

Optimization of Stereo Disparity Estimation Using the Instantaneous Frequency

M. Hansen, K. Daniilidis, G. Sommer
Institut für Informatik und Prakt. Mathematik
Christian-Albrechts Universität Kiel
Preusserstr. 1-9, 24105 Kiel, email: mha@informatik.uni-kiel.de

Abstract

The use of phase differences from Gabor filter responses is a well established technique for the computation of stereo disparity. It achieves the subpixel estimation of disparity without applying a correspondence search. However, the success of the phase approach has not been thoroughly analyzed. Here, we study the effects of varying filter frequency on the disparity estimation and we compare it to the use of the instantaneous frequency. The analytical results on several models of intensity and disparity variation yield disparity estimates robust against variations in the filter-wavelength and a new technique for frequency tuning of the quadrature filter pairs that considerably reduces the problem complexity.

Categories: stereo, low level processing

Optimization of Stereo Disparity Estimation Using the Instantaneous Frequency

Summary page

(a) What is the original contribution of this work?

This paper studies the use of local phase differences for stereo disparity computation. The robustness of phase methods has been repeatedly shown in experiments but has not been thoroughly analyzed yet. Furthermore, its application presumes the selection of the appropriate central frequency of the Gabor-filters involved. Here, we analytically study the behavior of the disparity estimate with respect to the central frequency and the associated bandwidth. We will show that the instantaneous frequency enables a disparity computation independent of the central frequency over a wide range of scales. The practical effects are the reduced number of Gabor-filters covering the frequency space and the robustness against spatial disparity variation and disparity differences in the left and the right image.

(b) What is the most closely related work by others and how does this work differ?

The most closely works are the studies of Maki et al. [10], Langley et al. [8] and Fleet-Jepson [4].

Maki et al. [10] studied the discrepancy between actual phase of the signal and the local phase computed from Gabor-responses. However, they did not derive an analytical expression for a simple signal like the one in our work and in [8]. The focus of the study was the effect of bandwidth size on the disparity error. The instantaneous frequency issue was not addressed.

Langley et al. [8] found a similar formula like ours for the above mentioned discrepancy with a sinusoidal input. However, they delved into the quadrature approximation without studying the relation between scale, disparity and instantaneous frequency.

Fleet and Jepson [4, 3] studied extensively the scale behavior of the local phase and its derivative (instantaneous frequency). The stability of phase was analyzed with respect to singularities and the spatial linearity of the local phase. However, the emphasis was on the singularities whereas our findings concern the global variation of disparity error and instantaneous frequency with respect to scale. Last but not least, our study leads to a practical implication which substantially reduces the algorithmic complexity.

Optimization of Stereo Disparity Estimation Using the Instantaneous Frequency

Abstract

The use of phase differences from Gabor filter responses is a well established technique for the computation of stereo disparity. It achieves the subpixel estimation of disparity without applying a correspondence search. However, the success of the phase approach has not been thoroughly analyzed. Here, we study the effects of varying filter frequency on the disparity estimation and we compare it to the use of the instantaneous frequency. The analytical results on several models of intensity and disparity variation yield disparity estimates robust against variations in the filter-wavelength and a new technique for frequency tuning of the quadrature filter pairs that considerably reduces the problem complexity.

1 Introduction and related work

The computation of stereo disparity is a necessary step before depth reconstruction from two views of the environment. The majority of the approaches is based on searching for the maximum correlation between an area in the left image and the area in the second image (cf. [2] for an extensive review) that corresponds to the projection of the same scene point. If the cameras are calibrated and the epipolar geometry is known this search is one-dimensional along the epipolar line in the right image corresponding to a point in the left image. Area-based approaches find the best correlation between two areas in the original or band-pass filtered images. Feature-based approaches first extract characteristic gray-value structures in both images which are then matched by applying a similarity criterion. Both groups suffer under the necessity of a search which increases the complexity and makes an explicit analysis intractable.

The paradigm of active vision introduced new prospects in stereoscopic vision. Real-time constraints on the complexity of the disparity computation are necessary in order to achieve a reactive behavior. Thus, techniques requiring a search for every image position are not appealing for real-time responses. On the other hand, active vision enables the control of the mechanical degrees of freedom of a stereo set-up. The vergence control of a camera yields a decrease on the disparity magnitude in a considerable area around the fixation point. Small disparities can now be computed with local techniques using only the responses of appropriate filters. It was shown that the local disparity variation can be computed from the intensity derivatives [5, 6]. However, a more general filter-based approach exhibited the most robust performance until now: the phase-based approach introduced in [11, 12] and further developed and studied in [3, 8, 13, 14, 10].

The phase-based disparity computation may be apparently justified by the shift-theorem for the Fourier-transform of an image. However, the shift-theorem is exact only for a globally

constant disparity. To capture varying disparity, all approaches apply the local spectral representation of the complex-valued Gabor-filter responses. As we will describe later, the shift-theorem for the phase difference of the Gabor responses is not valid even in the case of simple sinusoidal input. A more plausible formulation of the shift-theorem [12, 3] assumes the preservation of the local phase in the left and the right image. Application of the Taylor expansion to the local phase yields the disparity as the ratio of phase-difference and the first spatial derivative of the phase if the higher order disparity terms and derivatives are neglected. The first derivative of the local phase is well known as the definition of the instantaneous frequency of a signal [1].

The phase matching aspect of the local shift-theorem still involves the selection of the appropriate frequency to compute the local phase as well as its derivative. The Gabor-filters applied have coupled bandwidth and central frequency like wavelets so that the filter support is small for high central frequencies and large for low ones. Thus, frequency tuning is equivalent to the scale selection in the scale-space framework. Applying a bank of Gabor-filters with different scales yields a disparity estimate for each scale. We could choose the central frequency of the maximal energy response and compute the phase from the response of the Gabor filter with the same central frequency. However, this is not valid for multi-component signals and it is computationally very expensive to densely sample the scale (or frequency) space. Furthermore, the left and the right image can have different frequency contents due to perspective effects so that the search for the maximum response must be done for both images.

The main contribution of this study is the result that the instantaneous frequency remains almost the same for a wide range of scales. This fact is analytically exhibited for prototype signals like a sinusoid and an edge. The practical implications are extremely useful: the instantaneous frequency can be computed from a small filter-bank by sparsely sampling the frequency-space. In a second step the disparity can be computed with a filter tuned to the instantaneous-frequency. We show, that even if the second step is omitted the disparity error is still low if the instantaneous frequency and a phase difference from an almost arbitrary tuned Gabor-filter are used. We analyze the behavior of disparity with respect to the relative bandwidth and we show that the above mentioned strategy is valid for the sizes of the filter-supports used in real-time constrained applications. The robustness of the instantaneous frequency is shown in stereo views with spatially varying frequency in the left and the right images as well as in images with spatially varying scale due to perspective foreshortening.

The dependence of local phase and instantaneous frequency on scale has been extensively studied by Fleet and Jepson [4, 3]. The emphasis was first on the effects of spatial deformations between the left and the right image on the preservation of scale. Second, they found out that sensitivity of disparity to position and scale appears due to singularities in phase as a function of position and scale.

Maki et al. [10] studied the discrepancy between actual phase of the signal and the local phase computed from Gabor-responses. However, they did not derive an analytical expression for a simple signal like the one in our work and in [8]. The focus of the study is the effect of narrow or wide bandwidth on the discrepancy factor. The instantaneous frequency issue is not addressed.

Langley et al. [8] found the same formula for the above mentioned discrepancy with a sinusoidal input. However, they delved into the quadrature approximation without studying the relation between scale, disparity and instantaneous frequency.

We give a brief introduction of phase based disparity estimation with a instantaneous and constant frequency model in section 2. The analytical study of the local phase and the estimated disparity for two gray-value variations is subject of sections 3 and 4. In section 5 we show that we can apply the analytical results of chapters 3 and 4 in a real world experiment.

2 Phase-based disparity estimation

Stereo disparity has a horizontal and a vertical component but only one of them suffices to compute the depth. As it is usual and justified in active binocular systems [7] we will use only the horizontal disparity. The filters we apply have a two-dimensional support with a smoothing component in the vertical direction. Therefore we restrict our analysis to disparity estimation in 1D-signals.

Let us suppose that the disparity d is constant over the image. Then, according to the shift-theorem of the Fourier transform a shift in spatial domain transforms to a modulation in frequency domain:

$$f(x) \text{ --- } \bullet \text{ --- } F(\omega) \quad f(x+d) \text{ --- } \bullet \text{ --- } F(\omega)e^{i\omega d}. \quad (1)$$

To recover the disparity as spatial shift the phase differences have to be computed:

$$d(x) = \frac{\phi_l(x) - \phi_r(x)}{\omega}. \quad (2)$$

Unfortunately disparity is hardly ever globally constant so that some sort of local spectral representation must be extracted. We apply here complex Gabor-filters with coupled bandwidth and central frequency. Since a bank of Gabor filter responses samples the frequency space we are asking whether the shift theorem still holds for the local phase extracted from the complex Gabor responses. In this case, the mean frequency of the left and the right image could be used in order to apply (1).

A slight variation of the shift-theorem can be obtained if we assume that local phase is preserved in the left and the right image [9, 3]

$$\phi_l(x+d) - \phi_r(x) = 0. \quad (3)$$

A Taylor series expansion of $\phi_l(x+d)$ leads to :

$$\phi_l(x+d) = \phi_l(x) + \phi'_l(x) \cdot d + \mathcal{O}(d^2) \quad (4)$$

The spatial shift d of the signal in the left image is defined in a first order approximation as:

$$d(x) = \frac{\phi_l(x) - \phi_r(x)}{\phi'_l(x)} \quad (5)$$

The derivative $\phi'_{l/r}(x)$ of the local phase is well known [1] as *instantaneous* or *local* frequency of the signal. Often the instantaneous frequency ω_{inst} is defined as the average of $(\phi'_l(x) + \phi'_r(x))/2$. We will next elucidate the importance of the phase derivative in the task of disparity estimation.

We use the following definition of the Gabor impulse response

$$G(x; \sigma, \omega_g) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma^2} e^{-i\omega_g x} \quad (6)$$

The bandwidth factor $t = \frac{1}{\omega_g \sigma}$ of the filter relates its support σ with the frequency ω_g . For example a bandwidth factor $t = 0.33$ yields a full bandwidth of one octave. Normally we choose the bandwidth factor t in the interval $(0.2, 0.7)$ in order to obtain from four down to one oscillations of the Gabor-harmonic inside the filter support.

We denote the Gabor filter responses by $f_g(x) = G(x; \sigma, \omega_g) * f(x)$. The magnitude or local energy of the right image response is $r(x) = |f_{g,r}(x)|$ and that of the left image is $l(x)$. The phases are denoted $\phi_r(x) = \arg[f_{g,r}(x)]$ for the right and $\phi_l(x)$ for the left image.

The results in disparity estimation with a phase based approach using Gabor filters depends strongly on the instantaneous frequencies of the signals due to the bandpass characteristic of the Gabor filter. We treat here two cases, the case of a sinusoid and the case of an edge modelled as a superposition of sinusoidal functions.

In the sinusoidal case the left and right signals read

$$\begin{aligned} T_r(x) &= \sin(\omega_0 x) \\ T_l(x) &= \sin(\omega_0(x + d_{tr})) \end{aligned}$$

where d_{tr} is the true disparity. The edge signals are given by following expansions

$$\begin{aligned} E_r(x) &= \sin(\omega_0 x) + \frac{\sin(3\omega_0 x)}{3} + \frac{\sin(5\omega_0 x)}{5} \\ E_l(x) &= \sin(\omega_0(x + d_{tr})) + \frac{\sin(3\omega_0(x + d_{tr}))}{3} + \frac{\sin(5\omega_0(x + d_{tr}))}{5} \end{aligned}$$

We will study the case of a linear ground-truth disparity $d_{tr}(x) = Ax + B$ analytically and the case of a quadratically varying disparity in simulations. Last we will apply our theory on real stereo images with known ground-truth.

3 Analysis of a sinusoidal pattern

In the case of a sinusoidal pattern the local phase $\phi_{l/r}(x)$ and its derivative $\phi'_{l/r}(x)$ are functions of the Gabor-filter parameter ω_g, t , the signal frequency $\omega_0 = 2\pi/\lambda_0$ and the parameters A, B of the disparity variation. Our goal is to study the effects of the above parameters on the instantaneous frequency ω_{inst} and the disparity estimate $d(x)$.

The convolutions of the harmonic functions $T_{r/l}(x)$ with the Gabor filter $G(x; \omega_g, t)$ yield responses with phases $\phi_{r/l}(x) = \text{arg}[G * T_{r/l}(x)]$ e.q:

$$\phi_r(x) = \arctan\left(-\frac{\tanh(\sigma^2 \omega_0 \omega_g / 2)}{\tan(\omega_0 x)}\right) \quad (7)$$

The corresponding first derivative of the local phase with respect to x reads:

$$\phi'_r(x) = \omega_0 \frac{\tanh(2\sigma^2 \omega_0 \omega_g)}{\left(1 - \frac{\cos(2\omega_0 x)}{\cosh(2\sigma^2 \omega_0 \omega_g)}\right)} \quad (8)$$

The same expressions are obtained for the left signal with $\tilde{\omega}_0 = \omega_0(A + 1)$ instead of ω_0 and the phase shifted by a constant offset $\omega_0 B$.

Without loss of generality we assume that $\omega_0^{max} \geq \omega_g^{max}$ for the case of a sinusoid, which means that the signal wavelength is smaller than the filter wavelength. The violation of this assumption is shown in Fig. 1.

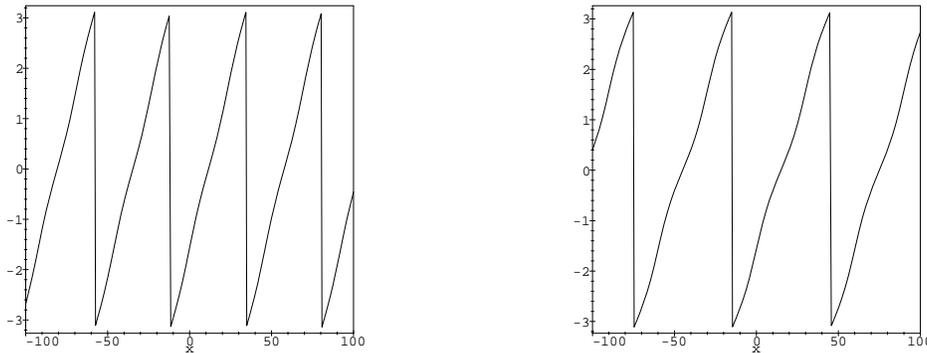


Fig. 1: Nonlinear behavior of the local phase corresponding equation to (7) if $\omega_0 < \omega_g$ instead of our assumption $\omega_0^{max} > \omega_g^{max}$. left: $\phi_l(x + Ax + B)$ and right: $\phi_r(x)$ ($\omega_0 = 1/6, \omega_g = 1/3, \sigma = 1.9$ and $A = 0.3, B = 0$.)

To improve the intuitive understanding of the above expressions we use wavelengths measured in pixels instead of frequencies with $\lambda_0 = 2\pi/\omega_0$ for the signal and $\lambda_g = 2\pi/\omega_g$ for the Gabor-filter, respectively.

For the disparity computation in eq. (5) we use the average of the left and right signal

$$\lambda_{inst}(x) = \frac{4\pi}{(\phi'_l(x) + \phi'_r(x))} \quad (9)$$

as the instantaneous wavelength $\lambda_{inst}(x)$ of the Gabor response [3]. Langley et al. [9] defined a magnitude weighted sum of the local phase derivatives instead of the average.

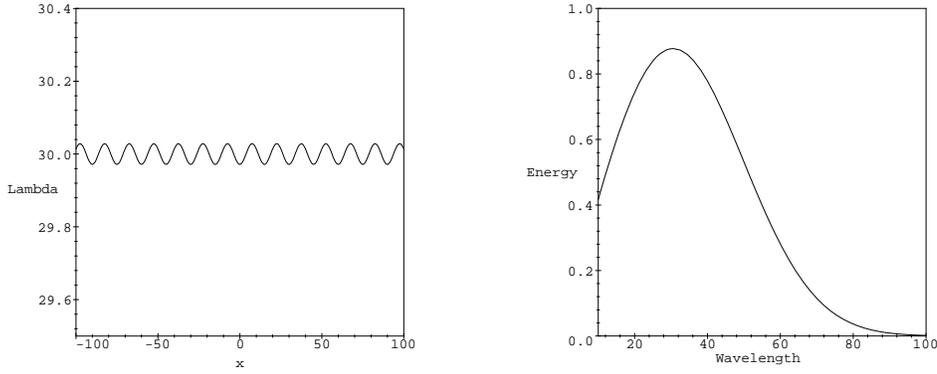


Fig. 2: Behavior of the instantaneous frequency $\lambda_{inst}(x)$ (left) and the local energy as a function of the Gabor wavelength λ_g (right) ($t = 0.66, \lambda_g = 50$ pixel) for a sinusoidal signal with $\lambda_0 = 30$ pixel.

We observe in eq. (8) and in Fig. 2 that the instantaneous wavelength spatially oscillates with a frequency $2\omega_0$ due to the fraction $\frac{\cos(2\omega_0 x)}{\cosh(2\sigma^2 \omega_0 \omega_g)}$ in eq. (8). In the special case $\lambda_g = \lambda_0 t^2$ the instantaneous wavelength oscillates between $\lambda_0 - 25\% < \lambda_{inst}(x) < \lambda_0 + 33\%$. However, the spatial average over a local region (Fig. 2, left) is still equal to λ_0 , as expected according to the theory on instantaneous frequency.

We observe in eq. (8) that ω_{inst} is equal to the signal frequency ω_0 only if $\tanh(2\sigma^2 \omega_0 \omega_g)$ is approximately one. This condition can be achieved if

$$\lambda_g > \lambda_0 t^2. \quad (10)$$

If this inequality is satisfied the instantaneous wavelength represents the signal wavelength $\lambda_{inst} \approx \lambda_0$ of the signal over a wide region in the Gabor parameter space (λ_g, t) (see Fig. 3).

In Fig. 2 (right) we observe as expected that the maximum of the local energy achieves its maximum if the Gabor-filter wavelength λ_g at a given pixel x coincides with the signal wavelength λ_0 . The practical effect is crucial: If we rely on a maximum response analysis in order to find the optimal scale we have to filter with different wavelengths to capture the variation of the responses in Fig. 2 (right). If we compute the instantaneous frequency just one filter response suffices to obtain the optimal scale with a negligible error if not exact when $\lambda_g > \lambda_0 t^2$ is satisfied. We will also see in the next section that the influence of the wavelength-bandwidth ratio t has also minor effects. The instantaneous wavelength is constant over the major part of the parameter space (λ_g, t) .

The computation of the instantaneous wavelength leads directly to an estimate of the optimal scale without any maximum search over different filter responses which is computational expensive especially for real time implementations.

3.1 Accuracy of disparity estimation in sinusoids

We address the accuracy of the disparity estimation for two models: the *constant* model using the filter frequency in eq. (2) and the *instantaneous* model using eq. (5). We choose the true disparity $d_{tr}(x) = Ax + B$ with $A = 0.1$ and $B = 1$. The choice of the disparity gradient $A \neq 0$ yields a scaling by $\frac{1}{(A+1)}$ of the signal wavelength λ_0 in the left view.

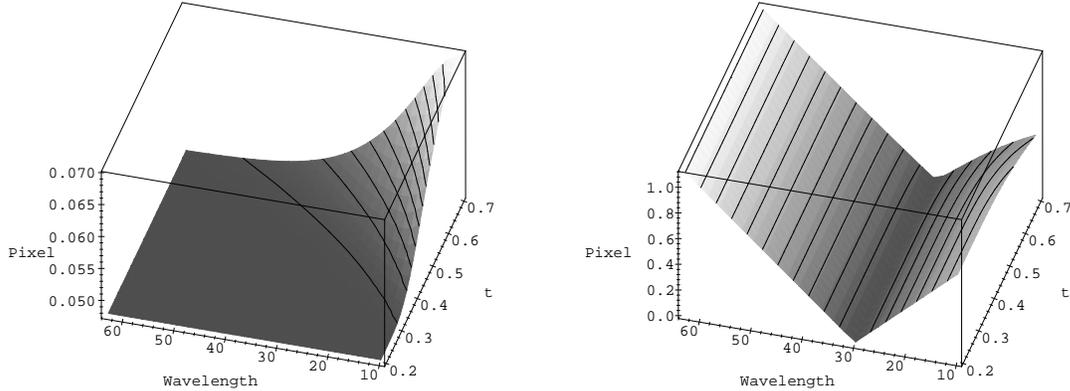


Fig. 3: Absolute error (geom. in pixel) of disparity estimation at position $x=0$ as a function of bandwidth factor t and Gabor wavelength λ_g . The signal wavelength is $\lambda_0 = 30$ pixel and the true disparity is one pixel. (left) The error increases for the instantaneous model, if $\lambda_g \approx \lambda_0 t^2$. (right) The error for the constant model is generally much more higher with the exception, if $\lambda_0 \approx \lambda_g$.

We analytically compute the disparity for both models and use the absolute error $\Delta d_{abs}(x) = |d_{tr}(x) - d_{est}(x)|$ to compare both models for a given position $x = 0$ where $d_{tr}(0) = 1$. Fig.3 shows the variation of the absolute disparity error at $x = 0$ as a function of the Gabor-filter parameters (λ_g, t) . The range $(\lambda_g, t) = (10 \dots 63, 0.2 \dots 0.7)$ of these parameters is according to the implementation in a hierarchical scheme in a real-time architecture.

In Fig.3 (left) we observe that the absolute disparity error is very low if the instantaneous frequency is used and condition (10) is satisfied. The phase difference $\phi_l(x) - \phi_r(x)$ and the instantaneous wavelength are nearly independent of the Gabor-filter wavelength λ_g and strongly dependent on the signal wavelength λ_0 for a wide variation of the bandwidth factor t . The constant model Fig.3 (right) is not able to deal with phase differences independent of λ_g and this leads directly to an error $\Delta d_{abs}(x) \sim |\lambda_g - \lambda_0|$ in disparity estimation.

The use of the instantaneous model (5) definitely leads to better results than the constant model (2). The relative error varies from 5% to 7% in the instantaneous and 0.1% to 100% in the constant model. The error of the instantaneous model is due to the disparity gradient A and will be analyzed next.

In Fig. 4 we show the disparity error as a function of filter wavelength and spatial position for the instantaneous model (left) and the constant model (right). For the majority of filter

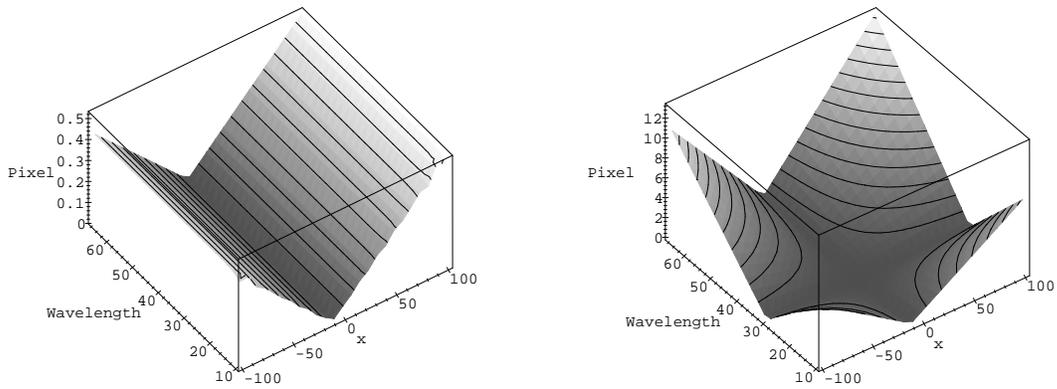


Fig. 4: Absolute error in disparity estimation for the instantaneous model (left) and the constant model (right) as a function of position x and Gabor-filter wavelength λ_g . The disparity is linear $d_{tr} = Ax + B$ with $A = 0.1, B = 1$ and the bandwidth factor is $t = 0.33$. The signal wavelength is $\lambda_0 = 30$ pixel.

wavelengths the instantaneous model outperforms the constant model. However, we notice that the absolute error in the constant model is very low independently of x if $\lambda_g \approx \lambda_0$.

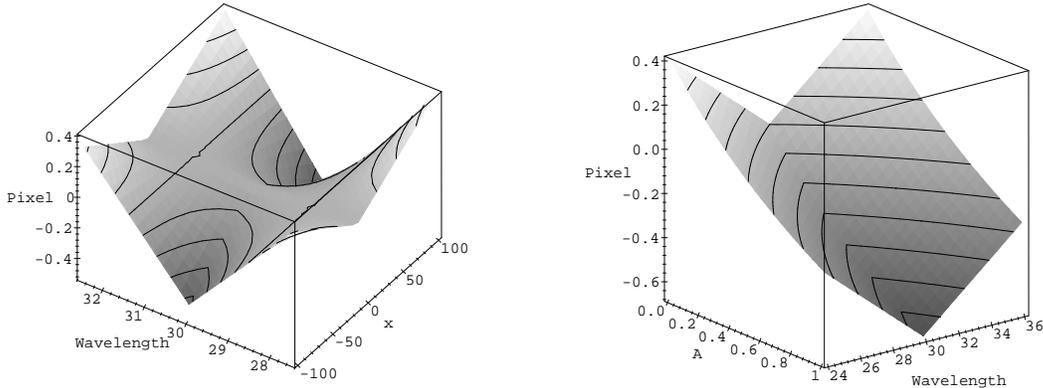


Fig. 5: Difference $\Delta d_{const.} - \Delta d_{inst.}(x)$ of the absolute errors in disparity estimation using the instantaneous and the constant model as a function of the filter wavelength and the spatial position x (left) and as a function of the disparity gradient A and the Gabor-filter wavelength λ_g (right). Negative values indicate that the constant model achieves a lower error than the instantaneous model for this small region near $\lambda_g \approx \lambda_0$.

We will inspect closer a small neighborhood at $\lambda_g \approx \lambda_0$. In Fig. 5 (left) we show the difference in the errors between the constant and the instantaneous models where negative

values indicate a lower error. We see that for the area between the two straight iso-contours the constant model exhibits lower error than the instantaneous one. The more the tuning wavelength differs from the ground-truth frequency the better becomes the instantaneous model. The variation with respect to x is coupled with the disparity gradient A . In Fig. 5 (right) the error difference is plotted as a function of the disparity gradient A and the wavelength range in the immediate neighborhood of the signal wavelength. The error in the instantaneous model increases with the disparity gradient A . Due to the decreasing signal wavelength $\tilde{\lambda}_0 = \lambda_0/(A + 1)$ in the left image, the instantaneous wavelength as the average of the right and left instantaneous wavelength (see definition. 9) is only an approximation for small disparity gradients in the signals. The absolute error for the instantaneous model can be estimated by $\Delta d_{abs}(x) \sim \frac{A}{A+1}$ derived from the difference between the instantaneous wavelengths in the left and the right signals.

4 Analysis in case of an edge

In order to guarantee that our superposition edge model (7) acts as an edge during filtering we restrict the filter wavelengths to be smaller than half the maximum wavelength of the superposition model: $\lambda_g \leq \frac{\lambda_0}{2}$. This restriction is illustrated in Figure 6. In fact, what we are doing is to sample the edge spectrum $I(\omega) \approx \frac{1}{\omega}$ at the positions ω_0 , $3\omega_0$ and $5\omega_0$.

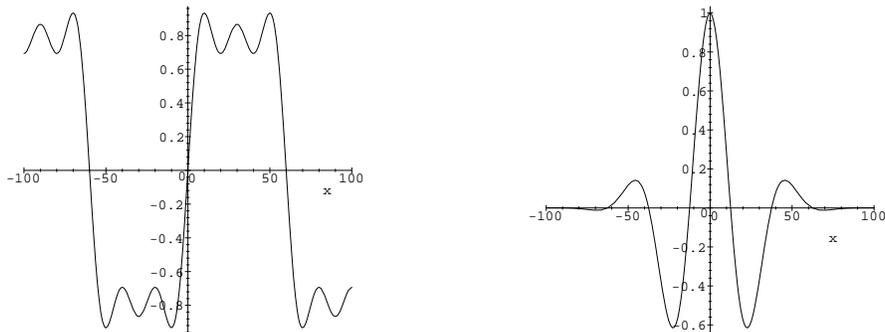


Fig. 6: The edge function with wavelength $\lambda_0 = 120$ pixel (left) and the odd part of the Gabor-filter with $\lambda_g = 50$ pixel $< \lambda_0/2$, $t = 0.33$ (right).

Convolving symbolically ¹ the Gabor-filter with the edge functions $E_{r/l}(x)$ leads to much longer analytical expressions for the local phase $\phi_l(x)$, $\phi_r(x)$ and the instantaneous wavelength λ_{inst} than in the sinusoid case. We will show the plots instead of bothering the reader with page-long formulas.

¹Using the symbolic package MAPLE.

We show in Fig. 7 the instantaneous wavelength and the energy as a function of the filter-wavelength λ_g . The inherent frequencies of the signal are manifested as local plateaus and local maxima in the instantaneous wavelength and the local energy, respectively.

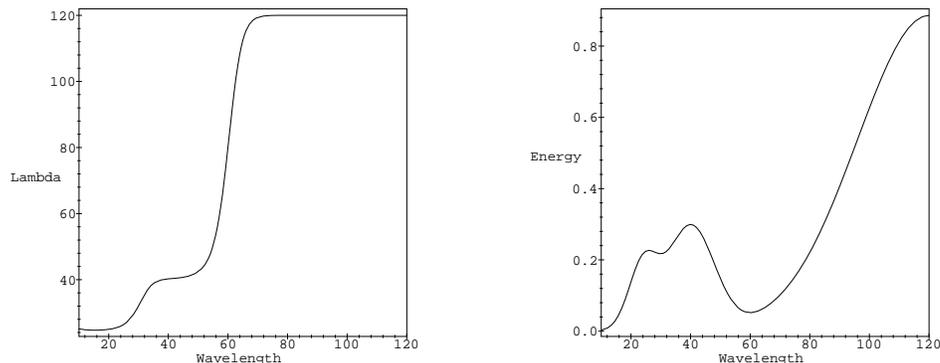


Fig. 7: (left) Behavior of the instantaneous wavelength λ_{inst} with $\lambda_0=120$ pixel and (right) the local energy at an edge as functions of the Gabor-filter wavelength λ_g at position $x = 0$.

4.1 Accuracy of disparity estimation at an edge

The assumptions and definitions are applied as in the case of an edge. The behavior of the disparity error as a function of filter-wavelength λ_g and t is the same as in the sinusoidal case (Fig. 3). To avoid replication we show the effect of filter-bandwidth on the disparity error in the instantaneous model (Fig. 8 left) and in the constant model (Fig. 8 right), respectively.

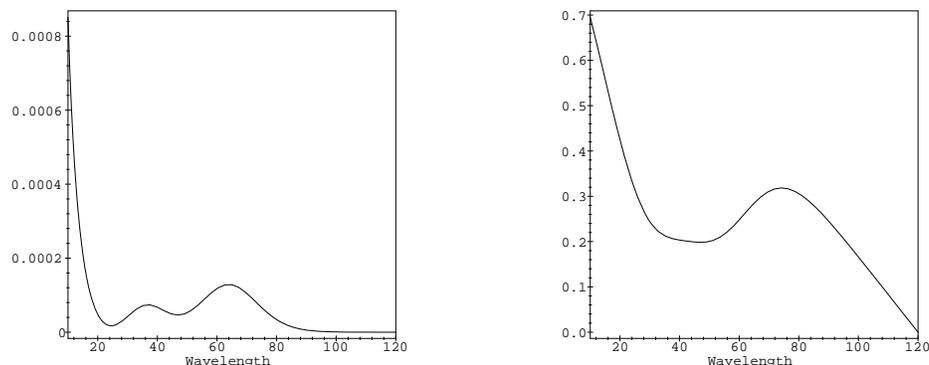


Fig. 8: Absolute error $\Delta d_{abs}(x)$ at position $x=0$ as a function of Gabor-filter wavelength λ_g for the instantaneous (left) and the constant model (right). ($t=0.4$ edge wavelength is $\lambda_0=120$ pixel, $d_{tr}=1$ pixel).

Using the instantaneous model the relative error decreased from 10% - 70 % to $\approx 0\%$ - 0.1 % in comparison to the constant model.

We next study the reasons for the high error in the constant model which increases even more in the neighborhood of an edge. The local phase does not behave linearly in the neighborhood of an edge. Although the disparity may be constant the phase difference is not constant and varies with $\lambda_0/2$, the wavelength corresponding to the considered signal region (Fig. 9 left).

On the other hand, the phase derivative is also oscillating because the inequality $\lambda_g \geq \lambda_0 t^2$ concerning the sinusoidal case is no more valid. The filter wavelengths are always larger than λ_0 due to the superposition model applied. Therefore, we observe the oscillating pattern in the instantaneous frequency (Fig. 9 right) with the same period $\lambda_0/2$, as that of the phase difference (Fig. 9 left). When the phase difference is divided by the phase derivative to obtain the disparity the two oscillating patterns cancel because they have almost the same offset regarding x . This cancelation explains the low error in the instantaneous model and is the reason for the high error of the constant model, too. The disparity errors still oscillates with high magnitude because the phase difference is divided by a spatially constant frequency.

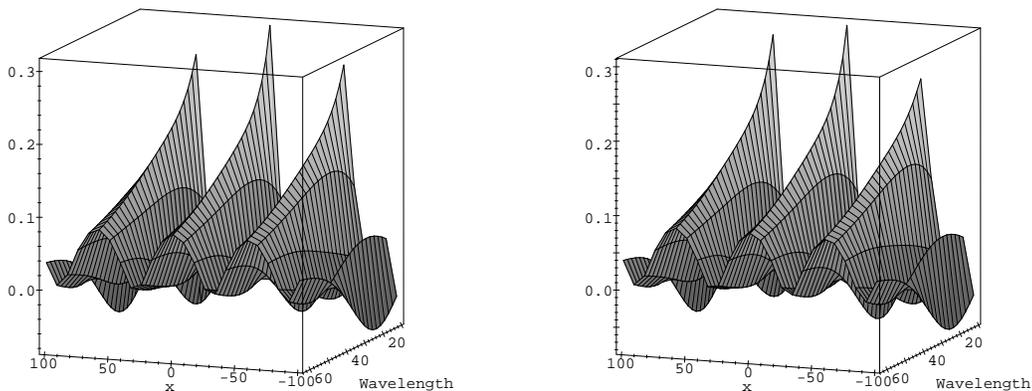


Fig. 9: The oscillations of the phase difference (left) and the phase derivative (right) with a frequency ($\approx 2\omega_0$) as a function of the Gabor-filter wavelength λ_g and the position x . ($t = 0.66$, $\lambda_0 = 120$ and $B = 1$ pixel.)

5 Simulations with perspective foreshortening

In the previous sections we could treat all expressions analytically because we assumed spatially linear disparity. In this section we will study a quadratic disparity model in order to describe the perspective effects of a viewed slanted plane. We will again find out that the expressions concerning the simple sinusoid in section 3 are still useful if we assume nonlinear disparity. Fixating to a point of a flat surface leads to an approximately second order disparity model $d(x) = Cx^2 + Ax$. By fixation the disparity near the center is vanished

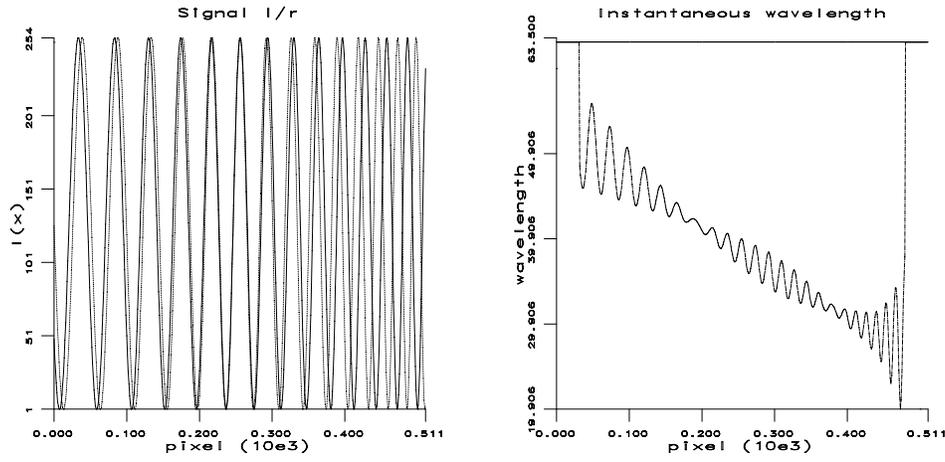


Fig. 10: On the left of the figure we show superimposed the signal of the left view (dotted line) and that of the right view (continuous line). The reader can observe the vanishing disparity in the center due to fixation. On the right we show the instantaneous wavelength (average of the left and right phase derivatives) as a function of spatial position x with a Gabor-filter of wavelength $\lambda_g = 63$ pixel (left), computed in a simulation.

and the wavelengths in both signals (left/right) are similar. The obtained stereo signals are shown on the left of Fig. 10. The measured instantaneous wavelength λ_{inst} for one of the Gabor-filter wavelengths is shown in Fig. 10 (right). The decrease of λ_{inst} is overlaid by an oscillation in x due to the small violation of the inequality constraint (10).

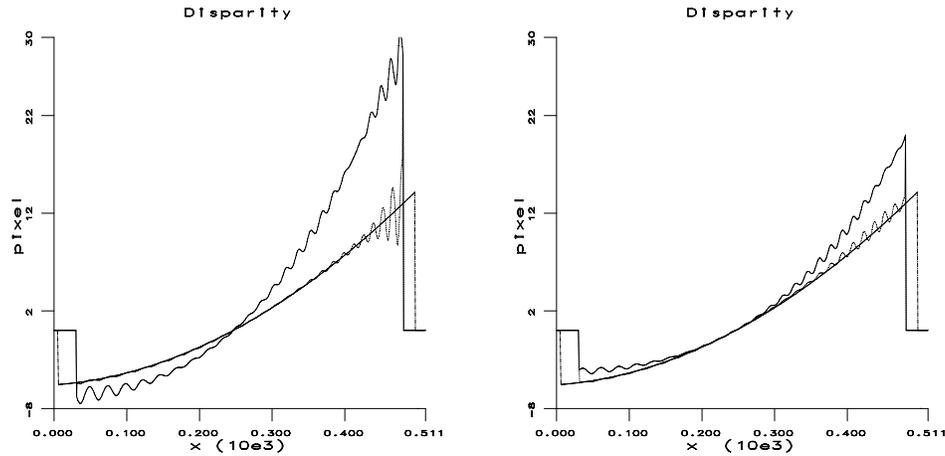


Fig. 11: True disparity (continuous smooth line), disparity estimate for the instantaneous model (dotted line) and disparity estimate for the constant model (continuous line far from the ground truth) for a Gabor-filter $\lambda_g=63$ pixel are shown on the left. On the right we show the same curves but for a Gabor filter tuned to the center instantaneous frequency $\lambda_{inst}(x = 256)$

This oscillation affects also the error in the disparity plotted on the left of Fig.11 obtained by both models for a Gabor filter with wavelength $\lambda_g=63$ pixels. We repeat the same curves on the right of Fig.11 but this time obtained from filtering with a Gabor with wavelength equal to the instantaneous wavelength at the center. We observe that the constant model

performs better but does not reach the performance of the instantaneous model which is an order of magnitude closer to the ground-truth at almost every position in the image. The results (Fig.11) emphasize that the instantaneous model provides good results without a special filter tuning.

5.1 Application to real images

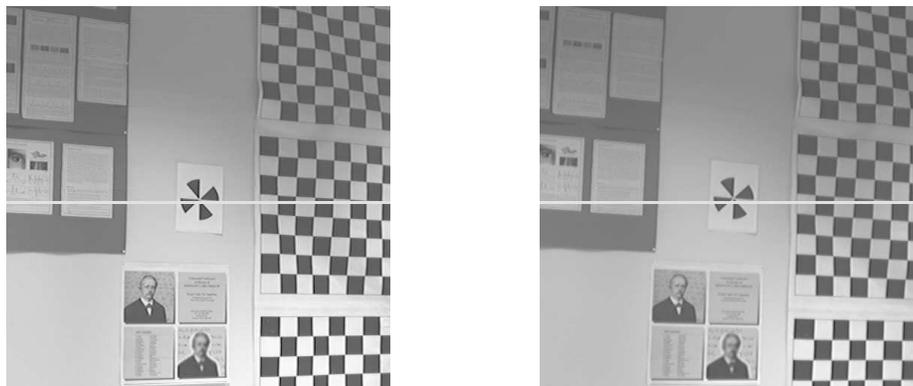


Fig. 12: Stereo images of our lab with some well structured regions at the wall. The system fixates the black cross. The white line indicates the 1D-signal used in further plots.

We will apply the methods of the previous section at a *real* stereo image pair which contains edges and textured areas (Figure 12). The stereo system fixates the black cross at the wall. The *true* disparity in Fig. 13 is computed from the simulation by calibrating the extrinsic and intrinsic camera parameter in relation to the wall. Due to the errors resulting from the erroneous calibration an exact quantitative comparison is not possible. In the following we will use this calibration result as ground-truth.

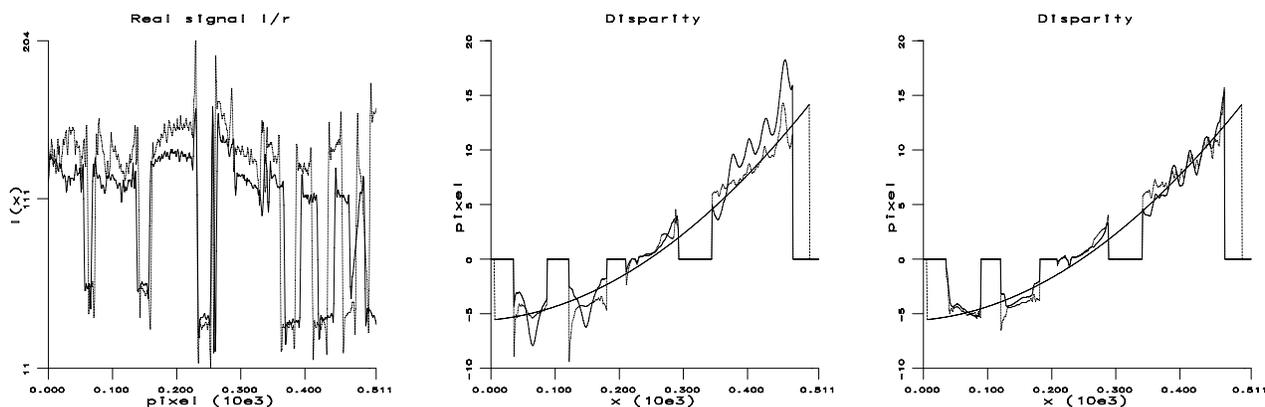


Fig. 13: Signals (l/r) of the real images with different amplitudes at row 260 through the image center (left). True and estimated disparity for a Gabor-filter ($\lambda_g=63$ pixel, $t=0.66$) (middle) and a Gabor-filter tuned ($\lambda_g=55$ pixel, $t=0.33$) to the instantaneous wavelength $\lambda_{inst}(x = 256)$ (right). The smooth line represents the true disparity.

The results in Fig. 13 for this row (and others) are nearly the same as in the simulation. The instantaneous model provides a good approximation of the *true* disparity with both

Gabor-filters (Fig. 13 middle). The constant model (Fig. 13 right) has to be tuned to the measured instantaneous wavelength $\lambda_{inst}(256)=55$ of the row center in order to get a good local approximation near the center. The zero disparities outside the center are regions of unreliable phase information where the Gabor-filter response is below a chosen threshold (15% of the magnitude maximum).

Both models were also applied in the entire image without tuning. The disparity map (Fig. 14) on the left is the ground truth masked with a confidence map of areas of stable phase information which is always necessary in such approach. The instantaneous and the constant disparity model are shown on the middle and left of Fig. 14, respectively.

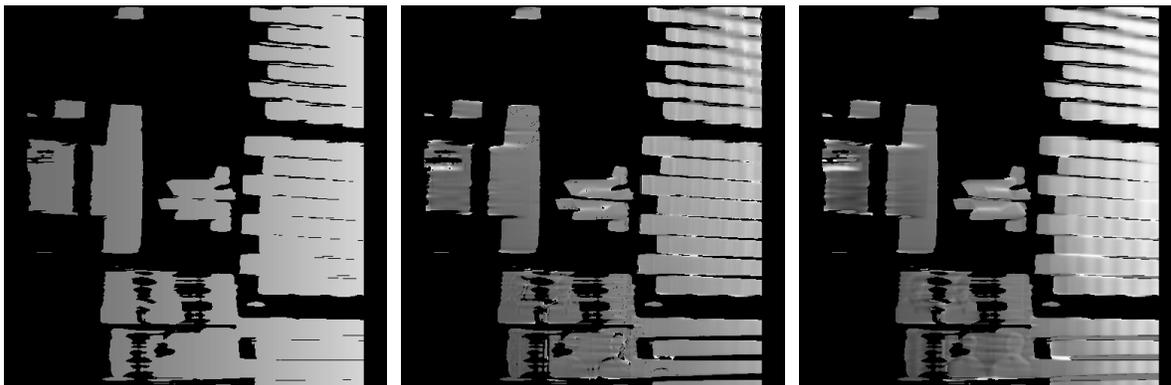


Fig. 14: Qualitative results in disparity estimation: (left) the *true* disparity map of the images pair. (middle) Disparity map estimated with the instantaneous model, (right) disparity map estimated with the constant model ($\lambda_g=63$ pixel, $t=0.33$).

6 Conclusions

We have analytically studied the behavior of the constant and the instantaneous model for disparity estimation using complex Gabor- filters. The instantaneous model has a superior performance over the entire spectrum of filter-wavelengths whereas the constant model results are not useful at all over the most wavelengths. Until now the procedure was to optimize by tuning near the frequency of maximal response. This procedure could be computationally expensive and can fail if such an unique maximum does not exist. Tuning the filters to the instantaneous frequency improves their performance, a step that in most cases is not even necessary. Taking the phase differences over the instantaneous frequency performs perfectly for most filter-wavelengths.

References

- [1] B. Boashash. Estimating and interpreting the instantaneous frequency of a signal-part 1: Fundamentals. *Proceedings of the IEEE*, 80:520–539, 1992.
- [2] U.R. Dhond and J.K. Aggarawal. Structure from stereo - a review. *IEEE Trans. Syst. Man. and Cyber.*, 19, 1489-1510 1989.
- [3] D. J. Fleet, A. D. Jepson, and M. R. M. Jenkin. Phase-based disparity measurement. *CVGIP: Image Understanding*, 53(2), 3 1991.
- [4] D.J. Fleet and A.P. Jepson. Stability of phase information. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15:1253–1268, 1993.
- [5] J. Garding and T. Lindeberg. Direct computation of shape cues using scale-adapted spatial derivative operators. *International Journal of Computer Vision*, 17:163–191, 1996.
- [6] D.G. Jones and J. Malik. Computational framework for determining stereo correspondence from a set of linear spatial filters. *Image and Vision Computing*, 10:699–708, 1992.
- [7] E. Krotkov and R. Bajcsy. Active vision for reliable ranging: Cooperating focus, stereo, and vergence. *International Journal of Computer Vision*, 11:187–203, 1993.
- [8] K. Langley, T.J. Atherton, R.G. Wilson, and M.H.E. Larcombe. Vertical and horizontal disparities from phase. In *Proc. First European Conference on Computer Vision*, pages 315–325. Antibes, France, Apr. 23-26, O.D. Faugeras (Ed.), Lecture Notes in Computer Science 427, Springer-Verlag, Berlin et al., 1990.
- [9] K. Langley, T.J. Atherton, R.G. Wilson, and M.H.E. Larcombe. Vertical and horizontal disparities from phase. In *Proc. 1st ECCV*, pages 315–325, 1990.
- [10] A. Maki, L. Bretzner, and J.-O. Eklundh. Local fourier phase and disparity estimation: An analytical study. In *V. Hlavac et al. (Ed.), Proc. Int. Conf. Computer Analysis of Images and Patterns CAIP, Prag*, pages 868–874, 1995.
- [11] W.M. Miller. Video image stereo matching using phase-locked loop techniques. In *Proc. IEEE Int. Conf. on Robotics and Automation*, pages 112–117, 1986.
- [12] T.D. Sanger. Stereo disparity computation using gabor filters. *Biol.Cybernetics*, 59:405–418, 1988.
- [13] W. M. Theimer and H. P. Mallot. Phase-based binocular vergence control and depth reconstruction using active vision. *CVGIP: Image Understanding*, 60(2):343–358, 12 1994.
- [14] C.J. Westelius, H. Knutsson, J. Wiklund, and C.F. Westin. Phased-based disparity estimation. In J.L. Crowley and H.I.Christensen, editors, *Vision as Process*, pages 157–178. Springer Verlag, Berlin et al., 1994.