



A VLSI-Compatible Computer Vision Algorithm for Stereoscopic Depth Analysis in Real-Time

BERND PORR

*Department of Psychology, University of Stirling, Centre for Cognitive and Computational Neuroscience,
Stirling, FK9 4LA, Scotland, UK*

BERND NÜRENBERG

I-to-I GmbH, Department of 3D-Stereo, D-21079 Hamburg, Germany
sales@i-to-i.de

FLORENTIN WÖRGÖTTER

*Department of Psychology, University of Stirling, Centre for Cognitive and Computational Neuroscience,
Stirling, FK9 4LA, Scotland, UK*
worgott@cn.stir.ac.uk

Abstract. The analysis of the depth coordinates of objects in a visual scene is of vital importance for animals as well as in technological applications like autonomous robot navigation or product quality control. In this article we describe a phase-based algorithm for stereoscopic depth analysis which utilizes IIR-filters.¹ This algorithm is especially well suited to be built into dedicated VLSI-hardware and can therefore, also be used as a fast real-time front end in any more general image processing system. Example movies which demonstrate the real-time capabilities of this algorithm can be found at: <http://www.cn.stir.ac.uk/Real-Time-Stereo>.

Keywords: stereo-disparity, phase-based, causal processing, electronic filter

1. Introduction

In general there are several strategies of how to retrieve depth information from a sequence of images, like depth from motion (flow-field analysis), depth from shading and depth from stereopsis, on which we concentrate in this article. In a conventional stereoscopic approach usually two cameras are mounted with a horizontal distance between them. As a consequence objects displaced in depth from the fixation point are projected onto image regions which are shifted with respect to the image center. The horizontal (lateral) component of this shift is called (*lateral*) disparity and it can be used to determine the depth of the object. Due to the geometry of the optic system it is thereby sufficient to restrict disparity analysis to the projection of

corresponding linear segments (lines) in the left and right eye (epipolar line constraint). As a direct consequence of this epipolar line constraint, it is, therefore, not necessary to try to solve the disparity estimation problem in two dimensions (Faugeras, 1993), which raises computational complexity. To do so, however, can improve the results.

In the most straightforward approaches that address the problem of depth from stereo, the disparity is computed by searching the maximum of the cross-correlation between image windows along the epipolar lines of the left and right image (Haralick and Shapiro, 1992). Similarly this can be done by trying to match discernible image features (token-based matching). Fourier techniques can also be used to calculate the disparity. One of the first examples of such

an approach is the Kuglin-Hines method (Kuglin and Hines, 1975) which utilizes the phase shift theorem of the Fourier transform. Really applicable became this idea with the introduction of localized frequency filters called gabor filters. This method computes the convolution between Gabor kernels (Eq. (1)) and the left and right image parts.

$$G(x - x_0) = e^{\left(\frac{-(x-x_0)^2}{2\sigma_x^2}\right)} \cdot e^{i\omega_G(x-x_0)} \quad (1)$$

Each filter response is represented by its amplitude and its phase value. The disparity is computed from the difference of the two phase values obtained from the left and right images divided by the filter tuning frequency ω_G . The amplitude of the filter response can be used to estimate the reliability of the obtained result: the bigger, the more reliable is the phase difference (Sanger, 1988). If the amplitude is zero, obviously, the phase is ill-defined. Since this idea was introduced a large body of literature has been devoted to these approaches (Sanger, 1988; Fleet et al., 1991a, 1991b; Langley et al., 1991), which are commonly called *phase-based stereo algorithms*. Fleet et al. (1991a, 1991b) have shown that the results can be improved dividing by the local spatial frequency of the response instead of the tuning filter frequency. The same authors also provide a detailed theoretical analysis of the phase based stereo approaches (Fleet and Jepson, 1993), which shall not be discussed here. An error analysis considering real and artificial images is given by Cozzi et al. (1997).

Correlation techniques and phase based stereo algorithms are acausal in the sense that data acquisition of at least parts of the image needs to be completed before the computation of the disparity can start. Thus, these approaches are—strictly speaking—never real time techniques. In addition, the spatial convolution operations necessary in these algorithms require considerable computational effort. Thus, despite of the tremendous increase in computational power over the last years, close to video real-time performance can still only be achieved for a rather limited image resolution (Kanade et al., 1995, see also Internet Publication, SVS).

Furthermore, the retrieval of disparity information is usually a very early step in image analysis. It requires stereotyped processing of dense input data (every single pixel enters the computation) and only later steps in image analysis start to make use of data reduction strategies by concentrating on higher level image features (like junctions, surfaces, etc.) which are less

dense but require more complex algorithms. Therefore, it seems advisable to consider the use of special hardware for the calculation of the disparity or other such early image descriptors (like edges, optic flow, color, contrast, etc.). This way multi-level systems can be designed with such hardware at the front-end followed up by conventional processors which are then free to be used for “more intelligent” image analysis tasks. A similar strategy of task splitting and parallel processing is also pursued in the brain and, therefore, such multi-level systems are currently also of great interest to the computer vision community.

For this reason, in this study we describe a novel VLSI-compatible, phase-based algorithm to determine the disparities in two stereo images which uses causal filters and thereby operates in real-time regardless of the image size. The central idea behind this approach is to transpose the spatially-defined problem of disparity estimation into the temporal domain and compute the disparity simultaneously with the incoming data flow. This can be achieved realizing that in a well-calibrated fronto-parallel camera arrangement (no torsion, same focal length, etc.) the epipolar lines are horizontal and thereby identical to the camera scan-lines. Alternatively horizontal epipolar-lines can be obtained by image-rectification techniques (see Faugeras (1993) for a textbook treatment of this) employed as a front-end to our algorithm. This way both epipolar-lines can be simultaneously fed into a chain of causal electronic filters which compute the disparity. Hence, the structure of the algorithm is such that the elaborate and well-consolidated theory of causal electronic filters can be immediately applied and that it can be easily built in silicon. We will show that our one-dimensional real-time algorithm reaches a performance which is similar to many two-dimensional approaches.

2. Description of the Algorithm

As mentioned, we assume an ideal fronto-parallel camera arrangement. Let us first restrict ourselves to a set of images which contain only a single contrast step at the x -coordinates x_l in the left image and x_r in the right image, $x_l < x_r$. These images shall be sampled by two CCD cameras such that the contrast steps occur at times t_l and t_r for each scan line, where $t = 0$ represents the start of the scan line. The difference between the temporal coordinates t_l and t_r is, therefore, directly equivalent to the difference in the spatial coordinates which is identical to the disparity.

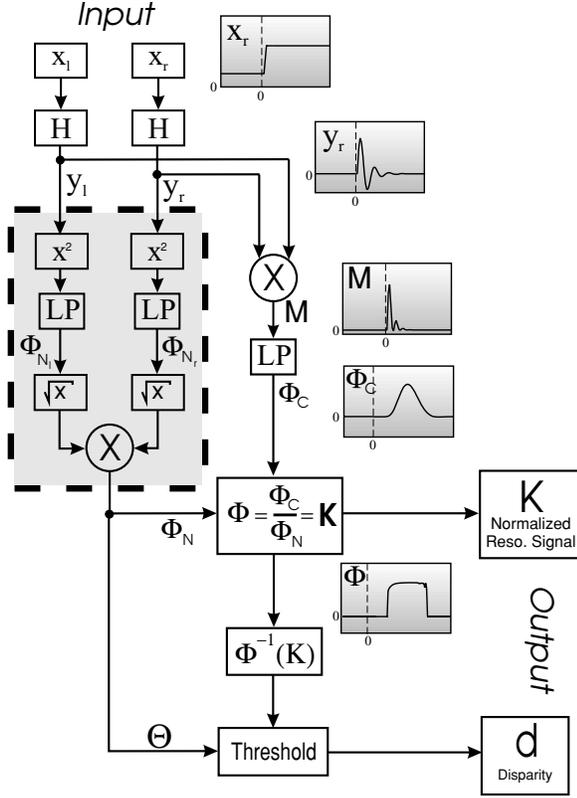


Figure 1. Scheme of the basic algorithm and results obtained from two step function inputs with a disparity of one pixel.

The two contrast step functions $x_l(t) \leftrightarrow X_l(s)$ and $x_r(t) \leftrightarrow X_r(s)$ are defined in the Laplace domain by (Fig. 1):

$$X_l(s) := \frac{1}{s} e^{-t_l s}, \quad \text{and} \quad X_r(s) := \frac{1}{s} e^{-t_r s}, \quad (2)$$

These inputs are then fed into two resonator circuits with transfer function:

$$H(s) = \frac{s}{(s - s_\infty)(s - s_\infty^*)} \quad (3)$$

where s_∞ is a filter pole and specifies the filter characteristic defined by f_0 and the filter quality Q , which determines the attenuation (see definition of the components, Section 3). The “*” denotes the complex conjugate.

We compute the convolution of signal and resonator for the right image as:

$$Y_r(s) = X_r(s)H(s) = \frac{s}{(s - s_\infty)(s - s_\infty^*)} \frac{1}{s} e^{-t_r s} \quad (4)$$

A similar convolution is performed for the left image. We define $a := (s_\infty - s_\infty^*)^{-1}$, then inverse Laplace transformation of $Y_r(s)$ yields (Fig. 1, y_r, y_l):

$$y_r(t) = \begin{cases} a e^{s_\infty(t-t_r)} + a^* e^{s_\infty^*(t-t_r)} & \text{if } t \geq t_r \\ 0 & \text{if } t < t_r \end{cases} \quad (5)$$

The temporal resonator signal $y(t)$ reflects an attenuated sine-wave with frequency f_0 (Fig. 1, y_l, y_r). The number of full cycles until the signal fades is roughly represented by the value of Q . Note that after the resonator the signal is DC clean. The DC usually poses an additional problem in all spatial filter approaches (Cozzi et al., 1997).

Finally, disparity is determined from the phase difference between the resonator signals from both images. This is done in a very conventional way by multiplication of the two signals in the time domain and subsequent low-pass filtering (Fig. 1, \times, LP).

Multiplication yields (Fig. 1, M):

$$M(t) = y_l(t)y_r(t) = \begin{cases} g_C(t) + \phi_C(t) & \text{if } t \geq t_r \\ 0 & \text{if } t < t_r \end{cases} \quad (6)$$

with:

$$g_C(t) = \underbrace{a^2 e^{s_\infty(2t-t_l-t_r)} + a^{*2} e^{s_\infty^*(2t-t_l-t_r)}}_{\text{double frequency term}} \quad (7)$$

and

$$\begin{aligned} \phi_C(t) &= \underbrace{|a|^2 e^{-s_\infty t_l - s_\infty^* t_r + (s_\infty + s_\infty^*)t} + |a|^2 e^{-s_\infty^* t_l - s_\infty t_r + (s_\infty^* + s_\infty)t}}_{\text{phase}} \\ &= \underbrace{\cos[(t_r - t_l)\text{Im}(s_\infty)]}_K \cdot 2|a|^2 e^{\text{Re}(s_\infty)(2t-t_r-t_l)} \quad (8) \end{aligned}$$

Note, the subscript C denotes the central signal pathway as opposed to the normalization pathways (subscript N , see below). The term $g_C(t)$ represents an oscillation with $2f_0$. It will be eliminated by low-pass filtering with low cut-off (Fig. 1, LP). In order to keep the equations simple we will first assume that we have an ideal low-pass filter with rectangular amplitude cut-off and without any phase distortions. The much more complicated correct solution for a realistic first order low-pass is given in the Appendix.

Applying an ideal low-pass filter to the signal from Eq. (6) we get (Fig. 1, ϕ_C):

$$\Phi_C(t) = \begin{cases} \phi_C(t) & \text{if } t \geq t_r \\ 0 & \text{if } t < t_r \end{cases} \quad (9)$$

The function $\phi_C(t)$ represents the phase between the two signals and contains an exponential relaxation term, which will be eliminated through normalization (see below), and a constant term K , which encodes the true disparity.

$$K = \cos[(t_r - t_l)\text{Im}(s_\infty)] \quad (10)$$

The disparity which is the spatial equivalent of $t_r - t_l$ can be computed by inverting Eq. (10) and it is obtained immediately at the second contrast step (i.e., for $t = t_r$), after which the signal relaxes to zero. This relaxation behavior which originates from the characteristic of the resonator assures temporal (viz. spatial) locality. Otherwise only the average phase (viz. disparity) of each image line could be computed, which would be meaningless in the context of local disparity analysis.

The disparity given by $t_r - t_l$ could, in principle, be read from the peak in Fig. 1, Φ_C . So far, however, the output signal is still ill-defined, because it is not normalized with respect to the absolute gray levels within the images. In order to solve this problem two additional filter chains are arranged in parallel to the main signal pathway, which are used to normalize the signal (shaded box in Fig. 1). Normalization is achieved by first squaring each resonator signal. This is essentially identical to computing the zero-disparity reference signal of each resonator. Thereby we re-obtain Eq. (6). Then both normalization signals are low-pass filtered using the same low-pass filter characteristic as in the main signal pathway. Again we assume an ideal low-pass and we get for the two normalization functions after the low-pass:

Left path, let $t_r := t_l$ (Fig. 1, Φ_{N_l}):

$$\Phi_{N_l}(t) = \begin{cases} \phi_{N_l}(t) = 2|a|^2 e^{\text{Re}(s_\infty)(2t-2t_l)} & \text{if } t \geq t_l \\ 0 & \text{if } t < t_l \end{cases} \quad (11)$$

and similarly with $t_l := t_r$ for the right path (Fig. 1, Φ_{N_r}).

Two normalization functions must be used to account for the causal behavior of the system and also

to compensate for luminance differences between both images. Normalization signals for the left and right path are in general not identical which can be seen for instance at the first moment when a true disparity value can be read which is at $t = t_r$. For this we get:

$$\phi_{N_l}(t_r) = 2|a|^2 e^{\text{Re}(s_\infty)(2t_r-2t_l)} \quad \text{and} \quad \phi_{N_r}(t_r) = 2|a|^2 \quad (12)$$

Because we have two normalization signals we need to take the square root of both before we can perform the normalization by dividing the main signal (Fig. 1, Φ_N).

$$\Phi_N(t) = \sqrt{\Phi_{N_l}(t)\Phi_{N_r}(t)} \quad (13)$$

The final output signal is (Fig. 1, Φ):

$$\Phi(t) = \begin{cases} 0 & \text{if } t < t_r \text{ or } \Phi_N(t) = 0 \\ \frac{\Phi_C(t)}{\Phi_N(t)} = \phi(t) & \text{if } t \geq t_r \text{ and } \Phi_N(t) \neq 0 \end{cases} \quad (14)$$

where we call $\phi(t)$ the *characteristic output function* of this filter module given by:

$$\begin{aligned} \phi(t) &= \frac{\phi_C(t)}{\sqrt{\phi_{N_l}(t)\phi_{N_r}(t)}} \\ &= \frac{2K|a|^2 e^{\text{Re}(s_\infty)(2t-t_r-t_l)}}{\sqrt{2|a|^2 e^{\text{Re}(s_\infty)(2t-2t_l)} 2|a|^2 e^{\text{Re}(s_\infty)(2t-2t_r)}}} = K \end{aligned} \quad (15)$$

Thus, the *characteristic output function* $\phi(t)$ is—as desired—now time independent and identical to the disparity dependent but otherwise constant term K .

The normalized disparity signal Φ for the two input step functions is shown in Fig. 1 and the unwanted curvy characteristic is replaced by a reasonably flat step function.

In the computer implementation the condition $\Phi_N(t) \neq 0$ in Eq. (14) is modified to $\Phi_N(t) > \Theta$, where Θ is a threshold value (Fig. 1, Threshold). Thresholding improves the reliability of the data, because all values will be blanked out for which the resonator responses are too small, but thereby the density of the disparity map is also reduced.

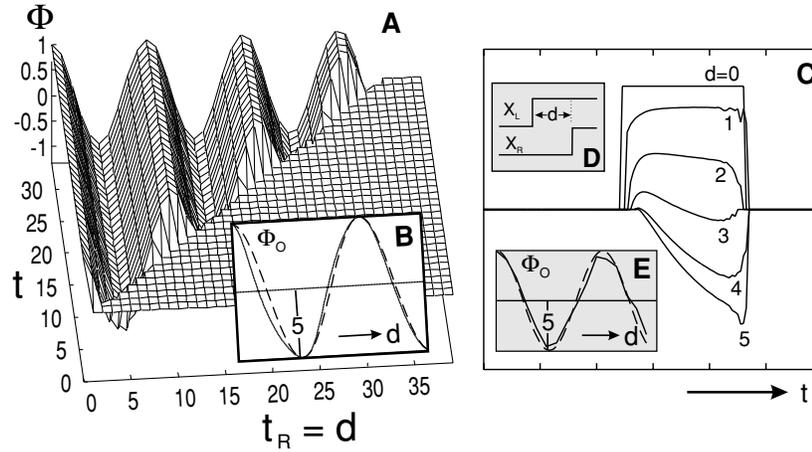


Figure 2. Theoretical and experimentally obtained curves for the characteristic function ϕ for $f_0 = 0.1$ and $Q = \pi$ for a first order low-pass with $\omega_0 = f_0 = 0.1$. (A) Theoretical curve ϕ obtained by plotting Eq. (38). (B, solid line) One-dimensional cut along the $t_r = d$ -axis through the curve in (A), (B, dashed line) pure cosine function with the respective frequency $\text{Im } s_\infty \approx 2\pi f_0$ (see Section 3). Dashed and solid curves are very similar showing that the cosine in Eq. (10) is a good approximation of Eq. (38). (C) Time-dependent shape of the characteristic function (compare (A)) along the t -axis determined by a computer implementation of the algorithm for different input disparities. Parameters: $f_0 = 0.1$, $Q = 1$, $\Theta = 0$, $n = 4$, $\omega_0 = f_0$. (D) Example of the input time-functions x_l and x_r . (E, solid line) Maxima of the curves in (C) determined between $d = 0$ and $d = 15$. (E, dashed line) pure cosine function as in (B).

In all following sections we will give all values for f_0 in pixel^{-1} . Equation (38) in the Appendix gives the correct but rather complicated solution for the characteristic output function by using the transfer function of a first order low-pass filter in all pathways. Figure 2(A) shows a 3-D plot of Eq. (38) as a function of time t and disparity (viz. t_r). The function is zero for $t < t_r$ and oscillates in t_r thereafter but stays constant in t . This means read-out of a disparity value obtained from a step function input can in principle take place at any time $t \geq t_r$. The oscillation is undamped in t_r with a period of about 10 pixel. This period is similar but not identical to the inverse of the characteristic frequency f_0 for resonator and low-pass, which we had set to 0.1 in this case (see Section 3 for the correct relation between $\text{Im}(s_\infty)$ and f_0). The oscillation is similar to the pure cosine which we would have obtained from Eqs. (9) and (10) in the case of an ideal low-pass filter (compare curves in inset B). It should be noted that this similarity becomes even more pronounced for higher filter orders (e.g., $n = 4$, which we usually use), because a higher order low pass filter is a better approximation of an ideal low pass. Therefore, for all practical purposes it is sufficient to assume a cosine characteristic for ϕ .

The periodic behavior of the output is a direct reflection of the phase wrap-around problem which is

present in all phase-based stereo algorithms. Here this problem is expressed by the fact that the characteristic output function ϕ can be inverted without ambiguities only between 0 and ≈ 5 pixels ($\approx \frac{1}{2} \frac{1}{f_0}$) disparity. In part C of Fig. 2 the results from a computer implementation of the algorithm are shown using simulated IIR-filters for the different components with the same parameters as for the theoretical curve. Noise free step functions x_l and x_r with identical amplitude but different disparities were presented to the system (inset D, top). At the output also step functions are obtained but the slope of the rising flank becomes increasingly shallow for larger disparities. As predicted the output remains stable for an extended duration after which it deteriorates due to the limited numerical accuracy (unsigned integer) of the computer implementation. The maxima of the output step functions were plotted against the disparities in the bottom inset E. This small diagram shows that the undamped oscillatory characteristic of ϕ is also re-obtained in the simulation (compare to dashed line which shows the theoretical curve). Numerical inaccuracy, however, increases for large disparities and the results cannot be trusted anymore beyond a certain disparity value. Numeric and stochastic effects will be treated in greater detail below also demonstrating the limitations of the theoretical derivation within a more practical context.

3. Specification of the Components

In the following section all components are discussed as analog circuitry. The computer implementation used to obtain the results in Section 2 relied on IIR-filter design which is essentially equivalent to analog circuitry. Implementation of the circuitry in FIR²-design leads to a problem at the low pass filter, which will be briefly discussed below. For the following it may help to disregard all implementation details at first and think of all filters as classical analog electronic filters.

3.1. Resonator

The two resonators are realized by basic band-pass filters. Resonance frequency is adjusted to f_0 and the *quality* is given by Q . These values are related to the transfer function of the resonator given in Eq. (3) by:

$$\operatorname{Re}(s_\infty) = -\pi f_0 / Q \quad (16)$$

$$\begin{aligned} \operatorname{Im}(s_\infty) &= \sqrt{(2\pi f_0)^2 - (\operatorname{Re}(s_\infty))^2} \\ &= \pi f_0 \sqrt{4 - \frac{1}{Q^2}} \end{aligned} \quad (17)$$

The value of Q determines the approximate number of oscillation cycles until the resonance has faded. The total duration ΔT until the oscillation has faded is thus approximately:

$$\Delta T \approx Q / f_0 \quad (18)$$

To assure temporal (viz. spatial) locality Q should be set to low values with a hard limit of $Q > 1/2$.

The resonance frequency f_0 determines the range of disparities $d = t_r - t_l$ which can be measured by:

$$0 \leq f_0 < \frac{1}{d \sqrt{4 - \frac{1}{Q^2}}} \quad (19)$$

above the upper limit *phase wrap-around* occurs. For $Q > 1$ this can be approximated by the limit of $Q \rightarrow \infty$ and relation 19 turns into:

$$0 \leq f_0 < \frac{1}{2d} \iff 0 \leq d < \frac{1}{2f_0} \quad (20)$$

which is now exactly half the resolution which can be attained with regular acausal Gabor filters (Cozzi

et al., 1997). This is due to the fact that our algorithm cannot detect negative disparities. This shall be mended in Section 5.

In the Fourier domain the bandwidth of the resonator Δf is roughly:

$$\Delta f \approx f_0 / Q \quad (21)$$

In analog circuitry this Q and f_0 are given as:

$$Q = R \sqrt{\frac{C}{L}}, \quad f_0 = \frac{1}{\sqrt{LC}} \quad (22)$$

3.2. Low-Pass Filter

Intriguingly the design of the low-pass filter is much more critical than that of the resonator. Several requirements have to be fulfilled without which a good performance of the algorithm cannot be assured.

1. The filter order n should be set to a high-enough value in order to eliminate the non-constant components as well as possible.
2. The filter characteristic must assure that the phase relation within the signal will not be distorted. Thus, a highly linear phase-characteristic is required. Otherwise the disparity values would be smeared out and the output would be garbled.
3. The impulse response must be strictly positive which prevents the usage of FIR-filters.

The second requirement restricts the choice for possible low-pass filters to the group of Bessel filters defined by:

$$H(s) = \epsilon \frac{1}{B_n\left(\frac{s}{\omega_0}\right)} \quad (23)$$

where B_n is a Bessel polynomial of order n , ω_0 is the cut-off frequency and ϵ a scaling factor. Furthermore, the impulse response of a Bessel filter is indeed strictly positive which is required for the square-root operation in the normalization pathway. In principle it would also be possible to implement the low-pass filter as an FIR-filter, because FIR-filters can be custom designed to obtain a linear phase relationship. FIR-filters, however, always produce a negative undershoot in the impulse response which is not tolerable because of the square-root operation in the normalization pathway. In FIR-chip-design complicated workarounds exist to avoid

this problem but it is more straight-forward to directly use IIR-filters.

The requirements for the low-pass filter also explain the apparently awkward way of normalizing the signal by a $()^2 \rightarrow LP \rightarrow \sqrt{\quad}$ chain (Fig. 1, light shading) instead of using a rectifier on the front-end of this chain. The reason for this is that rectification leads to a distortion of the phase relations between main signal flow and the normalization pathways. Squaring the signal, on the other hand, is equivalent to the multiplication operation performed in the main path. Thus, main- and normalization signals are treated identically keeping their intrinsic phase relations the same.

4. Quantification of the Performance

In this section we will derive empirically based limits for the different adjustable parameters and the applicable ranges of the algorithm. Results are obtained using the algorithmic framework as described above. Thus, the measurement of negative disparities is currently excluded. We will, however, extend the algorithm in Section 5 in order to also measure negative disparities and to be able to combine several stereo modules in order to improve the results. All simulation results were obtained from a digital implementation of the algorithm as described above using a unsigned integer precision IIR filter implementation for all filters on a Pentium PC, which has also been used for the chip implementation (FPGA) of our algorithm. This implementation can be ported 1:1 to a hardware system based on digital signal processors (e.g., Texas Instruments DSP TMS320 C80).

A quantification of the performance will be given using the Pentagon test-image (see Fig. 11) or R. Henkel's "2" (Henkel, 1994) (Fig. 3).

4.1. Measuring the Average Characteristic Output Function $\bar{\phi}$

In order to measure the accuracy of the algorithm we have determined the average characteristic output function $\bar{\phi}$ as a function of the input disparity much in the same way as for the step functions in Fig. 2(C). To this end a certain rightward shift by d pixels was introduced to all (!) pixels of the left Pentagon image. Then we used the un-shifted image as the left and the shifted image as the right image of the stereo pair. All pixels of the image were analyzed with the exception of a left margin of 50 pixels in order to avoid border effects. In Fig. 4(A) we plotted the measured value of $\bar{\phi}$ as a function of the input disparity (i.e., the shift) for $f_0 = 0.1$ and $Q = 1, 3, 10$ as well as for $f_0 = 0.2$ and $Q = 3$. As expected oscillatory curves obtain which start at 1.0. However, quite opposite to the theoretical curves for the (un-averaged) characteristic function ϕ , the diagrams for $\bar{\phi}$ display a strongly attenuated characteristic. For large input disparities the curve approaches zero and the average measured output disparity is, thus, roughly equal to $d = \frac{1}{4f_0}$, which is half of the resolution limit for this filter.³ Attenuation is less pronounced for larger values of Q . In the limit case of $Q \rightarrow \infty$ an undamped oscillation would be obtained. The standard deviation for zero input disparity is zero, thus at zero disparity the algorithm works error-free. For larger disparities the standard deviation increases continuously.

Intriguingly the simulated noise-free step functions (Fig. 2(C)) did not produce an attenuated characteristic, whereas artificial disparity steps introduced in a real image lead to this rather strong effect. This indicates that the attenuation must be an inherent feature of the real images. Similar to Fig. 2(C) in Fig. 4(B) we show the response for a pair of input step functions with different disparities when introducing a small signal

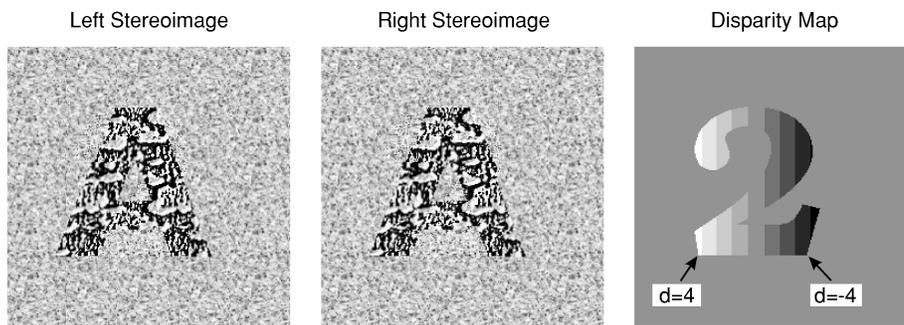


Figure 3. Test images and ground-truth disparity map.

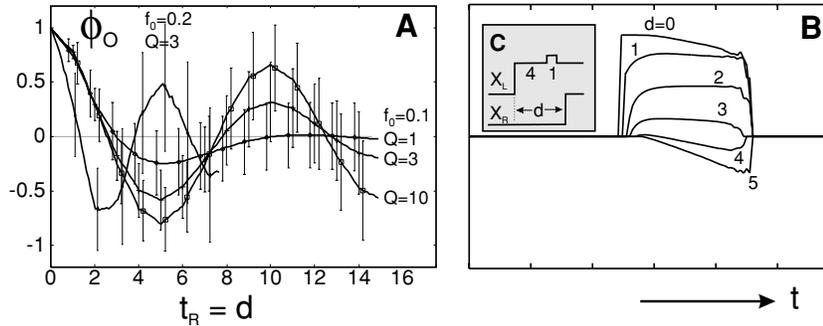


Figure 4. Empirically determined characteristic functions ϕ for different settings of f_0 and Q using a fourth order low pass with $\omega_0 = f_0$ and $\Theta = 0$. (A) The left Pentagon image was shifted by d pixels and used as the right image of the stereo pair. The input disparity d was varied between 0 and 15 in steps of 0.1 pixels. Fractional disparities were obtained by linear interpolation of the gray levels between adjacent pixels. A total of 52736 pixels (256^2 minus a border of 50×256 pixels) was analyzed for each disparity step. The period of ϕ is approximately 10 pixels for $f_0 = 0.1$ and 5 pixels for $f_0 = 0.2$ and phase-wrap around occurs above ≈ 5 (≈ 2.5) pixels. Error bars are only plotted for some data points. (B) Response of the system to a noise distorted step function shown in (C). Parameters: $f_0 = 0.1$, $Q = 1$, $\Theta = 0$, $n = 4$, $\omega_0 = f_0$. As opposed to Fig. 2(C) the responses are now attenuated and not anymore symmetrical with respect to the x -axis. (C) The input to the system consisted of two step functions with different disparities between them. An additional small secondary pulse was added to the left step function at disparity 4.

distortion in the left signal (C). Now the responses are not symmetrical anymore (compare to Fig. 2(C)) and an attenuated characteristic results. Obviously an averaging effect occurs between the primary step and the secondary pulse leading to a wrong disparity estimate at the output. This effect becomes more pronounced for larger input disparities because the distance between primary step and secondary pulse increases and, thus, the pulse occurs at a location where the left resonator oscillation is already rather small. This problem has an immediate practical consequence: Ideally, only small disparities should be measured because only their error is small. This aspect will be taken into account in Section 5.

It is tempting to use the invertible part of the average characteristic function $\bar{\phi}$ in order to try to correct the actually measured values of ϕ back to their theoretically expected value. Such a procedure, however, is only applicable if the underlying probability distribution of ϕ were Gaussian. Figure 5 shows that this is not the case. The actual distributions of the measured values of ϕ for artificial image shifts of 1–5 pixels are shown. Zero shift would lead to a needle-shaped distribution with all values found at $\phi = 1.0$. For larger shifts the peak of the distributions are still at their theoretically expected cosine-value until the distribution becomes almost flat for a shift of 5 pixels. A flat distribution bounded between -1 and 1 is the limit case. This demonstrates that the expectation value $\bar{\phi}$ is in general not identical to the peak of the distribution and prevents the use of the average characteristic function for value correction.

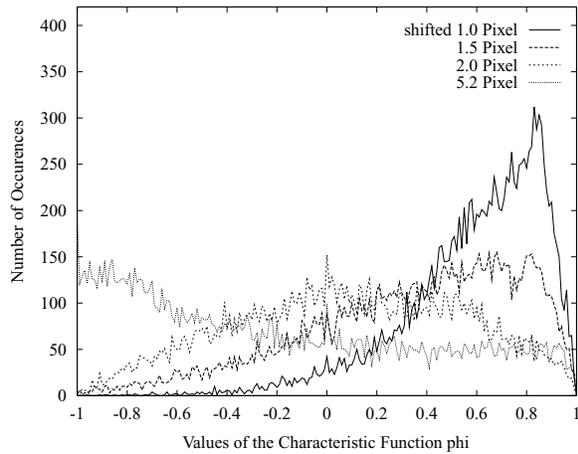


Figure 5. Distributions of the ϕ for different shifts d . As before, the left Pentagon image was shifted by d pixels and used as the right image of the stereo pair. Parameters: $f_0 = 0.1$, $Q = 1$, $\Theta = 0$, $n = 4$, $\omega_0 = f_0$. The same number of pixels was analyzed as in Fig. 4. Bounded distributions with a strong skewness are observed, but the peaks are at the expected cosine-values.

4.2. Error Analysis

In order to judge the quality of the algorithm several aspects have to be considered. *Accuracy* of the obtained disparity estimates is one of them, but also the *Spatial Localization*, which means how accurately these disparity estimates are associated to the actual pixels from which they originate. This tradeoff between the spatial localization and the accuracy of the disparity

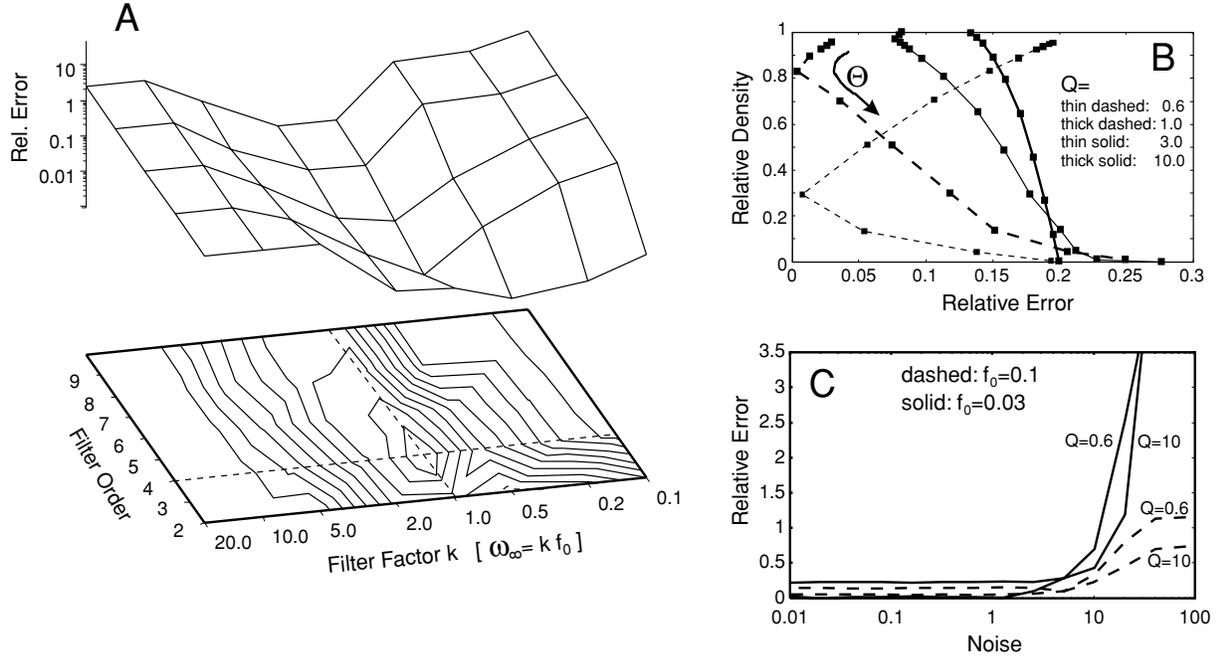


Figure 6. Estimating the algorithmic performance for different parameters. For all diagrams f_0 was set to 0.1 and we shifted the left Pentagon image by 1 pixel in order to use it as the right image of the stereo pair. (A) The relative error was determined for different values of ω_0 and the filter order n . Q was set to 1.0 and Θ to 16. (B) Relative error and relative density as a parametric function of the threshold Θ . Curves start at the top and the threshold was increased between 1 and 2^{16} as indicated by the arrow. Filter order was 4 and $\omega_0 = f_0$. (C) Relative error as a function of the input noise. White noise was added according to: $h_{I/r}(t) = \frac{R}{100}h_{Noise}(t) + \frac{100-R}{100}h_{Image}(t)$, $R \in \{0, \dots, 100\}$ where h_{Image} represents the pixel values of the image, h_{Noise} white noise of the same total amplitude. Thus, the relative ratio of signal to noise is adjusted by varying R . Filter order was 4, $\omega_0 = f_0$ and Θ was 128.

estimates is related to the Gabor uncertainty relation (Gabor, 1946) and cannot be avoided. Additional aspects which need to be considered in order to judge the algorithmic performance are the *sensitivity to phase wrap-around*, the *sensitivity to noise* and the *density of the disparity map*.

As before numerical quantification of the algorithmic performance are based on the left Pentagon-image in which artificial disparity steps were introduced. For visualization, however, we used the “2” of R. Henkel (Henkel, 1994) (Fig. 3), because a well-defined ground truth map exists for this image. Note, the “2” contains negative disparities which will be rectified by the current version of the algorithm (but see below).

All following results were obtained with unsigned integer precision. The motivation behind this is the chip implementation which has been performed with a XLINX FPGA.

4.2.1. Threshold Θ . Increasing the threshold Θ up to a certain point improves the total accuracy of the

disparity map for small values of Q (Fig. 6(B)). For even higher thresholds the error increases again. This unexpected effect is a direct consequence of the limited accuracy of the spatial localization of each response. The responses of the resonators do not immediately fade. Thus, usually several subsequent pixels elicit super-threshold responses but only the first pixel carries a correct disparity value. The following pixels represent the fading resonator response and their disparity values can be incorrect. For high thresholds the resulting disparity map consists mainly of such pixel clusters and the relative number of wrong disparity estimates increases continuously. This effect of smearing (compare also Fig. 7(C)) is dominant for $Q = 10$ and the corresponding curve does not anymore have a minimum (Fig. 6(B)). Most of the disparity maps in this study have been computed using $f_0 \approx 0.1$ and $Q = 1$. For this curve (thick dashed) the optimum is an error of 0.004 at a density of 0.83 for a threshold of $\Theta = 128$ at a total numerical range of 65536 (unsigned integer). Empirically we found the observation confirmed that

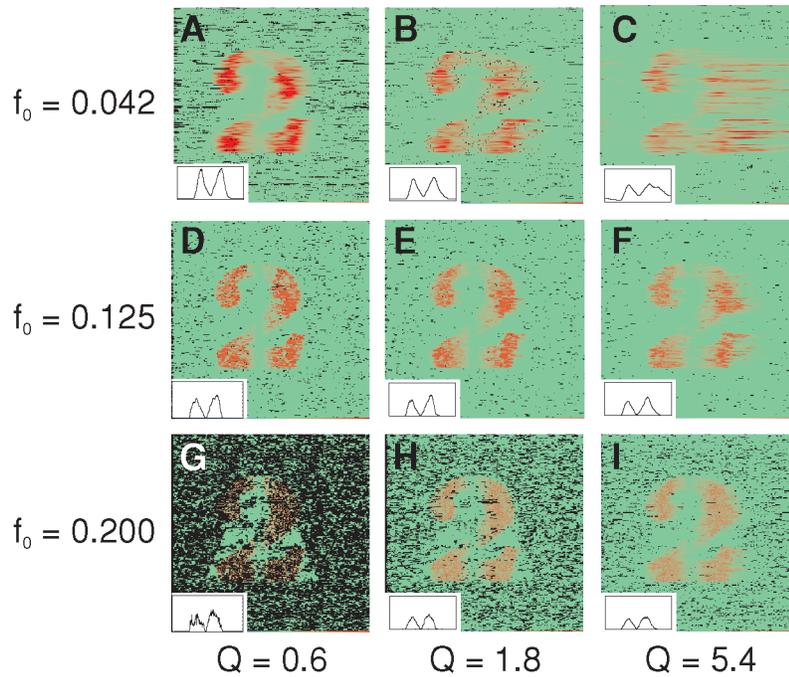


Figure 7. Analysis of Henkel’s “2” by the algorithm outlined in Fig. 1, which rectifies the results. Different settings of f_0 and Q are used as indicated. The threshold was fixed at $\Theta = 128$, filter order was 4 and $\omega_0 = f_0$. Disparities are color coded, black is used for pixels which were sub-threshold. In the top two rows (A–F) the parameters were adjusted such that $\Delta T = Q/f_0$ is the same for (A, E) and (B, F). In the third row (G–I) f_0 represents a limit case, because the characteristic functions contain too high standard deviations for values bigger than 0.2. The little insets at the bottom left corner of each diagram represent the average of 20 image lines at the base of the “2”. Insets were scaled between disparities $0 \leq d \leq 3$.

a threshold of about 100 would yield the best results in almost all cases. Note, however, that in a double-precision implementation the threshold can be set to almost zero (e.g. 10^{-20}) for optimal performance.

equal to f_0 (see Eq. (23)). It is obvious that low filter orders and/or high cut-off frequencies result in a less accurate behavior of the system. It is, however, not necessarily clear why these two parameters should lead to

4.2.2. Low-Pass Filter Order n and Cut-Off ω_0 . The curve in Fig. 6(A) shows that the relative error has a minimum with value 0.0144 for a filter order of $n = 4$ when setting the cut-off frequency ω_0 at the same time

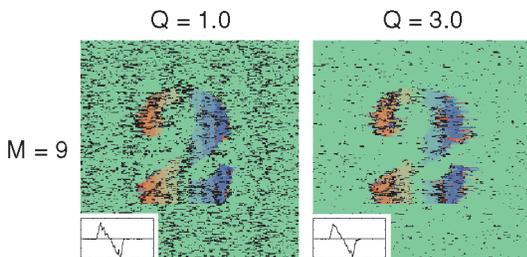


Figure 9. Analysis of Henkel’s “2” by the complete algorithm outlined in Fig. 8. Nine detectors were used, thus $\tau = 1$ pixel. Other parameters: $f_0 = 0.1$, $\Theta = 128$, $n = 4$, $\omega_0 = f_0$.

Original Right Stereo Images

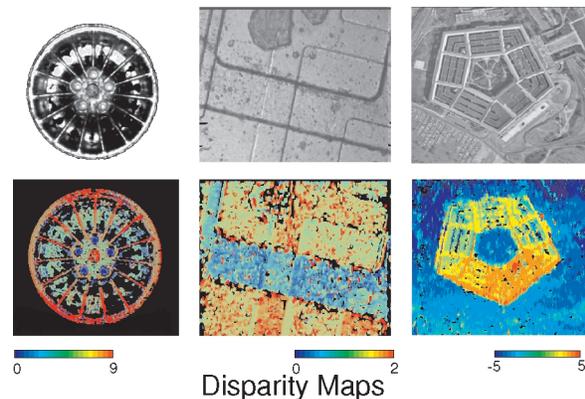


Figure 11. Two examples of industrially relevant applications and the pentagon test image. Parameters: $f_0 = 0.1$, $Q = 1$, $\Theta = 128$, $n = 4$, $\omega_0 = f_0$. Wheel: $M = 11$, Wafer: $M = 3$, Pentagon: $M = 11$.

an optimum curve such that filter orders above 4 and cut-off frequencies below f_0 will be suboptimal. This, however, can be explained by the limited amplitude of the signals. As soon as the filter order n is increased and/or the cut-off decreased the signal amplitude after the low-pass will become increasingly smaller. Due to the limited numerical accuracy of the computer implementation the relative error will increase in these cases leading to an optimum curve in Fig. 6(A). Note, in order to make this minimum clearly visible the threshold Θ was set to 16. For visualization zero threshold would even be better, but noise effects become too strong.

4.2.3. The Resonator Parameters f_0 and Q

4.2.3.1. Accuracy. The average accuracy of the disparity estimates can immediately be judged from the average characteristic functions $\bar{\phi}$ plotted in Fig. 4(A). The actual deviation from the expected cosine shape together with the standard deviations directly reflect the different accuracy for the different parameter settings. We expect that the accuracy of the algorithm should improve by using large values of Q and small values of f_0 and this is confirmed by the curves. In the case of $f_0 = 0.2$, $Q = 3$ the cosine shape is better preserved than for $f_0 = 0.1$, $Q = 3$ but the standard deviations are by far larger for $f_0 = 0.2$. Taken together the accuracy for $f_0 = 0.2$, $Q = 3$ is much worse than for $f_0 = 0.1$, $Q = 3$, as expected.

4.2.3.2. Spatial Localization and Phase Wrap-Around. In Fig. 7 we show the results obtained for the “2” for different settings of f_0 and Q at a fixed threshold of $\Theta = 128$. The quality of spatial localization can be judged from these diagrams. It follows roughly but not entirely the theoretically predicted relation of $\Delta T \approx Q/f$ (Eqs. (18) and (21)). Consider the two diagonal pairs (A, E) and (B, F). Their values of ΔT are identical. Visual inspection, however, shows that E and F are slightly more in focus than A, B. Thus, f_0 has a stronger effect on the accuracy of spatial localization than Q , but both influence the result significantly. The value of $f_0 = 0.2$ represents a limit case. The exceedingly high variance of the average characteristic function explains the bad performance in this case. Values higher than $f_0 = 0.2$ up to the theoretical Nyquist limit cannot be used at all. Figure 7(G) shows that the resonators are much stronger driven by the more dominant contrast changes inside of the “A” than outside. Most responses outside remain sub-threshold and the outline of the “A”

becomes visible. Phase-wrap around occurs for $d > 1/(2f_0)$ and it is faintly visible in (H) and (I). The other parameters do not affect the phase-wrap around.

4.2.3.3. Sensitivity to Noise. Figure 6(C) shows that, similar to the findings for acausal Gabor filters (Cozzi et al., 1997), the relative error abruptly and strongly increases above a certain noise level of roughly 7% and it becomes unacceptably high above this value. Below 7% the relative error is relatively constant and independent of the noise. This is the acceptable working range and we find that Q and the filter order n have a slight influence on the noise susceptibility while the influence of f_0 is negligible. As expected a strong influence is observed for the threshold Θ and a high threshold suppresses the noise efficiently (not shown).

4.2.4. Summary of the Results from the Error Analysis. The results for unsigned integer precision show that all algorithmic parameters can be fixed at a single given value while still (almost) optimal performance is obtained in most cases. This is a major advantage concerning an electronic implementation. In most cases *Accuracy* and *Quality of Spatial Localization* are the two sine-qua-non conditions for the analysis of stereo images. Thus, good performance is achieved by $f_0 = 0.1$ and $Q = 1$. The filter order n should be 4 and the cut-off frequency to $\omega_0 = f_0$. A threshold of around 100 should also almost always suffice.

5. Algorithmic Extensions

5.1. Multi-Detector Coupling

The section on error analysis (Section 4.2) focused on the adjustable algorithmic parameters. However, one should keep in mind that the actual to-be-measured disparity values introduce by far the largest error source, as explained in the Section 4.1: Only small (zero) disparities are measured with small (zero) error. In order to enforce this, in the next section we will introduce an extension, which firstly allows us to measure negative disparities (e.g., in the case of converging cameras) and secondly assures that only small disparities are measured (adopted from Henkel (1994)). To this end we extend the algorithm by creating several parallel operating detectors. Figure 8 shows the complete system. It is built from $2m + 1$ detectors each of which consists of the algorithmic structure described in Fig. 1. Both images are used as input to each detector, but the

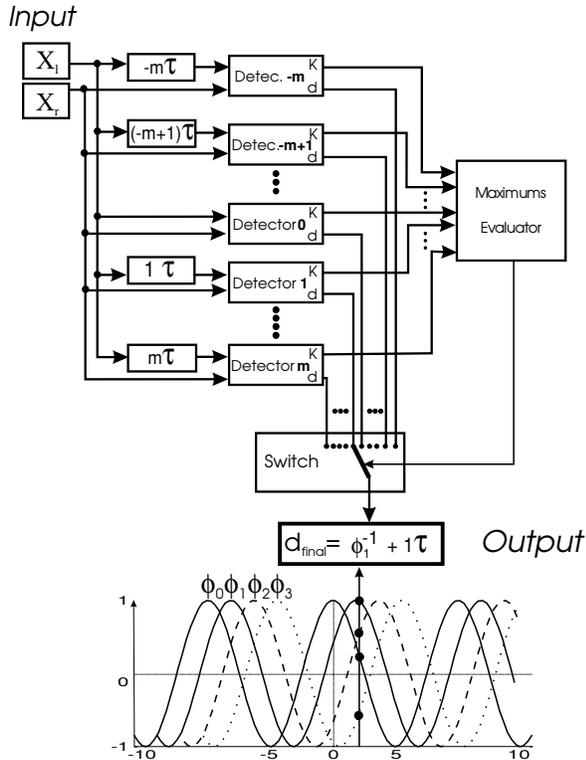


Figure 8. Scheme of the complete algorithm containing $2m + 1$ detectors as shown in Fig. 1 and (bottom) behavior of four characteristic functions at a given point in time.

left image is shifted in steps of $j\tau$, where τ usually equals one pixel and $-m \leq j \leq m$ (Fig. 8). Since shifts in both directions are introduced, this algorithmic structure immediately allows to measure positive and negative disparities. The reason for this is that the problem of measuring negative disparities in principle roots back to the ambiguity of the cosine-shaped characteristic function ϕ , which is symmetrical around zero. Thus, within one cosine-period, every value of ϕ can be attributed to one positive or symmetrically one negative disparity value. In a system with more than one detector this ambiguity can be resolved by looking at the values of ϕ obtained from the neighbors of the detector. Their values of ϕ will be different depending on the sign of the regarded disparity.⁴ In a system based on acausal Gabor filters both neighbors could be used. In our system only one neighbor can be used to determine the sign, namely the one for which the image was shifted “into the future”. The neighbor on the other side which was shifted back in time has still not “seen” the disparity of the currently regarded pixel.

Detector coupling furthermore assures that only small disparities will have to be measured, because the existing true disparities d are altered by shifting the input images and (provided enough detectors exist) for every pixel there will be one detector for which the resulting disparity $d_{res} = d - j\tau$ will fall in the range $-\tau/2 < d_{res} < \tau/2$. We will call this detector the *minimal disparity detector*. For all other detectors larger absolute disparities occur at this particular moment.

Ideally all detectors would now be able to measure these different disparity values (which of course represent the same initial disparity). The likelihood of a correct measurement, however, will be highest only for the minimal disparity detector, because its output value of ϕ occurs close to the origin where the variances of the average characteristic function $\bar{\phi}$ are small.

Initially it is unknown which of the shifted detectors is actually the minimal disparity detector. The cosine characteristic of the characteristic function ϕ , however, immediately allows to detect the most likely detector as the one which produces the largest positive output value. This is depicted at the bottom of Fig. 8 where for graphical reasons only four characteristic functions are drawn for positive shifts $j\tau$, $j = 0, 1, 2, 3$. The four marked points on the functions shall represent the output of these four detectors at a given moment. It can be seen that ϕ_1 yields the largest output value and at the same time we find that ϕ_1 also belongs to the minimal disparity detector with $d_1 = \phi_1^{-1} = \min_j(\phi_j^{-1})$. The disparity d_{final} is, thus, in this example computed as $d_{final} = \phi_1^{-1} + 1\tau = d_1 + 1\tau$. By adding different delays $j\tau$ ideally the inverted output values of the other characteristic functions ϕ_j^{-1} should also lead to the same disparity estimate d_{final} . However, the output values of their characteristic functions are found at locations distant from the origin where the variance is large.

The image of the “2” contains disparities between -4 and $+4$ and we used 9 detectors in Fig. 9 to cover $[-4.5, +4.5]$ in single disparity steps τ between detectors. The results $Q = 1$ are excellent given that this is a one-dimensional method. With larger Q more smearing is observed but the results are still acceptable.

5.2. Reducing the Correspondence Problem Through Correlations

The multi-modular arrangement has led to a tremendous improvement of our algorithm. Still, the correspondence problem remains unresolved by this setup.

There is, however, a simple way of how to utilize the multi-modular arrangement in order to strongly reduce the likelihood of a false match.

Let us introduce the problem by discussing 2-D cross-correlation techniques. A correspondence problem occurs if two image windows are “matched” which do not belong together. In real data, however, wrongly matched windows are almost always not as accurately fitting to each other than the one correct match. Therefore, increasing the window size will efficiently reduce the correspondence problem.

In our case, the *minimal disparity detector* gave us the best estimate for the disparity of a given pixel. If this estimate is correct then the detectors adjacent to the selected detector will with a rather large probability detect a very similar disparity value. In case of a false match, however, the likelihood that the two adjacent detectors will vote for the same disparity is much lower, because the variability in the data at wrong matches is larger than at the right match.

Thus, correlating three detectors reduces the correspondence problem as efficiently as increasing the window size in cross-correlation techniques.

6. Comparison to Other Algorithms

In Fig. 10 is shown a quantitative summary of the comparison of different stereo techniques plotting density against mean error. The synthetic stereo pair “2” of

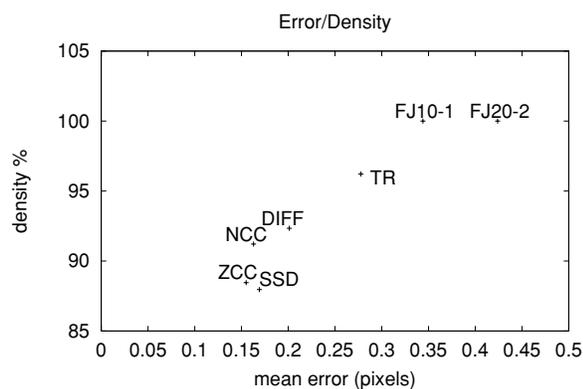


Figure 10. Error and density of different disparity estimation techniques. Results for the Henkel’s “2” synthetic stereo pair. The labels have the following meaning: TR is our algorithm (“Temporal Resonance”), NCC is Normalized Cross-Correlation, ZCC is Zero-mean Cross-Correlation, SSD is Sum of Squared Differences, FJ is the Fleet and Jepson algorithm, with two different Gabor filter’s parameters: 10 and 20 pixels of tuning period, 1 and 2 octaves of bandwidth, respectively.

R. Henkel (Fig. 3) is a good test for performances of the algorithms in a nearly-optimal case: without noise and smooth changes in depth. We have performed these tests by setting all relevant parameters of the different techniques to standard values taken from the literature.

From the plot can be seen that the algorithm proposed in this study (dubbed *Temporal Resonance*, TR) achieves a good trade-off between density (95.9%) and mean error (0.277 pixels). Fleet and Jepson’s algorithm is very sensitive to the choice of the Gabor filter parameters (Cozzi et al., 1997), achieving the worst results in precision; in both cases with a density of 100%, which is due to the fact that the test image pair is fully textured. The correlation-based and differential-based stereo techniques cluster near to each other, with densities around 90% and a mean errors of 0.15–0.20 pixels. It should be remembered that the Temporal Resonance algorithm is line-oriented. Despite of this it produces results of a quality comparable with that of more traditional approaches which operate on 2-D windows or with 2-D Gabor filters. Thus, we get good results with a much smaller computational load and with an algorithm even suited for electronic implementation.

7. Applications

Currently the system has been tested and/or used in several industrial applications, two of which are shown Fig. 11. The left part shows a depth profile of an aluminum wheel used to discriminate the burrs from the true wheel structure (courtesy Fa. Seifert, Ahrensburg, Germany). The right part shows a depth profile of a (dirty) silicon wafer recorded with a special 3-D stereo-microscope (courtesy Zeiss, Jena, Germany). The accuracy for maneuvering test needles to the test points on the wafer could in principle be improved by continuous depth monitoring. Currently this is done “by hand” and often leads to an unrecoverable damage of the wafer. The third example shows the results from the often used pentagon test image. Example movies of the real-time operation of the software version and of the FPGA implementation can be found at: <http://www.cn.stir.ac.uk/Real-Time-Stereo>.

8. Discussion

In this study we have designed and analyzed a novel, phase based algorithm, which uses causal electronic

filters, for the measurement of disparities in stereo image pairs. This approach relates to studies which utilize such filters for the analysis of optical flow (Langley and Fleet, 1992; Fleet and Langley, 1993, 1995). The strength of using a similar approach for disparity analysis lies in its simplicity, that also allows for an electronic implementation as a relatively small ASIC. From first estimates we expect real-time processing up to more than CCD-camera frame rate from this chip. Currently a working FPGA implementation of this exists already. The comparison of our approach with more traditional methods showed that only advanced 2-D cross-correlation techniques supersede our algorithm in performance at the cost of a much higher computational load.

8.1. Theoretical Aspects

The algorithm described here is the causal version of the classical phase-based approaches based on Gabor filters (Sanger, 1988; Fleet et al., 1991a). Therefore, our algorithm behaves in many ways similar to these older approaches. For example, our algorithm is robust against luminance or contrast differences between the two camera images, because the band-pass filtering removes the DC component from the input signal. Gabor filter based approaches obtain the same robustness as soon as the filters are DC-corrected, which is a simple procedure (Barron et al., 1994). The spatial resolution (Eqs. (19) and (20)) is half of that of a regular acausal Gabor filter with the same bandwidth. This is to be expected, because the causal filters cannot measure negative values.

We have solved the equations of our system for a luminance step function and a first order low-pass filter (Appendix). For a higher order low pass the solution can probably still be computed; However, it will certainly be overly complicated and not very illuminating. A luminance step function is the most relevant case for digitized camera images, because the smallest possible change is a one-bit step and all larger changes can be decomposed into these single steps. The complete solution (and also the approximate solution, Eqs. (9) and (10)) for our system are much more complex than the solutions described by Sanger (1988) using Gabor filters, or even the more advanced versions of Fleet et al. (1991a). The actual implementation of our algorithm based in IIR-filters, however, remains simple. Several well-known theoretical conjunctions and problems also hold for our algorithm.

- (1) The correspondence problem remains unresolved if only one stereo-detector is used. Correlation of several detectors reduces this problem very strongly but the disparity maps are also less dense. The reason is that several correctly estimated disparity values will also be rejected because the correlation with the adjacent detectors failed. This is due to the high variance of the average characteristic function $\bar{\phi}$ at values larger than zero (Fig. 4(A)), such that in many cases the vote of the adjacent detectors cannot be trusted anymore. A vote shared by all three adjacent detectors is, therefore, relatively rare and, in practice, we had to restrict the voting procedure to two out of three, otherwise the maps would be too sparse.
- (2) The steep drop of the average characteristic function is the reflection of another theoretically grounded problem which unequivocally occurs for all phase-based stereo-algorithms. Disparity is different from zero as soon as a spatial displacement between two corresponding contrast steps occurs. As a consequence the corresponding parts in both images of the scene in a small region which surrounds the regarded disparity change are not necessarily identical. Thus, any convolution result of a filter which covers this region will very likely be different for both images. This problem demands that only small disparities should be measured which can be achieved by vergence movement of the cameras or in a mechanically more robust way by shifting the images such as in our multi-detector coupling approach.
- (3) As a third problem we note that “phase wrap-around” occurs like in all other phase-based approaches. The multi-detector coupling procedure, however, completely eliminated this problem because the operational range of every detector is now far below the wrap-around disparity limit.

8.2. Performance of a Single Stereo Detector

The performance of a single detector which implements our algorithm with any given parameter combination ($f_0, Q, n, \omega_0, \Theta$) remains rather limited. With this respect our approach is also very similar to the older approaches (Cozzi et al., 1997). These performance limitations have so far prevented the use of the class of phase-based algorithms in high precision depth measurements (e.g. in industrial applications).

The reason for these strong performance limitations are the high susceptibility of all phase-based approaches to noise (Cozzi et al., 1997) and the above discussed unavoidable and tremendous performance decay when trying to measure larger disparities.

8.3. Improving the Performance by Multi-Detector Coupling

A central breakthrough for the performance of our algorithm was the invention of the multi-detector coupling (Henkel, 1994; Fleet, 1994), which we adopted from Henkel (1994). Henkel's main idea is the combination of several disparity-detectors. Each detector has a range of disparity-values within which the detected disparity results have low error. This range is called the 'working-range' of a detector (Henkel, 1994). The idea is to combine disparity detectors with overlapping working-ranges. In order to find the detectors which are within their working-ranges a coincidence detection was used in Henkel approach. Detectors within the working-range should produce equal disparity-values.

The cosine shape of the characteristic function proved to be a very nice feature in our algorithm by which the correct choice of the minimal disparity pair became automatically possible through maximum detection. This was a draw-back of Henkel's approach in which equal disparities (minimum difference) had to be found in a much more difficult way.

The multi-detector coupling finally made the algorithm also industrially applicable. The accuracy, however, is still slightly inferior to commonly used 2-D stereo methods (Sanger, 1988; Fleet et al., 1991a; Barnard and Fischler, 1982; Dhond and Aggarwal, 1989; Haralick and Shapiro, 1992; Faugeras, 1993) but the speed—even on conventional computers—is much higher such that a few industrially relevant applications could already be approached with our algorithm.

8.4. Practical Problems

Camera calibration is the most tedious practical problem. Our algorithm operates optimally with exactly aligned, parallel cameras because only by such an arrangement is assured that the epipolar lines are

horizontal and thus identical to the camera scan lines. In practice, we found that the calibration procedure required quite some effort. On the other hand, we observed that in the central part of the images still reasonable performance was obtained even for slightly converging cameras, where the epipolar lines are not anymore horizontal. The reason for this is, that the line-by-line differences along the Y-axis of the images are, usually, not very large such that a slightly distorted epipolar-line geometry is still tolerated. The image of the car wheel (Fig. 11) was taken with converging cameras. A general solution of the calibration problem would be *image rectification* (Faugeras, 1993), which would have to be performed by a pixel-warping procedure as a front-end to our algorithm.

8.5. Hardware Implementation

The central advantage of our algorithm is its simplicity from the viewpoint of an electronics engineer. The analytical equations are inelegantly complex, but the actual wiring diagram is, on the other hand, exceedingly simple. All components can be built as standard digital circuitry (IIR-filters, multipliers, etc.). Preliminary tests have shown that the necessary bit-depth can probably be restricted to 8 bit with the exception of the paths after the multiplication/squaring operation, where 16 bit are required. Most electronic operations remain restricted to adding or multiplying with a constant (IIR-filters), very few true integer operations exist (multiplication, division, squaring). The square root operation and the inversion of the cosine can be done by lookup tables. As a consequence this algorithm has been successfully implemented on a XLINX FPGA. An ASIC implementation is foreseen for the future.

Appendix

Here we derive the correct solution (Eq. (38)) for the characteristic function ϕ for the two input step functions assuming a first order low-pass at all stages. We start at Eq. (6):

$$M(t) = y_l(t)y_r(t) = \begin{cases} g_C(t) + \phi_C(t) & \text{if } t \geq t_r \\ 0 & \text{if } t < t_r \end{cases}$$

which is spelled out as:

$$M(t) = \begin{cases} \left(e^{-s_\infty t - s_\infty^* t_r + (s_\infty + s_\infty^*)t} + e^{-s_\infty^* t - s_\infty t_r + (s_\infty^* + s_\infty)t} \right. \\ \left. - e^{s_\infty(2t - t_l - t_r)} - e^{s_\infty^*(2t - t_l - t_r)} \right) \frac{1}{4\text{Im}^2(s_\infty)} & \text{if } t \geq t_r \\ 0 & \text{if } t < t_r \end{cases} \quad (24)$$

In the following we will only consider the part for $t \geq t_r$ abbreviated (as before) as ϕ_C which will first be Laplace-transformed to:

$$\phi_C(s) = \frac{1}{4\text{Im}^2(s_\infty) \left[\frac{e^{-(s_\infty t_l + s_\infty^* t_r)}}{s - s_\infty - s_\infty^*} + \frac{e^{-(s_\infty^* t_l + s_\infty t_r)}}{s - s_\infty - s_\infty^*} - \frac{e^{-s_\infty(t_l + t_r)}}{s - 2s_\infty} - \frac{e^{-s_\infty^*(t_l + t_r)}}{s - 2s_\infty^*} \right]} \quad (25)$$

We use a first order low-pass with the following transfer function:

$$H(s) = \frac{1}{1 + \frac{s}{\omega_\infty}} \quad (26)$$

where ω_∞ is the pole of the low-pass. After the low-pass we get:

$$\phi_C(s) = \frac{\omega_\infty}{4\text{Im}^2(s_\infty)} \left[\frac{e^{-(s_\infty t_l + s_\infty^* t_r)} + e^{-(s_\infty^* t_l + s_\infty t_r)}}{(s - s_\infty - s_\infty^*)(s - \omega_\infty)} - \frac{e^{-s_\infty(t_l + t_r)}}{(s - 2s_\infty)(s - \omega_\infty)} - \frac{e^{-s_\infty^*(t_l + t_r)}}{(s - 2s_\infty^*)(s - \omega_\infty)} \right] \quad (27)$$

To allow for inverse Laplace-transform this equation is modified to:

$$\phi_C(s) = \left[\frac{e^{-(s_\infty t_l + s_\infty^* t_r)} + e^{-(s_\infty^* t_l + s_\infty t_r)}}{s_\infty + s_\infty^* + \omega_\infty} \times \left(\frac{1}{s - (s_\infty + s_\infty^*)} - \frac{1}{s + \omega_\infty} \right) - \frac{e^{-s_\infty(t_l + t_r)}}{2s_\infty + \omega_\infty} \left(\frac{1}{s - 2s_\infty} - \frac{1}{s + \omega_\infty} \right) - \frac{e^{-s_\infty^*(t_l + t_r)}}{2s_\infty^* + \omega_\infty} \left(\frac{1}{s - 2s_\infty^*} - \frac{1}{s + \omega_\infty} \right) \right] \times \frac{\omega_\infty}{4\text{Im}^2(s_\infty)} \quad (28)$$

and inverse Laplace-transformation yields:

$$\phi_C(t) = \left[\frac{(e^{-(s_\infty t_l + s_\infty^* t_r)} + e^{-(s_\infty^* t_l + s_\infty t_r)})(e^{(s_\infty + s_\infty^*)t} - e^{-\omega_\infty t})}{s_\infty + s_\infty^* + \omega_\infty} - \frac{e^{-s_\infty t_r}(e^{2s_\infty t} - e^{-\omega_\infty t})}{2s_\infty + \omega_\infty} - \frac{e^{-s_\infty^* t_r}(e^{2s_\infty^* t} - e^{-\omega_\infty t})}{2s_\infty^* + \omega_\infty} \right] \frac{\omega_\infty}{4\text{Im}^2(s_\infty)} \quad (29)$$

$$\phi_C(s) = \frac{1}{4\text{Im}^2(s_\infty) \left[\frac{e^{-(s_\infty t_l + s_\infty^* t_r)}}{s - s_\infty - s_\infty^*} + \frac{e^{-(s_\infty^* t_l + s_\infty t_r)}}{s - s_\infty - s_\infty^*} - \frac{e^{-s_\infty(t_l + t_r)}}{s - 2s_\infty} - \frac{e^{-s_\infty^*(t_l + t_r)}}{s - 2s_\infty^*} \right]} \quad (25)$$

Let $t_l = 0$:

$$\phi_C(t) = \frac{\omega_\infty}{4\text{Im}^2(s_\infty)} \left[\frac{(e^{-s_\infty^* t_r} + e^{-s_\infty t_r})(e^{(s_\infty + s_\infty^*)t} - e^{-\omega_\infty t})}{s_\infty + s_\infty^* + \omega_\infty} - e^{-s_\infty t_r} \frac{e^{2s_\infty t} - e^{-\omega_\infty t}}{2s_\infty + \omega_\infty} - e^{-s_\infty^* t_r} \frac{e^{2s_\infty^* t} - e^{-\omega_\infty t}}{2s_\infty^* + \omega_\infty} \right] \quad (30)$$

For the two normalization pathways we have accordingly:

$$\phi_{N_l}(t) = \frac{\omega_\infty}{4\text{Im}^2(s_\infty)} \left[\frac{2(e^{(s_\infty + s_\infty^*)t} - e^{-\omega_\infty t})}{s_\infty + s_\infty^* + \omega_\infty} - \frac{e^{2s_\infty t} - e^{-\omega_\infty t}}{2s_\infty + \omega_\infty} - \frac{e^{2s_\infty^* t} - e^{-\omega_\infty t}}{2s_\infty^* + \omega_\infty} \right] \quad (31)$$

and

$$\phi_{N_r}(t) = \frac{\omega_\infty}{4\text{Im}^2(s_\infty)} \left[\frac{2e^{-(s_\infty + s_\infty^*)t_r} (e^{(s_\infty + s_\infty^*)t} - e^{-\omega_\infty t})}{s_\infty + s_\infty^* + \omega_\infty} - e^{-2s_\infty t_r} \frac{e^{2s_\infty t} - e^{-\omega_\infty t}}{2s_\infty + \omega_\infty} - e^{-2s_\infty^* t_r} \frac{e^{2s_\infty^* t} - e^{-\omega_\infty t}}{2s_\infty^* + \omega_\infty} \right] \quad (32)$$

In order to make the final result look nicer we need to define a few abbreviations:

$$\begin{aligned} & be^{-ns_\infty t_r} + b^* e^{-ns_\infty^* t_r} \\ &= 2\text{Re}(be^{-ns_\infty t_r}) \\ &= 2e^{-n\text{Re}(s_\infty)t_r} [\text{Re}(b) \cos(\text{Im}(s_\infty)nt_r) \\ &\quad + \text{Im}(b) \sin(\text{Im}(s_\infty)nt_r)] \\ &= 2|b|e^{-n\text{Re}(s_\infty)t_r} \sin(\text{Im}(s_\infty)nt_r + \angle b) \end{aligned} \quad (33)$$

$$\text{with } b = \frac{e^{2s_\infty t} - e^{-\omega_\infty t}}{2s_\infty + \omega_\infty} \quad (34)$$

$$\text{and } \angle b = \arctan \frac{\text{Re}(b)}{\text{Im}(b)} \quad (35)$$

Furthermore we set:

$$c = \frac{e^{2\text{Re}(s_\infty)t} - e^{-\omega_\infty t}}{2\text{Re}(s_\infty) + \omega_\infty} \quad (36)$$

The non-zero part of the characteristic function for $t \geq t_r$ which is ϕ is obtained as:

$$\phi = \frac{\phi_C}{\sqrt{\phi_{N_I} \phi_{N_R}}} \quad (37)$$

and we get Eq. (38):

$$\phi(t) = \frac{\cos(\text{Im}(s_\infty)t_r) - \frac{|b|}{c} \sin(\text{Im}(s_\infty)t_r + \angle b)}{\sqrt{1 - \frac{|b|}{c}(\sin(2\text{Im}(s_\infty)t_r + \angle b) + \sin \angle b) + \frac{|b|^2}{c^2} \sin(2\text{Im}(s_\infty)t_r + \angle b) \sin \angle b}} \quad (38)$$

Notes

1. IIR = infinite impulse response.
2. FIR = finite-impulse-response, IIR = infinite-impulse-response.
3. To get this one has to solve Eq. (10) for $\phi = 0$ assuming $Q \rightarrow \infty$.
4. This procedure is similar to the evaluation of a sine-cosine quadrature detector pair, only that we use cosine-cosine pairs.

References

Barnard, S. and Fischler, M. 1982. Computational stereo. *ACM Comput. Surveys*, 14(4):553–572.
 Barron, J.L., Fleet, D., and Beauchemin, S. 1994. Performance of optical flow techniques. *IJCV*, 12(1):43–77.

Cozzi, A., Crespi, B., Valentinotti, F., and Wörgötter, F. 1997. Performance of phase-based algorithms for disparity estimation. *Mach. Vision Appl.*, 9(5/6):334–340.
 Dhond, U.R. and Aggarwal, J.K. 1989. Structure from stereo: A review. *IEEE Trans. Man Mach. Cybern.*, 19(6):1489–1510.
 Faugeras O.D. 1993. *Three-Dimensional Computer Vision*. MIT Press: Cambridge, MA.
 Fleet, D. 1994. Disparity from local weighted phase-correlation. In *IEEE Int. Conf. Image Process.*, San Antonio, pp. 48–56.
 Fleet, D., Jepson, A., and Jenkin, M. 1991. Phase-based disparity measurement. *Comp. Vision, Graphic & Image Process.*, 53(2):198–210.
 Fleet, D., Jepson, A., and Tsotsos, J. 1991. Techniques for disparity measurement. *Comp. Vision, Graphic & Image Process.*, 53(1):14–30.
 Fleet, D.J. and Jepson, A.D. 1993. Stability of phase information. *IEEE Trans. PAMI*, 15(12):1253–1268.
 Fleet, D.J. and Langley, K. 1993. Toward real-time optical flow. In *Proc. Conf. on Vision Interface*, Toronto, Canada, pp. 116–124.
 Fleet, D.J. and Langley, K. 1995. Recursive filters for optical flow. *IEEE Trans. PAMI*, 17:61–67.
 Gabor, D. 1946. Theory of communication. *J. Institution of Electrical Engineers*, 93 Part III:429–457.
 Haralick, R.M. and Shapiro, L.G. 1992. *Computer and Robot Vision*, vol. 2. Addison Wesley: Reading, MA.
 Henkel, R.D. 1994. Segmentation with synchronizing neural oscillators. Technical Report 4, Zentrum für Kognitionswissenschaften, Universität Bremen.

Kanade, T., Kano, H., and Kimura, S. 1995. Development of a video-rate stereo machine. In *Proc. Int. Robotics and System Conf. (IROS-95)*, Pittsburg, PA.
 Kuglin, C.D. and Hines, D.C. 1975. The phase correlation image alignment method. In *Proc. IEEE Int. Conf. on Cybernetics and Society*, pp. 163–165.
 Langley, K. and Fleet, D. 1992. Recursive filters for phase based optic flow. In *Proc. 9th Isr. Symp. AI Comput. Vision, Ramat San Conf.*, pp. 255–264.
 Langley, K., Atherton, T.J., Wilson, R.G., and Larcombe, M.H.E. 1991. Vertical and horizontal disparities from phase. *Image and Vision Comp.*, 9(5):296–302.
 Sanger, T.D. 1988. Stereo disparity computation using gabor filters. *Biol. Cybern.*, 59:405–418.
 For a commercially available stereo-system see: <http://www.videredesign.com/svs.htm>