2010 Fifth International Conference on Digital Telecommunications

Stereo Correspondence Matching using Multiwavelets

Pooneh Bagheri Zadeh and Cristian V. Serdean
Department of Engineering, Faculty of Technology,
De Montfort University
Leicester, UK
E-mail: pbz@dmu.ac.uk , cvs@dmu.ac.uk

***Abstract*—this paper presents a novel multiwavelet-based stereo correspondence matching technique. A multiwavelet transform is first applied to a pair of stereo images to decorrelate the images into a number of approximation (baseband) and detail subbands. Information in the basebands is less sensitive to shift variability of the multiwavelet transform. Basebands of each input image carry different spectral content of the image. Therefore, using the basebands to generate the disparity map is likely to produce more accurate results. A global error energy minimization technique is employed to generate a disparity map for each baseband of the stereo pairs. Information in the resulting disparity maps is then combined using a Fuzzy algorithm to construct a dense disparity map. A filtering process is finally applied to smooth the disparity map and reduce its erroneous matches. Middlebury stereo test images are used to generate experimental results. Results show that the proposed technique produces smoother disparity maps with less mismatch errors compared to applying the same global error energy minimization technique to wavelet transformed image data.**

Keywords- Multiwavelets, Correspondence matching, Disparity estimation, Stereo vision.

I. INTRODUCTION

Stereo correspondence matching is defined as the process of finding the best correspondence points between a stereo image pair. The disparity map generated from the correspondence matching process, along with the stereo camera parameters are then used to compute the depth map and produce a 3D view of the scene. However, the accuracy of the correspondence map, which is crucial in generating a precise 3D view of the scene, is limited due to a number of problems such as occlusion, ambiguity, illumination variation and radial distortion [1].

Area-based (local) and energy-based (global) correspondence matching algorithms are the two most common types of algorithms used in the literature to generate disparity maps. In area-based methods a disparity vector for each pixel within a window search area is calculated using a matching algorithm, while in energy-based methods, the disparity vector is determined using a global cost function minimization technique. Area-based methods are fast but produce descent results, while global methods are more time consuming and in turn generate more accurate results.

Muhlman et al [2] presented an area-based matching technique for RGB stereo images. This algorithm uses left-to-right consistency and uniqueness constrains to generate the initial disparity map. The resulting disparity map is then further smoothed by applying a median filter. Another area-based scheme was proposed by Stefano et al [3]. Stefano's algorithm is based on the uniqueness and constraint but it relies on the left-to-right matching phase. Yoon et al [4] introduced a local correlation-based 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 [5] proposed another local-based algorithm, which uses different supporting weights based on the color similarity and geometric distances for each pixel in the search area to reduce the ambiguity errors.

Kim et al [6] reported a global-based technique for stereo correspondence matching. This algorithm first generates a dense disparity map using a region dividing technique based on Canny edge detection. It then further refines the disparity map by minimizing the energy function using a Lagrangian optimization algorithm. Ogale and Aloimonos [7] proposed another global-based correspondence matching algorithm, which is independent of the contrast variation of the stereo images. This algorithm relies on multiple spatial frequency channels for local matching and a fast non-iterative left/right diffusion process for the global solution. An energy-based algorithm for stereo matching, which uses a belief propagation algorithm, was presented in [8]. This algorithm uses hierarchical belief propagation to iteratively optimize the smoothness of the disparity map. It delivers fast convergence by removing redundant computations. Choi and Jeong [9] proposed an energy-based stereo matching technique, which models the intensity differences between the two stereo images using a uniform local bias assumption. This local bias assumption is less sensitive to intensity dissimilarity between the stereo images when using normalized crosscorrelation matching cost functions. The resulting information from the cost function and the fast belief propagation algorithm are combined to generate a

978-0-7695-4071-9/10 $26.00 © 2010 IEEE
DOI 10.1109/ICDT.2010.37

IEEE computer society

smooth disparity map.

Over the past years much research has been done to improve the performance of the correspondence matching techniques. Multiresolution-based stereo matching algorithms have received much attention due to hierarchical and scale-space localization properties of the wavelets [10, 12]. This allows for correspondence matching to be performed on a coarse-to-fine basis, resulting in decreased computational costs. Sarkar and Bansal [12] presented a multiresolution-based correspondence generation technique using a mutual information algorithm. They showed that the multiresolution technique produces significantly more accurate matching results compared to correlation-based algorithms at much lower computational cost.

Multiwavelets offer some notable advantages compared to scalar wavelets, such as possessing orthogonality (preserving length), symmetry (good performance at the boundaries via linear-phase) and a high approximation order at the same time [11], which could potentially increase the accuracy of the correspondence matching techniques. However, in spite of this, the application of the multiwavelets in stereo correspondence matching has been little investigated in the literature so far.

In this paper, a novel multiwavelet-based stereo matching algorithm using a global error energy minimization technique is presented. An unbalanced multiwavelet is first applied to the input stereo images to decompose them into a number of subbands. The global energy minimization algorithm is then employed to generate a disparity map using the coarse subbands. A fuzzy algorithm is used to combine the disparity maps and generate a dense disparity map. The rest of the paper is organized as it follows. Section II presents a brief review of the multiwavelet transform. The proposed stereo matching technique is discussed in Section III. Experimental results are presented in Section IV and Section V is dedicated to the conclusions.

II. MULTIWAVELET TRANSFORM

Multiwavelet transforms are in many ways similar to scalar wavelet transforms. Classical wavelet theory is based on the refinement equations given below:

$$\begin{aligned}\phi(t) &= \sum_{k=-\infty}^{k=\infty} h_k \phi(m\,t-k) \\ \psi(t) &= \sum_{k=-\infty}^{k=\infty} g_k \psi(m\;t-k)\end{aligned} \tag{1}$$

where $\phi(t)$ is a scaling function, $\psi(t)$ is a wavelet function, h_k and g_k are scalar filters and m represents the subband number. In contrast to wavelet transforms, multiwavelets have two or more scaling and respectively wavelet functions.

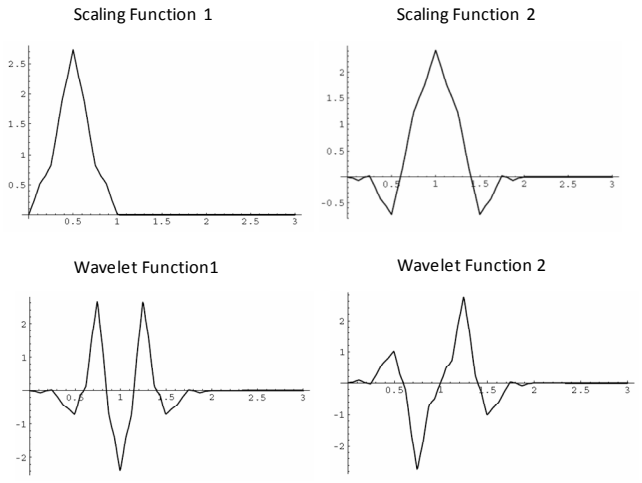


Figure 1. GHM multiwavelet with multiplicity r = 2 and approximation order 2.

The set of scaling and wavelet functions of a multiwavelet in vector notation can be defined as:

$$\begin{aligned}\Phi(t) &\equiv \begin{bmatrix}\phi_1(t) & \phi_2(t) & \phi_3(t) & \dots & \phi_r(t)\end{bmatrix}^T \\ \Psi(t) &\equiv \begin{bmatrix}\psi_1(t) & \psi_2(t) & \psi_3(t) & \dots & \psi_r(t)\end{bmatrix}^T\end{aligned} \tag{2}$$

where $\Phi(t)$ and $\Psi(t)$ are the multiscaling and respectively multiwavelet functions, with r scaling- and wavelet functions. In the case of scalar wavelets $r=1$, while multiwavelets support $r \geq 2$. To date, most multiwavelets are restricted to $r=2$. Such multiwavelets poses two scaling and two wavelet functions and can be represented as [13]:

$$\begin{aligned}\Phi(t) &= \sqrt{2}\sum_{k=-\infty}^{k=\infty} H_k \Phi(m\,t-k) \\ \Psi(t) &= \sqrt{2}\sum_{k=-\infty}^{k=\infty} G_k \Psi(m\,t-k)\end{aligned} \tag{3}$$

where H_k and G_k are $r \times r$ matrix filters and m is the subband number. Figure 1 shows an example of GHM multiwavelet basis functions with multiplicity $r=2$ and approximation order 2 [11].

Due to their multiple filters, multiwavelets can possess symmetry, orthogonality and approximation orders higher than 1 simultaneously, while scalar wavelets do not allow this extra degree of freedom.

Since multiwavelets generate four different approximation subbands (basebands) from the input image, this can be exploited to increase the accuracy and reduce the number of erroneous matches in the disparity maps. More information about the multiwavelet transform and its applications can be found in [11, 13].

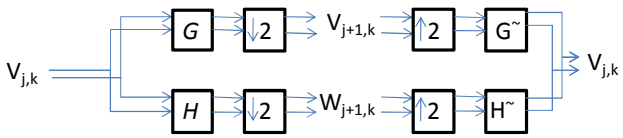


Figure 2. Analysis/synthesis stage of one level multiwavelet transform.

Similar to wavelet transforms, multiwavelets can be implemented using Mallat's filter bank theory [10]. Figure 2 shows one level of the analysis and synthesis stage for a 1D multiwavelet transform, where blocks G and H are low- and high-pass analysis filters and $G^{\sim}$ and $H^{\sim}$ are low- and high-pass synthesis filters. 2D multiwavelet transforms are separable and hence the 2D transform can be calculated via two 1D transforms. Therefore, for one level of decomposition and multiplicity 2, the 2D multiwavelet transform generates sixteen subbands, where four of them are approximation subbands. Figure 3 provides a visual comparison of the resulting subbands for a 2D scalar wavelet based on the Antonini 9/7 filter (Figure 3(a)) and respectively for a 2D multiwavelet based on unbalanced GHM filters (Figure 3(b)). As it can be seen from Figure 3, the multiwavelet transform generates four subbands instead of each subband generated by a scalar wavelet and these four subbands carry different spectral content of the input image due to multiwavelets' filter properties.

III. MULTIWAVELET-BASED STEREO MATCHNING TECHNIQUE

Figure 4 shows a block diagram of the proposed multiwavelet-based stereo matching technique using the global error energy minimization algorithm. A pair of stereo images is input to the system. The images are first rectified to suppress the vertical displacement. An unbalanced multiwavelet transform is then applied to the stereo images to decorrelate them into their subbands. Figure 5 shows the resulting subbands after applying a one level multiwavelet transform.

Due to the unbalanced nature of the multiwavelet, each resulting baseband (L_1L_1 , L_1L_2 , L_2L_1 and L_2L_2) is an approximation of the input image with different spectral content of the input image, while the remaining 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 multiwavelets. In this paper, an unbalanced multiwavelet with multiplicity *r=2* is used and as such the multiwavelet transform of each input image contains four basebands. The same basebands in the two images are then passed to a regional-based stereo matching block. The matching algorithm uses a global error energy minimization technique [14] to generate a disparity map between the two input subbands. This global error energy minimization technique is briefly described in sub-section A. The output of the matching process gives four disparity maps. These maps are then combined using a Fuzzy algorithm to generate a dense disparity map with less erroneous matches.

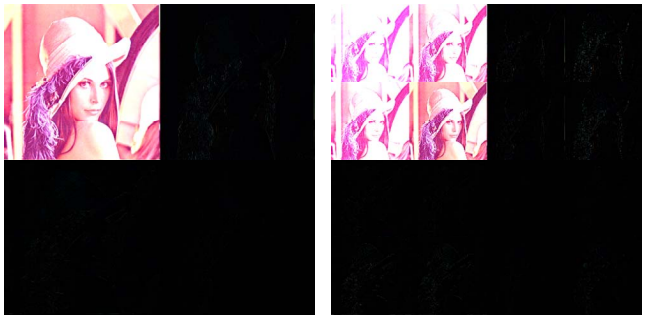


Figure 3. Single level decomposition of Lena test image (a) Antonini 9/7 wavelet transform (b) Unbalanced GHM multiwavelet transform.

A. *Global Error Energy Minimization technique*

The Global Error Energy Minimization (GEEM) technique [14] calculates a disparity vector for each pixel. It searches for the best match for each pixel in the correspondence search area of the other image using an error minimization criterion. For RGB images, the error energy criteria can be defined as:

$$Er_{en}(i,j,w_x,w_y) = \frac{1}{3}\sum_{k=1}^{3}\left(I_1(i+w_x,j+w_y,k)-I_2(i,j,k)\right)^2$$
$$-d_x \le w_x \le d_x \quad and \quad -d_y \le w_y \le d_y$$
$$i = 1, \ldots m \quad and \quad j = 1, \cdots n \qquad (4)$$

where I_1 and I_2 are the two input images, $Er_{en}(i,j,w_x,w_y)$ is the difference energy of the pixel $I_2(i,j)$ and pixel $I_1(i+w_x,j+w_y)$, d_x is the maximum displacement around the pixel in x direction, d_y is the maximum displacement around the pixel in y direction, m and n are the image size and k represents the three components of an RGB image.

In order for the GEEM algorithm to determine the disparity vector for each pixel in the current view, it first calculates Er_{en} of each pixel with all the pixels in its search area in the correspondence image. For every disparity vector (w_x,w_y) in the disparity search area, error energy is calculated using Equation 4 and placed into a matrix. Each of the resulting error energy matrices is first filtered using an average filter to decrease the number of incorrect matches [15]. The disparity index of each pixel is then determined by finding the disparity index of the matrix, which contains the minimum error energy for that pixel. In order to increase the reliability of the disparity vectors around the object boundaries, which is the result of object occlusion in images, the generated disparity map undergoes a thresholding procedure as it follows:

$$\tilde{d}(i,j) = \begin{cases} d(i,j) & Er_{en}(i,j) \le \alpha \times Mean(Er_{en}) \\ 0 & Er_{en}(i,j) > \alpha \times Mean(Er_{en}) \end{cases} \qquad (5)$$

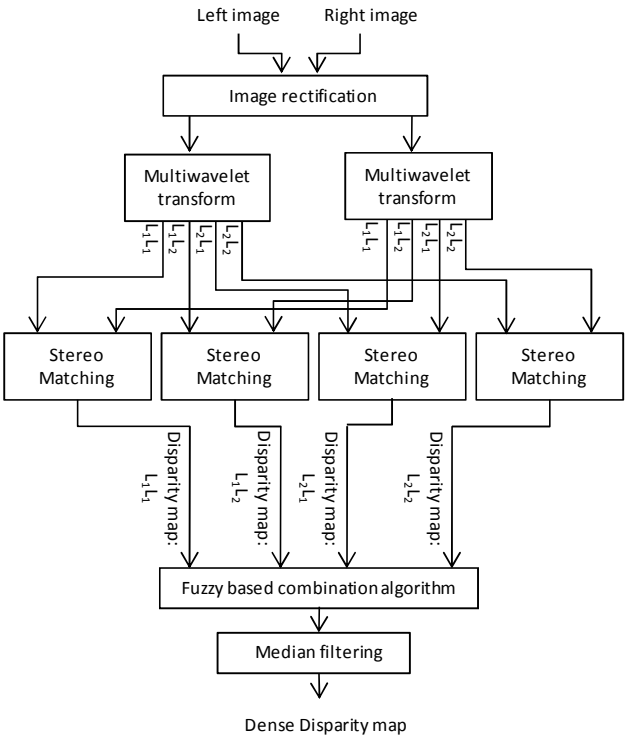


Figure 4. Block diagram of multiwavelet-based stereo matching technique using the global energy minimization algorithm.

L_1L_1	L_1L_2	L_1H_1	L_1H_2
L_2L_1	L_2L_2	L_2H_1	L_2H_2
H_1L_1	H_1L_2	H_1H_1	H_1H_2
H_2L_1	H_2L_2	H_2H_1	H_2H_2

Figure 5. One level of decomposition for a 2D multiwavelet.

(a) "Art" stereo test images

(b) "Cones" stereo test images

Figure 6. The left image and the ground truth of the (a) "Art", and (b) "Cones" stereo test images.

where $\tilde{d}(i,j)$ is the processed disparity map, $d(i,j)$ is the disparity map, α is a tolerance reliability factor and $Er_{en}(i,j)$ is the minimum error energy of the pixel (i,j) calculated and selected in the previous stage. Finally a median filter is applied to the processed disparity map $\tilde{d}(i,j)$, to further smooth the resulting disparity map.

IV. Simulation Results

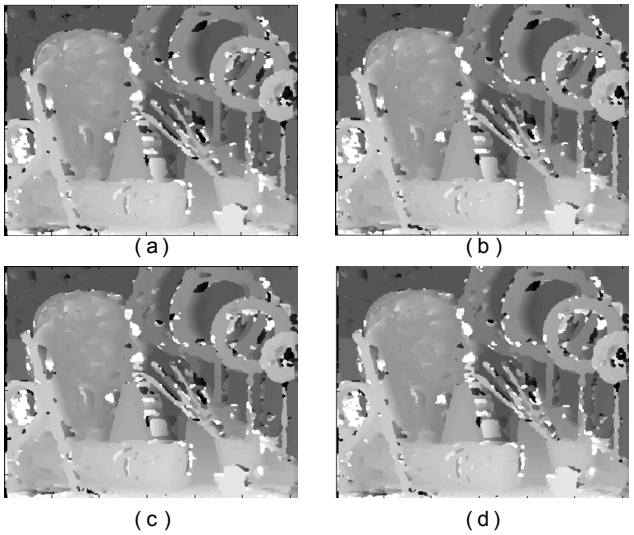
In order to evaluate the performance of the proposed technique, the multiwavelet-based GEEM algorithm and a similar wavelet-based technique were applied to the "Art" and "Cones" stereo test images from the Middlebury stereo database [16]. Figures 6(a) and 6(b) show the left image and the ground truth of the "Art" and "Cones" test images, respectively. The experimental results were generated using the GHM unbalanced multiwavelet and the Antonini 9/7 scalar wavelet. Figures 7(a) to 7(h) show the resulting disparity maps using the multiwavelet subbands L_1L_1, L_1L_2, L_2L_1 and L_2L_2 for the "Art" and "Cones" stereo test images. In order to give a visual comparison, the resulting disparity maps for the "Art" and "Cones" stereo test images, obtained for both the proposed multiwavelet-based algorithm and a similar wavelet-based algorithm, are illustrated in figures 8(a) to 8(d). In these figures areas with intensity zero represent unreliable disparities. From Figure 8, it is obvious that the disparity map of the multiwavelet-based algorithm is more accurate and smoother than that of the wavelet-based technique. This can be explained by the fact that basebands of the multiwavelet transformed images carry different spectral content of the input images, which enables the global error energy minimization algorithm to generate more reliable matches from the four combined multiwavelet basebands than from a single scalar wavelet baseband.

V. Conclusion

In this paper a new multiwavelet-based stereo matching technique using a global energy minimization algorithm was presented. An unbalanced multiwavelet transform with multiplicity of 2 decomposes the input stereo images into a number of subbands. The resulting four basebands of the two views were then used to generate the disparity map using the global error energy minimization algorithm described in this paper. The resulting four disparity maps were then combined using a Fuzzy algorithm to generate a dense disparity map. Results show that the proposed technique produces a disparity map with significantly less mismatch errors compared to the scalar wavelet-based algorithm.

Acknowledgment

This work was supported by EPSRC under the EP/G029423/1 grant.



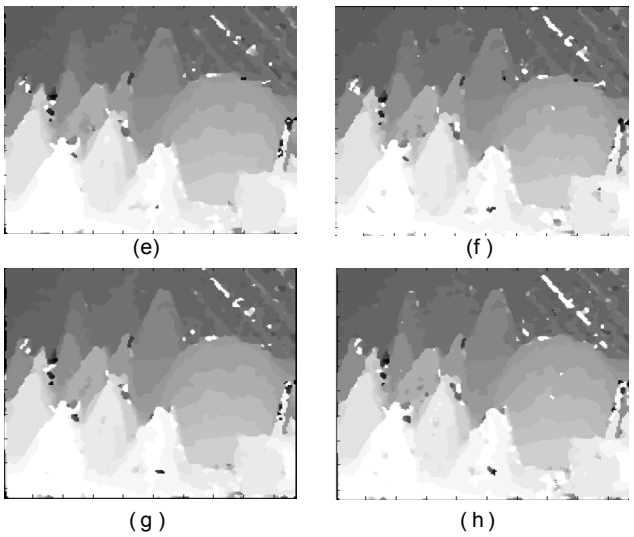


Figure 7. Disparity map using the multiwavelet basebands of "Art" stereo test image, a) L_1L_1 , b) L_1L_2 , c) L_2L_1 and d) L_2L_2 and "Cones" stereo test image, e) L_1L_1 , f) L_1L_2 , g) L_2L_1 and h) L_2L_2 .

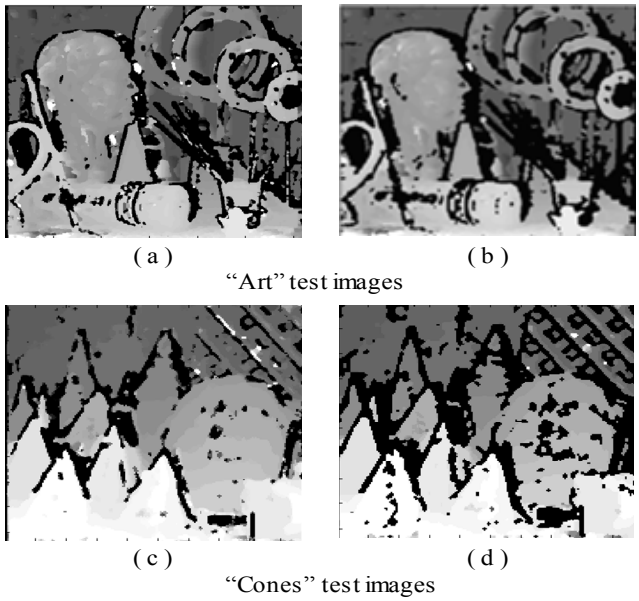


Figure 8. Disparity map using "Art" stereo test image, a) the proposed multiwavelet-based algorithm and b) the wavelet-based algorithm and "Cones" stereo test image, c) the proposed multiwavelet-based algorithm and d) the wavelet-based algorithm.

REFERENCES

[1] D. Scharstein and R. Szeliski, “A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms,” International Journal of Computer Vision, vol. 47, pp. 7-42, April 2002.

[2] 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), 2001, pp. 30-36 , doi: 10.1109/SMBV.2001.988760.

[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, March 2004.

[4] S. Yoon, S. K. Park, S. Kang and Y. Keun Kwak, “Fast correlation-based stereo matching with the reduction of systematic errors,”Pattern Recognition Letters, vol. 26 , pp. 2221-2231, October 2005.

[5] 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.

[6] 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.

[7] 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.

[8] Q. Yang, L. Wang, R. Yang, S Wang, M. Liao, and D. Nister, “Real-time global stereo matching using hierarchical belief propagation", In Proceedings of the British Machine Vision Conference (BMVC), 2006.

[9] I. Choi and H. Jeong, “Fast Belief Propagation for Real-time Stereo Matching,” International Conference on Advanced Communication Tech., ICACT 2009, pp. 1175-1179, Feburary 2009.

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

[11] V. Strela and A.T. Walden, “Signal and image denoising via wavelet thresholding: orthogonal and biorthogonal, scalar and multiple wavelet transforms,” In Nonlinear and Nonstationary Signal Processing, pp. 124-157, 1998.

[12] 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.

[13] V. Strela, “Multiwavelets: threory and applications,” PhD thesis, MIT, 1996.

[14] B. B. Alagoz, “Obtaning depth maps from colour images by region based stereo matching algorithms,” OncuBilim Algorithm and System Labs, vol. 08, Art.No:04, 2008.

[15] R.C. Gonzalez, R.E. Woods, and S.L. Eddins, Digital Image Processing second edition, Prentice Hall, pp. 75-142, 2002.

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