Machine Vision and Applications Manuscript N?
(will be inserted by hand later)

Performance of Phase-based Algorithms for Disparity Estimation

A. Cozzi[1] *, **B. Crespi**[2] **, **F. Valentinotti**[3] ***, **and F. Wörgötter**[1] †

[1] Ruhr-Universität, Dept. of Neurophysiology, Bochum, Germany
[2] Istituto per la Ricerca Scientifica e Tecnologica (IRST), Trento, Italy
[3] ENEA, Roma, Italy

August 9, 1996

Abstract. We characterize the performances of phase-based disparity estimators, giving quantitative measures of their precision and their limits, and how changes in contrast, imbalance, and noise modify the attainable accuracy. A first result is that the theoretical range of measurable disparities (one period of the modulation of the filter) is not attainable: the actual range is circa 2/3 of a period of the modulation of the filter. The phase-based disparity estimators have shown to be robust to changes in contrast of 100% and more and to well tolerate imbalances of luminosity between the images composing the stereo pair up to 400%. Our tests show that the biggest cause of error is noise: the error increases linearly with the increase of the noise level. We conclude studying the influence of the spectra and the luminosity of the input images on the error surface, both for artificial and natural images.

1 Introduction

Stereoscopy is a technique to extract depth information from two images of a scene taken from different view points. This information is derived from the relative positions of the projections of an object in the two images that compose the stereo pair. The difficult part of the technique is to find the pixels that are images of the same physical object.

In a first approximation the positions of corresponding pixels on the two images are related by a one-dimensional shift, the *disparity*, along the direction of the epipolar lines. For parallel optical axes, the epipolar lines are parallel to the line joining the optical centers of the cameras. In the following we consider the most simple case, where epipolar lines coincide with the x-axis of the images. In this configuration the two-dimensional search problem is reduced to a one-dimensional problem, that we can solve considering each pair of scan-lines independently.

* cozzi@neurop2.ruhr-uni-bochum.de
** crespi@irst.it
*** valenti@dsxis3.casaccia.enea.it
† worgott@neurop.ruhr-uni-bochum.de

We denote the one-dimensional signal of each corresponding pair of scan-lines as $f_{\mathrm{R}}(x)$ and $f_{\mathrm{L}}(x)$, where the subscript indicates that the scan-line comes from the right or the from left image of the stereo pair.

In the phase-difference method [2], [1], disparity is computed from the phase difference between the convolutions of the two stereo images with local bandpass filters. Since the two signals, $f_{\mathrm{R}}(x)$ and $f_{\mathrm{L}}(x)$, are *locally* related by a shift $\delta(x_0)$, i.e. in the vicinity of each point x_0

$$f_{\mathrm{L}}(x+\delta(x_0)/2) \approx f_{\mathrm{R}}(x-\delta(x_0)/2) \tag{1}$$

the *local* k_0 Fourier components of $f_{\mathrm{L}}(x)$ and $f_{\mathrm{R}}(x)$

$$\widehat{f}_{\mathrm{L/R}}(k_0) = \int e^{-\mathrm{i}\,k_0\,x}\, f_{\mathrm{L/R}}(x)\,dx = \rho(x)_{\mathrm{L/R}}\, \mathrm{e}^{-\mathrm{i}\phi(x)_{\mathrm{L/R}}}$$

are related by a phase difference equal to $\Delta\phi(x) = \phi_2(x) - \phi_1(x) = k_0\,\delta$.

We can extract the local Fourier components by convolving the images with the Gabor filters

$$\begin{aligned} F_{\mathrm{L/R}}(x,k) &= \int G(x-y)\exp(\mathrm{i}\,k\,(x-y)) f_{\mathrm{L/R}}(y)\,dy \\ &= \rho_{\mathrm{L/R}}(x)\ \exp(\mathrm{i}\,\psi_{\mathrm{L/R}}(x)) \end{aligned} \tag{2}$$

where $G(x-y)$ is the Gaussian function,

$$G(x) = \frac{1}{\sqrt{2\,\pi}\,\sigma}\ \exp(-\frac{x^2}{2\sigma^2})$$

As a function of the spatial position, the phase of the filter response, $\psi(x)$, has a quasi *linear* behavior dictated by the center k_0,

$$\psi(x) \approx \psi^{'}(x_0)\,(x-x_0) \approx k_0\,(x-x_0). \tag{3}$$

The local frequency, i.e. the derivative of the phase $\psi(x)$, is generally close to the value of the center frequency k_0. In fact, the Gabor filter is a bandpass filter around k_0.

In the Fleet, Jepson and Jenkin algorithm [1], the disparity is extracted from the phase difference, $\Delta\psi(x) = \psi_{\mathrm{L}}(x) - \psi_{\mathrm{R}}(x)$, by expanding $\Delta\psi(x)$ to the second order in δ,

$$\delta(x) \approx 2\,\frac{[\Delta\psi(x)]_{2\,\pi}}{\psi^{'}_{L}(x) + \psi^{'}_{R}(x)}. \tag{4}$$

The phase is not defined when the amplitude vanishes, i.e. when $\rho(x) = 0$ (singularity). Around these singular points the phase is very sensitive to spatial or scale variations. As a consequence, approximation (4)

fails and the calculation of disparity in the neighborhood of a singularity is unreliable. The neighborhoods of singular points can be detected [1] by means of

$$S(x) = \sigma \sqrt{(\psi' - k_0)^2 + \left(\frac{\rho'}{\rho}\right)^2} \leq T_1 \quad (5)$$

$$\rho(x)/\rho^* > T_2 \quad (6)$$

where T_1 and T_2 are opportunely chosen constants, and ρ^* denotes the maximum value of the amplitude. The first term of (5) measures the difference between the peak frequency, k_0, and the local frequency, $\psi'(x)$, in relation to the width of the filter $1/\sigma$. The second term of (5) measures local amplitude variations with respect to the spatial width σ. The relation (6) measures the "energy" of the response. The calculation at point x is accepted only if the above relations are satisfied. Usually, T_2 is set to $\approx 5\%$, and $T_1 \approx 1.25$.

The algorithm of Sanger [2] can be considered as a simplified version of the above calculation in which the derivatives of phase and amplitude are not used: 1) the disparity is estimated by phase difference divided by the filter's tuning frequency k_0, and 2) points where the stereo information is unreliable are detected using only the threshold, T_2, on the amplitude of the filter's output.

2 Sources of error

The performance of the phase-based algorithm is affected by different sources of error. We distinguish between 1) error sources that are intrinsic to the signals or are due to the basic assumptions on which the computational procedure of stereopsis is founded, and 2) sources that are due to the approximations introduced in the mathematical procedure.

Regarding the first issue, note that stereopsis is based on the assumption that a pair of stereo images are locally related by a one-dimensional shift. However, since the stereo images are 2D projections of the same 3D scene, taken from slightly different perspectives, we can expect

- presence of occlusions, i.e. visual features that are visible in one of the images but are not in the other,
- difference in the scale and form of corresponding visual features.

Furthermore, errors sources are introduced by calibration errors or difference in the optical parameters (illumination, contrast, etc.) of the cameras.

This work is focussed on the evaluation of the performance of the phase-difference-based algorithm. Therefore, we will concentrate on error sources of type 2.

For a proper evaluation of the algorithm performance, the different sources of error have to be isolated and independently analyzed. In order to separate error sources of type 1 from type 2, we test the algorithm on synthetic (random dots) images that are obtained by a constant 1D shift of the same image, therefore avoiding occlusions and changes in perspective and scale.

3 Experimental results

We tested experimentally both versions of the phase-based disparity estimators described in [2] and [1].

A further improvement possibility concerns the Gabor filter: the Gabor filter has a little residual sensitivity to the DC component of the input, but it is possible to modify the shape of the filter in order to eliminate this sensitivity. The kernel of the modified Gabor filter is

$$Z(x) = \frac{1}{\sqrt{2\,\pi}\,\sigma} \exp\left(-\frac{x^2}{2\sigma^2} + \mathrm{i}\,k_0\,x\right) - \exp\left(\frac{-(\sigma k_0)^2}{2}\right)$$

In summary, we studied four distinct combinations:

1. Sanger with a regular Gabor filter.
2. Sanger with a DC-clean Gabor filter.
3. Fleet, Jepson & Jenkin with a regular Gabor filter.
4. Fleet, Jepson & Jenkin with a DC-clean Gabor filter.

3.1 Error measures

Three kind of error measures were used:

The absolute disparity error, defined as

$$E = |g - d| \quad (7)$$

where g is the estimation and d is the actual disparity value.

The mean relative error, defined as

$$\epsilon = \frac{|g - d|}{|d|} \quad (8)$$

The percentage relative error, defined as

$$e = 100\epsilon \quad (9)$$

In the following sections we will always refer to the mean value of these errors on a set of samples: the average error is defined as the normalized sum of the absolute value of the difference between true (TD) and computed (D) disparity map at non-singular points,

$$\mu = \frac{1}{N}\sum_{i,j} |D_{i,j} - TD_{i,j}| \quad (10)$$

where i and j run over the nonsingular points and N is the total number of the nonsingular points. Similarly, the mean percentage relative error is defined as:

$$M = 100\frac{1}{N}\sum_{i,j} \frac{|D_{i,j} - TD_{i,j}|}{|TD_{i,j}|} \quad (11)$$

Δ is the deviation of the average error:

$$\Delta = \frac{1}{N}\sum_{i,j} |\mu - |D_{i,j} - TD_{i,j}||. \quad (12)$$

We did not perform any kind of interpolation on the estimated disparity map obtained from the disparity detectors: the points marked as unreliable are simply discarded and not taken in account in the computation of the error. Nevertheless we also measured the density of the estimate (i.e. the number of points marked as reliable divided by the size of the test image).

The tests have been performed using a large data set for each experimental condition: 64 scan-lines 256 pixels long. The pixels affected by the borders were cut away, and not considered in the error computation.

3.2 Range test

The working range of a disparity detector is the range of disparities that it can reliability measure. To measure the disparity range a sequence of test images of increasing disparity is presented to the disparity detector, until the detector is not able to detect it reliably anymore. Our test images were composed of a rectangular patch covered with a random dots texture with values in the range $[0, 1]$. The patch moved backwards, with negative disparities, starting from a disparity of zero pixels to $-\frac{\lambda}{2} - 1$.

The theoretical working range of a phase-based disparity detector is $[-\frac{\lambda}{2}, +\frac{\lambda}{2}]$ pixels, where λ is the period of the tuning frequency of the Gabor filter ($\lambda = \frac{2\pi}{k_0}$).

Our tests have evidenced that all the four considered cases have a plateau where the relative error is approximatively constant, for the disparities in the range $[-\frac{\lambda}{3}, +\frac{\lambda}{3}]$ pixels (Fig.1 and Fig.2).

The Fleet algorithm has always an evident advantage over the Sanger algorithm as seen by the mean percentage error, due to the better approximation of the local frequency used and due to the detection of phase instabilities (see Sec. 4).

To show the deviations, we draw the error with the error bars in the case of the Fleet and of the Sanger algorithm in Fig. 3.

3.3 Contrast test

This test probes the effects of the image contrast on the relative error: we want to see if the measurement of more luminous object "smears" on neighboring darker objects. The test image is divided vertically into two patches covered with a random dot texture, initially with the same mean value. The left patch has a disparity of -2 pixels, the right one of 2 pixels. The contrast between the two parts is progressively increased adding a factor variable from 0 (contrast 0%) to 1 (contrast 100%) to the pixel values in the left patch.

This test shows that all four techniques are not very sensitive to differences in contrast. For filters of narrow bandwidth ($\beta < 1$ octave), there was no difference between the techniques that use a DC-cleaned Gabor filter and techniques that use a normal Gabor filter (Fig. 4). For filters of bigger bandwidth, the error was not affected at all in the case of the techniques that use a DC-cleaned Gabor filter. There was a degradation of about 50% over a change in contrast of 100% for the algorithms using the non DC cleaned Gabor filters (Fig.5).

3.4 Camera sensitivity imbalance test

A grave source of problem for most of stereo algorithms is an imbalance of luminosity between the left and right image: in real conditions often one of the two images of the stereo pair is brighter or darker that the other, leading the same object to have differing intensity values in the two images.

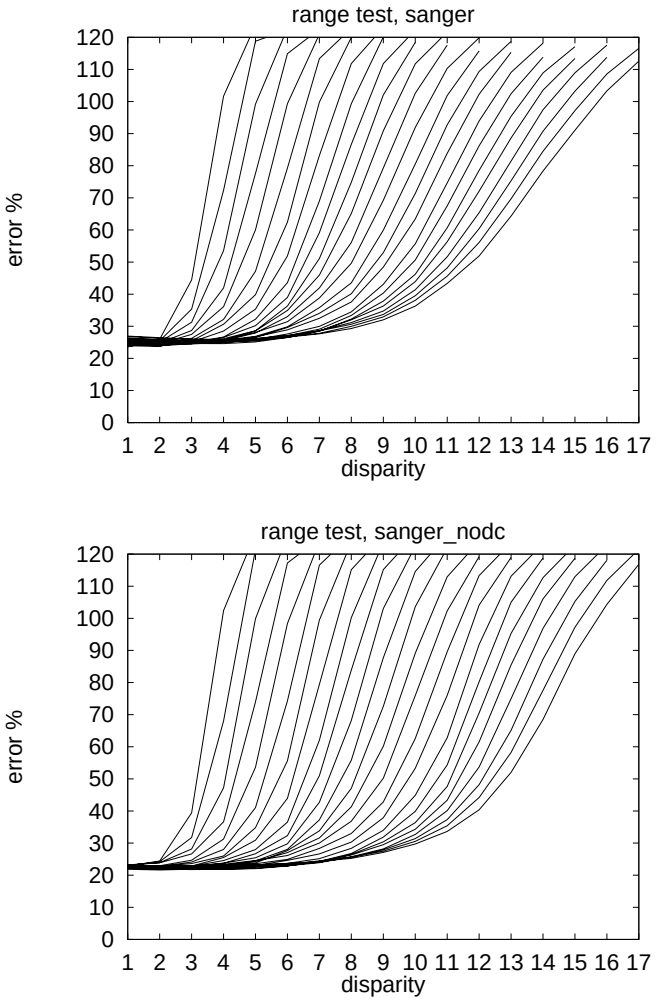


Fig. 1. Sanger algorithm, range test. Each line is the response of the phase detector for a different value of λ, starting form 8 (left) to 32 (right). For each line a plateau results, where the relative error remains constant.

We tested the sensitivity of the algorithms to this kind of problem by adding a constant factor to the intensity values contained in the left image of the test stereo pair. The test image was a random-dots textured image with a disparity value of 2 pixels. The texture values varied from 0 to 1. The imbalance factor varied from 0 (0%) to 4 (400%).

The algorithms showed to be rather robust in this respect: imbalances of up to 200% did not affect the performances too much, and the versions using the DC-cleaned filters or with β less than one octave were nearly insensitive to this effect (Fig. 6).

3.5 Noise test

The noise sensitivity of the algorithm was measured using a stereo pair corrupted by Gaussian random noise of progressively increasing variance. The test image is a random-dot textured plane with a disparity of 2 pixels, the texture on the plane assumes real values in the interval $[0, 1]$. The noise was of the additive Gaussian kind,

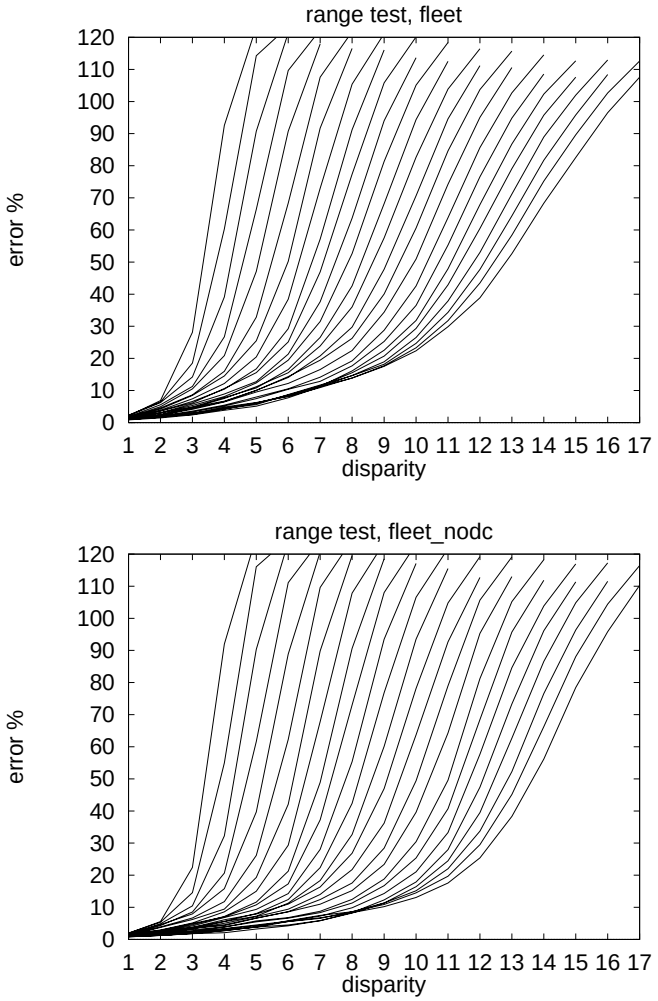


Fig. 2. Fleet algorithm, range test. Each line is the response of the phase detector for a different value of λ, starting form 8 (left) to 32 (right)

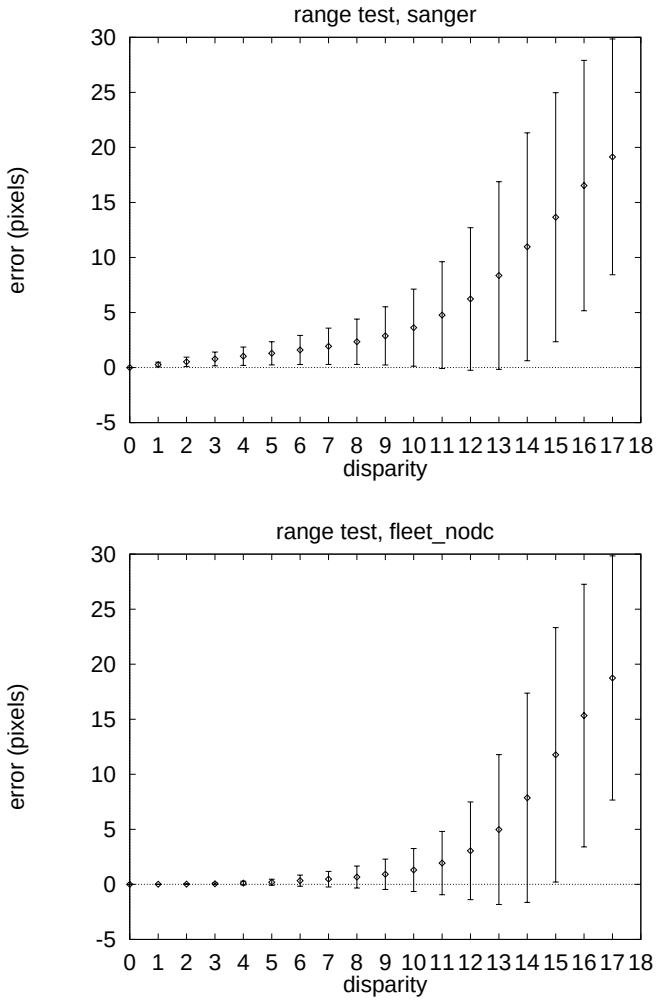


Fig. 3. Error and deviation for the Sanger and for the Fleet DC-cleaned algorithm. Parameters: $\lambda = 32, \beta = 1$

with a standard deviation varying from 0 (no noise) to 1 (100% noise).

All the algorithms are quite sensitive to noise: there is a linear relation between the noise percentage and the error (Fig.7). The sensitivity to noise is similar for both the Sanger and the Fleet algorithms. The filter's parameters have no significant influence on the noise sensitivity.

4 Choice of the thresholds T_1 and T_2

As explained in Section 2, around singular points the algorithm is not reliable and disparity computation is discarded. The extension of the neighborhoods of singular points is determined by constants T_1 and T_2 defined by (5) and (6). Therefore, the error and the density (number of non-singular points divided by the total number of points) are greatly affected by the choice of these two parameters.

Fig. 8 and Fig. 9 illustrate the behavior of error and density as a function of the two thresholds T_1 and T_2. The error points are obtained by plotting the error, $e(x, y; \lambda)$, for each point of the image and for different values of the wavelength, $\lambda = 2\pi/k_0 \in [8, 64]$, against $S = S_L + S_R$ in Fig. 8 and against ρ/ρ^* in Fig. 9. The density is obtained by integrating the number of points with values satisfying condition (5) or (6). The results show that the error sharply increases in the neighborhood of singular points; this behavior, obtained using a natural image, has a general validity. Using the results the values of T_1 and T_2 can be chosen according to the required accuracy. For example, choosing $T_1 = 2$, the error is almost always lower than 5% and the density is approximatively 60%. Furthermore, we notice that the constraint (5) is more effective than constraint (6).

5 Spectral Content

Since the Gabor filter is a band-pass filter, only frequencies contained in a neighborhood of the tuning frequency contribute to the disparity calculation. Image of different type have a different frequency spectrum: for example, random dots images have a flat spectrum (up to the Nyquist frequency) whereas the spectrum of natural images is characterized by a typical $1/f$ decay. Gabor filters are especially suited for $1/f$ spectra because the width

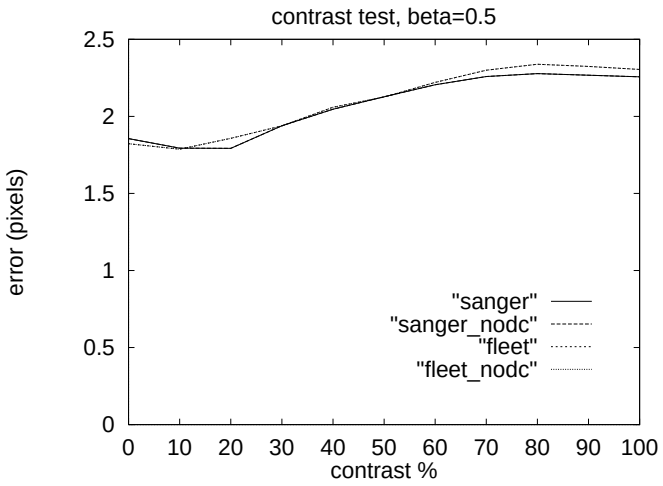


Fig. 4. Error at different level of contrast in the image. Filter with parameters $\lambda = 32$, $\beta = 0.5$.

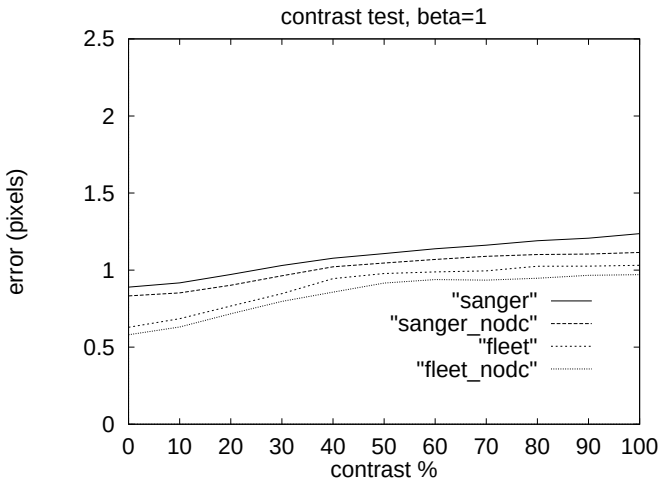
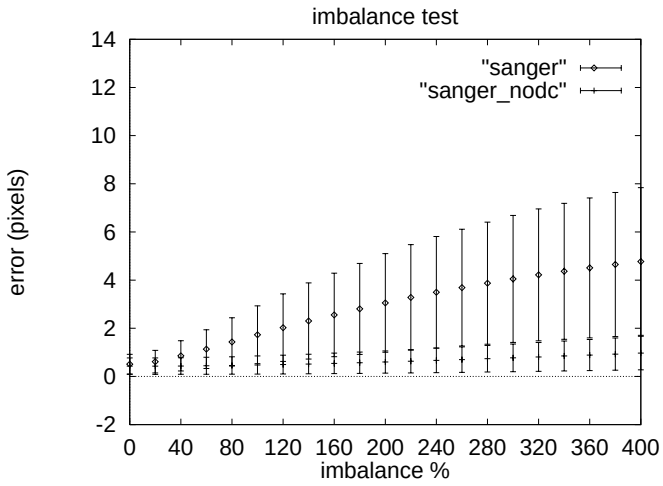


Fig. 5. Error at different level of contrast in the image. Filter with parameters $\lambda = 32$, $\beta = 1$



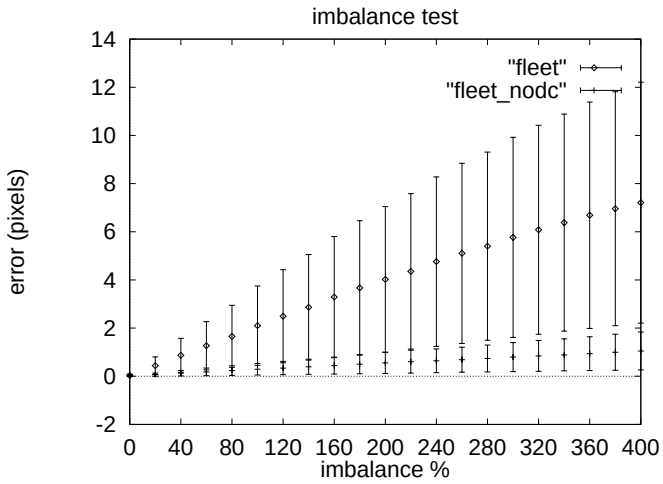


Fig. 6. Error and deviation for imbalanced intensity in the stereo pair. Filter with parameters $\lambda = 32$, $\beta = 1$

of the filter increases linearly with the tuning frequency, $f_0 = k_0/2\pi$, thus ensuring a constant "energy" of the filtered signal.

In the previous section random-dots images were used to measure the algorithm's performances. The results obtained have a general validity because different types of spectra do not modify strongly the general behavior of the algorithm. Fig. 10 shows the results of running the range test on random-dots images with different spectral distributions ($flat$, $1/f$, $1/f^2$) and natural images. The result show that natural images have a behavior comprised between the $1/f^2$ spectrum and the flat spectrum.

6 Discussion and Conclusion

The main objective of this work was to analyze the intrinsic accuracy of the phase-based disparity detectors, unrespectively from general problems of stereopsis, such as occlusion and deformation due to perspective. At this end we made use of computer-generated images of known disparity and spectrum.

Our tests show that the phase-based disparity estimators can measure with constant relative error disparities in the range $[-\lambda/3, \lambda/3]$ and that these techniques are robust to contrast changes and to imbalance between the images composing the stereo pair, in particular when making use of the modified Gabor filter without DC-sensitivity. Our tests have revealed that these algorithms are quite sensitive to noise, making them unsuitable for applications where the noise-to-signal ratio is bigger than 0.05–0.1.

On our artificial images, where every point has the same disparity, the error decreases with the filter width. In real stereo images disparity is not constant but varies in a way that can not be foreseen. Whereas the use of large filters reduces the error in region of constant disparity, it increases the error when it interpolates between different disparity values [3].

The results presented allow to estimate the performances that can be expected in applying the algorithms on real images, at least as a best-case situation. Another way to use the results is to design stereo system with

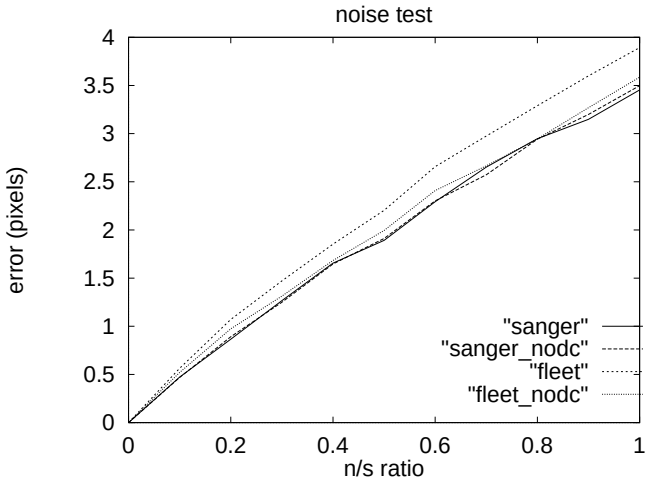


Fig. 7. Error level for different n/s ratios measured at zero disparity. Filter with parameters $\lambda = 32$, $\beta = 1$

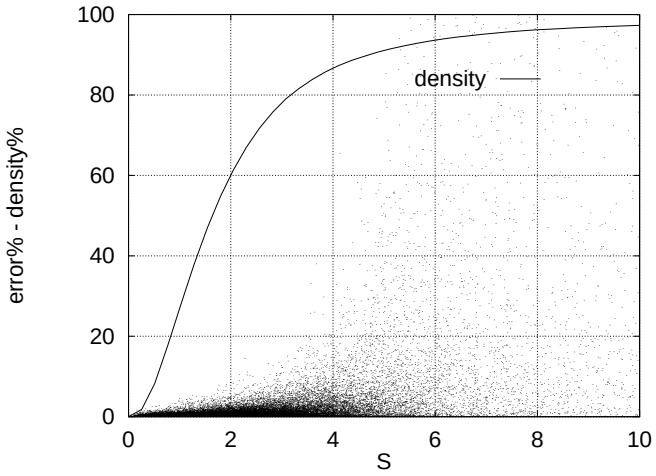


Fig. 8. Error and density vs S for $T_2 = 0$

known performances: given an estimate of the necessary accuracy, it is possible to choose the algorithm's parameters such that the requested precision will be attainable.

Bibliography

1. D. J. Fleet, A. D. Jepson, and M. Jenkin. Phase–based Disparity Measurement. *CVGIP: Image Understanding*, 53(2):198–210, March 1991.
2. T. D. Sanger. Stereo disparity computation using Gabor filters. *Biol. Cybern.*, 1988.
3. F. Valentinotti, G. Di Caro, B. Crespi, Real–time Parallel Computation of Disparity and Optical Flow Using Phase Difference, *Machine Vision and Applications*, to appear in Vol 9, issue 2.

This article was processed by the author using the LaTeX style file *cljour2* from Springer-Verlag.

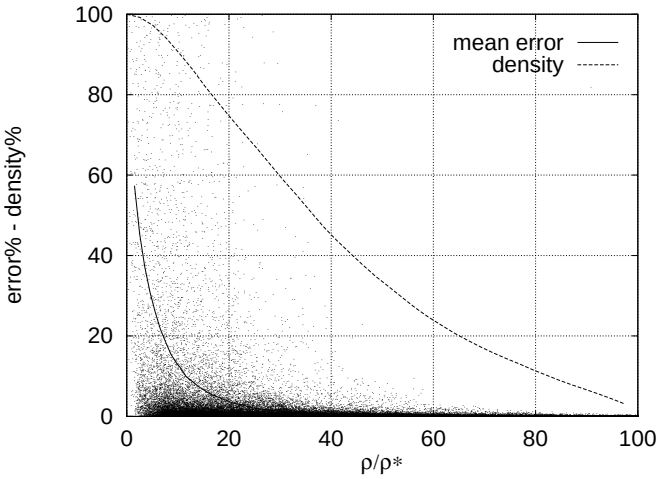


Fig. 9. Error and density vs ρ/ρ^* with $T_1 = 0$

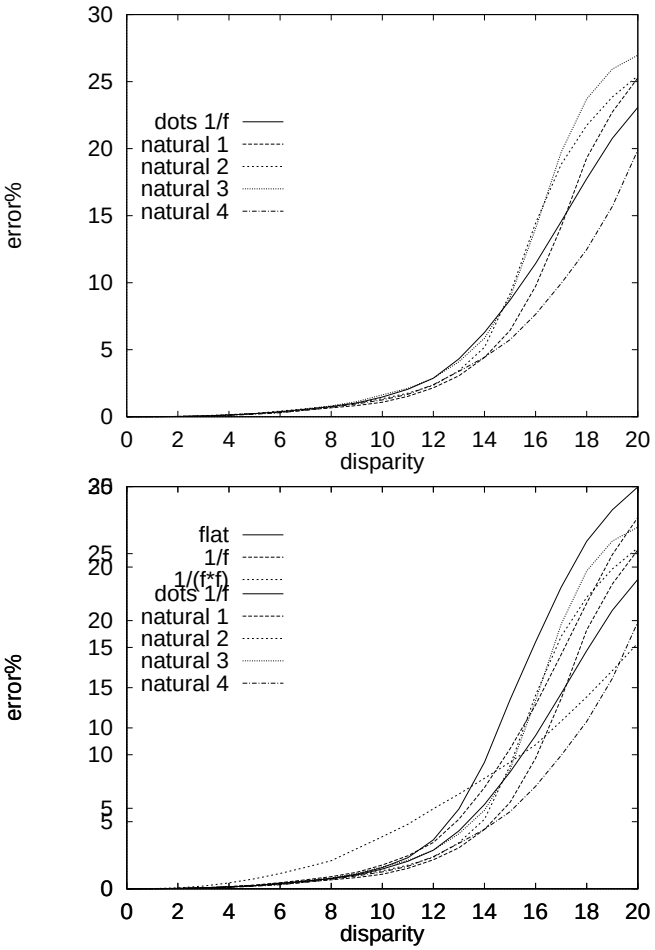


Fig. 10. Error vs. disparity for $\lambda = 32$ for natural images and random-dots images with $flat$, $1/f$, and $1/f^2$ spectra.