A Self Adaptability based Stereo Correspondence of Face Images using Wavelets

Prabhakar C.J
Department of Studies in Computer Science
Kuvempu University, Shankaraghatta,Karnataka, India.

Jyothi.K
Department of IS&E J.N.N. College of Engineering
Shimoga,Karnataka, India

ABSTRACT

This paper proposes a discrete wavelet-based stereo matching technique. A discrete wavelet transform is first applied to a pair of stereo images to decorrelate the images into a number of approximations. Information in the basebands is less sensitive to shift variability of the wavelet transform. A self-adapting dissimilarity measure is employed to generate a disparity map of the stereo pairs. Results show that the proposed technique produces smoother disparity maps with less computation cost.

Keywords

Stereo matching, Discrete Wavelet Transform , disparity.

1. INTRODUCTION

The stereo correspondence is process of finding the corresponding points in the stereo image pair. The estimated disparity between the stereo images is then used to extract depth information of the scene. This is used in the area of 3D object reconstruction, 3D face reconstruction and industrial automation systems.

In computer vision, estimation of the stereo correspondence between the stereo image pair is one of the critical problems. The stereo correspondence is mainly categorized into two types, those are local method and global method. Local stereo correspondence uses the information within the small area. Some of the local methods [1-8] such as block matching, gradient based optimization and feature matching are very efficient. While global methods optimize some global (energy) function.

The computation time of the area-based (local) methods are very less, but produces less accurate results.While energy-based (global) consumes more time for computation and generate more accurate result. The most commonly used matching algorithms are the sum of absolute differences (*SAD*), the sum of squared differences (*SSD*) and the normalized *SSD.*

Muhlman et al [4] introduced a local-based matching technique for RGB stereo images. This method uses left-to-right consistency and uniqueness constrains to produce the disparity map. The resulting disparity map is smoothed by applying a median filter. Yoon et al [5] proposed a correlation-based local stereo correspondence matching technique, which uses a refined implementation of the Sum of Absolute Differences (SAD) criteria and a left-to-right consistency check. This algorithm uses a variable correlation window size to reduce the errors in the areas containing blurring or mismatch errors. Yoon and Kweon [6] introduced another area-based algorithm, which uses adoptive weights based on the color similarity and geometric distances for each pixel in the search area to reduce the ambiguity errors. Yong and Lee reported area based method, which is insensitive to radiometric variation between stereo images using adaptive normalized cross correlation.

Kim et al [7] proposed an energy-based method for stereo correspondence matching, which generates a dense disparity map using a region dividing technique based on canny edge detection. The resulted disparity map is refined by minimizing the energy function using a Lagrangian optimization algorithm. Ogale and Aloimonos [8] reported new global-based correspondence matching algorithm, in which matching process is independent of the contrast variation of the stereo images. This algorithm based on multiple spatial frequency channels for local matching and a fast non-iterative left/right diffusion process for the global solution.

Current research in stereo correspondence estimation has attracted a lot of focus on multiresolution techniques, based on wavelets/multiwavelets scale-space representation and analysis [9]. However, very little work has been reported in this regard. Wavelet based stereo matching algorithms have received much attention due to scale-space localization properties of the wavelets [9, 10, 11 and 12].

In this paper, a wavelet-based stereo matching algorithm using a self-adapting dissimilarity measure [13] technique is presented. A discrete wavelet transform is applied to the rectified stereo images to decompose them into subbands. The self-adapting dissimilarity measure is used to generate a disparity map.

The remaining sections of the paper are organized as follows. The proposed wavelet based stereo matching method is discussed in section 2. The experimental results are presented in section 3. Finally, the conclusion is given in Section 4.

2. STEREO MATCHING

Figure 1 shows a block diagram of the proposed wavelet-based stereo matching technique using the self-adapting dissimilarity algorithm. A pair of stereo images is input to the system. The images are first rectified to suppress the vertical displacement. A 2-D discrete wavelet transform (DWT) is used to decompose the given image into four subbands, namely: 1) LL; 2) LH; 3) HL; and 4) HH. Approximation of the input image has different spectral content of the input image, while the detail subbands mainly contain a mixture of horizontal, vertical and diagonal details of input image. In addition to this, the information in the basebands is less sensitive to the shift variability of the wavelets. Approximation subband of reference image is segmented using mean shift method. The self-adapting dissimilarity measure [13] is used for stereo matching, that combines sum of absolute intensity differences and a gradient based measure to generate a disparity map. In our method instead of intensity and gradient, wavelet coefficients are used for matching between the two input subbands.

2.1Wavelets Fundamentals

A wavelet transforms are based on small waves, called wavelets of varying frequency and limited duration. In discrete wavelet transform (DWT), an image is decomposed into a set of band limited components, called subbands. Wavelet theory is based on the refinement equations are given below

$$\phi(t) = \sum_h \mathrm{w}_h \phi(Mt - h) \quad , \tag{1}$$

$$\varphi(t) = \sum_h c_h \phi(Mt - h) \quad . \tag{2}$$

Where c_h and W_h represents the scaling and wavelet coefficients. ϕ (x,y) is scaling function and φ (x,y) is wavelet function. Discrete wavelet transform is consisting of a scaling function and three wavelet functions. Wavelet transforms results in the information belonging to the approximation space. A_k and detail space D_k possessing the property

$$\mathbf{A_{k-1} = A_k \ominus D_k} \quad . \tag{3}$$

Mallat [9] introduced one of the most widely used discrete wavelet transform. In Mallat's representation, the details space (*Dk*) consists of three components, which are horizontal, vertical and diagonal. In general terms, the wavelet transform modulus maxima (WTMM) can be described at any point(s, k) .The vector coefficients of the wavelet transform modulus (WTM) can be expressed as

$$\mathrm{WTM}_{s,k} = \mathrm{W}_{s,k} \angle \Theta_{\mathrm{W}_{s,k}} \quad . \tag{4}$$

Where $|WT_{s,k}|$ represents the magnitude of the WTM and Θ_W the phase of the WTM, where as represents the scale and k for the coefficient under consideration. Furthermore, the magnitude of expressed in terms of horizontal and vertical details space as

$$W_{s,k} = \sqrt{\left|D_{h,s,k}\right|^2 + \left|D_{v,s,k}\right|^2} \quad , \tag{5}$$

where *Dh,s,k* and *Dv,s,k* are the *k*th horizontal and vertical detail components at scale *s*.Furthermore can be expressed as

$$\Theta_{W,s,k} = \begin{cases} \alpha(s,k) & if \;\; D_{h,s,k} > 0 \\ \pi - \alpha(s,k) & if \;\; D_{h,s,k} < 0 \end{cases} \quad . \tag{6}$$

2.2. Matching Strategy

We are used self adapting dissimilarity measure to find the disparity map of stereo image pair. In self adapting dissimilarity measure, that combines sum of absolute intensity differences and a gradient based measure that are defined as

$$CSAD(x,y,d) = \sum_{(i,j)\in N(x,y)} f_1(i,j) - f_2(i+d,j). \tag{7}$$

$$Gradient\;(x,y,d) = \sum_{(i,j)\in N_x(x,y)} \left|\nabla_x C_1(i,j) - \nabla_x C_2(i+d,j)\right| + \sum_{(i,j)\in N_y(x,y)} \left|\nabla_y C_1(i,j) - \nabla_y C_2(i+d,j)\right|. \tag{8}$$

Fig 1: Rectified left and right Images

Where $N(x,y)$ is a 3x3 window at position (x,y), $N_x(x,y)$ surrounding window without the right most column, $N_y(x,y)$ a surrounding window without the lowest row, the forward gradient to the right and the forward gradient to the bottom. Color images are taken into account by summing up the dissimilarity measure for all channels.

An optimal weighting w between C_{SAD} and C_{GRAD} is determined by maximizing the number of reliable correspondences that are filtered out by comparing left to right and right to left disparity maps in conjunction with a winner take all optimization. The resulting dissimilarity measure is given below,

$$C(x,y,d) = (1-\omega) * CSAD(x,y,d) + \omega * Gradient(x,y,d) \quad . \tag{9}$$

3. EXPERIMENTAL RESULTS

3.1 Experiments on Stereo Face Images

In order to evaluate the performance of the proposed method the wavelet based self-adapting dissimilarity measure algorithm was applied to the stereo face images, which were taken from the stereo face database [14]. The stereo images are first rectified to suppress the vertical displacement. Figure.1 shows the rectified stereo face images. The biorthogonal discrete multiwavelet transform is then applied to rectified stereo images to decorrelate them into their subbands. Figure.2 shows the resulting subbands after applying one level wavelet decomposition.

Figure.3 shows the result obtained after applying disparity map resulted from our proposed method and the disparity map produced by refined SAD algorithm. The results show that disparity map obtained by our algorithm is better than SAD algorithm.

3.2 Experiments on Stereo Images of Middlebury database

Experiment is also performed on some of the stereo test images from the Middlebury stereo database [15] and compared with the corresponding ground truths. We have used stereo images viz. Teddy, Cones and Aloe (First Column of Figure 5) of Middlebury database. The corresponding ground truth disparity map is shown in second column of Figure 5. The third column shows disparity map obtained using proposed method. To evaluate the proposed method, bad pixel map is estimated between result of proposed method and corresponding images ground truth. The fourth column of Figure 5 shows the bad pixel map. The figure shows that mapping of bad pixel is very less.

Fig 2: Wavelet Decomposition at level 1 of left face image shown in Figure 1

Root mean square method is quantitative way to estimate the quality of the computed correspondences. The RMS error is calculated between the computed depth map $d_c(x,y)$ and the ground truth map $d_T(x,y)$, that is,

$$R = \left(\frac{1}{N} \sum_{(x,y)} \left| d_c(x,y) - d_T(x,y) \right|^2 \right)^{\frac{1}{2}} \quad (9).$$

Where N is the total number of pixels. The Table 1 shows the root mean square error estimated for results obtained for proposed method and SAD algorithm.

Table 1. Root Mean Square error

	Teddy	Cones	Aloe
Proposed method	0.0710	0.0694	0.0374
SAD	0.0745	0.0823	0.045

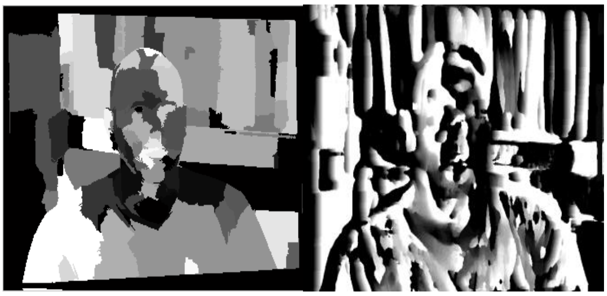

Fig 3: Disparity map of proposed method and SAD

4. CONCLUSIONS

In this paper a new wavelet-based stereo matching technique using a self-adapting dissimilarity measure algorithm was presented. A discrete wavelet transform decomposes the input stereo images into a number of subbands. The resulting subbands of the two views were then used to generate the disparity map using the self-adapting dissimilarity measure algorithm with wavelets described in this paper. Results show that the proposed technique produces a disparity map with significantly less mismatch errors.

5. REFERENCES

[1] G. Vander Val, M. Hansen, M. Piacentino, “The ACADIA Vision processor”, Proceedings of 5th International Workshop on Computer Architecture for Machine Perception, Padova, Italy, pp. 31–40,2001.

[2] L. Di Stefano, S. Mattoccia, “Fast Stereo Matching for the VIDET System using a General Purpose Processor with Multimedia Extensions”, Proceedings of 5th International Workshop on Computer Architecture for Machine Perception, Padova, Italy, pp. 356–362,2000.

[3] L. Di Stefano, M. Marchionni, S. Mattoccia, and G. Neri ,“A Fast Area-Based Stereo Matching Algorithm”, Image and Vision Computing, vol. 22, pp. 983-1005,2004.

[4] K. Muhlmann, D. Maier, R. Hesser, and R. Manner, “Calculating dense disparity maps from color stereo images, an efficient implementation,” Proc. IEEE Workshop on Stereo and Multi-Baseline Vision (SMBV 2001), pp. 30-36, 2001.

[5] L. Di Stefano, M. Marchionni, S. Mattoccia, and G. Neri, “A Fast Area-Based Stereo MatchingAlgorithm,” Image and Vision Computing, vol. 22, pp. 983-1005, March 2004.

[6] K.J. Yoon and I. S. Kweon, “ Adaptive support-weight approach for correspondence search,” IEEE

Transactions on Pattern Analysis and Machine Intelligence, vol. 28, pp. 650-656, April 2006.

[7] H. Kim, S. Yang, and K. Sohn, "3D reconstruction of stereo images for interaction between real and virtual worlds," Proc. IEEE International Conference on Mixed and Augmented Reality, October 2003.

[8] A.S. Ogale and Y. Aloimonos, "Robust Contrast Invariant Stereo Correspondence," Proc. IEEE International Conference on Robotics and Automation, ICRA 2005, pp. 819-824, April 2005.

[9] S. Mallat, "A Wavelet Tour of Signal Processing", Academic Press, 1999.

[10] I. Sarkar and M. Bansal, "A wavelet-based multiresolution approch to solve the stereo correspondence problem using mutual information", IEEE Transaction on System, Man, and Cybernetics, vol. 37, pp. 1009-1014, August 2007.

[11] P Begheri and C.V Sedan "Stereo correspondence matching using multiwavelets" , Fifth International conference on Digital Telecommunication, 2010.

[12] A.Bhatti, S.nahavandi, M.Hossny, "Wavelets/ Multiwavelets bases and correspondence estimation problem; An analytic study", 11th International Conference on control, automation, robotics and vision, Dec 2010.

[13] A.Klaus, M Sormann and Konrad Karner "Segment-Based Stereo Matching Using Belief Propagation and a Self-Adapting Dissimilarity Measure ", 2006.

[14] http:// cvlab.epfl.ch/data/strecham VS/ , 2010.

[15] http://vision.middlebury.edu/stereo/, January 2010.

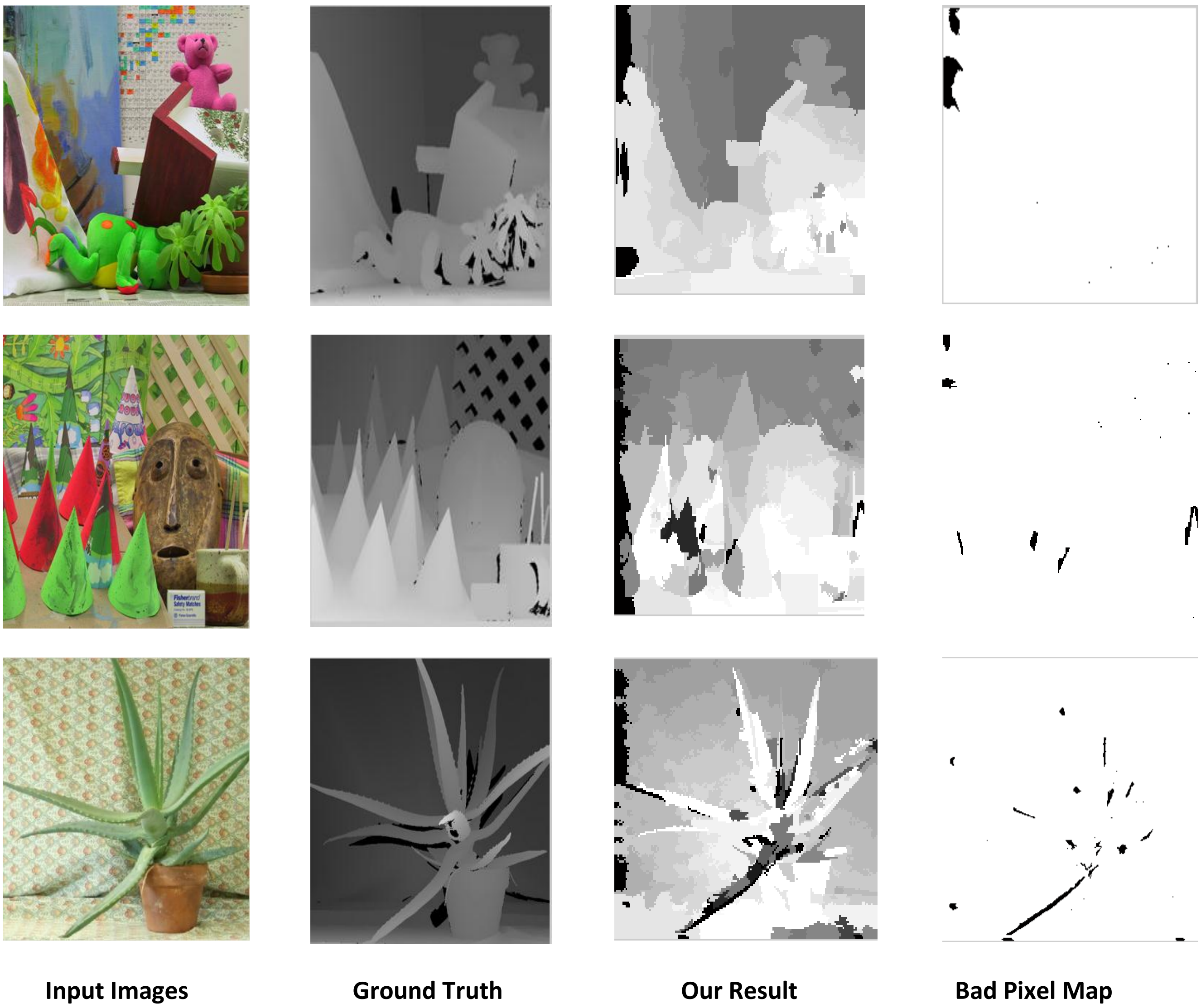

Fig 4: Results on Middleburry datasets. From top to bottom: Teddy, Cones,alloe. From left to right: reference images, ground truth disparities, the results of the proposed algorithm and the error images where the black regions represent the erroneous pixels.