 6: Disparity between the left and right views in Fig. 5 detected by Sanger's and shift-trials' methods respectively.

The second set of data is a 1-D section of a random dot stereogram[5]. The left and right views and the disparities detected by the shift-trial and Sanger's methods (at wavelength = 14) are illustrated in Fig. 7 and Fig. 8 respectively.

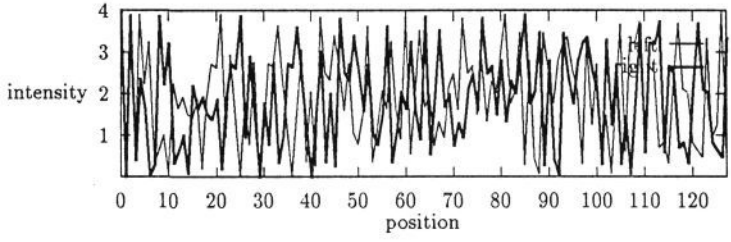

Figure 7: The left and right views of a 1-D random-dot stereogram (target width = 120, image width = 128).

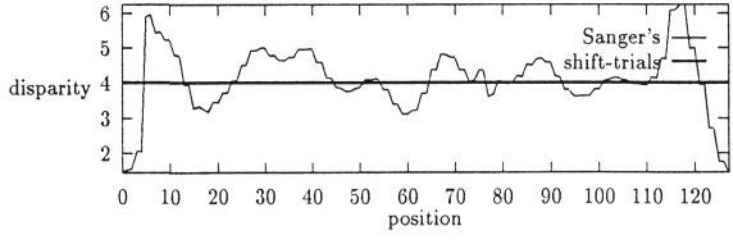

Figure 8: Disparities of the 1-D random-dot stereogram detected by Sanger's and shift-trials' methods respectively.

The third set of data is a pair of stereo images of a shell as shown in Fig. 9. The disparity detection runs for their 1-D section at the scanline = 127 as shown

in Fig. 10. The tuning wavelength $\lambda = 85$ is the peak wavelength indicated by a spectral power analysis. Results of both Sanger's and shift-trials' methods are illustrated in Fig. 11.

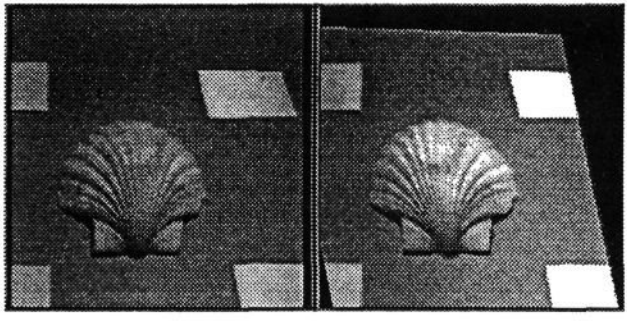

Figure 9: The left and (rectified) right views of a shell.

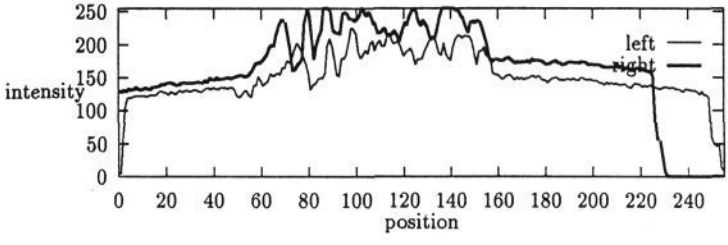

Figure 10: The left and right sections (scan-line=127) in Fig. 9.

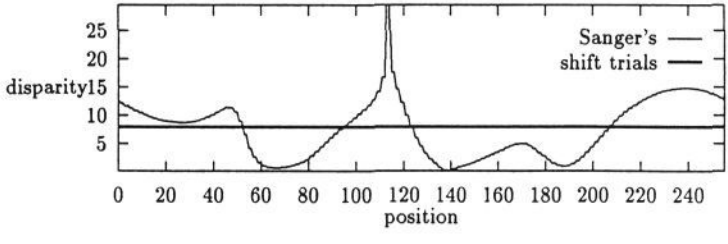

Figure 11: Disparity of the the above sections (scan-line=127) detected by Sanger's and shift-trials' methods respectively.

4 Discussion

So far, the detection of the disparity is up to the precision of one pixel, which is the unit of the sampling interval of the given data. Hence, a sub-pixel precision can be made when a denser data set is produced by using interpolation. However, given that the errors resulting from data noise could be larger than one pixel, a method of one-pixel precision seems enough for target detection in practice.

Note that the disparity detected by the shift-trial method is a global quantity according to the definition of the norm function in Eq. (14). While it is an advantage for shifted harmonic data and real data as shown in Fig. 6 and Fig. 11, it could meet difficulty in some cases of random dot stereograms if the target width is far from the image width, leading to a significant inconsistency between the target's disparity and the background's disparity that is assumed to be zero. As a result the shift-trials method will yield an intermediate value of the above two disparities. However, this problem can be resolved by using a smaller sampling window whose width is close to the size of object width plus margins to account for the expected shift. Thus the effect from the inconsistency can be ignored as shown in Fig. 7.

Acknowledgements

This research was funded by a grant of the ESPRIT VOILA project.

References

[1] Cai, L.D., 1991: A note on Sanger's and Fleet's disparity estimation methods. AIVRU Mem., University of Sheffield, October, 1991.

[2] Fleet, D.J., Jepson, A.D., and Jenkin, M.R.M., 1989: Phase-Based Disparity Measurement. RBCV-TR-89-29, University of Toronto, Nov. 1989.

[3] Gabor, D., 1946: Theory of Communication. *IEE* **93**, 429-459.

[4] Jenkin, M.R.M., Jepson, A.D., and Tsotsos, J.K., 1987: Techniques For Disparity Measurement, RBCV-TR-87-16, University of Toronto, September 1987.

[5] Julsez, B., 1971: *Foundations of cyclopean perception.* University of Chicago Press, Chicago, 1971.

[6] Marr, D., and Poggio, T., 1979: A computational theory of human stereo vision. *Proc. of Royal Society, London* **B204**, 301-328 (1979).

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

[8] Sanger, T.D., 1988: Stereo Disparity Computation Using Gabor Filters. *Biological Cybernetics* **59**, 405-418 (1988).

[9] Westelius, C-J, Knutsson H. and Wiklund J., 1992: Robust Vergence Control Using Scale-Space Phase Information. LiTH-ISY-I-1363, Linköping University, 1992.