# A fast technique for phase-based disparity estimation with no explicit calculation of phase

F. Solari, S.P. Sabatini, and G.M. Bisio

Department of Biophysical and Electronic Engineering

University of Genova-Via all'Opera Pia 11/A-16145 Genova-ITALY

ph: +39 10 3532092, fax: +39 10 3532289, e_mail: fabio@dibe.unige.it

**Abstract**

A simpler and faster technique for depth estimation, based on phase measurements of disparity, is presented. The technique provides direct evaluations of phase differences, avoids explicit calculations of single phases and the attendant problem of phase wrapping, and is suitable for efficient software and hardware implementations.

**Indexing terms:** stereo image processing, phase-based approach, depth map

## 1 Introduction

Binocular depth perception, that is the capability of estimating the distances of the scene's objects from the observer that looks at the scene from two slightly different vantage (view) points, has important implications in many real-world visual application domains, such as autonomous robot navigation and grasping tasks. This problem has been tackled using various approaches based, in general, on the comparison of the image pairs taken from the left and right views of a (binocular) stereo system [1]. In contrast, to feature correspondence and correlation techniques, in the last decade a phase-based approach has been proposed [2] [3] as an interesting alternative, mainly because of the locality of its operations (because of its robustness, its biological motivation), and of the resulting dense depth maps (it provides). This approach relies upon the fact that the depth information from the band-passed versions of the left and right raw images can be related to the local phase difference between them. However, because of the phase periodicity, specific devices must be used to bring down the phase difference in the $(-\pi, \pi]$ range and to avoid the otherwise attendant wrap-around error in depth estimation. In this letter, we present a technique that provides a direct measure of the phase difference, thus avoiding the explicit computation of local phases. It is computationally less expensive than the modulo $2\pi$ phase manipulations required by conventional techniques.

## 2 Phase-based stereopsis

Depth perception derives from the differences in the positions of corresponding points in the stereo image pair projected on the two retinas of a binocular system. In a first approximation, the positions of corresponding points are related by a 1-D shift, the *disparity*, along the direction of the epipolar lines. Formally, the left and right intensity patterns observed by the two eyes, respectively $I^L(x)$ and $I^R(x)$, result related as: $I^L(x) = I^R[x + \delta(x)]$ where $\delta(x)$ is the binocular disparity. According to the *Fourier Shift Theorem*, the spatial shift $\delta$ in an image pattern causes a phase shift $k\delta$ in the Fourier domain, where $k$ is the spatial frequency. On the basis of this property, disparity is estimated in terms of phase differences in the spectral components of the stereo image pair. Spatially-localized phase measures can be obtained by filtering operations with complex-valued Gabor filters: $h(x; k_0) = \mathrm{e}^{-x^2/\sigma^2}\mathrm{e}^{ik_0x} = h_c(x; k_0) + ih_s(x; k_0)$, where $k_0$ is the peak frequency of the filter and $\sigma$ determines its spatial extension. The resulting convolutions can be expressed as $Q(x) = I * h(x; k_0) = \rho(x)\mathrm{e}^{i\phi(x)} = C(x) + iS(x)$, where $\rho(x)$ and $\phi(x)$ denote their amplitude and phase components, and $C(x)$ and $S(x)$ are the responses of the quadrature filter pair. Binocular disparity is related to the phase difference by [3]:

$$\delta(x) = \frac{\lfloor \phi^L(x) - \phi^R(x) \rfloor_{2\pi}}{k(x)} = \frac{\lfloor \Delta\phi(x) \rfloor_{2\pi}}{k(x)} \tag{1}$$

where $k(x) = [\phi^L_x(x) + \phi^R_x(x)]/2$, with $\phi_x$ spatial derivative of phase $\phi$, is the average spatial frequency of the bandpass signal observed at point $x$, that only under a linear phase model can be approximated by $k_0$; and where $\lfloor \cdot \rfloor_{2\pi}$ denotes the principal part of its argument, i.e., $\lfloor \cdot \rfloor_{2\pi} \in (-\pi, \pi]$ (see Fig. 1).

# 3 Direct phase difference calculation

It is worthy to note that in Eq. 1 what is important is the phase difference: the disparity estimation does not require the explicit calculation of left and right phase. Therefore, we can compute directly such a difference in the complex plane using the following identities:

$$\begin{aligned} \lfloor \Delta\phi \rfloor_{2\pi} = \lfloor \arg(Q^L Q^{*R}) \rfloor_{2\pi} &= \operatorname{atan2}\left(\operatorname{Im}(Q^L Q^{*R}), \operatorname{Re}(Q^L Q^{*R})\right) \\ &= \operatorname{atan2}\left(C^R S^L - C^L S^R, C^L C^R + S^L S^R\right) \end{aligned} \tag{2}$$

where $Q^*$ denotes complex conjugate of $Q$.

This formulation is computationally simple, because it is composed primarily of algebraic combinations of the filter outputs. Moreover, it embeds the calculation of the principal part of phase differences, without explicit manipulations of the two phases of the left and right images. In this way, it takes into account the periodicity of the phase without incurring in the "wrapping" effects on the resulting depth map (see Fig. 1).

Furthermore, following [5], with reference to the expression of the average spatial frequency, to eliminate the need for an explicit calculation of phases and, consequently, the problems arising from phase unwrapping, we use the identities:

$$\phi_x = \frac{\operatorname{Im}[Q^* Q_x]}{\rho^2} = \frac{S_x C - S C_x}{\rho^2} = \frac{S_x C - S C_x}{C^2 + S^2} \tag{3}$$

where $Q_x$, $C_x$, $S_x$ are the spatial derivatives of $Q$, $C$, $S$.

# 4 Results

The performances of the proposed technique to recover the depth from phase information are compared (see Fig. 2) with those of a conventional technique, for which we consider the following algorithmic steps to implement the wrap-around compensation logic block:

```
tmp=mod(atan2(SL,CL)-atan2(SR,CR),2*Pi);
if (sin(tmp)!=0)
    Delta_phi=tmp-Pi*(1-sign(sin(tmp)));
else
    Delta_phi=0;
```

where $\mathtt{mod}(\cdot,\cdot)$ denotes the signed remainder after division of the first argument by the second one, and $\mathtt{sign}(\cdot)$ is the signum function.

Both the conventional and the proposed have been tested on the $128\times 128$ random dot stereogram pair shown in Fig. 2A. Note that the simple difference of phases, that can be obtained through inverse trigonometric functions in the conventional approach, does not provide reliable information about the disparities in the image (see Fig. 2C). Further processing is necessary in the conventional technique to obtain results comparable to those of our technique (see Fig. 2D). The computations and the data flows involved in our technique are simpler and cheaper than those employed by a conventional approach, thus resulting more convenient and suitable for both software or hardware implementation.

Moreover, we tested our technique on real-world images. Fig. 3 shows a $233\times233$ stereo pair of a real scene together with the recovered depth map.

The simulations on both synthetic and real-world image sequences yield to excellent performances, resulting in correct discrimination between nearby and distant objects in the scenes. Points where phase information are unreliable are discarded according to a confidence measure that is related to the local energy of the quadrature filter pair output [3].

# 5 Discussion

The assets of our technique can be considered respect to both computations and implementation.

From a computational point of view, our technique is less CPU-intensive than a conventional technique. If we compared the two techniques using the built-in functions provided by MATLAB 5.3, we observed that our technique requires about half floating point operations to obtain a depth map with the same degree of reliability.

From a implementation point of view, our technique better suits hardware realization, since the procedure is based on simple operations (i.e., algebraic operations) and only one inverse trigonometric function on the outputs of a filtering stage that can be directly implemented in analog VLSI, as demonstrated by recent prototypes of our group [6].

## Acknowledgments

This work was partially supported by the *UNIGE-2000 Project "Spatio-temporal Operators for the Analysis of Motion in Depth from Binocular Images"*.

## References

[1] M.R.M. Jenkin, A.D. Jepson, and K. Tsotsos. Techniques for disparity measurement. *CVGIP: image understanding*, 53:14–30, 1991.

[2] A.D. Jenkin and M. Jenkin. The measurement of binocular disparity. In Z. Pylyshyn, editor, *Computational Processes in Human Vision*. Ablex Publ., New Jersey, 1988.

[3] T.D. Sanger. Stereo disparity computation using Gabor filters. *Biol. Cybern.*, 59:405–418, 1988.

[4] D. J. Fleet, H. Wagner, and D. J. Heeger. Neural encoding of binocular diparity: energy models, position shift and phase shift. *Vision Research*, 17:345–398, 1996.

[5] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal of Computer Vision*, 1:77–104, 1990.

[6] L. Raffo, S.P. Sabatini, G.M. Bo, and G.M. Bisio. Analog VLSI circuits as physical structures for perception in early visual tasks. *IEEE Trans. Neural Net.*, 9(6):1483–1494, 1998.

# Figure captions

**Figure 1:** (top) The left (solid line) and right (dotted line) phase of quadrature filter pair responses to a random dot stereogram. (bottom) The phase difference of such responses. The plots are normalized to $\pi$. Note the relative position of the dotted and solid lines: in the middle the right phase has a lead over the left one, and the opposite occurs at the periphery, corresponding to regions before and beyond the fixation point, respectively. The periodic jumps in the phase difference signal evidence that the mere subtraction of the two computed phases is not sufficient to obtain a reliable information about disparity.

**Figure 2:** A schematic comparison between the conventional technique and the one proposed in this letter. (A) The left and right random dot stereogram together with its ground truth. The disparity range for this image is [-2,2] pixels. (B) The common filtering stage with quadrature filter pairs to obtain the bandpass signals. (C-D) Further processing stages, intermediate result for the conventional technique before wrap-around compensation, and the final depth map estimated by the two techniques with respect to the peak frequency $k_0$ (a single picture is shown since the two results coincide).

**Figure 3:** Example of depth estimation from a natural stereo image pair. The disparity map evidences in lighter gray the objects closer to the viewer, and in darker gray the objects in the background.

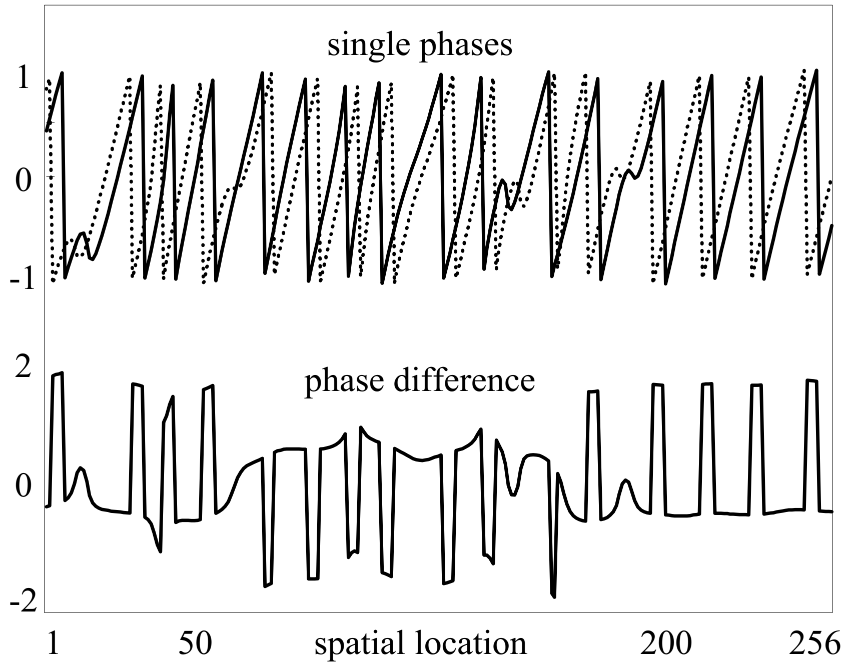


Figure 1:

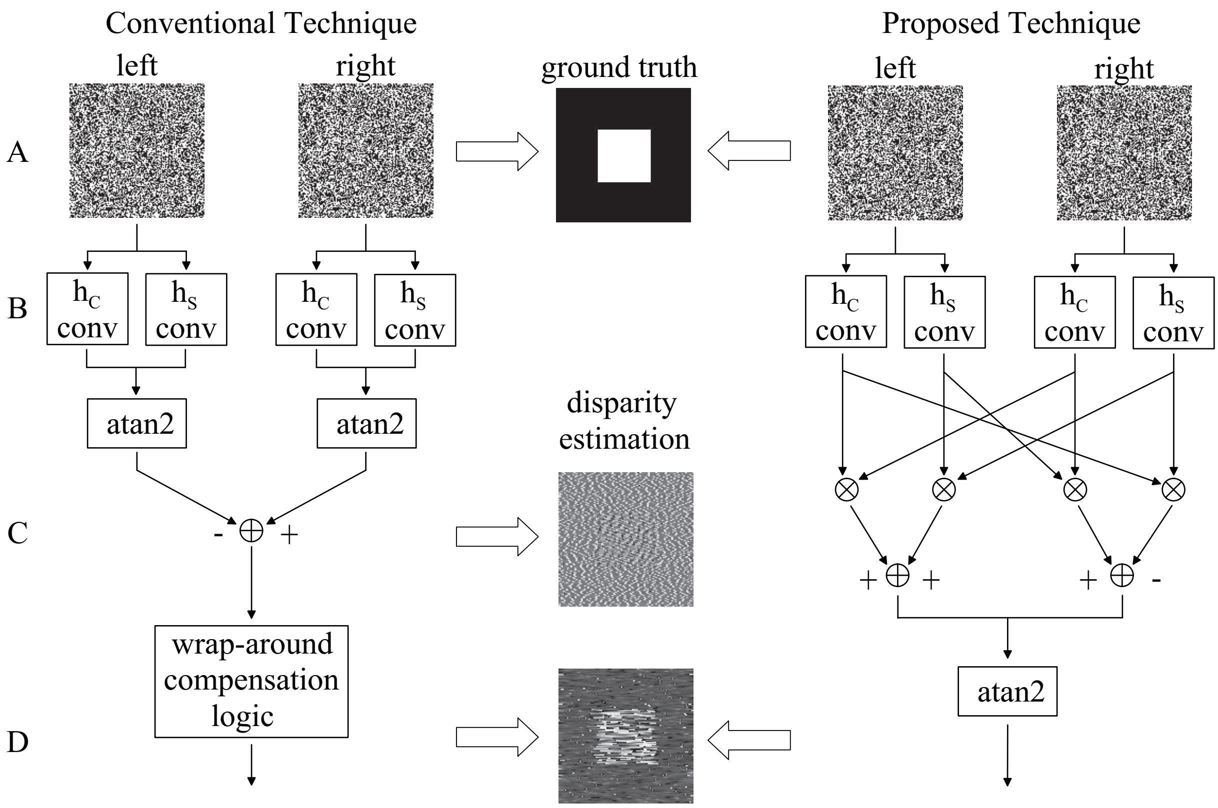


Figure 2:

left image　　　right image　　　depth map

Figure 3: