Density and Accuracy Improvement of Phase-Based Disparity

M. H. Ouali *† D. Ziou † C. Laurgeau *

†Dépt. de Mathématiques et d'Informatique
Université de Sherbrooke
Sherbrooke, PQ Canada J1K 2R1

*Centre de Robotique
École des Mines de Paris
60, bd Saint-Michel, 75006 Paris, France

Abstract

This paper tackles the recurrent problem of disparity estimation. The measurement of image disparity is a fundamental precursor to binocular depth estimation. The mapping from disparity to depth is well understood, while the automatic disparity extraction is still subject to errors. We propose to use the image derivatives with the phase-based approach to overcome the tuning problem of the filter. Moreover, we propose a quadratic model for the singularities neighborhood detection and the phase quasi-linearity will be revisited. The approach is characterized by the simplicity of its implementation. It also provides dense and accurate disparity maps. A numerical error analysis against a ground-truth shows that the results are very satisfactory.

1 Introduction

This work aims at the perception and the understanding of three-dimensional scenes in computer vision. 3D information is widely used in a large number of practical applications: robot navigation, aerial or satellite mapping, medical imaging, and part inspection to name a few. Stereo is one of the 3D structure extraction techniques. Stereo allows the computation of the depth using two images of the same scene, taken with two cameras having different viewpoints. The primary task of stereo is to locate pairs of pixels that are images of the same point of the scene.

Many approaches have been proposed for disparity measurement. These approaches differ from one another in the matching primitives, the density of the results, the accuracy of the estimates and the underlying computation time. The most reported methods in the literature are feature-based, correlation-based and energy minimization-based matching. Recently, several works proposed phase-based techniques to disparity estimation. One of the first applications to exploit phase was the Kuglin-Hines method [1]. The method, which is conceptually simple, uses only the Fourier phase and assumes that the two images are related by pure translation. Sanger [2] proposed a method using complex Gabor filter. He extracts the phase from the complex response of Gabor filter, and uses the Fourier shift theorem to compute the disparity. Other earlier algorithms expressed the computational task as a nonlinear differential equation that must be solved at each image point [3]. The solving of a differential equation at a large number of image points and disparities makes the algorithm unsuitable for real time computer applications. Many other authors have used Gabor filter since the work of Sanger [4, 5]. The work of Fleet [6] and Fleet *et al.* [7] improved this approach. Fleet [7] employs a coarse-to-fine strategy, first used in this application area by [8], in order to allow the filter's frequency to vary. However, the disparity estimates are not accurate when phase at coarse scales is unavailable. Weng [9] used the windowed Fourier phase. He proposes the use of a set of variable-size windows to compute the disparity.

In this paper, we will motivate the choice of the band-pass filter which will be used for disparity estimation. Furthermore, a theoretical study showed that for some functions the phase is either linear or quasi-linear [10]. This finding allows us to assume without loss of generality that the phase of the response is linear or locally linear. We propose to use the images and their derivatives instead of a set of filters to overcome the tuning problem. Furthermore, the linear model used in the singularity neighborhood detection is improved by the quadratic model which seems to better approximate the data. In section 2, we present the disparity as Fourier phase difference and motivate the use of the Gabor filter. The phase behavior, the singularities detection and the algorithm are subject of section 3. In section 4, we show the experimental results.

2 Disparity as phase difference

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

$$f(x - \Delta x) \circ\!\!-\!\!\bullet\, \mathcal{F}\{f\}(\omega)\, e^{-j\omega\Delta x} \qquad (1)$$

where f is the image and $\mathcal{F}\{f\}$ is its Fourier transform.

The fundamental idea of the phase-based approach to disparity measurement is to recover the disparity as phase difference observed in the Fourier domain. Let $l(x)$ and $r(x)$ be the left and right images of the pair. They are related to each other by the shift $l(x) = r(x - \Delta x)$, and thus in the frequency domain by $L(\omega) = e^{-j\omega\Delta x}R(\omega)$, where $L(\omega)$ and $R(\omega)$ denote the Fourier transforms of the image pair. In terms of phase, we have the property

$$\phi_l(\omega) = \phi_r(\omega) - \omega\Delta x \tag{2}$$

where $\phi_l(\omega) \equiv \arg L(\omega)$ and $\phi_r(\omega) \equiv \arg R(\omega)$. Given a frequency ω_0, we can then recover the disparity Δx, using the difference of phases, $\phi_l(\omega_0)$ and $\phi_r(\omega_0)$, at that frequency:

$$\Delta x = \frac{\phi_r(\omega_0) - \phi_l(\omega_0)}{\omega_0} = \frac{\Delta\Phi}{\omega_0} \tag{3}$$

Unfortunately, the recovered disparity is global and cannot be assigned to a particular region in the image. This is due to the Fourier transform: the phase terms in the equation (3) are computed from all the points of the signal. In order to have a local disparity estimate, it is necessary to define a window inside which the points are picked up to compute the phase terms. Unluckily, according to the scaling property, the Fourier transform of $f(ax)$ is $\frac{1}{|a|}\mathcal{F}\{f\}(\frac{\omega}{a})$ where $\mathcal{F}\{f\}(\omega)$ denotes the Fourier transform of $f(x)$. In other words, if the analysis window is localized in the time domain, we will be unable to know precisely which frequency is present in this window because there will not be only a single frequency but a large frequency range. In this case, the term ω_0 in the equation (3) has to be identified.

Time-frequency decomposition suggests a trade-off between the extreme (space and frequency) representations. It will be possible to know the frequency content of a signal around each image point with an uncertainty. If we want the decomposition coefficients to have a direct space and frequency information, we must define a localized transform in the time and frequency domains simultaneously. The resolutions in both domains must be acceptable. Gabor [11] proposed to use an exponential modulated Gaussian as a decomposition function. As a matter of fact, the Gaussian function is well localized in both space and frequency domains. More precisely, it achieves the minimum uncertainty. We measure the antagonism between the space and frequency supports by the uncertainty principle. The proof is beyond the scope of this paper but can be easily verified [11, 10].

Gabor filters are, by far, the most commonly used tool in joint space-frequency analysis. They are chosen for the minimum space-frequency uncertainty product and for the separability of center frequency and bandwidth [11, 12]. A Gabor filter, tuned to a frequency ω_0, is created in the spatial domain by modulating a Gaussian envelope function, with standard deviation σ, by a complex harmonic with frequency ω_0. The design parameters of Gabor filters are the standard deviation σ of the Gaussian envelope and the center frequency ω_0 of the complex harmonic. The definition in the spatial domain is given by

$$g(x;\sigma,\omega_0) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{-\frac{x^2}{2\sigma^2}}\, e^{j\omega_0 x} \tag{4}$$

and, therefore, in the frequency domain by

$$\mathcal{G}(\omega;\sigma_f,\omega_0) = \mathcal{F}\{g\}(\omega;\sigma,\omega_0) = e^{-\frac{(\omega-\omega_0)^2}{2{\sigma_f}^2}} \tag{5}$$

where $\sigma_f = 1/\sigma$. The real (even) and imaginary (odd) components of the filter are shown in figure (1) as well as its frequency spectrum.

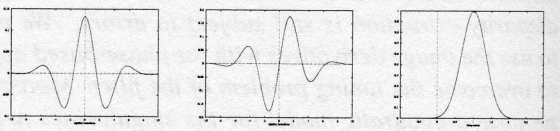

Figure 1: The complex Gabor filter. Left: real component. Middle: imaginary component. Right: its frequency spectrum.

3 Disparity measurement: from the signal or from its derivatives?

Standard phase-based techniques use the original signal to estimate the disparity. However, a single filter with a given tuning is not sufficient to estimate the disparity adequately because of the large frequency range of the image: the filter will loose a lot of frequency information. Instead of considering a set of filters, we will use the signal derivatives to deal with a large frequency range. In the following, $l(x)$, respectively $r(x)$, will denote the left, respectively the right, signal or its derivatives.

Let us assume that there is a constant disparity across some parts of the image and that the phase of the Gabor response is linear. The linearity assumption will be discussed later. Then, in that regions we have $l(x) = r(x - \Delta x)$. Because the filter is complex, so is the response to a given image. For any point x_0 and particular Gabor filter with center frequency ω_0, we can form the convolution products:

$$\begin{cases} C_R(x_0,\omega_0) &= \int r(x)g(x_0 - x;\sigma,\omega_0)\, dx \\ C_L(x_0,\omega_0) &= \int l(x)g(x_0 - x;\sigma,\omega_0)\, dx \\ &= \int r(x - \Delta x)g(x_0 - x;\sigma,\omega_0)\, dx \end{cases} \tag{6}$$

and the complex phase difference may be written as:

$$\begin{aligned}\Delta\Phi(x_0,\omega_0) &= \arg[C_L(x_0,\omega_0)] - \arg[C_R(x_0,\omega_0)] \\ &= \phi_l(x_0,\omega_0) - \phi_r(x_0,\omega_0) \end{aligned} \quad (7)$$

Following the same reasoning as in the beginning of section 2, we approximate $\Delta x(x_0,\omega_0)$ by $\Delta\Phi(x_0,\omega_0)/\omega_0$ as in equation (3). The local phases, $\arg[C_L]$ and $\arg[C_R]$, are derived from the output of the Gabor bandpass filter. This local phase approximation, and then the disparity estimation, remains strictly valid only for filters of infinitesimal bandwidth, as follows directly from the Fourier shift theorem. As we are with non-infinitesimal bandwidth filters, the frequency ω_0 in the denominator is no longer adapted to compute the disparity. Then, we have to estimate this frequency [6]. To show this, figure (2) illustrates an example with two signals, each one shifted to each other. The spatial positions of the maxima of the signals are x_1 and x_2. In the phase representation, we see that the displacement $\Delta x = x_1 - x_2$ is related to the phase difference and to the slope $\tan\alpha$ of the phase curve. Since the phases are assumed to be locally linear, we have $\tan\alpha = d\phi/dx$, which is the local spatial frequency. Therefore, the computation of the disparity estimates is straightforward:

$$\Delta x(x_0,\omega_0) = \frac{\Delta\Phi(x_0,\omega_0)}{\left(\frac{d\phi}{dx}\right)(x_0,\omega_0)} \quad (8)$$

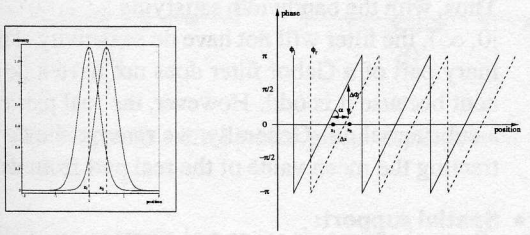

Figure 2: Disparity from phase differences and local spatial frequency.

In practice, the phase signal is not too linear and for those regions the disparity estimates are not reliable. We must define the areas where the measurements should not be trusted [2, 7]. In the following subsections, we will see the global behavior of the Gabor output phase, the detection of the singularities neighborhood as well as the algorithm and its implementation requirements.

3.1 Phase quasi-linearity

Although the output phase depends on the nature of the input, nevertheless we can expect a global behavior. For our purposes, it is helpful that the phase exhibits a linear behavior. An investigation shows that the output phase to an input Gaussian function, which may represent a line in the image, with standard deviation σ_1, has the form

$$\phi(x) = \frac{\omega_0\sigma^2}{\sigma_1^2+\sigma^2}\,x \quad (9)$$

and therefore is linear in x, σ and ω_0 being the filter's parameters. However, the output phase to a sine function, with frequency $2\pi/\omega$, is given by

$$\phi(x) = \arctan\left(-\frac{\tanh(\sigma^2\omega\omega_0)}{\tan(\omega x)}\right) \quad (10)$$

and seems to be not linear. But if we examine its derivative, we find

$$\phi'(x) = \omega\frac{\tanh(\sigma^2\omega\omega_0)}{1-\frac{\cos^2(\omega x)}{\cosh^2(\sigma^2\omega\omega_0)}} \quad (11)$$

The function $\cos^2$ is bounded by 1. In contrast, the function $\cosh^2$ has a fast growth. The term $\cos^2(\omega x)/\cosh^2(\sigma^2\omega\omega_0)$ can therefore be neglected for some combination of the parameters, see figure (3). Even when the analytical expression is not linear, it is possible to find a trade-off of the filter's parameters so that the global behavior is quasi-linear. From now on, we assume that the output phase is generally quasi-linear.

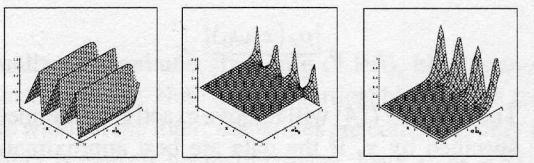

Figure 3: The phase is quasi-linear for $\sigma^2\omega_0 \geq 4$. Left: the phase. Middle: its derivative. Right: the term $\frac{\cos^2(\omega x)}{\cosh^2(\sigma^2\omega\omega_0)}$ vanishes for $\sigma^2\omega_0 \geq 4$.

3.2 Singularity neighborhood

To derive equation (8), we have assumed that the phase is linear. In the regions where the phase is not linear, the disparity will not be computed accurately. Furthermore, when the phase nonlinearity occurs, the magnitude is generally weak, see figure (4).

Fleet *et al.* [7] have proposed two criteria to remove such areas. The first criterion (12) removes all the areas where the instantaneous frequency deviates from the tuning frequency when the second criterion (13) removes the neighborhood adjacent to singular points, *i.e.* where the magnitude is very weak:

$$\frac{|\phi_x(x,\omega_0)-\omega_0|}{\sigma_f} \leq \tau_\phi \quad (12)$$

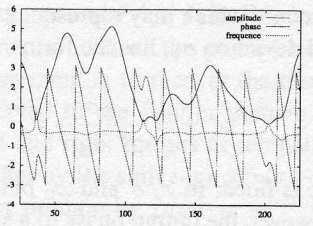

Figure 4: Relationship between low magnitude and nonlinear phase.

$$\sigma \frac{|\rho_x(x,\omega_0)|}{\rho(x,\omega_0)} \leq \tau_\rho \tag{13}$$

where $\rho(x,\omega_0)/|\rho_x(x,\omega_0)|$ in equation (13) approximates the distance from the singularity to x with a linear magnitude model. However, near a singularity, the magnitude behaves like a parabola. It follows that the criterion (13) removes the double of the area size which is intended to be removed in case where the data are best approximated by a parabola. Indeed, considering a quadratic model, near a singularity, the magnitude can be written as $\rho(x,\omega_0) = \alpha(x-x_s)^2$, where x_s is the singularity location. With this model, the distance from the singularity to x is approximated by $2\rho(x,\omega_0)/\rho_x(x,\omega_0)$. The criterion (13) is rewritten as:

$$\sigma \frac{|\rho_x(x,\omega_0)|}{2\rho(x,\omega_0)} \leq \tau_\rho \tag{14}$$

The criterion (14) will remove exactly the number of pixels specified by τ_ρ if the data are best approximated by a parabola. In singular locations, $\rho(x,\omega_0)$ and $\rho_x(x,\omega_0)$ tend toward zero and neither the criterion (13) nor (14) can be applied. We add a new constraint to remove the points where the magnitude is sufficiently small. The remaining pixels satisfy the constraint:

$$\rho(x,\omega_0) \geq \tau \tag{15}$$

3.3 Algorithm and implementation details

The algorithm consists in the following steps:

- convolve both images, and their derivatives, with the Gabor filter tuned to a given frequency

  for each derivative order $(0..n)$ do

  - compute the phase, the magnitude and the instantaneous frequency
  - compute the disparity
  - remove all the areas which do not satisfy the criteria (15), (12) and (14)

- combine the n disparity estimates, obtained for each pixel: the final disparity is that which minimizes the error

$$e = \min_{i\in[0..n]} (\Sigma_{\xi\in N} |left_i(\xi) - right_i(\xi + d_i(\xi))|)$$

  where i denotes the derivative order, N being the neighborhood of the location to be checked.

There are a number of requirements to be considered when designing phase-based disparity estimator filters:

- **Only one half plane of the frequency domain:**
  We illustrate this point by two examples. Let us consider that the signal is a Gaussian function with standard deviation σ_1. The phase of the Gabor response to this signal is $\phi(x) = \frac{\sigma^2\omega_0}{\sigma^2+\sigma_1^2}x$. Since the phase is assumed to be linear and either it is increasing or decreasing, it is clear that ω_0 must be either positive or negative. In the case of a sine function with frequency ω, $\phi(x) = \arctan\left(-\frac{\tanh(\sigma^2\omega\omega_0)}{\tan(\omega x)}\right)$. Here again, ω_0 must be either positive or negative. To conclude, we limit ourselves to the right half-plane.

- **No dc component:**
  Flat and constant surfaces in the image do not yield 3D information. These surfaces correspond to very low frequencies. The filter must not respond to such frequencies. We should have $\omega - \sigma_f > 0$, *i.e.* $\omega > \sigma_f$. Thus, with the bandwidth satisfying $[\omega-\sigma_f, \omega+\sigma_f] \subset]0,\infty)$, the filter will not have dc sensitivity. The imaginary part of a Gabor filter does not have a dc component because it is odd. However, the real part can have a substantial dc. Generally, we remove the dc by subtracting the mean value of the real part from itself.

- **Spatial support:**
  As the spatial support grows, the support in the frequency domain decreases and the bandwidth tends to be infinitesimal as for the Fourier transform. In this case, the filter can handle large disparities but the hypothesis of constant disparity in the window may be transgressed. Furthermore, the computational cost of the convolution increases, and the extracted disparity is global. With small spatial support, the bandwidth is large and the computational cost decreases but the filter can handle only small disparities.

4 Experimental results

We will show the performance of the algorithm on two test images. The first test example is an edge shifted by a 5 pixels disparity, see figure (5.a). The filter's parameters are $\sigma = 12.5$ and $\omega_0 = 0.24$. The filter is tuned to relatively

high frequency. For this example, since the region that has moved is small, the filter size must be small as well. Figure (5.b) shows the magnitude of the responses to the step edges of figure (5.a). Figure (5.c) shows the phase of the responses to the same step edges. From figure (5.c), the phase curves can be assumed quasi-linear. The figure (5.d) presents the disparity estimates. The disparity is not computed everywhere because of flat surfaces. In these areas the phases of the two signals are equal because the filter is convolved with the same homogeneous surfaces and the phase difference is zero. The disparity is recovered in the region that has moved. The most accurate disparity estimate corresponds to the maximum of both response magnitudes, left and right, cf. figure (5.b) and (5.d). Elsewhere, the estimates corrupt as the magnitudes decrease.

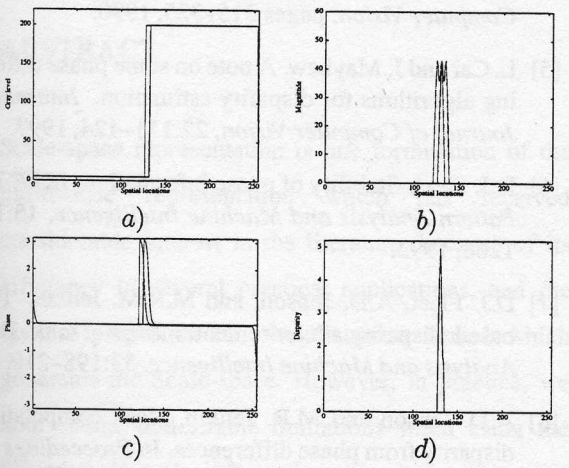

Figure 5: Step edge shifted with 5 pixels disparity, $\sigma = 2$ and $\omega_0 = 0.24$.

The next example is a scene that is made up of a chessboard and some pawns, see figure (6). The background of the scene has zero disparity, when the different pawns and the chessboard have disparities inversely proportional to their depth. The chessboard is an inclined plane with respect to the depth. So the disparity estimates are expected to be decreasing (from the bottom to the top of the image). The nearer pawn must have large disparity values.

The figure (7 left) shows the results obtained on this scene, using the images and their first and second derivatives ($n = 2$). The filter is tuned to a wavelength of 16 pixels, corresponding to the center frequency $\omega_0 = 0.39$ and the scale $\sigma = 7.63$ pixels. The criterion (12) is applied with $\tau_\phi = 1.1$ so that instantaneous frequencies are accepted up to 10% outside the nominal range of the filter. The criterion (14) discards the pixels within σ of a singularity ($\tau_\rho = 1.0$). Finally, τ is set to 0.04 in equation (15) removing all the magnitudes below of 4% of the maximal

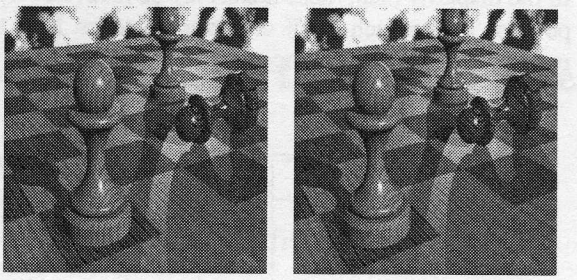

Figure 6: Chessboard scene. Left and right images (courtesy of J. Perrin).

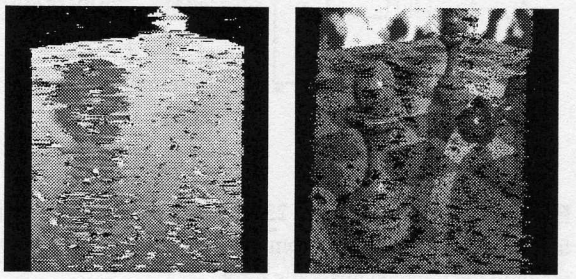

Figure 7: Left: disparity estimates. Right: image reconstruction.

normalized magnitude. In figure (7 left), black stands for the zero disparity. Note that the vertical black strips are due to the width of the filter: disparities are computed only when the filter is entirely inside the image. Dark areas are farther while the lighter ones are closer. The disparity is recovered successfully in the background (skyline). Figure (8) shows the disparity estimates using the original signal (8.a), its first derivative (8.b), its second derivative (8.c) and the combination of the estimates (8.d). The constraints used to suppress the wrong estimates are tight. This causes the removal of large areas. We note that the removed areas do not correspond. The combination of the estimates is intended to increase both the density and the accuracy. Table 1 shows the attainable densities and accuracies of the algorithm. It is obvious that the combination increases the density of the estimates. The error percentages are given with respect to the density, and the density with respect to the whole image. This means that the accuracy is also increased: 67.40% of 49.71% have a null error using the signal when 65.96% of 80.14% have a null error after the combination. The results reflect the general shape of the scene. From the left image, figure (6 left) and the computed disparity, figure (7 left), we have reconstructed the right image (each pixel in the left image was shifted with the corresponding disparity), figure (7 right). Black spots in figure (7 right) are due to the areas removed by the constraints in equations (12), (14) and (15).

In the disparity map, these areas may be recovered by interpolation to approach 100% of correct matches. Note that no preprocessing nor post processing was carried out.

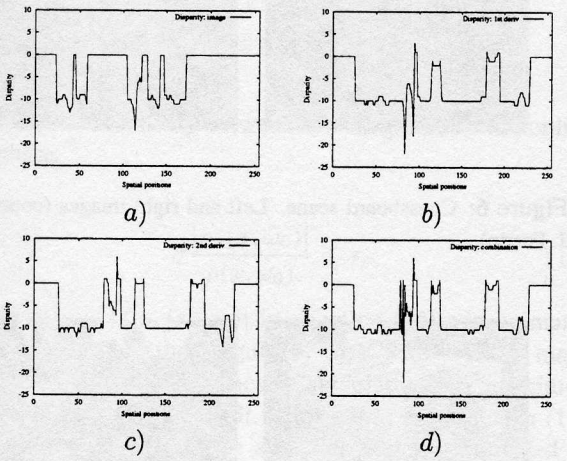

Figure 8: Disparity at line 235. a) from the original signal. b) from the 1^{st} derivative. c) from the 2^{nd} derivative. d) Combination yielding the minimal quadratic error.

Type/Err.	Density	0 pixel	1 pixel	2 pixels	3 pixels	Q. Err.
Signal	49.71%	67.40%	82.25%	87.72%	90.71%	0.03
1^{st} der	79.69%	62.28%	76.74%	83.01%	86.76%	0.23
2^{nd} der	78.12%	62.02%	75.89%	81.65%	85.25%	0.17
combin.	80.14%	65.96%	78.31%	83.31%	86.29%	0.02

Table 1: Comparison of the densities and the accuracies of the estimates.

5 Conclusion

Gabor phase is a useful tool providing a natural mechanism for binocular disparity estimation. It avoids many of the problems of correspondence algorithms. Its use is motivated with the Fourier shift theorem. An advantage of phase information is its stability with respect to contrast differences between the left and right views. In some regions, the phase signal has an unwanted behavior. These regions have a low magnitude, corresponding to a lack of signal. Disparity estimates should not be trusted in these regions. The criteria gave good results and the combination increased both the density and the accuracy of the estimates. The results obtained are very satisfactory and help improve our theoretical understanding.

References

[1] C.D. Kuglin and D.C. Hines. The phase correlation image alignment method. In *Proceedings of the IEEE International Conference on Cybernetics and Society*, pages 163–165, 1975.

[2] T.D. Sanger. Stereo disparity computation using gabor filters. *Biological Cybernetics*, 54:405–418, 1988.

[3] M.R. Jenkin, A.D. Jepson, and J.K. Tsotsos. Techniques for disparity measurement. *Computer Vision Graphics and Image Processing: Graphics Models and Image Understanding*, 53(1):14–30, 1988.

[4] K. Langley, T.J. Atherton, R.G. Wilson, and L.H.E. Larcombe. Vertical and horizontal disparities from phase. In *Proceedings of 1st European Conference on Computer Vision*, pages 315–325, 1990.

[5] L. Cai and J. Mayhew. A note on some phase differencing algorithms for disparity estimation. *International Journal of Computer Vision*, 22:111–124, 1997.

[6] D.J. Fleet. Stability of phase information. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 15:1253–1268, 1993.

[7] D.J. Fleet, A.D. Jepson, and M.R.M. Jenkin. Phase-based disparity measurement. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 53:198–210, 1991.

[8] A.D. Jepson and M.R. Jenkin. Fast computation of disparity from phase differences. In *Proceedings of the IEEE CVPR, San Diego, CA*, pages 398–403, 1989.

[9] J. Weng. Image matching using the windowed fourier phase. *International Journal of Computer Vision*, 11:211–236, 1993.

[10] M.H. Ouali, D. Ziou, and C. Laurgeau. Filtre de gabor: Principe d'incertitude et quasi-linéarité de la phase. Technical Report 221, Dept of Mathematics and Computer Science, University of Sherbrooke, Sherbrooke, Canada, 1998.

[11] D. Gabor. Theory of communication. *J. Inst. Electr. Eng.*, 93:429–441, 1946.

[12] J. Daugman. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *Journal of the Optical Society of America*, 2:1160–1169, 1985.