Generation of a Disparity Map Using Piecewise Linear Transformation

HYEJIN KIM, JAEWAN CHOI, HOWOOK CHANG, KIYUN YU
School of Civil, Urban & Geo-System Engineering
Seoul National University
Shinrim-Dong, Gwanak-Ku, Seoul
KOREA

Abstract: - The generation of a disparity map usually requires a pair of precisely rectified stereo images, which implies that the images have epipolar geometry. In many practical cases, it is not easy to obtain a rectified stereo pair without using specialized stereo camera systems. We have proposed a new approach for generating a disparity map from a random stereo pair, called a segment- based piecewise linear transformation. The basic concept relies on the fact that displacements from piecewise image to image registration approximate to the disparity. Using this approach, segmentation of the left image is carried out first, followed by extraction of conjugate points from the stereo pair. Finally, a set of linear transformation functions is determined using least squares method. By applying these functions, the displacement for each pixel is calculated to allow for the generation of a disparity map. To estimate the quality of the resulting disparity map, two stereo anaglyphs, one from the disparity map and the other from the original stereo pair, were produced and compared visually. The results show that the disparity map works well on both uniform and slanted disparity surfaces. An advantage of this approach is that it does not necessarily require stereo rectification, and that it is applicable to any set of stereo images that are not in epipolar geometry.

Key-Words: - Disparity map, Piecewise linear transformation, Segmentation, Mean-shift algorithm, Conjugate points, SIFT detector, Anaglyph

1 Introduction

Generation of disparity maps is one of the most active research topics in computer vision. The disparity indicates the difference in location of the conjugate pixels from a pair of stereo images, and it is represented as an inverse depth. Generating a disparity map implies solving the many problems that occur in textureless regions, disparity discontinuous boundaries, and occluded areas. Therefore, many techniques have been developed, and the major part of previous related works has been devoted to local window based algorithms, which have estimated the disparity at a given pixel, only depending on its intensity within a finite neighboring window [8, 13]. Recently, global algorithms using various energy minimization techniques, such as graph cuts, have attracted much attention [1, 7, 9, 12]. Regardless of the algorithms introduced in previous research efforts, there has been the assumption that precise rectified stereo images are available. This implies that the images are in epipolar geometry. However, in many practical applications, it is not always easy to obtain a pair of rectified stereo images without employing specially designed camera systems. This hinders the wider applications of the previously proposed algorithms [10, 11].

Therefore, we have proposed a new approach consisting of three steps.

First, segmentation of the left image is carried out to generate a series of small segments. Then, extraction of the conjugate points from the stereo pair is conducted, and a set of linear transformation functions is determined using the least squares technique. Once these functions are produced, the displacement for each pixel is calculated after applying the functions to the small segments that have been generated. This generates a disparity map. To examine the quality of the disparity map, two stereo anaglyphs, one from the disparity map and the other from the original stereo pair, are produced and compared visually. The entire workflow is illustrated in Figure 1.

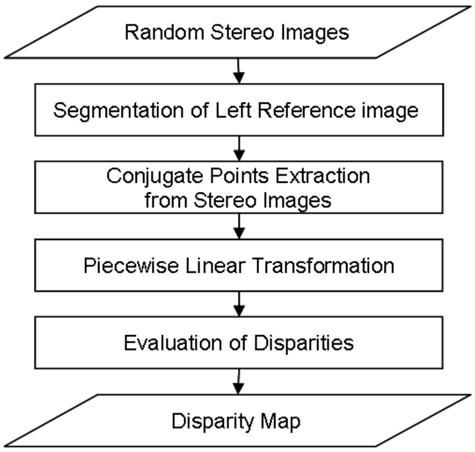

Fig.3 Flowchart of the proposed algorithm

2 Color Segmentation

Our approach is built on the assumption that large disparity discontinuities only occur at the boundaries of homogeneous color segments [7, 9]. For a qualified result, the number of segments must be handled carefully, because too many segments makes it difficult to obtain enough conjugate points, and too small number of segments result in ill merging between separate objects. A mean shift color segmentation algorithm was used in our implementation, as mean shift based segmentation is one of the most powerful and generally used segmentation techniques [6, 9].

3 Conjugate Point Extractions

To extract conjugate points from the left and right images, we detected feature points in two images using Scale-Invariant Feature Transform (SIFT) [3, 4, 5]. In the SIFT technique, the local image gradient orientation around the detected features was used as a feature descriptor. Two points with the most similar feature descriptors were selected as a corresponding pair. SIFT has some advantages over other matching algorithms for our experiments because it extracts a large number of key points and even small segments can include at least three matched key points as discussed in [3]. In addition, SIFT is invariant to scale change, translation, and rotation. It also yields the correct matching results despite illumination change and affine distortions. Moreover, it removes any key points located along the edges. Considering that the edges usually coincided with the segments boundaries, rejecting the key points around the edges has a favorable effect. Because points located along an edge are highly likely to be included in a wrong segment, then removing the key points along the edges enhances the accuracy of depth calculation in each segment.

4 Piecewise Linear Transformation

Throughout the procedures described above, a set of matched points for each segment was obtained for the piecewise linear transformation. Because the left image was the reference image, we estimated the linear transformational functions after registering the left image (x, y) to right image (x', y') using the least square method. The piecewise linear transformation function is expressed by the following equations :

$$x' = a + bx + cy, \quad y' = d + ex + fy \tag{1}$$

$$\begin{bmatrix} x'_1 \\ \vdots \\ x'_n \\ y'_1 \\ \vdots \\ y'_n \end{bmatrix}_k = \begin{bmatrix} 1 & x_1 & y_1 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_n & y_n & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_1 & y_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 1 & x_n & y_n \end{bmatrix}_k \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}_k \tag{2}$$

where k is the segment index, (x_i, y_i) is conjugate point coordinates of left image, (x_i', y_i') is conjugate point coordinates of right image, and n is the number of matched points within segment k

Equation (2) can be rewritten in brief :

$$L_k = A_k X_k \tag{3}$$

Therefore, the coefficients matrix, X, can be solved using the least squares method as follows :

$$X_k = (A_k^T A_k)^{-1} A_k^T L_k \tag{4}$$

Then, the displacement value, v, of each pixel in the entire image can be estimated using the transformation function of each segment :

$$v_{xy} = \sqrt{(x - x')^2 + (y - y')^2} \tag{5}$$

where v_{ij} is the displacement value at (x, y), (x, y) is the coordinates of left image, and (x', y') is the transformed coordinates.

Obtaining the affine transformation coefficients requires at least three points that are matched correctly in each segment. If a given segment contains fewer than three matched points, then we substitute the mean displacement values between the conjugate points for the displacement value in that segment to prevent a null-type value occurring.

5 Experimental Results

5.1 Data acquisition

To evaluate our approach, we took two stereo pair images of different scene using Nikon Coolpix 3700 digital camera, set to manual exposure and focus. The first scene taken was of a parking area, and the second scene taken was of a desk. The distance between the camera centers in each stereo pair was adjusted to be about 1/30 of the distance between the camera and the nearest object. The image size was 640 ×480 pixels and the color depth was 24bits.

5.2 Algorithm applications

We executed segmentation of the left image in each test pair. The number of segments in the left image of the first and second pairs was 42 and 30, respectively. Figure 2 shows the test images and segmentation results of these left images. Through extraction of the conjugate points, we obtained 422 and 413 matched points for each pair, respectively. The disparity maps were then generated using the piecewise linear transformation.

5.3 Result evaluation

Figure 3 shows the results of our experiment. Figures 3(a) and 3(d) show the disparity maps generated from the two test pairs, 3(b) and 3(e) show the anaglyphs from the disparity maps that we produced and their corresponding left images, and 3(c) and 3(f) show the reference anaglyph images obtained from the original stereo pairs. Our proposed algorithm generated the disparity relatively well within the entire scene, not only in the uniform plane, but also in the slanted disparity surfaces, such as along the sides of cars, along the ground, and around the surface of the desk. In addition, the discontinuous boundaries were expressed clearly.

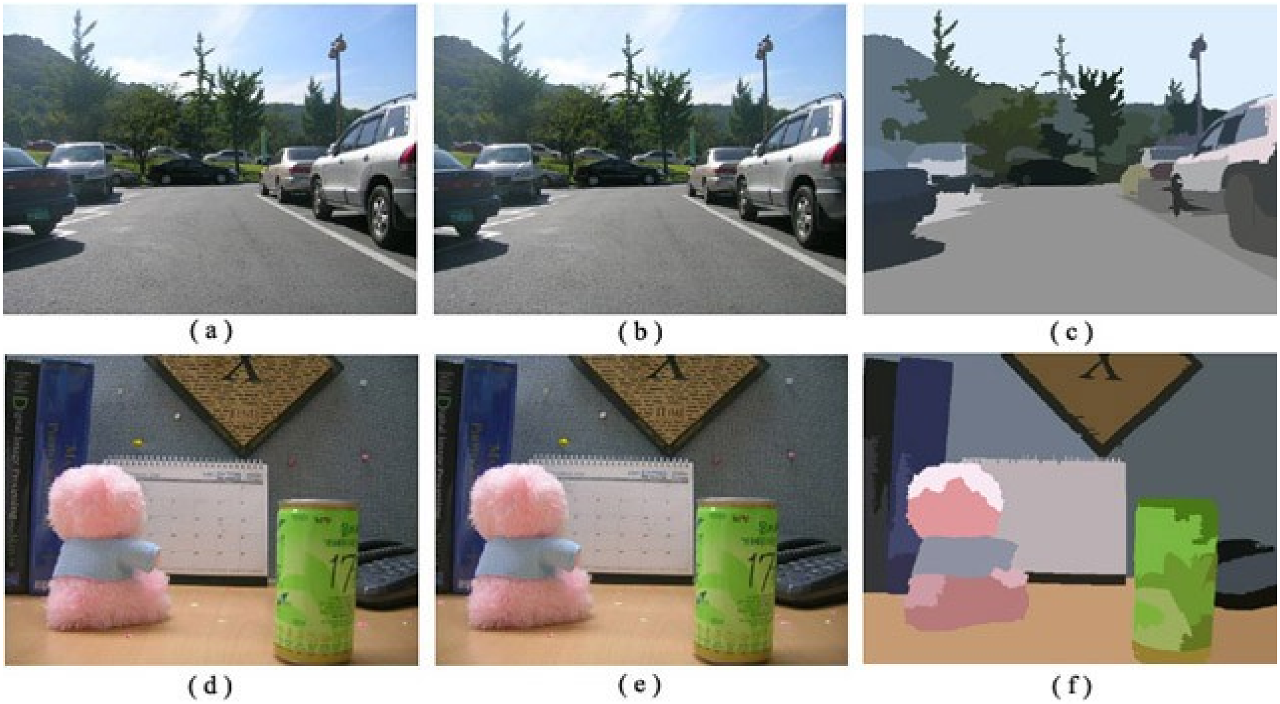

Fig. 3. Images (a), (b), and (c) show the left, right, and segmentation images of the first test pair ; images (d), (e), and (f) show the left, right, and segmentation images of the second test pair.
1

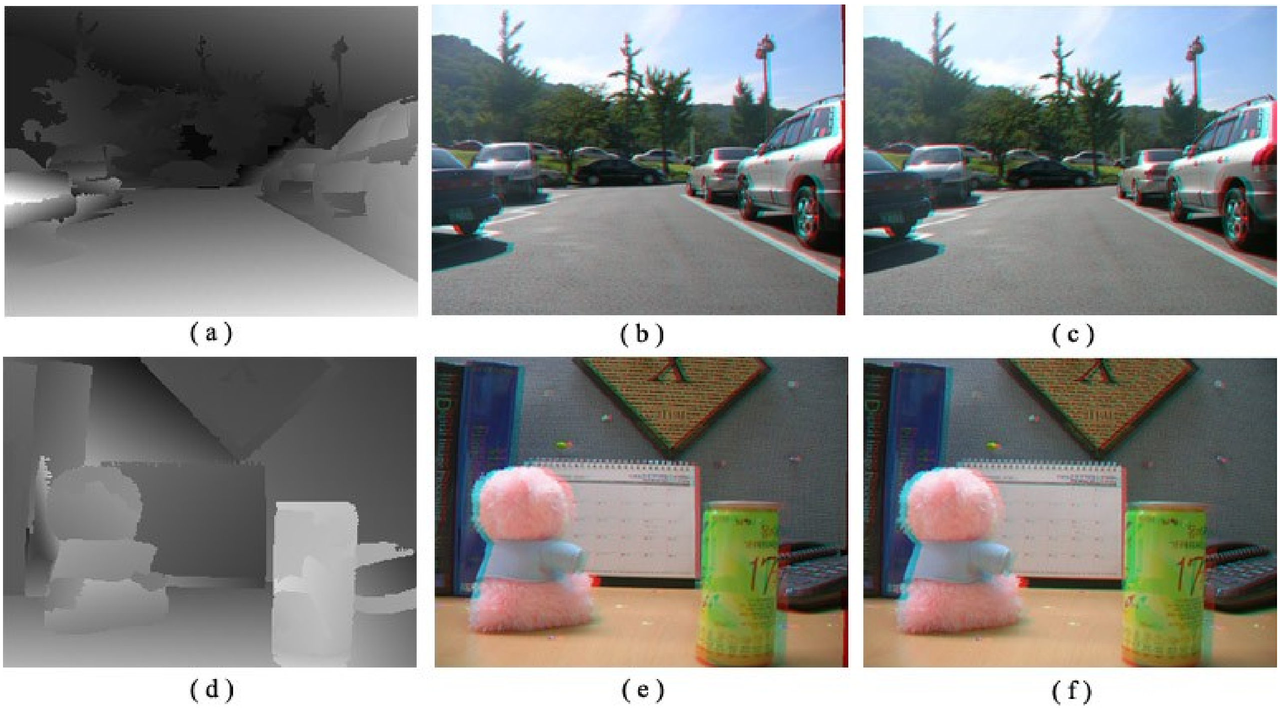

Fig. 3 Images (a) and (d) show the disparity maps for each test pair, images (b) and (e) show the anaglyphs from the disparity maps, and images (c) and (f) show the reference anaglyphs.

However, in the areas that had indefinite boundaries, such as mountains and trees, or few corresponding points, such as side of the telephone, satisfactory results were not obtained. To evaluate the generated disparity map, we compared our results with the original stereo pair by viewing the 3D stereo anaglyphs. Through this evaluation, the anaglyphs constructed using our proposed approach could express depth as well as the reference anaglyphs do, especially the close objects.

6 Conclusions

We have proposed a new approach that enables the generation of a disparity map from a random stereo pair where the stereo pair is not necessarily required to be in epipolar geometry. This approach adopts the piecewise linear transformation for the generation of a disparity map. In experiments, the resulting disparity map was shown to be suitable for expressing the uniform disparity plane and the slanted disparity surfaces. The advantages of this approach are that it does not necessarily require stereo rectification, and that it is applicable to any set of stereo images that are not in epipolar geometry. Apart from these advantages, it's not appropriate to express a curved surface, and the areas that had indefinite boundaries.

Our proposed approach was affected by the quality of the segmentation and corresponding point pair extraction, so it is important to determine the number of segments and matched points properly. This is now the topic of our ongoing research.

Acknowledgments

The authors gratefully acknowledge financial support form Seoul Research and Business Development (10540).

References:

[1] Andreas Klaus, Mario Sormann and Konrad Karner, Segment Based Stereo Matching Using Belief Propagation and a Self Adapting Dissimilarity Measure, Pattern Recognition. ICPR 2006, Vol.3, 2006, pp. 15-18.

[2] Bogdan Georgescu, Point Matching under Large Image Deformations and Illumination Changes, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol.26, 2004, pp. 674-688.

[3] David G. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision, Vol.60, 2004, pp. 91-100.

[4] David G. Lowe, Local Feature View Clustering for 3D Object Recognition, IEEE Conference on Computer Vision and Pattern Recognition, Vol.1, 2001, pp. 682-688

[5] David G. Lowe, Object Recognition from Local Scale-Invariant Features, International Conference on Computer Vision, Vol.2, 1999, pp. 1150-1157

[6] Dorin Comaniciu and Peter Meer, Robust Analysis of Feature Spaces: Color Image Segmentation, Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1997, pp. 750-755.

[7] Hai Tao, Harpreet S. Sawhney and Rakesh Kumar, A Global Matching Framework for Stereo Computation, Eighth International Conference on Computer Vision, Vol.1, 2001, p. 532

[8] Kuk-Jin Yoon and In-So Kweon, Locally Adaptive Support-Weight Approach for Visual Correspondence Search, Computer Vision and Pattern Recognition, Vol.2, 2005, pp. 924-931.

[9] Li Hong and George Chen, Segment-based Stereo Matching Using Graph Cuts, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.1, 2004, pp. 74-81.

[10] Marc Pollefeys, Reinhard Koch and Luc Van Gool, A Simple and Efficient Rectification Method for General Motion, International Conference on Computer Vision, Vol.2, 1999, pp. 496-501.

[11] Pravin Bhat, Ke Colin Zheng, Noah Snavely and Aseem Agarwala, Piecewise Image Registration in the Presence of Multiple Large Motions, Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, 2006, pp. 2491-2497.

[12] Qingxiong Yang, Liang Wang, Ruigang Yang, Henrik Stewénius and David Nistér, Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling, Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol.2, 2006, pp. 2347-2354.

[13] Vladimir Kolmogorov and Ramin Zabith, Multi Camera Scene Reconstruction via Graph Cuts, Lecture Notes in Computer Science, Vol.2532, 2002, pp. 82-96.