

Stereo Matching via Disparity Estimation and Surface Modeling

Jong Dae Oh, Siwei Ma and C.-C. Jay Kuo

University of Southern California, Los Angeles, CA 90089-2564

(jongoh, siweima, cckuo)@usc.edu

Abstract

Two new techniques are proposed to improve stereo matching performance in this work. First, to address the disparity discontinuity problem in occluded regions, we present a disparity estimation procedure, which consists of two steps; namely, a greedy disparity filling algorithm and a least-squared-errors (LSE) fitting method. Second, it is observed that the existing fronto-parallel model with color segmentation is built upon the piecewise constant surface approximation, which is however not efficient in approximating slanted or curved objects. We use a piecewise linear surface model to represent 3-dimensional (3D) geometric structure for better surface modeling. The proposed stereo matching system with these two new components is evaluated with Middlebury data sets with excellent quantitative and qualitative results.

1. Introduction

The goal of stereo matching is to estimate the depth information by finding correspondence pairs from a pair of stereo images. Even though it is a classical problem that has been extensively studied for several decades, it is still one of the most challenging problems in the field of computer vision and attracts a large amount of attention. We have seen new interests in accurate stereo matching in recent years due to emerging applications such as 3-dimensional (3D) TV, image and video based rendering, and stereo video coding.

One of the main challenges in stereo matching is to handle the disparity estimation in occluded areas. Since occluded areas are only visible in one image of a stereo image pair, there is no sufficient information to find the correspondence between the two input images. A disparity estimation procedure is proposed to address this issue.

Furthermore, there are many different object shapes in the real world. The traditional fronto-parallel model with color segmentation is inherently built upon a piecewise constant surface model. However, this model does not work well for highly slanted and/or curved surfaces. To overcome this shortcoming, we consider the use of a piecewise linear

surface model to represent 3-dimensional (3D) geometric structure for better surface modeling.

The stereo matching system with proposed disparity estimation and surface modeling is evaluated with Middlebury data sets. Excellent performance is achieved in terms of both subjective qualitative evaluation and objective quantitative measures.

1.1. Related Previous Work

We refer to [11] for a general survey on the history and concepts of stereo matching. In the following, we provide a review of previous work on disparity estimation in occlusion regions and surface modeling in the context of stereo matching.

To detect occluded regions and estimate disparity values in these regions have been recently studied in [6, 12] by new graph models with the uniqueness or the visibility constraint. The uniqueness constraint assumes the one-to-one pixel correspondence between every stereo image pair. Kolmogorov and Zabih [6] used an infinity penalty to reflect the uniqueness constraint in their energy cost function. The visibility constraint assumes that the non-occluded region between a stereo image pair has at least one match. Thus, this constraint allows one-to-many pixel correspondence. Sun *et. al.* [12] incorporated the visibility constraint in their energy cost function by treating left and right images equally. Then, to assign disparity values in detected occluded regions, the disparity of the closest visible pixel to the left direction is used in the left image. However, one main weakness of these approaches is to enforce the same discrete disparity value in one segmented region. This is a direct consequence of the fronto-parallel model, which means that all surfaces are in parallel with the image plane of stereo cameras.

To improve the fronto-parallel model for slanted and/or curved objects that are close to the stereo camera pair, more sophisticated prior models and matching algorithms are needed. Some approaches have been proposed before, and they can be classified into two categories as detailed below.

The first one uses a continuous prior term in the energy

equation. Ogale and Aloimonos [10] proposed a method to handle a horizontally slanted surface, which relies on unequally sampled correspondences and the 3D uniqueness constraint. Tsin and Kanada [13] proposed a correlation-based prior to represent the geometric model. Li and Zucker [7] proposed a new prior term based on differential geometric constraints for slanted and curved surface. While these algorithms estimate the smoothly varying disparity value inside objects, they do not work well with disparity discontinuity regions along object boundaries.

The second one decomposes disparity estimation into two steps: disparity optimization and surface fitting. Birchfield and Tomasi [2] used an affine system to model a slanted surface. Based on an initial disparity map obtained by the multiway-cut algorithm, they partitioned an image into non-overlapping regions and assumed that each region is a planar surface with a set of affine parameters to represent the displacement function. A bicubic B-spline model and a tensor voting framework were proposed by Lin and Tomasi [8] and Mordohai and Medioni [9], respectively, to compute the 3D geometric surface, where the 3D geometric information of each pixel is estimated by the 3D geometric function of neighbor pixels. This approach estimates an accurate disparity map inside an object while maintains sharp and precise disparity discontinuity regions along object boundaries. The main problem with the second approach is the size of the surface region is the same as the 3D geometric information. If the entire object is modeled by a planar surface, it would be difficult to represent a curved surface region. On the other hand, if the 3D coordinates of each pixel are determined based on the 3D geometric information of neighbor pixels, the model should work well for different types of surfaces but the associated complexity would be too high and the obtained result could be too noisy.

1.2. Contributions and Paper Organization

The current graph models work well in occlusion detection and disparity estimation when occluded regions are small or narrow. However, for large or complex occluded regions, their performance degrades substantially. Thus, more advanced techniques are needed for disparity estimation in occluded regions. After determining unreliable pixels that fail the correspondence check, a new disparity estimation technique for unreliable pixels is proposed in this paper, which uses the binocular and monocular information sequentially to fill in disparity data in big and complex occluded regions. This is the first major contribution of this work.

Various surface models as reviewed in Section 1.1 do not yield good disparity results due to the size of a surface segment. To address the problem, we propose a piecewise *linear* surface model that consists of linear planar patches

that are grouped by disparity similarity with disparity variation restriction. Each patch has 3 parameters to represent the 3D geometric structure locally. Thus, this model maintains both simple and good approximation capability. This is the second major contribution of this work. The proposed stereo matching scheme is evaluated by Middlebury data sets quantitatively and qualitatively. The performance evaluation shows that the proposed algorithm is among one of the top ranked algorithms.

The rest of this paper is organized as follows. An overview of the proposed algorithm is given in Section 2. The initial matching module is discussed in Section 3. The problem of unreliable pixel detection is address in Section 4. Then, the process of occlusion handling is presented in Section 5. All processing modules described in Sections 3-5 are based on the fronto-parallel surface model. Then, we propose a piecewise linear surface model in Section 6. Experimental results are given in Section 7 and concluding remarks are followed in Section 8.

2. Overview of Proposed Algorithm

As shown in Fig. 1, our stereo matching algorithm consists of the following four processing modules: (1) initial matching; (2) cross checking and outlier removing; (3) disparity filling; (4) surface modeling.

The four processing modules are roughly explained below.

- **Initial matching.** The inputs to the initial matching module consist of a pair of rectified stereo images with their associated color segmentation results. The α -expansion algorithm in [3], which is based on the technique of graph cuts, is used to construct initial dense left and right disparity maps in this module.
- **Cross checking and outlier removing.** The main purpose of this module is to identify pixels which have unreliable disparity. Cross checking is conducted based on the uniqueness constraint that assumes the one-to-one mapping between correspondences of every stereo image pair. The conventional uniqueness constraint cannot be applied properly to the stereo image pair with highly slanted object surfaces. To manage them, we need to generalize the criteria adopted by cross checking. For outlier removing, we allow a larger disparity range in one color segment, which is determined by the ratio of the group of disparity and unreliable pixels. Such flexibility makes the proposed overall algorithm more robust to color segmentation errors.
- **Disparity filling.** The main purpose of this module is to estimate the disparity for unreliable pixels identified by the module of cross checking and outlier removing.

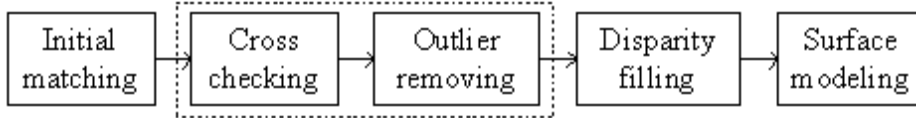


Figure 1. Overview of the proposed stereo matching algorithm.

Two algorithms are performed to explore the binocular and monocular information, respectively, so as to minimize the energy function locally.

- **Surface modeling.** We group neighboring pixels that have a similar disparity value into a set to form a slanted surface candidate by some criteria. After the grouping, a plane normal vector of each object is estimated by the random sample consensus (RANSAC) algorithm [5].

More details of the four processing modules will be given in Sections 3-6.

3. Initial Matching

The family of graph cuts algorithms, which is based on the Markov random field (MRF), has been shown to provide a robust result in stereo matching. One such algorithm, known as the α -expansion algorithm [3], is used in the initial matching module. This algorithm minimizes an energy cost function of the following form

$$E(D) = \min(E_{data}(D) + E_{smooth}(D)), \quad (1)$$

where D is the disparity map, $E_{data}(D)$ is the data term that computes the error cost between a stereo image pair, and $E_{smooth}(D)$ is the smoothness term that characterizes the smoothness between neighboring pixels. In our implementation, the data term is treated by following Birchfield-Tomasi's work [1] while Potts' model is employed for the smoothness term as done in [3]. This module yields initial left and right disparity maps.

The reason to use the α -expansion algorithm is due to its simplicity. Please note that it is a basic graph-cuts algorithm, which does not consider occlusion detection and disparity guessing in the occluded regions. It does not include any slanted surface model either. However, its performance will be gradually improved by modules that follow.

4. Unreliable Pixel Detection

Unreliable pixels will be detected by the module of cross checking and outlier removing. Once being detected, they will demand special treatment in later modules.

4.1. Cross Checking

A simple test in cross-checking is given by

$$|d_l(x, y) + d_r(x', y)| < 1, \quad (2)$$

where (x, y) and (x', y) are the correspondence pair (*i.e.* $x + d(x, y) = x'$) and $d_l(x, y)$ and $d_r(x', y)$ are the left and right disparities for points (x, y) and (x', y) , respectively. Despite its simplicity, the test in Eq. (2) is one of popular choices to determine consistency in stereo vision. It comes from the uniqueness constraint, which assumes the one-to-one correspondence between pixels of a stereo image pair.

It is however desirable to allow one-to-many mapping for highly slanted object surfaces. Thus, we consider a generalized test to cover the one-to-many mapping scenario as

$$|d_l(x, y) + d_r(x', y)| \leq t, \quad (3)$$

where $t \geq 1$. It means that an object with slanted surfaces whose disparities between the left and right images are within t pixels.

4.2. Outlier Removing

To remove outliers for each color segment, plane fitting is used widely. It is assumed that each segment is in the same plane so that pixels in the segment should have similar disparity values. The plane fitting algorithm adjusts disparities of outliers so that they still fall into a pre-specified disparity range. However, it is often that color segmented regions are so big that all pixels in one segment may not produce similar disparities.

In our implementation, the mean shift algorithm [4] is used for color segmentation. After that, outlier removal is used to cluster reliable disparities in a color segment into groups and identify unreliable disparity based on the following two measurements:

- the ratio of occlusion in a segment;
- the ratio of each disparity in a segment.

On one hand, if the ratio of occlusion is larger than a threshold, the disparity values of all pixels in the segment are ignored since all of them will be treated as unreliable pixels, which demand special treatment as discussed later. On the other hand, if the ratio of occlusion is smaller than the threshold, disparity values are clustered and the ratio of

each disparity cluster in the segment is calculated. If the ratio of a disparity class is larger than a threshold, their disparities are survived. Otherwise, they are treated as unreliable pixels, too.

5. Disparity Filling

Disparity values of unreliable pixels are estimated using both binocular and monocular information. Two algorithms are proposed; *i.e.*, a greedy disparity filling scheme and a least-squared-errors (LSE) fitting scheme. They are detailed below.

5.1. Greedy Disparity Filling

The basic assumption for the greedy disparity filling scheme is that the disparity of an unreliable pixel is the same as that of one of its neighbors in the same color segment. The binocular and the monocular image data are used sequentially. First, the algorithm using the binocular image data is stated with an example below.

1. We consider an unreliable pixel, denoted by p , and its 8 neighbor pixels labeled by $np1$ to $np8$ as shown Fig. 2 (a). The number in the bracket below each pixel represents a color value. For simplicity, only one of the RGB color values is considered in this example.
2. We sort neighbor pixels by the absolute color difference with respect to pixel p . The result of Step 2 is shown in Fig. 2 (b). The number in the parenthesis after $np1$ - $np8$ is the sorted number while the absolute difference value is also shown inside the parallel bar.
3. We begin with the smallest sorted number j . If the following two conditions are met
 - p and $np(j)$ are in the same color segment;
 - $|I_L(x, y) - I_R(x + d, y)| < s$, where I_L and I_R are the left and the right images, respectively, (x, y) is the position of pixel p and s is a threshold;

we assign the disparity value d of pixel $np(j)$ to pixel p . Otherwise, the sorted number is incremented by one, and Step 3 is repeated.

If $j = 8$ and the above two conditions are still not met, we will terminate the above binocular image processing step and try to fill the disparity value of p by another algorithm with the monocular information alone. For the monocular image processing, we replace the second condition in Step 3 by

$$|I_L(x, y) - I_L(x + m, y + n)| < s,$$

$np1$ (78)	$np2$ (69)	$np3$ (60)
$np4$ (76)	p (80)	$np5$ (83)
$np6$ (74)	$np7$ (79)	$np8$ (85)

(a) 8 neighbor pixels

$np1(2)$ 2	$np2(7)$ 11	$np3(8)$ 20
$np4(4)$ 4	p 3	$np5(3)$ 3
$np6(6)$ 6	$np7(1)$ 1	$np8(5)$ 5

(b) Sorted neighbor pixels

Figure 2. (a) The eight neighbor pixels and (b) the sorted neighbor pixels of unreliable pixel p .

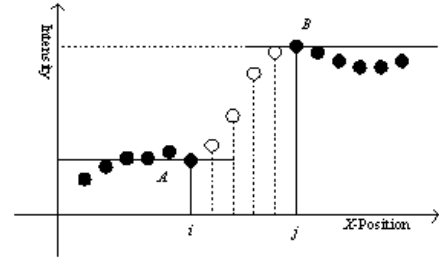


Figure 3. Disparities in a scanline, where black pixels have a disparity while white pixels do not have one.

where m and n take values among $-1, 0, \text{ or } 1$, but $(m, n) \neq (0, 0)$. Again, the processing order is determined by the sorted number, j .

After the greedy disparity filling procedure as described above, it is possible that the disparity map still has unreliable pixels which do not have a disparity value, since there is no guarantee at least one neighbor of a unreliable pixel may meet the conditions required by Step 3. This issue will be resolved by the least squared errors (LSE) method.

5.2. Least-Squared-Errors (LSE) Fitting

We attempt to make a guess for pixels which do not have a disparity up to now by the LSE scheme. With this scheme, we exploit the monocular information in the same scanline while ignore the boundaries of color segmentation. This can be explained by Fig. 3.

Black pixels in Fig. 3 represent pixels that have a disparity while white pixels have no disparity. The y coordinate represents intensity and the x coordinate stands for the pixel position. A (B) is a disparity value of black pixels i and j . Please note that the disparity of white pixels varies between A and B in this example. These are the closest disparity values that can be obtained using the greedy disparity filling scheme when being approached from the left and the right directions. This disparity gap over an interval makes disparity estimation a challenging problem.

To assign a disparity to white pixels, the LSE is calculated as a function of pixel position along the corresponding horizontal scanline of the left image. The white pixel

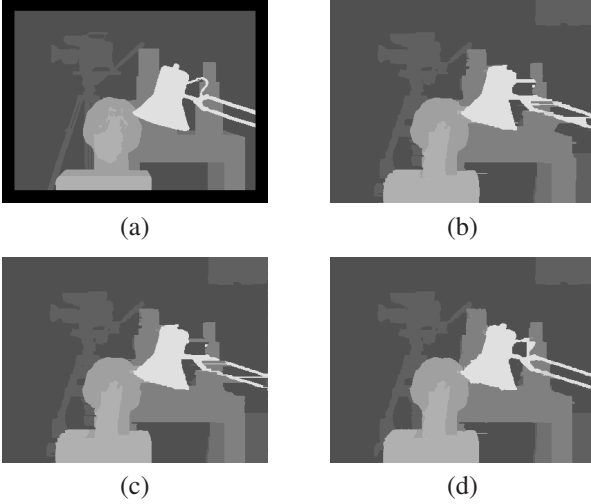


Figure 4. (a) Ground truth (b) Method 1, (c) Method 2, (d) Method 3 for test stereo image “Tsukuba”.

m^* that minimizes the LSE function can be computed via

$$m^* = \arg \min_{i \leq m \leq j} (f(m) + g(m)), \quad (4)$$

where

$$f(m) = \sum_{x=i}^m (I_L(x, y) - \text{Mean}(i, m))^2,$$

$$g(m) = \sum_{x=m+1}^j (I_L(x, y) - \text{Mean}(m+1, j))^2,$$

$$\text{Mean}(a, b) = \frac{\sum_{t=a}^b I_L(t, y)}{b - a + 1}.$$

Then, we assign disparity value A to white pixels $i + 1, \dots, m$ and disparity value B to white pixels $m + 1, \dots, j - 1$.

To compare results of the disparity filling algorithms, we show results of a test stereo image pair known as “Tsukuba” in Fig. 4 (a)-(d). Fig. 4(a) shows ground truth to compare results quantitatively. We consider the following three disparity filling methods.

- Method 1. The disparity of the closest reliable pixel to the left direction is assigned to unreliable pixels.
- Method 2. We check the closest reliable pixels to both the left and right directions and assign unreliable pixels the one of smaller disparity.
- Method 3. The proposed algorithm (*i.e.* the greedy disparity filling followed by LSE fitting.)

Results of Methods 1, 2 and 3 are shown in Figs. 4 (b), (c) and (d), respectively. The main improvement of Method 3 occurs around lamp’s clamp, which has a very complex shape. The error rate, defined by Scharstein and Szeliski in [11], measures the percentage of pixels in which the absolute difference between the result and the ground truth is greater than 1 pixel. It provides a useful measure for performance comparison of different algorithms qualitatively. The error rate results are shown in Table 1, where we compare error rates in the entire image (All), in the textureless region (Untex) and in the disparity discontinuity region (Dis). The rank indicates the rank performance among 41 different algorithms in the test. It is clear that the proposed disparity filling method (Method 3) provides an excellent performance improvement over the two conventional methods.

Table 1. Comparison of error rates of three methods

Method	All(%)	Rank	Untex.(%)	Rank	Dis.(%)	Rank
Method 1	2.17	21	1.09	18	12.17	23
Method 2	1.75	16	0.45	10	10.15	21
Method 3	1.05	3	0.17	1	6.01	3

6. Surface Modeling

All processing modules described in Sections 3-5 are based on the fronto-parallel surface model. However, this model has its limitation when an object has slanted local surfaces. Thus, a more generic surface model will be valuable. A method to approximate a 3D object using a local planar surface model is discussed in this section.

Based on the disparity filling results in Section 5, we can group neighboring pixels with similar disparity values as the projection of a local planar surface of an object onto the imaging plane. First, the smallest disparity group is assumed to be the background and the piecewise constant surface model is adopted for this group. Next, other pixels are grouped by comparing their disparity with 4-connective neighbor pixels. If the disparity difference is within 2 pixels, the pixel and its neighbor pixels are grouped into one object, and a local linear model is applied. In this model, one planar normal vector represents one object. However, for highly slanted or curved objects, it may not be sufficient to use one planar surface. To solve the problem caused by the lack of geometric information, we use the disparity variation range to partition a large surface into smaller segments and adopt the piecewise linear model for its approximation. The disparity variation range is set to 5 pixels. That is, if a new neighbor pixel has a disparity difference larger than 5 pixels in any pixel that is already in the same object, this pixel cannot be included in this group any longer. A new group is created to accommodate this new pixel.

After pixel grouping, we can determine the normal vector for each group using the the random sample consensus (RANSAC) algorithm [5]. The RANSAC algorithm is a robust algorithm widely employed in the field of computer vision. The RANSAC algorithm to calculate the normal vector for a local planar surface denoted by β is given below. For initialization, we select threshold r to determine inliners. Then, we perform the following iteration until the maximum number of inliners, m , or the maximum iteration number i .

1. Randomly select 3 pixels associated with β . They are denoted by (x_i, y_i) with $i = 1, 2, 3$.
2. Calculate the normal vector (a, b, c) by

$$\frac{1}{d_i} = ax_i + by_i + c, \quad (5)$$

where d_i is the disparity value of pixel (x_i, y_i) and $i = 1, 2, 3$.

3. Calculate the sum of distances between each pixel in β and the estimated plane. If the sum is smaller than threshold r , the 3 pixels used to estimate the normal vector is accepted as an inliner since a consensus is reached.
4. We repeat Steps 1-3 until one of the two stopping criteria is reached.

Finally, we use all obtained inliners to re-estimate the planar surface that has LSE, and the normal vector will be given accordingly. If the number of inliners is too small for robust normal estimation, we can increase the value of r and repeat the algorithm.

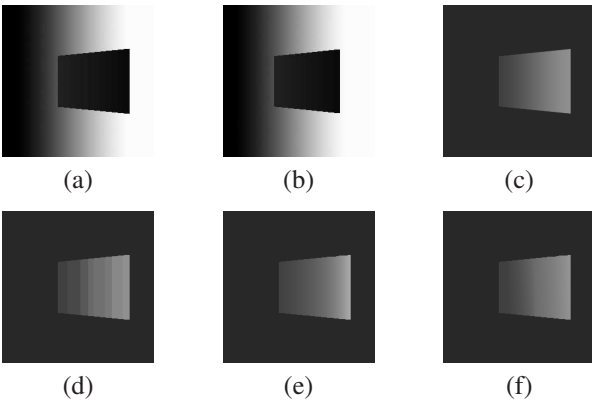


Figure 5. Results for the curved object test stereo pair: (a) the left image, (b) the right image, (c) the ground truth, (d) the fronto-parallel model, (e) the linear model and (f) the piecewise linear model.

Fig. 5 show results of the curved object from its stereo image pair. Fig. 5 (a) and (b) give input left and right images, respectively, and Fig. 5 (c) is the ground truth. As

shown in Fig. 5, the object in the middle of image is curved. Three surface models are compared: the fronto-parallel, the linear, and the piecewise linear model. The fronto-parallel model in Fig. 5 (d) is the result after the disparity filling module, which is used as the input to compute the linear and the piecewise linear models. The linear model fits the central region with one planar surface as shown in Fig. 5 (e). Thus, its result is poorer than the one obtained by the piecewise linear surface model.

To gain more insights, we show disparity values along scanline 127 of the curved object in Fig. 6. The proposed piecewise linear model (marked by “X”) has almost the same disparity as that of the ground truth (the solid line) while the linear model (marked by the square) has a significant error close to the right edge region of the object. The big discrepancy of the linear model is attributed to the fact that one plane is not sufficient to represent an object that has a large disparity variation. The fronto-parallel model (marked by the circle) has a stepwise disparity variation, which will result in artifacts in virtual view synthesis. Please note that the disparity is the reciprocal of the distance of the linear surface to the origin as shown in Eq. 5, which explains the curved shape of the disparity plot in the green line.

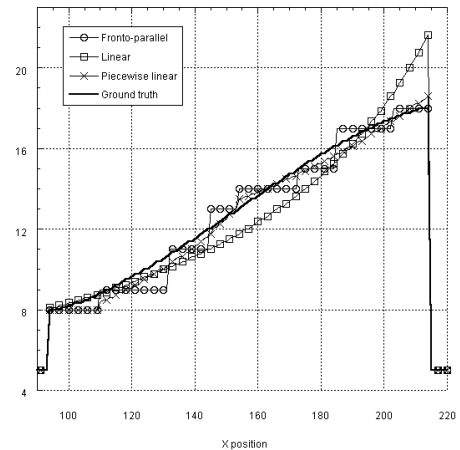


Figure 6. Disparity along scanline 127 of the curved object image

To compare the performance between the linear and the piecewise linear surface models for more complicated stereo images, test results of Venus image pair are shown in Fig. 7. Venus has several slanted objects. White pixels in error images are pixels whose disparity values matches with the ground truth. Other pixels are error pixels where the absolute difference between the result and the ground truth is greater than 1 pixel. We see that the disparity in the edge regions of the left-bottom object cannot be well represented by one planar surface. This problem is clearly resolved by the use of the piecewise linear surface model.

7. Experimental Results

7.1. Experimental Setup

The Middlebury data set is used to compare our algorithm with several benchmark algorithms [14]. The parameters used in the α -expansion algorithm for initial matching are chosen automatically [3]. For color segmentation, parameters of the mean shift segmentation are set by default values as given in [4]. Parameter t used in the cross-checking module is set to 2. For plane fitting, the first criterion (*i.e.*, the ratio of occlusion in the segment) is set to be $D/O < 0.5$, where D and O are the numbers of pixels without and with disparity values. The second criterion is chosen to be “if the percentage of the same disparity is smaller than 0.05%”, pixels with a disparity value are still set to unreliable pixels. For parameters in Section 5.1, $s = 5$ and r in the RANSAC algorithm is set to $SD(\beta) \times \frac{n}{f}$, where $SD(\beta)$ is the standard deviation of a local surface β , n is the total pixel number in β and $f = 100$. To increase r , we decrease the f value by 10 in each iteration until the consensus set is bigger than $S/10$, where S is the total number of pixels in β that have a disparity value.

7.2. Quantitative Evaluation

Quantitative evaluation is measured in terms of error rates. Error rates and ranking results for four test stereo image pairs for slanted surface models are given in Fig. 2, where the last three rows (called Prop. A, B and C) are all with the same new disparity estimation algorithm but different surface models. We see that the proposed disparity estimation procedure has boosted up the performance of our algorithm significantly. Among Prop. A, B and C, the piecewise linear surface model lowers the error rates for Sawtooth and Venus furthermore. We test Prop. A, B and C separately, and each of them ranks the 3rd among 41 methods in Nov. 2006.

7.3. Qualitative Evaluation

Fig. 8 shows results visual comparison. The test stereo image pairs are Tsukuba, Sawtooth, Venus, and Map (from left to right). The first row gives the ground truth. The second row is the result obtained by Boykov *et al.* in [3], which is used as the initial matching of the proposed algorithm. The third row is the result of our proposed algorithm. We see that a clear improvement of our work over the graph-cuts algorithm in [3].

8. Conclusion and Future Work

A disparity estimation procedure was developed to assign disparity values to unreliable pixels in occluded regions and a piecewise linear surface model was proposed to represent slanted or curved surfaces of objects in this

paper. It was demonstrated the disparity estimation procedure outperforms other existing disparity estimation techniques in handling large and/or complex occluded regions. It was also shown that the piecewise linear surface model outperforms the piecewise constant surface model and other higher-order surface models, since the latter does not exploit the piecewise surface decomposition. Both qualitative and quantitative evaluation of the proposed algorithm was conducted for the Middlebury data set. Our algorithm has been proved to be highly competitive with any state-of-the-art stereo matching algorithm. Although our algorithm is able to model the object surface well, further study is still needed for better surface decomposition when the algorithm is applied to a complex surface of an object.

References

- [1] S. Birchfield and C. Tomasi. A pixel dissimilarity measure that is insensitive to image sampling. *PAMI*, 20(4):401–406, April 1998. 3
- [2] S. Birchfield and C. Tomasi. Multiway cut for stereo and motion with slanted surfaces. *CVPR*, 1:489–495, Sept. 1999. 2, 8
- [3] Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *PAMI*, 23(11):1222–1239, Nov. 2000. 2, 3, 7, 8
- [4] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *PAMI*, 24(5):603–619, May 2002. 3, 7
- [5] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *ACM Trans. comm.*, 24:381–395, 1981. 3, 6
- [6] V. Kolmogorov and R. Zabih. Computing visual correspondence with occlusions using graph cuts. *ICCV*, 2:508–515, July 2001. 1
- [7] G. Li and S. W. Zucker. Surface geometric constraint for stereo in belief propagation. *CVPR*, 2:2355–2362, June 2006. 2
- [8] M. H. Lin and C. Tomasi. Surfaces with occlusions from layered stereo. *PAMI*, 26(8):1073–1078, Aug. 2004. 2, 8
- [9] P. Mordohai and G. Medioni. Stereo using monocular cues within the tensor voting framework. *PAMI*, 28(6):968–982, June 2006. 2, 8
- [10] A. S. Ogale and Y. Aloimonos. Stereo correspondence with slanted surfaces: critical implications of horizontal slant. *CVPR*, 1:568–573, July 2004. 2, 8
- [11] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *IJCV*, 47(1):7–42, 2002. 1, 5
- [12] J. Sun, Y. Li, S. B. Kang, and H.-Y. Shum. Symmetric stereo matching for occlusion handling. *CVPR*, 2:399–406, June 2005. 1
- [13] Y. Tsin and T. Kanade. A correlation-based model prior for stereo. *CVPR*, 1:135–142, June 2004. 2, 8
- [14] www.middlebury.edu/stereo. 7

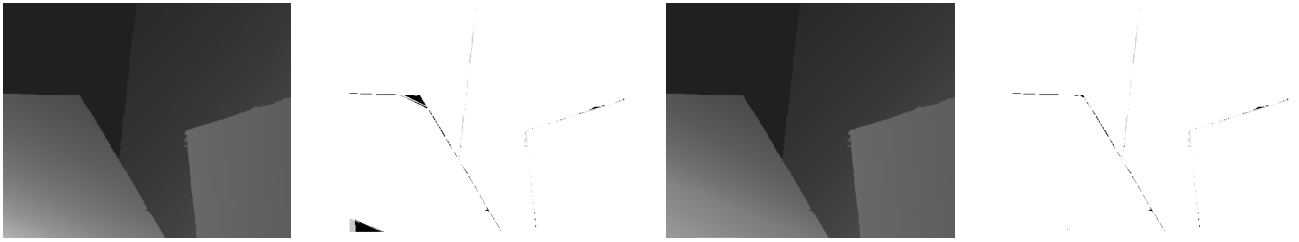


Figure 7. Results for test Venus stereo image pair (from left to right): the disparity result of the linear model, the error image of the linear model, the disparity result of the piecewise linear model and the error image of the piecewise linear model, where white and black pixels in error images indicate pixels with correct and erroneous disparity values, respectively.

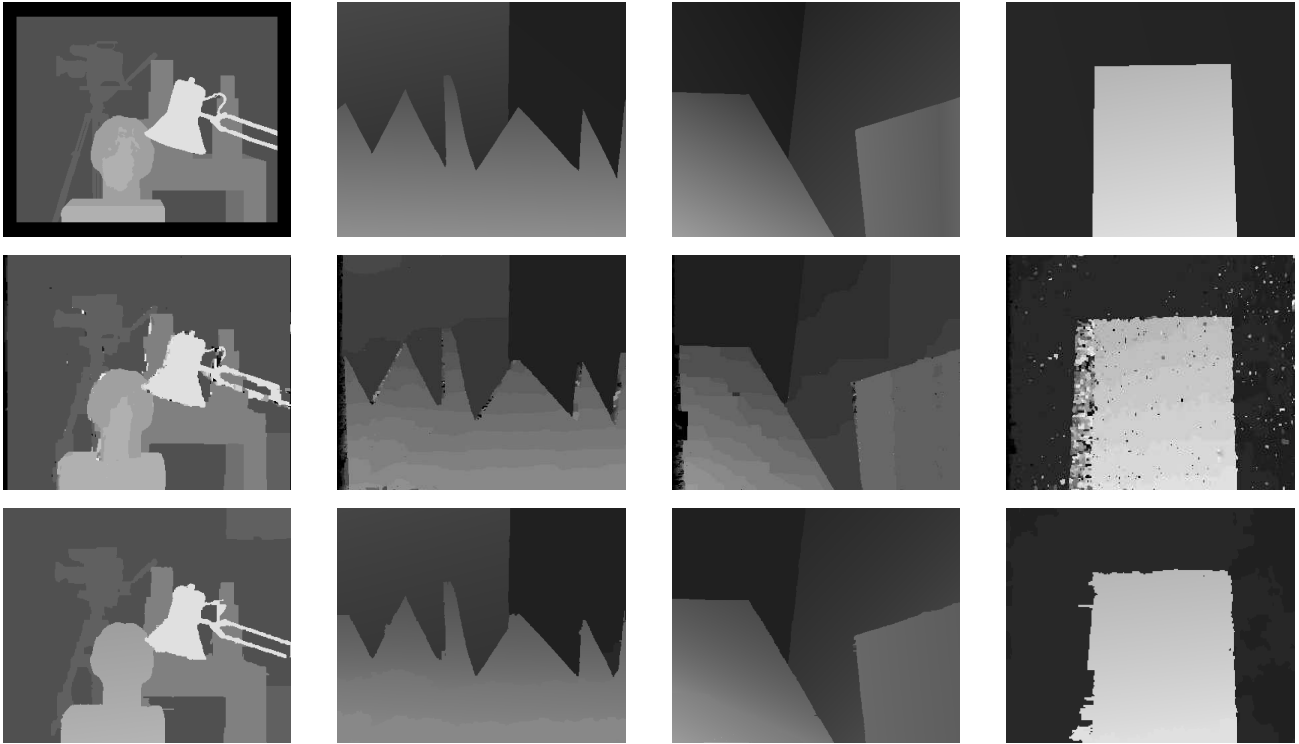


Figure 8. Qualitative comparison of the disparity map of two algorithms: the ground truth (the first row); the graph cuts algorithm [3] (the second row) and the proposed algorithm (the third row) for Tsukuba, Sawtooth, Venus and Map (from left to right).

Table 2. Quantitative evaluation of the proposed algorithm in terms of error rates and ranks. The table only compares algorithms of linear surface modeling. Error percentages and ranks in brackets are represented. The number in brackets shows the current rank in Nov. 2006. The * mark in brackets stands for the case that the current rank is not available.

Algorithm	Tsukuba			Sawtooth			Venus			Map	
	All	Untex.	Dis.	All	Untex.	Dis.	All	Untex.	Dis.	All	Dis.
Multi. cut [2]	8.08 (37)	6.53 (33)	25.33 (38)	0.61 (11)	0.46 (26)	4.60 (12)	0.53 (6)	0.31 (6)	8.06 (18)	0.26 (5)	3.27 (5)
Layered [8]	1.58 (14)	1.06 (17)	8.82 (15)	0.34 (7)	0.00 (1)	3.35 (8)	1.52 (19)	2.96 (29)	2.62 (5)	0.37 (13)	5.24 (13)
Tensor Voting [9]	1.51 (*)	2.02 (*)	7.96 (*)	0.70 (*)	0.50 (*)	4.35 (*)	1.09 (*)	1.39 (*)	13.95 (*)	1.31 (*)	11.47 (*)
Slanted Sur. [10]	1.82 (*)	1.09 (*)	9.47 (*)	0.72 (*)	0.24 (*)	6.00 (*)	3.25 (*)	5.73 (*)	8.51 (*)	0.22 (*)	3.10 (*)
Corr. Prior [13]	2.21 (*)	1.99 (*)	7.66 (*)	1.16 (*)	0.58 (*)	3.99 (*)	0.86 (*)	0.86 (*)	5.07 (*)	0.52 (*)	5.98 (*)
Prop. A (Fronto-Paral.)	1.05 (3)	0.17 (1)	6.01 (3)	0.36 (9)	0.00 (1)	3.46 (9)	0.50 (5)	0.23 (5)	3.95 (8)	0.44 (15)	5.33 (15)
Prop. B (Linear)	1.08 (3)	0.18 (1)	6.20 (3)	0.30 (7)	0.00 (1)	3.44 (9)	0.60 (7)	0.12 (5)	5.41 (11)	0.33 (12)	4.57 (13)
Prop. C (Piecew. linear)	1.08 (3)	0.18 (1)	6.20 (3)	0.30 (7)	0.00 (1)	3.44 (9)	0.18 (5)	0.03 (5)	3.24 (7)	0.33 (12)	4.57 (13)