

# An Area-Based Stereo Matching Using Adaptive Search Range and Window Size<sup>1</sup>

Han-Suh Koo and Chang-Sung Jeong

Department of Electronics Engineering, Korea University  
1-5ka, Anam-dong, Sungbuk-ku, Seoul 136-701, Korea  
E-mail: {esprit@snoopy, csjeong@chalie}.korea.ac.kr

**Abstract.** In area-based stereo matching algorithm, the proper determination of search range and window size are two important factors to improve the overall performance of the algorithm. In this paper we present a novel technique for area-based stereo matching algorithm which provides more accurate and error-prone matching capabilities by using adaptive search range and window size. We propose two new strategies (1) for determining search range adaptively from the disparity map and multiresolutional images of region segments obtained by applying feature-based algorithm, and (2) for changing the window size adaptively according to the edge information derived from the wavelet transform such that the combination of two adaptive methods in search range and window size greatly enhances accuracy while reducing errors. We test our matching algorithms for various types of images, and shall show the outperformance of our stereo matching algorithm.

**Keywords.** Visualization, Stereo Vision, Image Processing

## 1 Introduction

The stereo matching algorithm is a technique that analyses two or more images captured at diverse view points in order to find positions in real 3D space for the pixels of 2D image. The stereo matching methods have been used in various fields such as drawing the topographical map from aerial photograph and finding the depth information of objects in machine vision system. Nowadays, optical motion capture techniques using stereo matching algorithms are being developed for visual applications such as virtual reality or 3D graphics.

Stereo matching algorithms can be generally classified into two methods: feature-based and area-based ones. The feature-based method matches the feature elements between two images, and uses the interpolation to obtain the disparity information for the pixels other than the feature elements, while the area-based method performs matching between pixels in two images by calculating the correlations of the pixels residing in the search window. Area-based method cannot match feature element with more accuracy than feature-based method even though it can make more dense disparity map [1,2]. Moreover, it has more possibility of error in the area of insufficient texture information or depth discontinuities. In this paper we present a novel technique for area-based stereo

---

<sup>1</sup> This work was supported by KISTEP and BK21 Project.

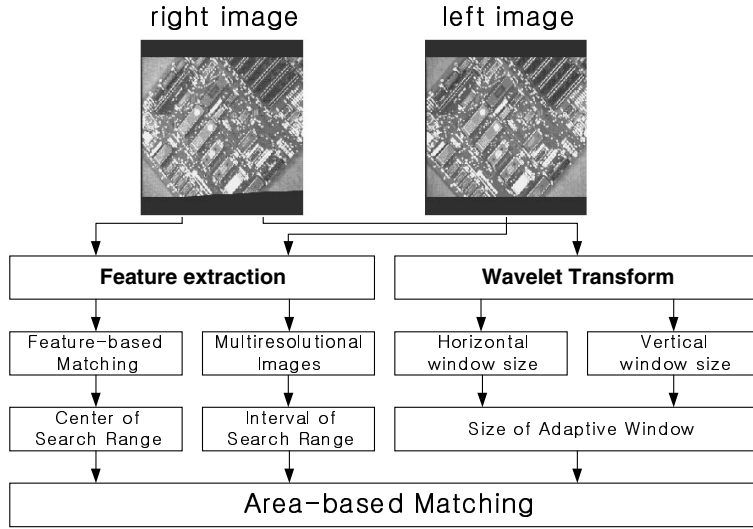
matching algorithm which provides more accurate and error-prone matching capabilities by using adaptive search range and window size. We propose two new strategies (1) for determining search range adaptively from the disparity map and multiresolutional images of region segments obtained by applying feature-based algorithm, and (2) for changing the window size adaptively according to the edge information derived from the wavelet transform such that the combination of two adaptive methods in search range and window size greatly enhance accuracy while reducing errors. We test our matching algorithms for various types of images, and shall show the outperformance of our stereo matching algorithm.

The paper is organized as follows. In section 2, we briefly describe about the existing stereo matching algorithms and review the merits and defects of each approach. In section 3, we present our area-based stereo matching algorithm. In section 4, we explain the experimental results of our matching algorithm. Finally, in section 5, we give a conclusion.

## 2 Stereo Matching Methods

In this section we classify stereo matching methods into area-based and feature-based matching methods and briefly describe about them. Area-based method uses the correlation of intensity between patches of two images to find, for each pixel in one image, its corresponding pixel in the other image based on the assumption that the disparities for the pixels in the neighborhood region of one pixel are nearly constant. It can match each pixel in the right image to its corresponding pixel in the left image with maximum correlation or minimum SSD (Sum of squared difference) between two windows in the left and right images[3,4]. However, most of the area-based methods are sensitive to noise and variation between images, and have a high probability of error in the area where there are insufficient texture information or depth discontinuities. Generally, if the window is too small, it gives a poor disparity estimate, because the signal(intensity variation) to noise ratio is low. If, on the other hand, the window is too large, then the position of maximum correlation or minimum SSD may not represent correct matching due to the different projective distortions in the left and right images. To overcome this defect, Kanade and Okutomi[5] proposed an adaptive window method in an iterative stereo matching algorithm. They selected the size and shape of a window by iteratively updating the disparity estimate in order to obtain the least uncertainty for each pixel of an image. Also, Sun[6] proposed a multiresolutional stereo algorithm which makes use of dynamic programming technique. Previous works with adaptive window require so much processing time because of the iterative routines, and restrict the matching window to a perfect square. Our algorithm provides more simple and flexible method for selecting the window adaptively.

Feature-based method extracts primitives such as corners, junctions, edges or curve segments as candidate features for matching by finding zero-crossings or gradient peaks. Many researchers use line segments frequently in the feature-based stereo matching algorithms[7,8] because they are abundant and easy to



**Fig. 1.** Process flow diagram of our algorithm

find in an image, and their characteristics such as length, gradient and orientation are easy to match compared with the other primitives. Hoff and Ahuja[9] integrate feature matching and interpolation by using local parametric fitting in a hypothesis-test manner. Their algorithm is computationally expensive, because it considers too many matching possibilities when fitting for edges by Hough transform. Maripane and Trivedi[10] uses multiple primitives such as regions, lines, and edges in order to reduce the ambiguity among the features. Feature-based method is more accurate and less sensitive to photometric variations than area-based method since feature elements are used for matching. However, it requires additional interpolation for sparse feature elements, and has some difficulties in deriving a general algorithm applicable to all type of images.

### 3 Algorithms

In this section, we present a new area-based stereo matching algorithm using adaptive search range and adaptive window. Our algorithm consists of three stages. In the first stage, we compute, for each pixel in the right image, its search range in the left image. The search range in the left image is a one dimensional range along  $x$  axis which is defined by its center and interval. The center of the search range is obtained by executing feature-based matching with the region segments as feature elements, and the interval of the search range by investigating the region segments in multiresolutional images. In the second stage, we determine, for each pixel in the right image, its search window based on the edge information derived from the wavelet transform. In the third stage,

we try to match each pixel in the right image with the pixels in the search range by moving the window along the pixels in the search range, and determine the disparity map by finding the pixel in the search range with the maximum correlation between two windows in the left and right images respectively. The diagram of the overall process flow of our algorithm is illustrated in figure 1. We shall describe about each stage more in detail in the following subsequent subsections.

### 3.1 Adaptive Search Range

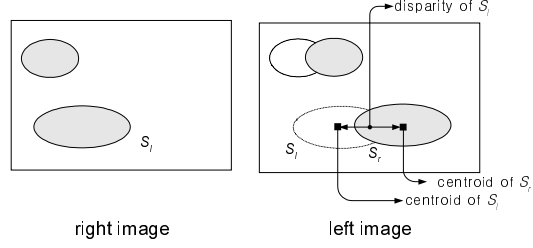
In this subsection we shall show how to find the center and interval of the search range.

**The center of the search range:** The center of the search range is determined from the disparity map by executing feature-based matching with region segments as features. Most of the existing feature-based algorithms use edges as feature elements since they can obtain a relatively precise matching result due to the abundant appearance of edges in an image. Instead, our algorithm adopts region segments as feature elements since the smaller number of regions segments in an image, along with the region-based primitive's high discrimination power, reduces the number of false targets and increases the accuracy of matching. Moreover, the region segments can provide information about the search range for larger portion of the image when executing area-based matching algorithm.

**1). Feature extraction:** Region segments are extracted by finding threshold values from gray-level histogram[11]. Three threshold values are calculated by executing threshold-based segmentation method iteratively, and the region segments are divided into four groups according to the threshold values. Then, a series of high, medium, and low resolution images are generated by reductions in size by a factor of two using a Gaussian convolution filter for each group. Each group is represented by a binary map where each pixel indicates whether it belongs to the group or not. The number of groups is determined from the experiment so that the size of region segments may not be too large or too small. Although thresholding is unstable against photometric variations, the results in this step are acceptable, because it helps to find candidate matches but not correct matches.

**2). Morphology filter:** There may exist excessively small region segments generated during the feature extraction. Therefore, we apply the opening operation in morphology filter to the low resolution images in four groups in order to prevent the errors which may arise from those small region segments.

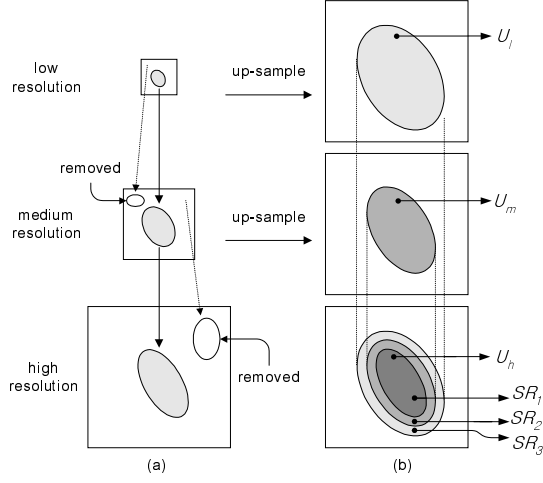
**3). Feature-based matching:** The disparity map is obtained by executing feature-based matching with region segments as feature elements. However, it is not the final disparity map, but the one which shall be used to compute the center of the search range and hence to find the more accurate disparity map by area-based matching. Therefore, we do not need the exact disparity map at this point. Thus, in our feature-based matching, low resolution images in four groups are used in order to ignore the small region segments and speed up the



**Fig. 2.** Disparity of region segment

matching process. We up-sample the low resolution images to the image with original size in order to calculate the disparities in actual distance. After up-sampling, we match region segments using geometric properties such as position, size, eccentricity, and orientation, and find, for each pixel in the right image, its disparity and center of the search range as follows: Let  $S_l$  be a region segment in the right image and  $S_r$  be its matched region segment in the left image. If a pixel belongs to any region segment  $S_l$ , its disparity is defined as the difference in  $x$  value between the centroids of  $S_l$  and  $S_r$ ; otherwise its disparity is set to 0. (See figure 2.) Then, for each pixel in the right image, the center of its search range is defined as the pixel in the left image which is apart from itself by its disparity. The center of the search range computed from the disparity map allows the more exact area-based matching through the enhanced selection of the search range.

**The Interval of Search Range** The interval of search range for a pixel in the right image is determined according to its closeness to the boundary of the region segment. The more a pixel is close to the boundary of the region segment to which it belongs, the larger the interval of its search range becomes to reflect the abrupt change in the boundary of the region segments. The closeness of a pixel to the boundary of the region segment is determined by investigating the inclusion property in the region segments of multiresolutional images as follows: First, we remove region segments which do not appear in low resolution image in order to eliminate small region segments which have insignificant influence but might incur errors. (See Figure 3.) Then, we up-sample each of multiresolutional images to the original high resolution size. For a region segment  $R_i$ , let  $U_l$ ,  $U_m$ , and  $U_h$  be its up-sampled region segments from low, medium, and high resolution images respectively. Then, the relation  $U_l \subset U_m \subset U_h$  usually holds, since during up-sampling, one pixel is expanded to its corresponding four neighbor pixels in the next higher resolution image. For the region segment  $R_i$ , we can classify three subregions:  $SR_1 = U_h$ ,  $SR_2 = U_m - U_h$ ,  $SR_3 = U_l - U_m$ .  $SR_1$  is the innermost subregion which corresponds to the original region segment, and  $SR_3$  is the outermost subregion.  $SR_2$  is the subregion which lies between  $SR_1$  and  $SR_3$ . Let  $I_i$  be the interval of the search range for a pixel  $p$  which lies in  $SR_i$ . Then, we determine, for each pixel  $p$ , its interval of the search range according to in which subregion it lies while satisfying  $I_1 < I_2 < I_3$ . In other words, the interval of the search range for a pixel becomes larger if it gets more close to the outer boundary of the region segments to which it belongs, since the edges in

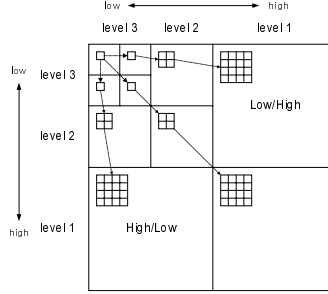


**Fig. 3.** The Interval of Search Range (a) Elimination of small region segment (b) Up-sampled regions  $U_l, U_m, U_h$  and subregions  $SR_1, SR_2, SR_3$

the boundary the region segments are liable to change abruptly. If pixels are not detected in any up-sampled region segments, the interval of its search range is set to maximum since no disparity information is acquired from region segment matching process.

### 3.2 Adaptive Search Window

The proper size of matching window is one of the important factors to improve the performance of the area-based matching algorithm. A large window is proper to flat surface, but it blurs disparity edge. On the other hand, a small window preserves a disparity edge well, but it produces much error in flat surface. For this reason, a window size need to be selected adaptively depending on edge information. We propose a new strategy to select an appropriate window by evaluating edge information derived from wavelet transform[12]. Edge existence can be determined by applying a threshold to the detail coefficients of Low/High and High/Low frequency band since high frequency component appears in the detail coefficients. (See figure 4.) Therefore the vertical and horizontal sizes of the matching window are determined as follows: First, we apply 3-level 2D wavelet transform to the right image, and then determine whether a pixel is an edge component or not at each level of 2D wavelet transform by comparing the detail coefficient with the threshold value which is obtained from the experiment. The edge strength of a pixel can be regarded as more strong if the pixel is detected as edge component at lower levels. If a pixel is detected as edge component at all levels, the edge strength becomes the largest at the pixel, and the size of the search window is set to minimum value in order to capture the disparity edge more precisely. If a pixel is not detected as edge component at higher levels, the edge strength becomes smaller at the pixel, and the size of the search window increases. If an edge is not detected at any level, the size of search window is



**Fig. 4.** Detail coefficients of Wavelet Transform

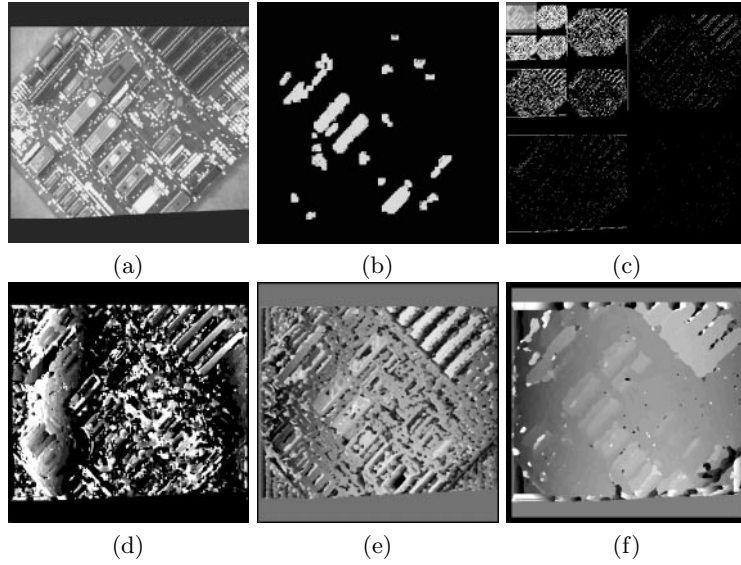
set to to maximum value. The horizontal and vertical sizes of search window are determined separately by using the strategy mentioned above for the Low/High frequency and the High/Low frequency bands respectively to adaptively reflect the edge information in the selection of window size.

### 3.3 Area-Based Stereo Matching

After computing search range and window size for each pixel in the right image in the first and second stages respectively, we now attempt to execute area-based matching in order to find the final disparity map in the third stage. We try to match each pixel in the right image with the pixels residing in its search range of the left image using the appropriate window selected during the second stage, and compute the disparity map by finding the pixel in the left image with the maximum correlation between two image pairs in the corresponding windows. Finally, we perform the refinement of the disparity map by applying median filter in order to satisfy the continuity constraint[6]. The continuity constraint need to be satisfied since a disparity of single pixel which is excessively different from its neighborhood can be regarded as an error rather than correct value[3].

## 4 Experimental Results

We tested our stereo-matching algorithms presented in this paper for three real stereo images: apple, fruit and pentagon. Figure 5 illustrates stereo matching analysis for apple stereo image. Figure 5(a) shows a  $512 \times 512$  right image for apple, and figure 5(b) shows the search range map where each pixel is represented by intensity according to its search interval. The brighter pixel represents the shorter search interval. Figure 5(c) shows the edge component map obtained by applying thresholding to the detail coefficients of wavelet transform for the right image, where the bright pixels at each level represent the existence of edge components. We used the daubechies basis vectors for wavelet transform. Based on the edge information, we determined the horizontal and vertical sizes of the search window. Figure 5(d) through (f) compare the disparity maps under various conditions. In the disparity map, the brighter pixel is the more close one to the camera. Figure 5(d) shows the disparity map obtained by applying



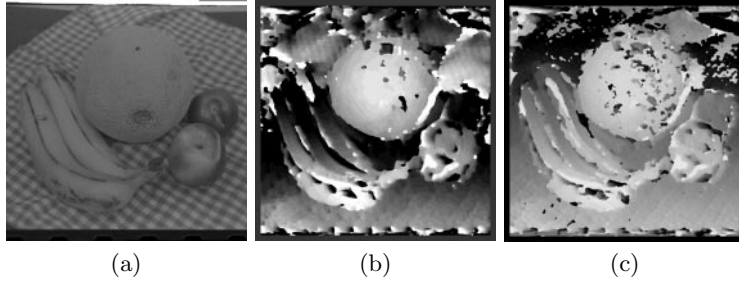
**Fig. 5.** Stereo matching analysis for apple image (a) Right image (b) Search range map (c) Edge component map obtained from wavelet transform (d) Disparity map of area-based matching algorithm using search range of 6 pixels and  $9 \times 9$  search window (e) Disparity map of area-based matching algorithm using adaptive search range algorithm but without adaptive window algorithm (f) Disparity map of our area-based matching algorithm using adaptive search range and window

general area-based matching algorithm with search range and the search window fixed to 6 pixels and  $9 \times 9$  respectively, and figure 5(e) shows the disparity map obtained by area-based matching algorithm with adaptive search range but with  $3 \times 3$  fixed window size. Figure 5(f) is the disparity map obtained by using our area-based matching algorithm using adaptive search range and window. In our algorithm, the search range varies from 6 pixels to 42 pixels, and the window size from 3 pixels to 21 pixels adaptively. Compared with figure 5(e) and (f), figure 5(d) shows more errors in disparities around the center of the image. Figure 5(e) clearly shows the improvement over figure 5(d) by using adaptive search range. However, figure 5(e) still has some errors around the center since  $3 \times 3$  search window is too small to get the accurate result for the flat surface around the center. Figure 5(f) shows the further improvement over figure 5(e) by making use of adaptive search range and window together.

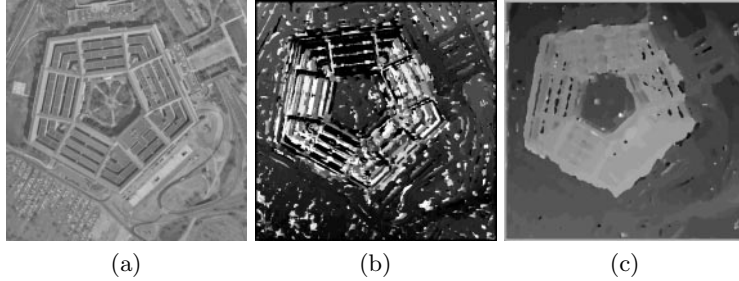
Figure 6 illustrates stereo matching analysis for  $512 \times 512$  fruit image shown in figure 6(a). Figure 6(b) shows the disparity map obtained by area-based matching algorithm with adaptive window and fixed search range of 20 pixels. Compared with figure 6(c), figure 6(b) shows more errors in a tablecloth and an apple behind. This fact tells that too large and uniform search range is not adequate for distant objects in the image.

Figure 7 illustrates stereo matching analysis for  $512 \times 512$  pentagon image shown in figure 7(a). We compare the disparity map for two cases: one when ap-





**Fig. 6.** Stereo matching analysis for fruit image (a) Right image (b) Disparity map of area-based matching algorithm using adaptive window algorithm but without adaptive search range algorithm (c) Disparity map of our area-based matching algorithm using adaptive search range and window



**Fig. 7.** Stereo matching analysis for pentagon image (a) Right image (b) Disparity map of area-based matching algorithm using search range of 20 pixels and  $9 \times 9$  window (c) Disparity map of our area-based matching algorithm using adaptive search range and window

plying area-based matching algorithm using fixed search range and window size, and the other using adaptive search range and window size as in our matching algorithm. Figure 7(b) shows the disparity map when using fixed search range of 20 pixels and fixed  $9 \times 9$  window. In figure 7(b) we can easily see that some errors occurs in upper portion of the pentagon building due to the large window size. Figure 7(c) shows the clear improvement over figure 7(b) by using adaptive search range and window size. The comparison of disparity maps shows that the combination of adaptive search range and window as in our matching algorithm greatly enhance the matching capability by appropriately changing the search range and window size according to the various surface conditions.

## 5 Conclusion

In this paper we have presented the new technique for area-based stereo matching algorithm which provides more accurate and error-prone matching capabilities by using adaptive search range and window size. We have proposed two new strategies for finding adaptive search range and window size respectively.

The center of the search range is obtained by executing feature-based matching with the region segments as feature elements. Adopting region segments as feature elements allows reduction in the number of false targets and increase in

the accuracy of matching due to the smaller number of region segments in an image, along with the region-based primitive's high discrimination power while providing information about the search range for larger portion of the image for area-based matching. The center of the search range computed from the disparity map using feature-based matching allows the significant enhancement in the selection of the search range. The interval of the search range is determined by investigating the up-sampled multiresolution images. Classifying the region segments in multiresolutional images into several subregions provides the elaborated criteria to the closeness to the boundary of the region segment and allows the more refined care for the abrupt change in the edge boundary. The appropriate window is selected adaptively according to the edge information derived from the wavelet transform by reflecting the edge information appearing in the detail coefficients of the wavelet transform. We have tested our matching algorithms for various types of images, and have shown that the combination of two adaptive methods in search range and window size greatly improves the performance of area-based matching algorithm.

## References

1. S. D. Cochran, G. Medioni, "3-D Surface Description From Binocular Stereo," IEEE Trans. PAMI, vol. 14, no. 10, pp. 981-994, Oct. 1992.
2. G. Wei, W. Brauer, and G. Hirzinger, "Intensity- and Gradient-Based Stereo Matching Using Hierarchical Gaussian Basis Functions," IEEE Trans. PAMI, vol. 20, no. 11, pp. 1143-1160, Nov. 1998.
3. O. Faugeras, "Three-Dimensional Computer Vision; A Geometric Viewpoint," pp. 189-196, Massachusetts Institute of Technology, 1993.
4. M. Okutomi and T. Kanade, "A Multiple-Baseline Stereo," IEEE Trans. PAMI, vol. 15, no. 4, pp. 353-363, Apr. 1993.
5. T. Kanade and M. Okutomi, "A Stereo Matching Algorithm with an Adaptive Window: Theory and Experiment," IEEE Trans. PAMI, vol. 16, no. 9, pp. 920-932, Sep. 1994.
6. C. Sun, "A Fast Stereo Matching Method," pp. 95-100, Digital Image Computing: Techniques and Applications, Massey University, Auckland, New Zealand, Dec. 1997.
7. K. L. Boyer and A. C. Kak, "Structural Stereopsis for 3-D Vision," IEEE Trans. PAMI, vol. 10, no. 2, pp. 144-166, Mar. 1988.
8. Shing-Huan Lee and Jin-Jang Leou, "A dynamic programming approach to line segment matching in stereo vision," Pattern Recognition, vol. 27, no. 8, pp. 961-986, 1994.
9. W. Hoff and N. Ahuja, "Surface From Stereo: Integrating Feature Matching, Disparity Estimation and Contour Detection," IEEE Trans. PAMI, vol. 11, no. 2, pp. 121-136, Feb. 1989.
10. S. B. Maripane and M. M. Trivedi, "Multi-Primitive Hierarchical(MPH) Stereo Analysis," IEEE Trans. PAMI, vol. 16, no. 3, pp. 227-240, Mar. 1994.
11. J. R. Parker, "Algorithms for Image Processing and Computer Vision," pp. 119-120, John Wiley & Sons, 1997.
12. S. E. Umbaugh, "Computer Vision and Image Processing," pp. 125-130, Prentice Hall, 1998.