IMPROVEMENT OF PHASE-BASED ALGORITHMS FOR DISPARITY ESTIMATION BY MEANS OF MAGNITUDE INFORMATION

Udo Ahlvers, Udo Zoelzer

Department of Signal Processing and Communications
Helmut-Schmidt-University / University of the Federal Armed Forces
Hamburg, Germany
udo.ahlvers@hsu-hamburg.de

ABSTRACT

Estimating depth information from two-dimensional stereoscopic image pairs is an important task in image processing. The disparity, i.e. the displacement of corresponding pixels in the image pair, can be estimated directly from local image phase in the fourier domain. In this paper a combined FFT-based approach for disparity estimation is proposed, where additional magnitude information is included in the disparity estimation process. Confidence criteria derived from the magnitude spectra are used to eliminate errors which makes this approach both easier and more robust compared to single phase-based algorithms. Results are presented for an artifical image pair. A stereo image coding approach is applied to compare the results in an objective manner.

1. INTRODUCTION

Depth information of a scene is necessary in various applications, e.g. obstacle detection, autonomous driving or virtual reality. On the other hand, a camera target produces only 2D-images, thus the depth information of the scene is lost [1]. This is to be solved by using stereoscopic image pairs in which the disparity can be estimated. With the well-known relation $A \sim 1/d$ it is possible to get object distances A from the disparity values d. Besides the classical correlation-based approaches [1, 2], especially phase-based approaches [3, 4, 5] have drawn attention due to their processing speed and inherent subpixel accuracy.
When estimating disparity in the Fourier domain, not only phase but also magnitude information is available. A combined approach, which takes phase **and** magnitude into account, should be promising for fast and robust disparity estimation. Furthermore, disparity maps can be used for stereo image coding [5, 6]. The quality of reconstruction is directly dependent to the quality of the disparity map. This also shows the necessity of improved disparity estimation algorithms.
In section 2 existing phase-based algorithms for disparity estimation are discussed. An extension for combined magnitude and phase evaluation, is presented in section 3. Section 4 shows simulation results for an artificial image pair.

2. PHASED-BASED ALGORITHMS FOR DISPARITY ESTIMATION

There are different algorithms available for disparity estimation. In contrast to correlation-based or feature-based approaches, which suffer from high computational load or classification problems [7], the approach in the Fourier domain, where disparity can be calculated directly from local phase differences, has been proven advantageous [3]. Furthermore, phase-based algorithms provide inherent subpixel-accuracy, as they yield information directly from the phase difference.

2.1. Gabor filters

First phase-based algorithms used Gabor filters [8] given by

$$g(n-n_0) = e^{-\frac{(n-n_0)^2}{2\sigma_n^2}} \cdot e^{j\frac{2\pi}{N}k_0(n-n_0)}, \qquad (1)$$

where the disparity was evaluated only for one frequency k_0. Such a Gabor filter can be interpreted as a discrete bandpass of mid-frequency k_0 and width σ_n^2. The Gaussian envelope has the effect of *windowing* the spatial signal before transformation. However, distortions at that particular frequency k_0, the periodicity of the phase and problems with setting the bandpass parameters lead to errors in the disparity estimation.

2.2. Using the FFT

As an enhancement to Gabor filters, a FFT-based algorithm for disparity estimation was introduced in [5]. The phase differencies between the two images are calculated at **several** frequencies k for each pixel.
To obtain dense disparity maps, the following calculations

0-7803-8554-3/04/$20.00 ©2004 IEEE.

have to be carried out for each pixel $px(n_1, n_2)$ with coordinates n_1, n_2 in the image pair. As common in image processing, algorithms are expressed with formulas only for one single pixel. In a calibrated camera-set, corresponding pixels lie in the same row (*epipolar constraint* [1]), so there exists only horizontal disparity. Due to that reason we have a one-dimensional problem with only the column index n_2 as a parameter. For simplification, we write $n = n_2$ for the column index from here on. Then $x(n)$ with $n = 0, 1, \ldots, N-1$ denotes a local area of a row of the image centered at $px(n_1, n_2)$. All terms in the Fourier domain will be expressed by capital letters $X(k)$ with the according discrete frequency index k.
Using the local discrete Fourier Transform

$$X(k) = \sum_{n=0}^{N-1} x(n) \cdot e^{-j\frac{2\pi}{N}kn} = |X(k)| \cdot e^{j\varphi(k)} \tag{2}$$

with $k = 0, 1, \ldots, N-1$ a general spatial shift yields

$$x(n-d) \quad \circ\!\!-\!\!\bullet \quad X(k) \cdot e^{-j\frac{2\pi}{N}kd} \tag{3}$$

$$= |X(k)| \cdot e^{j\varphi(k)} \cdot e^{-j\frac{2\pi}{N}kd}\,. \tag{4}$$

Applying this to two local areas of a stereoscopic image pair yields

$$x_L(n) = x_R(n-d) \tag{5}$$

$$\updownarrow$$

$$|X_L(k)| \cdot e^{j\varphi_L(k)} = |X_R(k)| \cdot e^{j\varphi_R(k)} \cdot e^{-j\varphi_D(k)} \tag{6}$$

For identical contents, the magnitude spectra $|X_L(k)|$ and $|X_R(k)|$ are identical, too. So from the phase terms

$$\varphi_L(k) = \varphi_R(k) - \varphi_D(k) \tag{7}$$

$$\Rightarrow \varphi_D(k) = \varphi_R(k) - \varphi_L(k) \tag{8}$$

and the formal definition of $\varphi_D(k)$ in Eq. (4) we get the frequency-dependent disparity

$$d(k) = \varphi_D(k) \cdot \frac{N}{2\pi k}\,,\ k = 0, 1, \ldots, N-1\,. \tag{9}$$

From this set of frequency-dependent values $d(k)$, a resulting disparity $d(n)$ has to be found for the observed area $x(n)$ in the image.
This FFT-based algorithm is shown as a part of Fig. 1. Up to now magnitude information has not been considered. This will be discussed in the next section.

3. INCLUSION OF MAGNITUDE SPECTRA

3.1. General Idea

Although the FFT-based approach in section 2.2 is promising, up to now only the **phase** information is evaluated to derive disparity values. On the other hand, performing a FFT yields not only phase $e^{j\varphi_D(k)}$ but also **magnitude** information $|X_{L,R}(k)|$ (see Eq. 2). So it is obvious to evaluate this magnitude spectra in *combination* with the phase information to further improve the quality and robustness of our algorithm. Fig. 1 shows the principle of this extended approach.

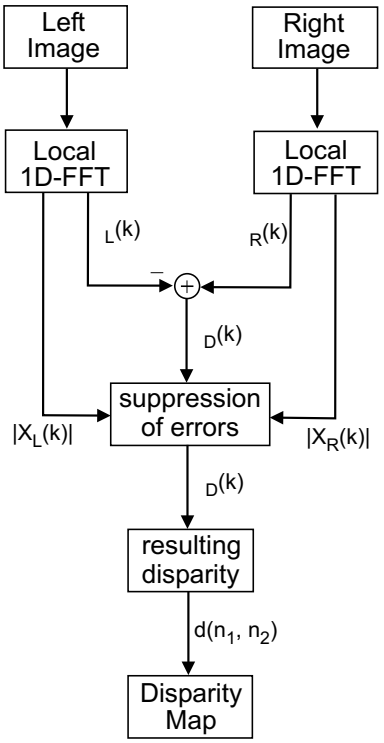


Fig. 1. *Principle of combined FFT-based disparity estimation. Besides the phase difference $\varphi_D(k)$ also the magnitude spectra $|X_{L,R}(k)|$ are taken into account.*

3.2. Criteria for Evaluation

It is a basic fact that only white noise produces a uniform distribution of magnitude in the fourier domain. In opposite to that, images will have significant distinctions in their magnitude distribution depending on texture, contrast or edges [9].
Derived from the shift theorem of the Fourier Transform (see Eq. 4), the magnitudes are identical. Assigned to a stereoscopic image pair, perfect identity will not occur e.g. due to luminance variations. But non the less, the magnitude spectra $|X_L(k)|$ and $|X_R(k)|$ should be *similar* (see Eq. 6):

$$|X_L(k)| \approx |X_R(k)|\,. \tag{10}$$

For an evaluation, criteria have to be found to check the magnitude spectra $|X_L(k)|$ and $|X_R(k)|$ concerning their

Nr.	Formula	Description
1	$\frac{\lvert X_L\rvert}{\max(\lvert X_L\rvert)}$	Normalization of single left magnitude spectrum
2	$\frac{\lvert X_R\rvert}{\max(\lvert X_R\rvert)}$	Normalization of single right magnitude spectrum
3	$\frac{\lvert\lvert X_R\rvert - \lvert X_L\rvert\rvert}{\lvert X_L\rvert}$	Magnitude difference normalized to the corresponding value of $\lvert X_L(k)\rvert$
4	$\frac{\lvert\lvert X_R\rvert - \lvert X_L\rvert\rvert}{\lvert X_R\rvert}$	Magnitude difference normalized to the corresponding value of $\lvert X_R(k)\rvert$

Table 1. *Criteria for evaluation of magnitude similarity.*

similarity. The most obvious approach is to search for local minima in the magnitude spectra. These minima denote frequencies with only low signal energy. So the phase information at these frequencies will not be reliable. Taking both magnitude spectra together into account, we found that only a normalization of the magnitude difference $||X_L| - |X_R||$ to both the corresponding values of $|X_L|$ and $|X_R|$ yields a robust criterion. Local maxima here denote a strong inequality between the two magnitude spectra. Tab. 1 lists four possible criteria.

3.3. Combined Analysis for Disparity Estimation

The further approach now is to detect "weak points" in the magnitude spectra and then to eliminate corresponding phase values for the final disparity estimation. This should make our FFT-based approach both more robust and easy, because now only phase values at **reliable** frequencies are taken into account. On the other hand too much information should not be suppressed. Therefore we only extract the *minima* derived from criteria 1 and 2 as well as the *maxima* obtained from criteria 3 and 4. Fig. 2 shows the results of this combined analysis for a local area of a stereoscopic image pair ($N = 32$). On the left hand side the magnitude spectra $|X_L(k)|$ and $|X_R(k)|$ derived from Eq. 4 are plotted. The criteria 1 and 2 are displayed by a threshold (dashed red lines). Below that the resulting phase difference $\varphi_D(k)$ is shown. The right hand side contains the criteria 3 and 4. Taking all criteria into account, it is easy to detect the frequencies $k = 6$ and $k = 13$ as "weak points". After the suppression of these critical phase values, the estimation of the resulting disparity gets easier. We now can perform a simple *unwrap*-operation. Here phase values are changed to their 2π-complement, if absolute jumps between consecutive phase values are greater than π. On the lower right of Fig. 2, the resulting phase difference $\varphi_D(k)$ after

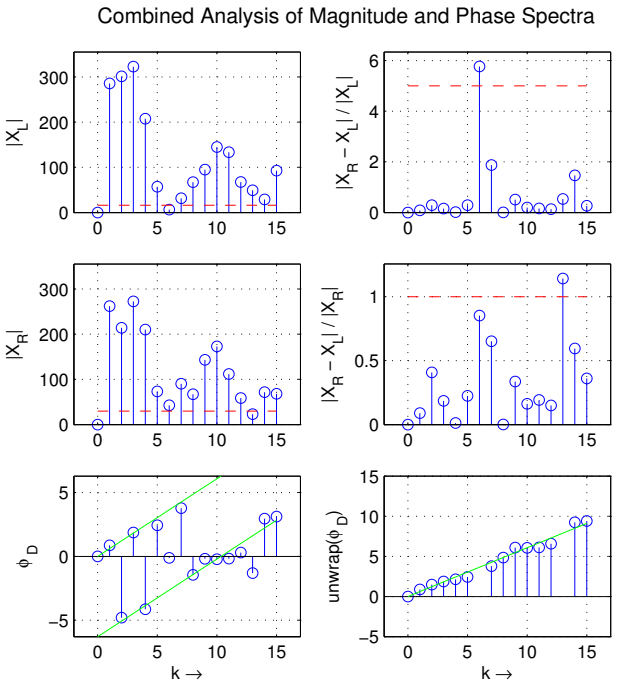


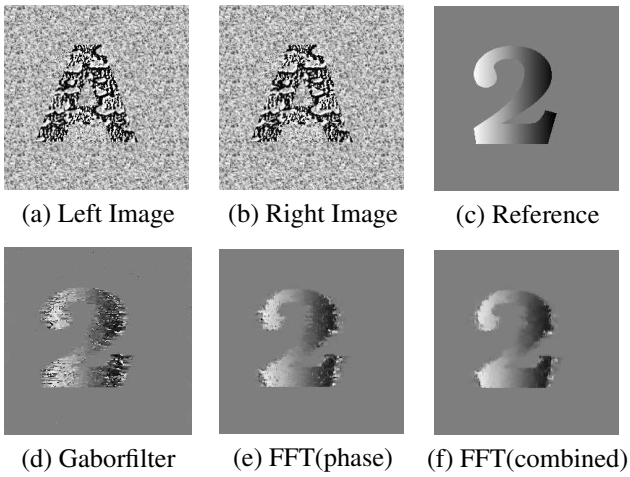
Fig. 2. *Evaluation of magnitude spectra for FFT-based disparity estimation. Due to symmetry of the FFT only $k = 0 \ldots N/2 - 1$ is plotted. Dashed red lines are the thresholds derived from the criteria in Tab. 1. Solid green lines show the resulting disparity as the gradient.*

suppression of critical values and execution of the *unwrap*-operation is shown. From the frequency-dependent disparity values $d(k)$ (see Eq. 9) the resulting disparity $d(n)$ for that particular observed image area is easily derived by

$$d(n) = \frac{1}{N/2} \sum_{k=0}^{N/2-1} d(k)\,. \tag{11}$$

4. RESULTS

Fig. 3 shows the results for different phase-based algorithms. This artifical image pair [10] was consciously chosen due to its variable disparity process (strong leaps as well as zero-disparity background) and slightly different intensity values between the two images.
In the upper row, the original image pair (a,b) and an existing reference disparity map (c) are shown. The result obtained with a Gabor filter is given in the lower left (d). Situated aside is the FFT-based output with exclusive phase evaluation, where $d(n)$ was gained as the gradient of the phase shift $\varphi_D(k)$ (see Eq. 9) with a linear regression procedure [11] (e). In the lower right, the result of the proposed combined magnitude and phase evaluation is shown (f). It



(a) Left Image (b) Right Image (c) Reference

(d) Gaborfilter (e) FFT(phase) (f) FFT(combined)

Fig. 3. *Results of phase-based disparity estimation for an artificial image pair [10]. Image size is* 256x256 *px for all images.*

Algorithm	*ErrorCrit*
Gabor-Filter	0.053
FFT (phase)	0.035
FFT (combined)	0.021

Table 2. *Error values of reconstruction. The values for* $ErrorCrit$ *were derived from Eq. 13.*

is clearly that our new approach produces the best results when compared to the reference.

For an **objective** comparison, we reconstruct the right image x_{RR} based on the left image x_L and the disparity map d according to

$$x_{RR}(n_1, n_2 - d(n_1, n_2)) = x_L(n_1, n_2)\,. \qquad (12)$$

The difference between original x_R and reconstruction x_{RR} yields the reconstruction error or *disparity compensated difference* [2, 6]. To express the whole error image by one scalar, we perform [5]

$$ErrorCrit = \frac{\sum_{n_1=0}^{N_1-1}\sum_{n_2=0}^{N_2-1} |x_R(n_1, n_2) - x_{RR}(n_1, n_2)|}{N_1 \cdot N_2 \cdot Res}\,. \qquad (13)$$

Here the parameter Res in the denominator stands for the resolution of the image ($Res = 2^8 = 256$ grayscale values in this case). Applying Eq. 13 to the results in Fig. 3, we get the error values listed in Tab. 2.

This means that with our combined FFT-based approach it is possible to reconstruct the right image according to Eq. 12 with a reliability of about 98%.

5. CONCLUSION AND OUTLOOK

We have presented a new approach for combining magnitude and phase information in FFT-based algorithms for disparity estimation. Additionally to the evaluation of the phase difference, different magnitude criteria are used to detect "weak points" in the frequency domain. After their suppression only reliable phase values remain for a robust estimation of the resulting disparity.

This combined approach outperforms classical approaches based exclusively on the evaluation of local phase information.

Future work focuses on the implementation of a cascaded approach. In the resulting image pyramid, the disparity is reduced from one level to another. Keeping a constant window length, even image areas with strong leaps in the disparity process (e.g. sharp object edges) should be processable.

6. REFERENCES

[1] C.-E. Liedtke. *Computer Aided Scene Analysis.* Script for Lecture, University of Hannover, 1998.

[2] W. Woo. *Overlapped Block Disparity Compensation with Adaptive Windows for Stereo Image Coding.* In IEEE Trans. CSVT, Vol.10, pages 861–867, 2000.

[3] D. Vernon. *Fourier Vision.* Kluwer Academic Publishers, 2001.

[4] M. Felsberg. *Disparity from Monogenic Phase.* In Pattern Recognition, Proc. DAGM, LNCS2449, pages 248-256, Springer, 2002.

[5] U. Ahlvers, U. Zoelzer, S. Rechmeier. *FFT-based Disparity Estimation for Stereo Image Coding.* In Proc. ICIP, Barcelona, Spain, 2003.

[6] N. Boulgouris. *Embedded Coding of Stereo Images.* In Proc. ICIP, Vol.3, pages 640–643, Vancouver, Canada, 2000.

[7] R. Henkel. *Fast Stereovision by Coherence Detection.* In Proc. CAIP, LNCS1296, pages 297–304. Springer, 1997.

[8] T. Sanger. *Stereo Disparity Computation using Gabor Filters.* Biological Cybernetics, 59, 1988.

[9] A. Jain. *Fundamentals of Digital Image Processing.* Prentice Hall, 1989.

[10] R. Henkel. *Rolf Henkels Homepage.* http://axon.physik.uni-bremen.de/ rdh/research/.

[11] E. Kreyszig. *Statistical Methods and their Applications.* Vandenhoeck & Ruprecht, 7. edition, 1985.