MVA '96 IAPR Workshop on Machine Vision Applications, November. 12-14, 1996, Tokyo, Japan

A Direct Disparity Estimation Technique for Depth Segmentation

Atsuto Maki *
TOSHIBA Kansai Research Laboratories
8-6-26 Motoyama-Minami-Cho, Higashinada-ku, Kobe, 658 Japan

Tomas Uhlin and Jan-Olof Eklundh
Computational Vision and Active Perception Laboratory (CVAP)
Royal Institute of Technology (KTH), S-100 44, Stockholm, Sweden

Abstract: *This paper concerns acquisition of dense depth maps in the context of image segmentation, which is fundamental in various vision applications. Focusing on stereoscopic vision as the methodology, our goal here is to develop a scheme which is computationally simple but still allows a dense disparity map. For the estimation of stereo disparity, for this reason, we employ an approach based on local Fourier phase obtained by complex bandpass filters. We consider the characteristics of this search-free and thus fast approach as suitable to compute stereo disparity as a basic cue for image segmentation. Within the framework of the phase-based algorithm, in this paper, two issues are discussed. One is the use of the derivative-based filters, and the other is the certainty-weighted disparity propagation.*

Keywords: *disparity, depth segmentation, phase-based algorithm, complex filters, disparity propagation*

1 Introduction

While different image features such as motion, color, or edge information provide cues, stereoscopic disparity also provides a strong cue for image segmentation as it carries depth information. A highly desirable property for depth segmentation is the availability of disparity estimations tightly connected to the spatial locations. In this article, we propose a technique to realize dense disparity maps while keeping computational simplicity employing the phase-based approach. The basic concept of the phase-based approach is to convolve the left and right stereo images with a complex filter, and then estimate the local disparity by computing the complex phase difference of the filter output. Since the stereopsis

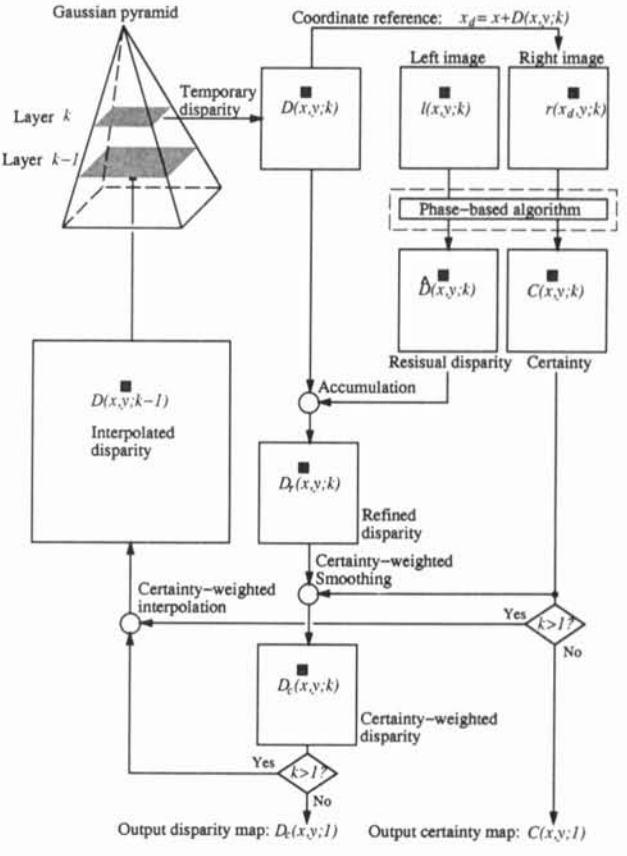


Figure 1: The framework of the disparity estimation algorithm including the hierarchical structure. The information flow in layer k is shown. The procedure in the dashed box corresponds to equation (2) and (5). Down at the original scale ($k = 1$), the disparity and the certainty maps are taken as the outputs of the algorithm.

*This report describes research done while the author was with the Computational Vision and Active Perception Laboratory (CVAP) at the Royal Institute of Technology (KTH), Sweden.

algorithm using output phase of Gabor filters was introduced as a disparity estimator [9, 11, 4], several works based on the technique have reported on its efficiency both by extensive analysis and applications [5, 10, 2, 8]. While the advantages of the phase-based method include computational simplicity, stability against varying lighting condition and especially direct localization of the disparity estimation, remaining issues concern the complex filters which need be carefully designed, and for image segmentaion the whole scheme must be constructed in such a way that disparity estimates are derived even in parts of input image where the intensity variance is limited.

In the following, we first introduce our framework of phase-based algorithm in Section 2. Motivating the advantages of the derivative-based filters in Section 3, we derive a dense disparity map by way of certainty-weighted disparity propagation in Section 4. The performance of the proposed scheme is exemplified in Section 5. Finally we summarize the work in Section 6.

2 Disparity from phase

Here we briefly introduce the principle of phase-based disparity estimation and describe the structure of our framework. In the following, $V_l(x,y)$ and $V_r(x,y)$ are the convolutions at a coordinate (x,y), obtained by a complex filter applied to the left and right images respectively. These complex functions are approximately related to each other by a phase shift, which arises from the spatial shift (i.e. disparity):

$$V_l(x,y) \simeq \exp\left[j\omega(x,y)D(x,y)\right] \cdot V_r(x,y). \quad (1)$$

$D(x,y)$ denotes disparity at (x,y) in the image and $\omega(x,y)$ represents some measure of local frequency of the image intensity function in the neighborhood of (x,y). As $\omega(x,y)$ we employ the so-called instantaneous frequency which is introduced in [5]. It is defined locally by the derivative of the phase function and therefore directly related to the local structure of the image intensity function. The relation in equation (1) then leads to a disparity estimate through computation of the complex phase difference:

$$D(x,y) \simeq \frac{\arg V_l(x,y) - \arg V_r(x,y)}{\omega(x,y)}. \quad (2)$$

The algorithm which yields the disparity and the certainty map (see Section 4) in our framework is constructed in multi-resolution hierarchy. Figure 1 schematically depicts the framework by outlining the procedure at one scale.

3 Derivative-based filters

Complex filters as disparity estimators are required to satisfy several existing constraints and different filters have been proposed accordingly [10]. Among different filter types, Gabor filters are a common choice as they minimize the product of spatial width and bandwidth [3]. They are defined by:

$$g(x;\sigma,\omega_0) = e^{\frac{-x^2}{2\sigma^2}} \cdot e^{j\omega_0 x} \quad (3)$$

where σ and ω_0 indicate the spatial half-width and the central frequency of the filter. In the literature [2, 10] Gabor filters with bandwidth close to 1 octave is a usual choice. We call them NB-Gabor filters. In our previous work [6] though, based on uncertainty analysis, we have pointed out the effectiveness of filters with larger bandwidth ($\sigma\omega_0 \simeq 1.3$; about 3 octave) in terms of both disparity localization and estimation accuracy. We call them WB-Gabor filters[1].

In order to apply the framework in a working system, the filters are required to be computationally simple. We here employ discrete approximations to the first and second derivatives which are first introduced by Westelius [10]; $t\cdot(-1,0,1)$ and $(1,0,-2,0,1)$[2]. Our goal here is to motivate the availability of the derivative-based filters as substitutes for WB-Gabor filters. The advantages of the derivative-based filters lie in their small spatial support and the normalization of the features. Smaller spatial support not only reduces the computational cost but at the same time allows better disparity localization. Furthermore, it is also convenient to have no DC-component, which is more or less inevitably involved in case of Gabor filters.

In the frequency domain, as known in Fourier theory, to take the first and the second derivative of the image intensity function $i(x)$ is equivalent to multiply the Fourier transform $I(\omega)$ by $j\omega$ and $(j\omega)^2$. Since they are not bandpass operators, as far as the continuous theory is concerned, preliminary smoothing of $i(x)$ is necessary so that the derivative operators become band limited. In the case of discrete operators, however, the smoothing is not a prerequisite as the discrete approach includes some smoothing implicitly. The Fourier transform of the derivative-based filters has the form:

$$D(\omega) = 2\ (1 + t\cdot\sin\omega - \cos 2\omega). \quad (4)$$

It contains some contributions in the negative frequency domain though they can be suppressed by adjusting t to some extent[3]. It also covers a rather wide bandwidth around the central frequency $\omega_0 = \pi/2$ as a natural feature derived from the small spatial support, i.e. 5 pixels. This number can be crudely associated with the spatial

[1] WB and NB are derived from wide-bandpass and narrow-bandpass.

[2] Integer values of the weighting factor t make it possible to implement the filter on a pipeline processor.

[3] The weighting factor $t \simeq 1.732$ is reported to minimize the negative frequency in a detailed investigation [10].

half-width σ (see equation (3)) as $2 \cdot 3\sigma \simeq 5$ pixels, which leads to $\sigma\omega_0 \simeq 1.3$. This indicates, in the context of the Gabor representation, that the bandwidth of the derivative-based filter is similar to that of the WB-Gabor filter, implying possible similarity in performance.

4 Disparity propagation

As is the case for any disparity estimators, the phase-based approach may also work improperly in parts of images where too little texture exists or stereo-correspondence is difficult. Hence, it is indispensable for an estimated disparity map to entail a certainty measure for evaluation of its reliability. There are several definitions of such certainty measures, simple or elaborate [10]. The magnitude of the filter outputs $|V_l|$ and $|V_r|$ are rather commonly used properties. It is based on the fact that the odd and even filters practically function as vertically oriented edge and line detectors, and above all, those properties are available directly from the filtering process. Our certainty measure is defined by combining that information:

$$C(x,y) \equiv \sqrt{|V_l|\ |V_r|} \cdot \frac{2\sqrt{|V_l|\ |V_r|}}{|V_l| + |V_r|}. \tag{5}$$

High certainty is reflected in the first factor of the definition which is large for strong filter outputs both in the left and the right images. The second factor captures the similarity between them, and lies in the range [0,1]. $C(x,y)$ will be kept low around the area where the disparity estimation is unstable due to bad-correspondence or lack of image intensity variance, thus helping to avoid singularities.

Based on $C(x,y)$ we compute a certainty-weighted disparity $D_c(x,y)$ with a Gaussian envelope $G(x,y;\sigma_w)$ (σ_w: the standard deviation):

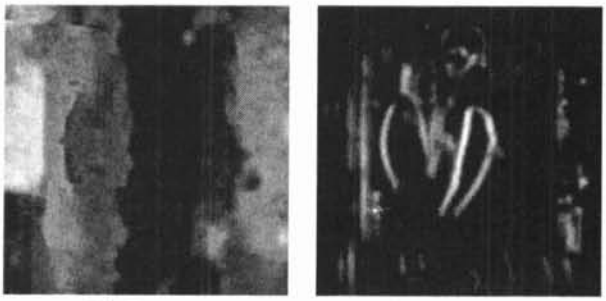
$$D_c(x,y) = \frac{\sum_{x',y'} G(x'-x,y'-y;\sigma_w)C(x',y')D(x',y')}{\sum_{x',y'} G(x'-x,y'-y;\sigma_w)C(x',y')}. \tag{6}$$

The certainty values are used as a weighting factor for the disparity estimates in a Gaussian region $G(x'-x,y'-y;\sigma_w)$ around each coordinate (x,y). Hence, disparities with higher certainty are propagated to the vicinity while those with lower certainty are suppressed. In a coarse-to-fine strategy, repeated use of this technique is especially useful, since erroneous estimates are attenuated at the early stage instead of causing recursive errors.

5 Experiments

The introduced scheme has been implemented on a Ultra-I 170/E. An example is shown in Figure 2 with three persons in a laboratory scene. The resulting disparity and certainty maps are obtained in a 4-layered coarse-to-fine framework using the derivative-based filters. It is seen that relative depth is recovered appropriately to serve as a cue for image segmentation. Figure 3 shows results of the same experiments but using NB-Gabor filters instead. It is observed that the localization is not as precise as in the case of using the derivative-based filters[4]. Especially in both ends of the image the derivative-based filters allows better localization because of the small spatial support.

(a) A sample stereo image pair (to be cross-fused)

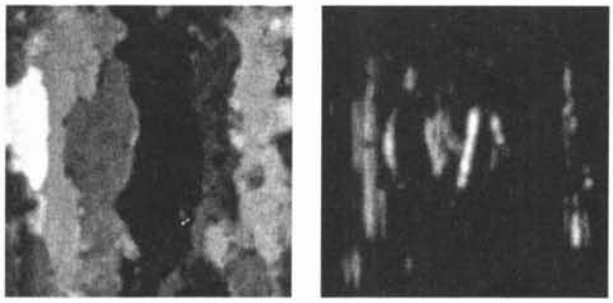


(b) Disparity *(c) Certainty*

Figure 2: Example of disparity and certainty maps I. The results are by the derivative-based filters with certainty-weighted disparity propagation. In (b) the higher the gray scale is, the larger the estimated disparity is (further away in the scene). In (c) higher gray scale represents higher certainty measure.

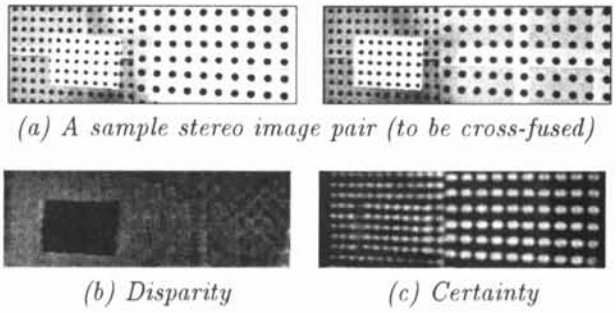
Examples of computational time are shown in Table 1. Listed are the time required for the computation by the derivative-based filters as well as by Gabor filters with and without the certainty-weighted disparity propagation, i.e. $\sigma_w = 0, 2.0$. The simplicity of the derivative-based filters is reflected in the computational cost. Though extra computation is needed when

[4] Our earlier work [8] includes a comparison between the derivative-based filters and Gabor filters. However, it is without disparity propagation.



(a) Disparity *(b) Certainty*

Figure 3: Example of disparity and certainty maps II. The results are by NB-Gabor filters ($\sigma = 3.7, \omega_0 = 1.0$) with certainty-weighted disparity propagation.



(a) A sample stereo image pair (to be cross-fused)

(b) Disparity *(c) Certainty*

Figure 4: Example of disparity and certainty maps III. In (a) a paper is hanging in front of a slanted background. The results are by the derivative-based filters with certainty-weighted disparity propagation.

the propagation process is included, it does not increase the cost substantially considering the effect it brings in the result.

Table 1: Computational Time on Ultra-I 170/E (CPU Time [sec]) using 4-layered hierarchical scheme on 256×256 images.

The standard deviation	$\sigma_w = 0$	$\sigma_w = 2.0$
Derivative-based filters	1.95	2.14
NB-Gabor filters	2.83	3.01

In Figure 4 another example with a repetitive pattern is shown. The input image is with slanted background and a paper rectangle in front. The disparity and certainty maps are obtained analogously using the derivative-based filters. It is seen in the disparity map that the scheme discriminates the paper from the background. Such a pair of dense disparity map and a certainty measure provides significant information for depth image segmentation. It should also be noted that the presented algorithm is equally suited to detect horizontal image velocity, by replacing the input stereo images with a pair of time consecutive images. Together with the disparity map, the computed motion field has been implemented to applications such as an attentional mechanism. For practical examples, see [7, 1].

6 Summary

In this article, we have considered the problem of disparity estimation using a phase-based approach to depth segmentation. We have (i) motivated the derivative-based filters in contrast to Gabor filters, and (ii) applied them in conjunction with a certainty-weighted disparity propagation with Gaussian envelope. In our framework of phase-based algorithm, throughout the coarse-to-fine scheme, the use of Gaussian envelope as an integral operator compensates the instability arising from the derivative operators. As the result, the proposed techniques improve the scheme in terms of (iii) the computational cost and (iv) the accuracy of the estimation localization. The efficiency has been confirmed through the experiments.

References

[1] J.-O. Eklundh, T. Uhlin, and P. Nordlund. Issues in active vision: attention and cue integration/selection. In *BMVC*, 1996. To appear.

[2] D. J. Fleet and A. D. Jepson. Stability of phase information. *IEEE-PAMI*, 15(12):1253–1268, 1993.

[3] D. Gabor. Theory of communication. *JIEE*, 93:429–459, 1946.

[4] A. D. Jepson and M. R. M. Jenkin. The fast computation of disparity from phase differences. In *CVPR*, pages 398–403, June 1989.

[5] K. Langley, T. J. A., R. G. Wilson, and M. H. E. Larcombe. Vertical and horizontal disparities from phase. In *1st ECCV*, pages 315–325, April 1990.

[6] A. Maki, L. Bretzner, and J.-O. Eklundh. Local Fourier phase and disparity estimates: an analytical study. In *6th CAIP*, pages 868–873, September 1995.

[7] A. Maki, J.-O. Eklundh, and P. Nordlund. A computational model of depth-based attention. In *13th ICPR*, volume IV, pages 734–739, 1996.

[8] A. Maki, T. Uhlin, and J.-O. Eklundh. Phase-based disparity estimation in binocular tracking. In *8th SCIA*, pages 1145–1152, May 1993.

[9] T. D. Sanger. Stereo disparity computation using Gabor filters. *Biological Cybernetics*, 59:405–418, 1988.

[10] C.-J. Westelius. *Focus of attention and gaze control for robot vision*. PhD thesis, Department of Electrical Engineering, Linköping University, 1995.

[11] R. Wilson and H. Knutsson. A multiresolution stereopsis algorithm based on the Gabor representation. *Proceedings IEE International Conference Im. Proc. and Applic., Warwick, U.K.*, pages 19–22, 1989.