

Research Article

Feature Based Stereo Matching Using Two-Step Expansion

Liqiang Wang,¹ Zhen Liu,¹ and Zhonghua Zhang²

¹ School of Instrumentation Science & Opto-Electronics Engineering, Beihang University, Beijing 100191, China

² National Institute of Metrology, Beijing 100029, China

Correspondence should be addressed to Liqiang Wang; gueij@163.com

Received 2 January 2014; Revised 19 June 2014; Accepted 21 July 2014; Published 18 December 2014

Academic Editor: Yi Chen

Copyright © 2014 Liqiang Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a novel method for stereo matching which is based on image features to produce a dense disparity map through two different expansion phases. It can find denser point correspondences than those of the existing seed-growing algorithms, and it has a good performance in short and wide baseline situations. This method supposes that all pixel coordinates in each image segment corresponding to a 3D surface separately satisfy projective geometry of 1D in horizontal axis. Firstly, a state-of-the-art method of feature matching is used to obtain sparse support points and an image segmentation-based prior is employed to assist the first region outspread. Secondly, the first-step expansion is to find more feature correspondences in the uniform region via initial support points, which is based on the invariant cross ratio in 1D projective transformation. In order to find enough point correspondences, we use a regular seed-growing algorithm as the second-step expansion and produce a quasi-dense disparity map. Finally, two different methods are used to obtain dense disparity map from quasi-dense pixel correspondences. Experimental results show the effectiveness of our method.

1. Introduction

Stereo matching is an international research focus of computer vision [1]. It can produce a disparity map from stereo images which are captured by cameras in different viewpoints. This technology is important in 3D reconstruction, virtual view rendering, and automatic navigation. It is a key point to know how to compute a precise disparity map in a complex environment by stereo matching. There is much excellent research to solve this problem. However, it still has some inherent challenges, such as unavoidable light variations, textureless regions, occluded areas, and nonplanar surface, that make the disparity estimation difficult [2–4].

To solve the inherent problems, numerous methods have been proposed in the past two decades. They consist of local and global methods [5, 6]. Local methods generally compute the correlation between these points and candidates over an adequate window and then use winner-takes-all (WTA) algorithm to find the best candidate to the point [7, 8]. They are fast to compute a disparity and flexible to model parametric surfaces within the neighborhood but have difficulties in handling poorly textured and ambiguous surfaces. Global methods are different from local approaches;

they commonly integrate prior constraints into optimization of the point correspondences to solve the poorly textured areas and lessen the matching ambiguities. They produce the disparity map by an energy minimization algorithm and have a better performance in poorly textured and textureless regions but are limited to model piecewise planar scenes [9]. Global methods have a goodish performance when the viewpoints are close [10] but do not handle well when the space of viewpoints becomes large [11, 12].

In large-scale stereo images, ambiguous areas exist more than their short-baseline counterpart. Whether the viewpoints are close or wide, there are always some significant features, such as points of interest, which are invariable. An alternative method uses reliable feature correspondences as seeds and expands these points by using a growing-like process to obtain more point correspondences [13–18]. The methods named seed-growing or region-growing can yield much better results in large perspective distortions and increased occluded areas than traditional ones. Seed-growing methods have a low computational complexity since they are not using global optimization but are sensitive to mismatches. To lessen the influence of wrong points, Cech and Sara [19] employed an optimal solution and introduced an improved

growing method which can handle many difficult instances, such as repetitive or complex textures. The method does not need each seed to be accurate in disparity map. However, seed-growing algorithms only generate a semidense disparity map because of sparse feature points.

To overcome drawbacks of traditional matching methods and seed-growing algorithms, the matched features are naturally integrated into state-of-the-art stereo methods as soft constraints [3, 20]. In these methods, a primary work is to find accurate point correspondences as GCPs (ground control points) [21]. GCP-based approaches improve stereo matching accuracy and correctness. However, GCP-based approaches need much time to obtain an accurate disparity map.

In this paper, a two-step expansion based robust dense matching algorithm is proposed based on the previous works [19, 22–24]. Sparse support points are obtained by state-of-the-art feature matching methods [22, 23]. Before two-step expansion, the segmentation-based prior [24] is used to encode the assumption that the region which has the same color is a 3D surface. The first-step is a feature expansion that is presented based on the invariant cross ratio of projective transformation. The basic idea is to match more features from initial support points in uniform region via cross ratio constraint. However, there is no ability to find enough matched pixels to obtain dense disparity map. To obtain more point correspondences, in the second-step, the matched features from the first-step are used as seeds to grow and build a quasi-dense disparity map which is denser than the feature correspondences of the first-step but not an absolutely dense disparity. About the process stage from quasi-dense disparity to dense disparity, the paper introduces two methods: (i) fitting process: a planar surface fitting is used to remove mismatches and can fill blank occluded areas in the uniform region and (ii) synthesized method: an optimal solution incorporates quasi-dense pixels into global energy methods to reduce the matching ambiguities.

This new work mainly focuses on the first-step that uses a feature-expanded algorithm for stereo matching. In the first step, we suppose that it is a set of sparse points whose coordinates are given in the same 3D surface, and the coordinates of the homologous image pixels satisfy projective geometry of 1D in horizontal axis. Our motivation comes from the theory that the points of axis satisfy 1D projective transformation and that the cross ratio is invariant. By using the invariance of cross ratio, the inhomogeneous coordinates of each corresponding pixel can be approximated. The accurate coordinates of the corresponding pixel are found by a search model that computes a correlation statistic for neighboring pixels. In addition, to solve the poorly textured regions, we employ a propagation algorithm to expand low feature pixels. Occluded areas can be filled by a fitting process or a synthesized method, and the fitting process method does not use cross-checking (checking and optimizing the disparity by computing the differences between left-to-right disparity and right-to-left disparity). Experimental results demonstrate that the method of two-step expansion has considerable performances over the existing ones. It can produce denser disparity than these existing seed-growing

algorithms, and it has a goodish result in short-baseline and wide-baseline stereo matching.

The paper is structured as follows: firstly, related work is discussed in Section 2. In Section 3, we introduce a support-point based expansion algorithm with cross ratio constraint. Then, a two-step expansion method is described, and it mainly presents the first-step about application of features expanded in Section 4. In Section 5, we describe two different methods to produce dense disparity map. Finally, we give the experimental validation supporting the feasibility of the method in Section 6. In Section 7, we give a conclusion and hint some future works.

2. Related Work

There are numerous literatures related to this work. Firstly, Scharstein and Szeliski [1] summarized dense stereo methods and established an early test bed for stereo matching algorithm. Then, Geiger et al. provided a newly outdoor challenge [25] for the quantitative evaluation of large-scale stereo matching. Seitz et al. [26] introduced a comprehensive study and made a comparison of stereo techniques. It included two main strategies for obtaining stereo correspondence: feature correspondences based local approaches and energy-minimum based global methods. In our method, the previous two-step expansion algorithm and the latter fitting process stage belong to the first strategy, and the later synthesized method falls in the second one.

Dense energy-minimum based global methods had a good performance in the past decade. Local stereo algorithms based on feature correspondences are speedy to estimate disparity [1, 27] but cannot effectively handle the blurry border and mismatches [7]. Hence, most excellent stereo matching algorithms rely first on using local approaches to find the pixel correspondences and then incorporate them into global constrains by dynamic programming (DP) [28–31], level sets [32], space carving [33], PDE [12, 34], EM [35], and voxel coloring [36]. Recently, two global methods based on Markov random fields (MRFs) are used as basic algorithms to be improved: Graph Cuts [37] and Belief Propagation [38]. Many works of research about both of them have achieved a desirable result [4, 39, 40]. Both methods are often used to be comparable data of the top contenders in the realm of dense stereo matching and are powerful tools to produce disparity map but intractable to finish the solution in wide-baseline stereo. In contrast, our method can lessen the matching ambiguities and is efficient to large-scale stereo matching.

Sparse local feature based approaches are robust to the large-scale images. Image features play an important role in computer vision. They have already been used in wide-baseline stereo matching [41–43]. In a wide-baseline setup, the inherent problems are perspective distortions and occlusion. Feature based matching methods are particularly effective because features are robust, distinctive, and invariant to various image and scene transformations [22, 23, 44–47]. However, the traditional methods based on feature matching produces only sparse pixel correspondences. To find more

matched points than features, a propagation algorithm from the matched points to their neighbors is introduced.

The rule of growing a region from primary seeds was used to segment image [48]. The seed-growing principle was originally introduced into stereo matching by Otto and Chau [49], O'Neill and Denos [50], and Kim and Muller [51] and used for photogrammetric community. Then Lhuillier and Quan [15, 52] employed the epipolar constraint and uniqueness constraint to greedily reproduce adjacent components in disparity blankness from corresponding seeds. The growth algorithm cannot achieve a good performance in the areas of repetitive patterns. The best first strategy as an optimal solution was used to replace the pixel-wise growth increments by Zeng et al. [17, 18]. And the optimization cannot be able to remove the previous match errors, especially in complex scenes. Kannala and Brandt [53] and Megyesi et al. [54] introduced a propagation algorithm by affine deformation of image similarity patches. But it had inaccurate affine parameters due to wrong initial seeds and made a bad propagation. Cech and Sara [19] introduced an optimal solution and presented a seed-growing method that could recover from errors in initial seeds. However, the method only produced a semidense disparity map. In contrast, our method can not only handle the difficult instances (e.g., repetitive texture, complex scene, and wrong initial seeds) but also produce denser point correspondences than the existing methods.

To compute an accurate dense disparity map, we incorporate quasi-dense pixel correspondences as GCPs into state-of-the-art global matching framework. In these literatures about stereo matching, GCPs-based methods can achieve a precise result. Bobick and Intille [2] used GCPs to optimize DP solution and reduce large occlusion. GCPs were used in preprocessing stage to guide the previous matching process and could reduce false matched points by using the method of Kim [3] and Wang et al. [20]. In [21], a GCPs-based regularization was incorporated into global method by using the Bayes optimization rule. In contrast, our method does not need provided special GCPs and can offer quasi-dense pixel correspondences as GCPs.

Geiger et al. proposed a generative probabilistic model ELAS [7] for wide-baseline stereo matching and offered a challenging KITTI dataset [25]. On KITTI dataset, these methods [55–57] that were used to compute optical flow had better results. In contrast, our method is a just strategy for stereo matching and receives a result compared with ELAS.

3. Efficient Expansion with Cross Ratio Constraint

3.1. Cross Ratio Constraint Model. In the epipolar geometry of two views, it can restrict the corresponding point on the polar line. To find the precise position of the corresponding point, traditional algorithms employ exhaustive search along the corresponding line and give a statistic for correlation of all candidates. To fasten the position estimation of the corresponding point on line, we introduce a new constraint based on 1D projective geometry.

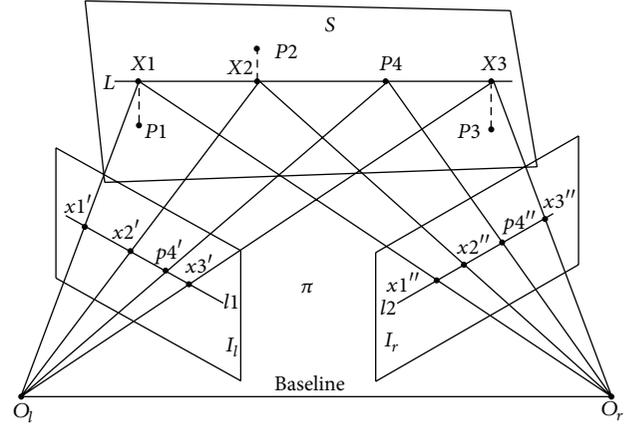


FIGURE 1: Two cameras are indicated by their centres O_l and O_r and their image planes I_l and I_r . There are 4 3D points $P1$, $P2$, $P3$, and $P4$ in uniform 3D planar surface S , and the point $P4$ projects to $p4'$ and $p4''$ in the images I_l and I_r , respectively. Line $l2$ in the right image and line $l1$ in the left image are epipolar lines separately with respect to points $p4'$ and $p4''$. Two camera centres, 3-space point $P4$, and its images $p4'$ and $p4''$ lie in an epipolar plane π . The intersection of the planes π and S determines the line L in 3D. The 3D points $X1$, $X2$, and $X3$ are the closest points on the line L to the 3D points $P1$, $P2$, and $P3$. Points $x1'$, $x2'$, and $x3'$ in the left image and points $x1''$, $x2''$, and $x3''$ in the right image are projected by 3D points $X1$, $X2$, and $X3$, and points $x1'$, $x2'$, and $x3'$ lie on the line $l1$, and points $x1''$, $x2''$, and $x3''$ lie on the line $l2$.

We assume there is a stereovision system as shown in Figure 1. It can be seen that there are three sets of four collinear points in the polar plane π . Each set is related to the others by a line-to-line projective transformation. Since the cross ratio is invariant under a 1D projective geometry, it has the same value as

$$\frac{X1P4 X2X3}{X2P4 X1X3} = \frac{x1' p4' x2' x3'}{x2' p4' x1' x3'} = \frac{x1'' p4'' x2'' x3''}{x2'' p4'' x1'' x3''}. \quad (1)$$

3.2. Estimation Model via Cross Ratio. The cross ratio constraint based on 1D projective geometry needs three or more known 3D points. Thus, it needs to obtain the known 3D points. We employ features matching algorithm as prior to produce reliable point correspondences which can be used to calculate the fundamental matrix. The proportional coordinates of the known 3D points can be estimated by the point correspondences and the fundamental matrix. It can produce more errors when the region including the known 3D points is not a planar surface.

To lessen the error, we know that the points on the same epipolar line satisfy 1D projective geometry whether the surface is plane or not and introduce a search strategy that uses image points near the epipolar line instead of the 3D points, as shown in Figure 2. Suppose the images have been rectified and the point correspondences lie on the same line in both images. We wonder how the corresponding point $p4''$ in the right image is found. Firstly, we can find the

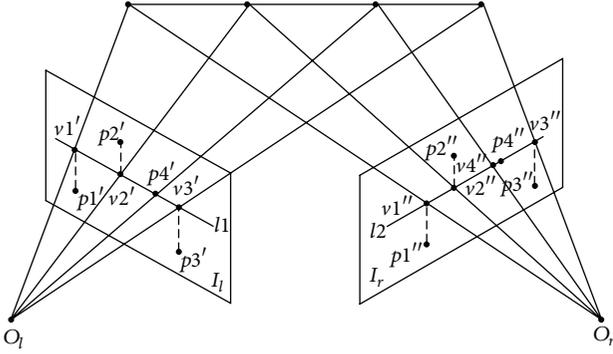


FIGURE 2: I_l and I_r are the images observed from the camera centres O_l and O_r separately. Given three sets of matched point pairs $(p1', p1'')$, $(p2', p2'')$, and $(p3', p3'')$, an unmatched point $p4'$ in the left image and the unknown corresponding point $p4''$ in the right image, $l1$ and $l2$ are the lines on which the corresponding points $p4'$ and $p4''$ lie. Points $v1', v2', v3'$ are the closest points on the line $l1$ to the points $p1', p2', p3'$ in the left image, and points $v1'', v2'', v3''$ are the closest points on the line $l2$ to the points $p1'', p2'', p3''$ in the right image.

corresponding polar lines $l1$ and $l2$. Then we can find a point $v4''$ satisfies the following equation:

$$\frac{v1' p4' v2' v3'}{v2' p4' v1' v3'} = \frac{v1'' v4'' v2'' v3''}{v2'' v4'' v1'' v3''}. \quad (2)$$

If the points $p1', p2', p3'$ and $p1'', p2'', p3''$ are the homologous points that the 3D points $P1, P2, P3$ project to two images separately in Figure 1, the points $v1', v2', v3'$, and $v1'', v2'', v3''$ are not the corresponding points projected by the 3D points $X1, X2, X3$ because of projective transformation. Hence, the point $v4''$ is not the corresponding point $p4''$, but adjacent to the point $p4''$. The distance points to the epipolar line are shorter; the position of $v4''$ is nearer to the point $p4''$. We employ a probabilistic search strategy to ensure the point $p4''$ in the $v4''$ has contiguous pixels along the line $l2$.

3.3. Search Strategy. The search strategy computes all correlations with the neighbors of the point $v4''$ and decides the position of the point $p4''$. As shown in Figure 3, a set of neighborhoods $\{c(1), \dots, c(m), v4'', c(m+1), \dots, c(2m)\}$ of size $(2 \times m + 1)$ whose centre is point $v4''$ is built as a set of candidate matches. The value m is the radius of search and is decided by the maximum d of the Euclidean distances from referenced points $p1'', p2'', p3''$ to the line $l2$. If the distance from reference point pi'' ($i = 1, 2, 3$) to the line $l2$ is di , then the maximum is $d = \max(di)$ and $m = \min(\text{ratio} \times d, r)$, where ratio, r , is nonzero constants for proportional d and fixed radius. We use sum of absolute differences (SAD) [58] on $N \times N$ window as image similarity statistic between the point $p4'$ and all candidate points in the right image, where N is a positive constant. Assuming the SAD value between the point $p4'$ and candidate points is $D(i)$, where $i = 1, 2, 3, \dots, 2 \times m + 1$, if the SAD value between $p4'$ and $v4''$ is $Dv = D(m + 1)$, and the n th point in candidate

ones except the point $v4''$ has the SAD minimum and $D(n) = \min\{D(i)\}$, where $i = 1, 2, \dots, m, m + 2, \dots, 2 \times m + 1$, then the corresponding point $p4''$ is defined as

$$p4'' = \begin{cases} v4'', & \text{if: } Dv < D \min \times \alpha, Dv < \beta, \\ c(n), & \text{if: } Dv > D \min \times \alpha, D \min < \beta, \\ -1, & \text{if: } Dv > \beta, D \min > \beta, \end{cases} \quad (3)$$

where α is a proportional constant and β is the threshold for the correct correspondence, if $p4'' = -1$ means no corresponding point.

4. A Two-Step Expansion Method

In this section, we describe a two-step expansion algorithm based on image features to compute quasi-dense point correspondences between two views. Our method is inspired by observing an instance where all points in the uniform surface satisfy 1D projective geometry in horizontal axis. And in 1D projective transformation, the cross ratio of the projected points is invariant. Our algorithm is arranged as follows: firstly, a sparse set of initial support points are found by excellent feature matching method. Then, in the first-step expansion, we use segmentation-based prior to partition the image into different regions and employ the invariance of cross ratio as a restrictive condition to find the more corresponding feature points from the support points in the same region. Finally, a regular seed-growing approach is used to obtain more pixel correspondences as the second-step expansion.

Suppose there exists a pair of images $I = \{I_l, I_r\}$, where I_l, I_r are the separately left and right images, this section aims at finding the quasi-dense disparity D_q corresponding to I_l . To expediently introduce our method, we suppose input images I_l and I_r are rectified, such that corresponding points lie on epipolar lines of two images.

4.1. Initial Support Points. Before expansion, we introduce how to establish a sparse set of feature correspondences as initial support points. Most algorithms which are used to extract image features can be categorized as either corner detectors (such as Harris and Stephens [22] and SUSAN [46]) or descriptor extraction (such as SIFT [23], SURF [44], and DAISY [47]). Recently, a regional feature detector [59] based on descriptor [23] had a good performance in dealing with the large-scale instance. In our method, we employed regular Harris method [22] to obtain initial support points. While in the presence of large disparity ranges, the number of the successfully matched Harris points is less than the threshold Np (which is decided by the number of the segmented regions; refer to Section 4.2), scale invariant feature transform (SIFT) algorithm is used to extract features, and the KD-tree with the best bin first (BBF) [60] algorithm is employed to index and match these features. We assume the $\{(pi', pi'')\}_i$ is matched point pairs by feature matching method, where $\{pi'\}$ and $\{pi''\}$ are points from two images.

4.2. The First-Step Expansion. At this stage, our objective is to compute all the possible feature point correspondences

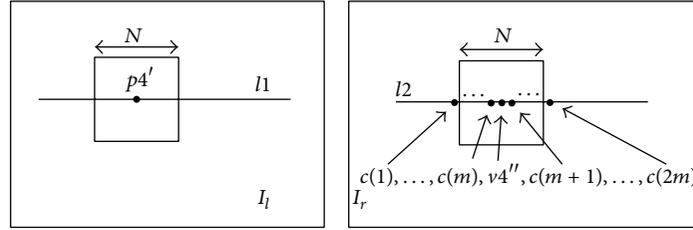


FIGURE 3: I_l and I_r , and l_1 and l_2 are separately the corresponding planes and lines. Given a pair of points (p_4', v_4'') whose v_4'' is estimated by cross ratio model, $\{c(1), \dots, c(m), v_4'', c(m+1), \dots, c(2m)\}$ is a set of neighborhoods of the point v_4'' . N is the size of window which is used to compute correlation.

through the initial support points in the uniform region. The first-step expansion is based on segmented regions; thus, we employ the mean-shift method to segment the reference image I_l before expanding feature points. The mean-shift algorithm which was successfully used to partition images by Comaniciu and Meer [24] can ensure our method estimates regions correctly and localizes depth boundaries precisely. The result of image segmentation will set a different label to each segmented region. If the number of the segmented regions is Kl , then the threshold Np in Section 4.1 is $Np = 3 \times Kl$.

In Section 3, we introduce an expansion model based on 1D projective geometry in the same planar surface. We now use the expansion model as the first-step expansion algorithm. More formally, let $L = \{l_1, l_2, l_3, \dots, l_k\}$ be a set of labels with respect to the different segmented regions of the left image. Each pixel pi is assigned to a corresponding label l_j where $l_j \in L$.

We assume the initial support points belonging to a label l_j construct a set of samples $Sa_j = \{(pi', pi'')\}_i$, where $pi' \in l_j$. In this step, we spread sweeping feature correspondences from initial support points in the same region. In our method, the feature means a point whose absolute value of gradient is bigger than 1. Hence, our prior is a process that computes gradient $G(x, y)$ of each pixel in the image and selects the pixels whose gradient $|G(x, y)| \geq 1$ as our candidate feature points. Suppose we have found all the feature points $\{xi'\}$ and each feature $xi' \in \{xi'\}$ is assigned to a corresponding label l_j . We will introduce how to find the corresponding point xi'' through Sa_j .

The expansion algorithm mainly is based on epipolar geometry and 1D projective transformation. Epipolar restraint has been used to rectify the image and restrict the corresponding points to the same lines in the images. We just need to find three support points to estimate the probabilistic position of the corresponding point. The number of the initial support points $\{pi\}$ in each region l_j is not fixed and can be sorted to two statuses: more than three points and no more than three points. This step mainly handles the first status.

When the number of $\{pi\}$ is more than three, as shown in Figure 4, we can consider each horizontal axis on which pixels lie is its corresponding epipolar because of image rectification. Suppose each point $pi' = (xi', yi')$, where $i = 1, 2, \dots, 8$, and three support points which would be found satisfy the following conditions: (i) three points have a minimum summation of the distances to the epipolar line

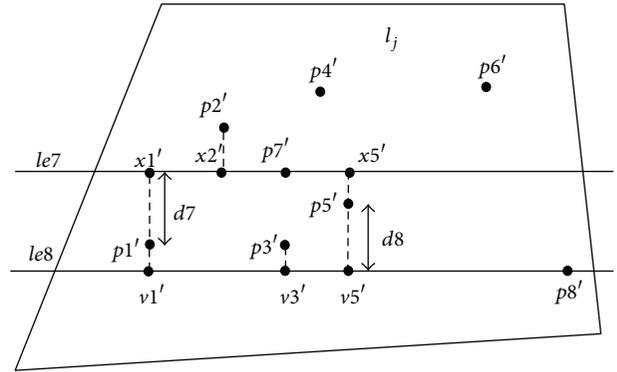


FIGURE 4: In the region l_j , a set of matched points $\{p_1', p_2', p_3', p_4', p_5', p_6'\}$ has been known, and p_7' and p_8' are the points which need to search the corresponding point. The horizontal axes le_7 and le_8 are the epipolar lines of the points p_7' and p_8' , separately.

at the same time and (ii) any two x -coordinates of all points should not be equal. For example, as shown in Figure 4, the support points of p_7' are p_1', p_2' , and p_5' and the support points of p_8' are p_1', p_3' , and p_5' , that satisfy the two conditions. The corresponding search radii are separately $m_7 = \min(\text{ratio} \times d_7, r)$ and $m_8 = \min(\text{ratio} \times d_8, r)$. Then the corresponding point can be found by the method of Section 3.

4.3. The Second-Step Expansion. The second-step employs a regular seed-growing method to obtain stable correspondences in poorly textured regions. Suppose the first-step produces a list of point correspondences $S = \{(pi', pi'')\}_i$. We regard the point correspondences S as seeds to grow corresponding patches. Despite the fact that the first-step can find more effective point correspondences, it inevitably introduces errors in complex areas. The traditional seed-growing algorithms do not handle wrong initial seeds well. To overcome the drawbacks, Cech and Sara [19] temporarily forwent uniqueness constraint, propagated most disparity ingredients, and then optimized them to remove these false disparity components. Hence, the second-step employs the method of Cech to obtain quasi-dense disparity D_q .

Cech method includes two phases: (i) growing and propagating as many seeds as possible regardless of their overlaps and (ii) optimizing these seeds of the first phase and

removing these false ones. The seed-growing method of Cech can keep accurate point correspondences and recover most disparities from false seeds. The detailed descriptions of the seed-growing method can be referred to the literature [19].

5. Obtaining Dense Disparity Map

The two-step expansion method cannot find all pixel correspondences in some regions because of occlusion and cannot produce completely dense disparity map. We introduce two different processes to compute dense disparity map from quasi-dense point correspondences. One is a filling process by regional 3D surface fitting; the other is a synthesized method that integrates quasi-dense pixel correspondences as GCPs into global optimization frameworks in a principled way.

5.1. Fitting Process. In Section 4.2, we have obtained different regions from the image I_l by the segmentation-based prior. The segmented regions may be with respect to different 3D surface. Now we assume each 3D surface is planar. In some regions of a quasi-dense disparity D_q , there may be only a few corresponding points and some piecewise patches which are built by unmatched points due to occlusion. A 3D planar surface fitting can be applied to fill the uncharted patches in the same region.

Assume there exists a set of pixel correspondences $\{(p_i', p_i'')\}_i$ in an arbitrary region l_j , and we use the regional data $\{p_i'\}$ to compute a 3D plane Pl_j . We describe each pixel of the quasi-dense disparity D_q as $V_i = (xi', yi', zi')$, where xi' and yi' are the coordinates of p_i' in the image, and zi' is a corresponding disparity. Then, we can use a set of points $\{V_i\} \in Pl_j$ to fit a 3D planar surface:

$$f(xi', yi', zi') = axi' + byi' + czi' + d, \quad (4)$$

where a , b , c , and d are the parameters which are used to describe a plane. These pixels belonging to the same area satisfy the 3D plane equation and can be computed the involved disparities.

5.2. Synthesized Method. Recently, a mixed stereo model which uses these known point correspondences as the GCPs to improve the result of global matching has a good performance in textureless areas and occlusion.

Synthesized method is inspired from the method of Wang [21] and formulates the stereo modal as a MAP-MRF problem. Assume the quasi-dense D_q is produced from a pair of images $I = \{I_l, I_r\}$ by two-step expansion. Based on Bayes' rule, the posterior probability of the disparity map D is expressed as $P(D | I, D_q) \propto P(I | D)P(D_q | D)P(D)$. Finding the maximum posterior cost means minimizing the corresponding negative log likelihood. Thus computing a disparity map D becomes the problem for minimizing the energy function:

$$E(D) = E_{\text{data}}(D) + E_{\text{smoothness}}(D) + E_{\text{quasi-dense}}(D), \quad (5)$$

where E_{data} is a function to estimate the probability for disparity map, $E_{\text{smoothness}}$ is a smoothness term to encourage

TABLE 1: Comparison of the results on number of corresponding points.

	St. Martin	Head	Larch
Cech 07	481733	293004	165218
Our method	749585	379153	195531

similar neighboring points in locally smooth region, and $E_{\text{quasi-dense}}$ is the energy of D_q to constrain the accuracy of disparity map D . The details can be referred to the literature [21].

5.3. The Overall Algorithm. The process of two-step expansion algorithm is summarized as in Algorithm 1.

6. Experiments

We took different experiments to demonstrate the validity of our method. In Section 6.1, we compared our approach to the seed-growing method of Cech on the real complex scenes [19]. It was a test on running time for different image resolutions in Section 6.2. We then separately evaluated fitting process and synthesized method on Middlebury benchmark short-baseline stereo images with known ground truth data in Section 6.3. In Section 6.4, we tested our algorithm on large-scale stereo image pairs.

Throughout all experiments we set ratio = 0.3, $r = 7$, $N = 5$, $\alpha = 1.2$, and $\beta = 4 \times 10^{-4}$, which were empirically determined. All experiments were operated on the computer with Intel core 2 duo CPU and 2.93 GHz clock frequency. Unless stated otherwise, we employed regular Harris method to obtain initial matched points and performed mean-shift image segmentation using EDISON code [61] implementation of Comaniciu's paper [24].

6.1. Computing Quasi-Dense Disparities. Firstly, we can obtain quasi-dense disparity map by the two-step spreading. We demonstrated the difference between the seed-growing method and our algorithm by comparing their performances on some real data. In known seed-growing algorithms, the method proposed by Cech and Sara [19] has a better performance, even in the presence of repetitive patterns. Hence, we compared our approach to the method of Cech. We tested different stereo images from the Cech dataset [62], that is, St. Martin, Head, and Larch. The relevant quasi-dense disparities of the images are shown in Figure 5. It can be seen that our algorithm can produce denser disparity map than the algorithm proposed by Cech. A comparison to the number of the corresponding points in different images is depicted in Table 1.

This experiment result demonstrates that our method can produce a quasi-dense disparity map via a sparse set of initial feature correspondences. Our method does not need too accurate matched features as seeds. In a repeated experiment, our method always found more point correspondences than Cech's method.

```

Input: A pair of rectified images  $I_l$  and  $I_r$  from different viewpoints of one scene;
      Set the values of ratio,  $r$ ,  $N$ ,  $\alpha$ , and  $\beta$ .
Output: The disparity map with respect to  $I_l$ .
Begin:
  Step 1. Finding initial point correspondences  $\{(p_i', p_i'')\}_i$  by using state-of-the-art matching method;
  Step 2. Using the mean-shift to partition the image  $I_1$  to different areas denoted as:  $L = \{l_1, l_2, l_3, \dots, l_k\}$ ;
  Step 3. Assigning the points  $\{p_i'\}$  into the corresponding area  $l_j$ ;
  Step 4. Removing the coarse mismatches in the area  $l_j$  by using regional affine transformation;
  Step 5. Computes gradient  $G(x, y)$  of each pixel in the image and selects the pixels  $\{x_i'\}$ 
          whose gradient  $|G(x, y)| \geq 1$  as our candidate feature points;
  Step 6. Assign each feature  $x_i'$  to a corresponding label  $l_j$ ;
  Step 7. The first step for the matched feature expansion
  Repeat:
    for:  $i = 1: \text{size}(\{x_i'\})$ 
      ensure  $l_j$  via to:  $x_i' \in l_j$ ;
      find a set of samples  $Sa_j = \{(p_i', p_i'')\}_i$ , where  $p_i' \in l_j$ ;
      if:  $\text{size}(\{p_i\}) > 3$ 
        compute the point  $x_i''$  by estimation model via cross ratio;
      end if
    end for
  Until  $\{l_j\}$  is empty
  Step 8. The second step expansion by using a regular seed-growing method
  Step 9. Obtain dense disparity map by using fitting process or synthesized method

```

ALGORITHM 1: Two-step expansion algorithm.

6.2. Running Time. The term of running time is relative to the elements, that is, image resolution, segmented regions, and initial support points. We changed the image resolutions for Tsukuba, Teddy, Cones, and Venus from Middlebury benchmark [63] and then took a running time statistic with respect to the three elements separately. We downscaled the images bicubically by 10%~90%, tested the running time in different resolutions, and recorded the corresponding numbers of regions and points. As shown in Figures 6(a), 6(b), and 6(c), it is related to the illustration about the running time of images in different resolutions, regions, and points. It can be seen that the numbers of the segmented regions and the initial points are more, and the corresponding running time is shorter in the same resolution. Figure 6(d) shows the corresponding relations between segmented regions and matched points in different image resolutions.

6.3. Short-Baseline Stereo Matching. We tested the fitting process and the synthesized method on several image pairs, that is, Tsukuba, Venus, Teddy, and Cones from the Middlebury benchmark [63]. The maximum of the disparity for the images is less than 100. Firstly, we used a two-step method to produce the different quasi-dense disparities of the images. Then, we found the corresponding disparity maps by the fitting process and synthesized method. When we implemented the fitting process, we restricted the maximal difference of disparity in the same region less than 10. In the synthesized method, we employed Graph Cuts [37] as our assistant global method. The goal is to compute a disparity map D by the function (5). In time statistics, the fitting process takes about 1.3 minutes to estimate a disparity and the synthesized method takes about 1.8 minutes. We demonstrate

in Figure 7 the results for the images of Tsukuba, Venus, Teddy, and Cones. As can be seen, these disparities produced by synthesized method have a clear structure and few blurry areas.

To evaluate the performance of our method, we used the quality measure method proposed in [1] with known ground truth data to evaluate the synthesized results. The matching results rank 87 and 62 with respect to 1 and 0.5 pixels error in Middlebury website. These competitive algorithms were based on these classical original methods and proposed an improved algorithm. They commonly integrated many technologies into their methods and had a better performance. Our method is first proposed by us without more refined technologies. To verify the validity of the method, we compared it with these classical original methods, that is, GC (graph cuts) [37], CSBP (constant-space belief propagation) [64], DP [29], and SO (scanline optimization) [1], as shown in Table 2, where the absolute error is more than 2 pixels. Quality evaluation uses the three performance measures: nonocc (bad pixels of nonoccluded), all (bad pixels of the entire image), and disc (bad pixels of near discontinuous).

It can be seen from Table 2 that our method has a better result than these of the traditional methods. The wider the baseline is, the more obvious the accuracy of the disparities is, for example, the scenes of Teddy and Cones. Because our method is based on feature matching which is efficient to large-scale images. On the Venus scene, the error of our method is not the lowest in these methods. Since global methods employ energy-minimum function to optimize the disparity map and have a better performance than feature-based local methods in short-baseline. The images of Venus scene belong to short-baseline. The global methods of GC

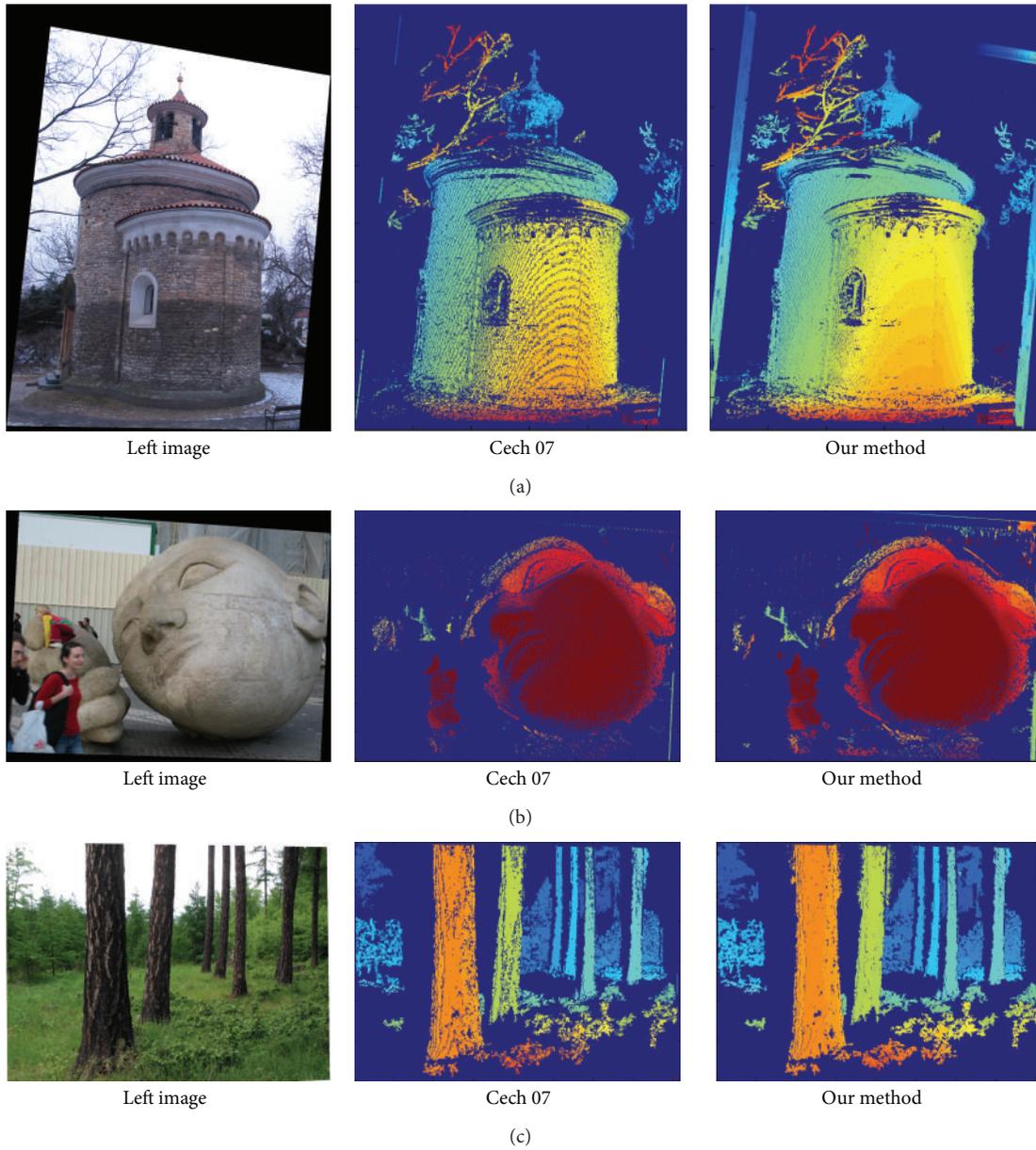


FIGURE 5: Results for quasi-dense disparities of Cech dataset are as follows: (a) St. Martin, (b) Head, and (c) Larch. Disparity maps are partitioned in different colors: colder color means smaller disparities, warmer color means larger disparities, and deeply blue areas are unassigned disparity.

TABLE 2: Comparative performance of stereo algorithms according to Middlebury methodology.

	Tsukuba			Venus			Teddy			Cones		
	Nonocc	All	Disc	Nonocc	All	Disc	Nonocc	All	Disc	Nonocc	All	Disc
Ours	2.57	3.13	10.8	3.34	3.49	13.1	5.44	7.84	12.1	3.41	8.30	9.21
GC	1.67	3.75	8.20	0.83	2.37	8.12	9.72	18.8	17.2	4.51	15.0	11.2
CSBP	1.74	3.84	9.10	1.09	2.52	12.9	8.18	17.3	20.6	4.07	14.2	11.3
DP	3.43	4.23	9.85	6.50	7.43	17.4	7.11	14.9	13.4	6.52	15.1	15.1
SO	4.23	6.21	10.7	5.14	6.58	17.0	11.3	20.2	18.6	8.76	18.8	16.1

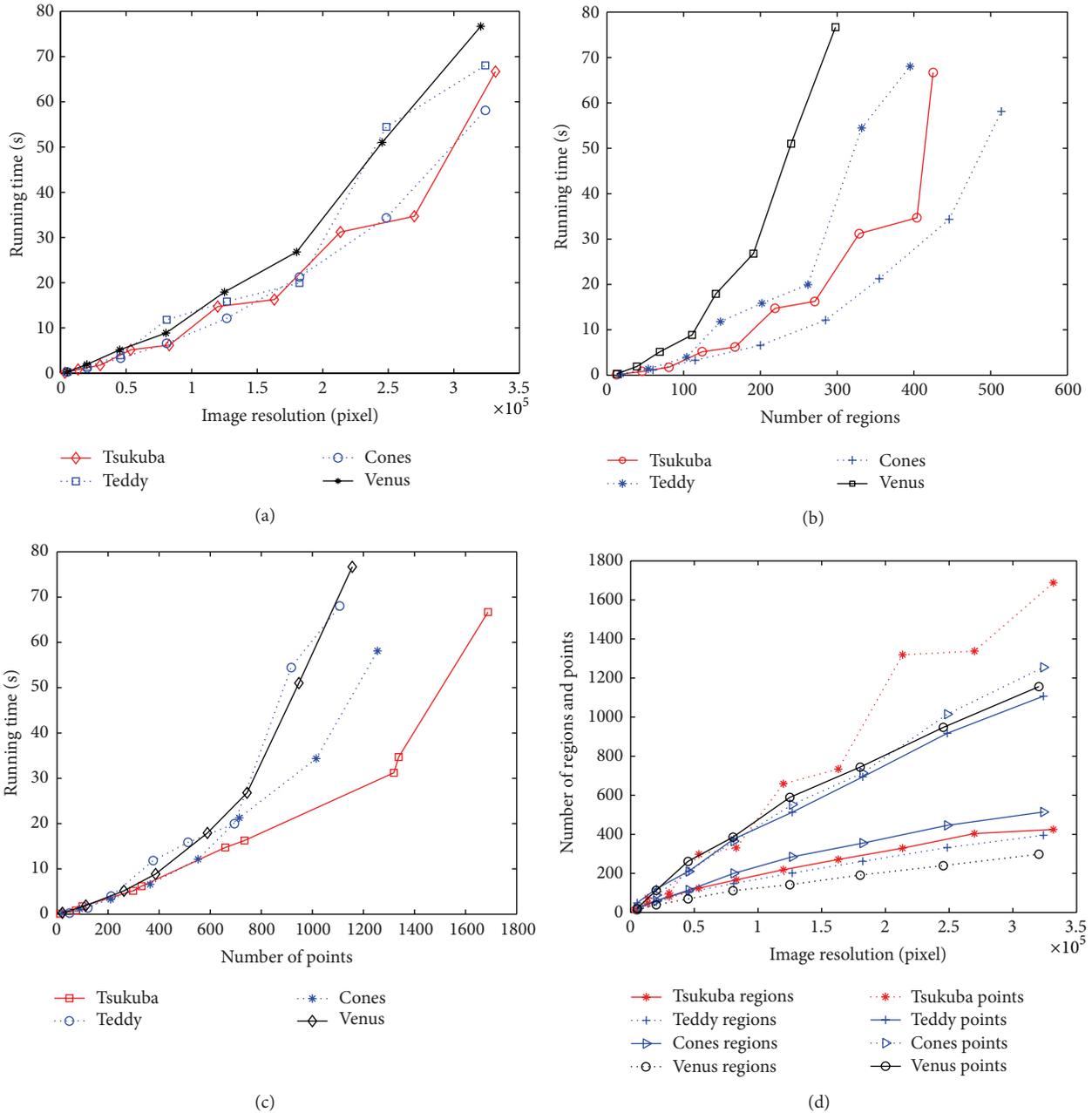


FIGURE 6: The relations of running time to (a) image resolution, (b) number of regions, (c) number of support points on the Tsukuba, Teddy, Cones, and Venus image pairs, and (d) the relevant segmented regions and corresponding points in different resolutions of the images.

and CSBP have more accurate results than our method on the Venus scene.

6.4. Large-Scale Stereo Matching. Though short-baseline stereo matching can yield accurately dense disparity, there is much more challenge in large-scale stereo images because of too much occlusion. In large-scale stereo images, we computed the disparity just via the fitting process and without the synthesized method. Firstly, we compared the fitting process on a wide range of baseline high resolution images, that is, Aloe and Raindeer from the Middlebury benchmark

[63] whose maximum of the disparity is bigger than 200. In particular, we compared our method against the method ELAS proposed by Geiger et al. [7], as shown in Figure 8. We compute all erroneous pixels of the entire image whose absolute error is more than 3 pixels. The error results of our method are 13.03% and 20.36% separately corresponding to the images Aloe and Raindeer, and the results of ELAS are 14.14% and 22.28%.

Then, we took a test on the KITTI dataset [25], which consists of 194 training and 195 test pairs of urban images. The training images with semidense ground truth disparities are used to adapt the parameters of stereo matching methods.

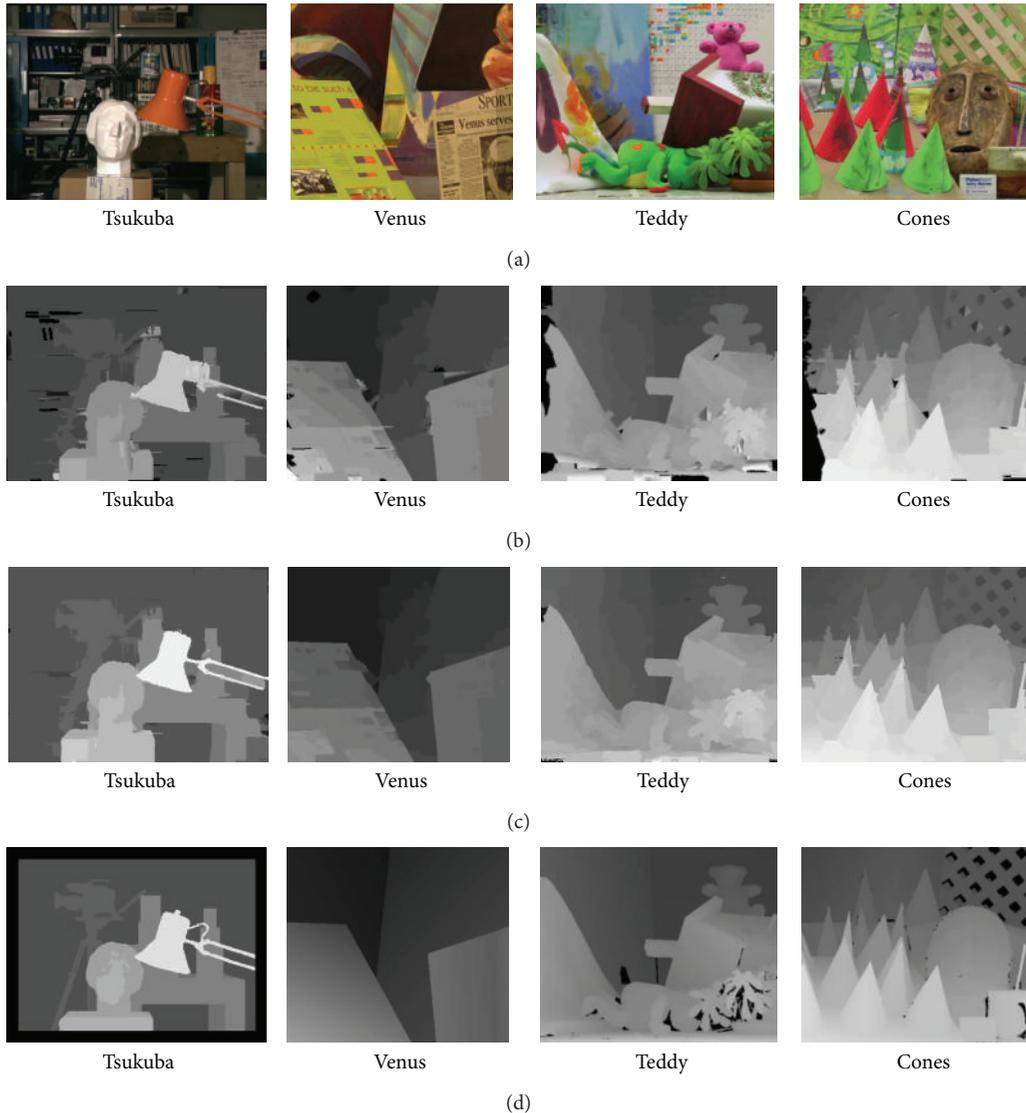


FIGURE 7: Results of our two different methods on short-baseline dataset. (a) Left images. (b) Results of fitting process. (c) Results of synthesized method. (d) Ground truth disparities.

There is no parameter to be trained and modified in our method. The test images without ground truth are used to evaluate participants in the challenge. On the dataset, the main problem is how to handle these textureless areas. We computed the disparity maps of these test images via the fitting process, and some results are shown in Figure 9. The average run time for computing a disparity map is about 4.7 minutes. The matching results rank 38 and 35 with respect to 3- and 5-pixel error in KITTI website. We compared our method to the similar methods, that is, ELAS [7], GCSF (growing correspondence seeds flow) [55], and GCS (growing correspondence seeds) [19], as shown in Table 3, where Out-Noc is the percentage of erroneous pixels in nonoccluded areas, and Out-All is the percentage of erroneous pixels in total. Avg-Noc is the ratio of average disparity or end-point error in nonoccluded areas. Avg-All is the ratio of average

disparity or end-point error in total. The qualitative results for this dataset are similar to the previous evaluation. We are able to robustly reconstruct large-scale images, which leads to low error rates on the street and on other slanted surfaces.

7. Conclusion

In this paper, we introduce a two-step expansion to produce precise disparity maps from stereo images whether the stereo baseline is short or large. Our method is based on feature matching and can cope with the difficult cases such as large perspective distortions, increased occluded areas, and complex scenes. Our experiments on Cech's dataset, the Middlebury benchmark, and KITTI dataset demonstrate that our method achieves good results in the real complex

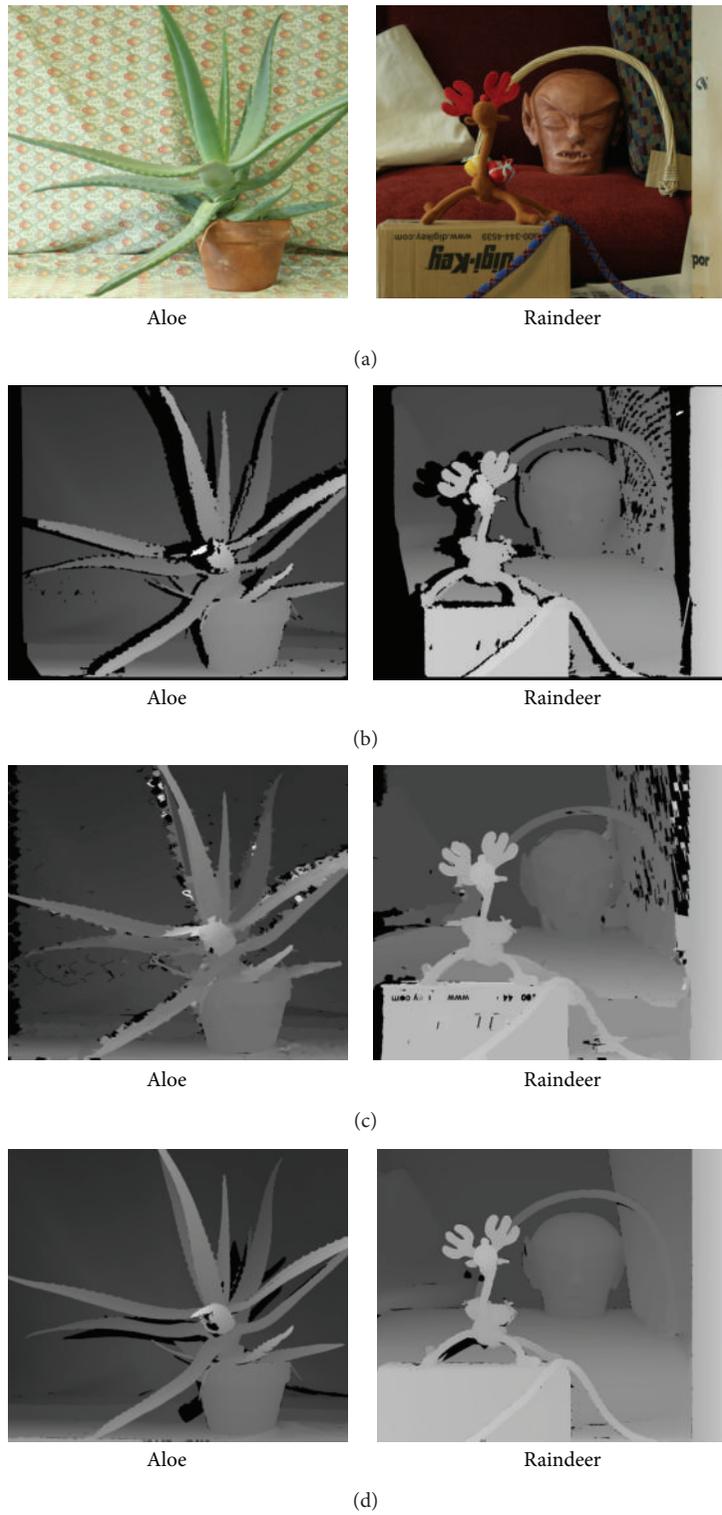


FIGURE 8: Comparison to Geiger's method on the Aloe and Raindeer image pairs. (a) Left images. (b) ELAS results. (c) Fitting process results. (d) Ground truth disparities.

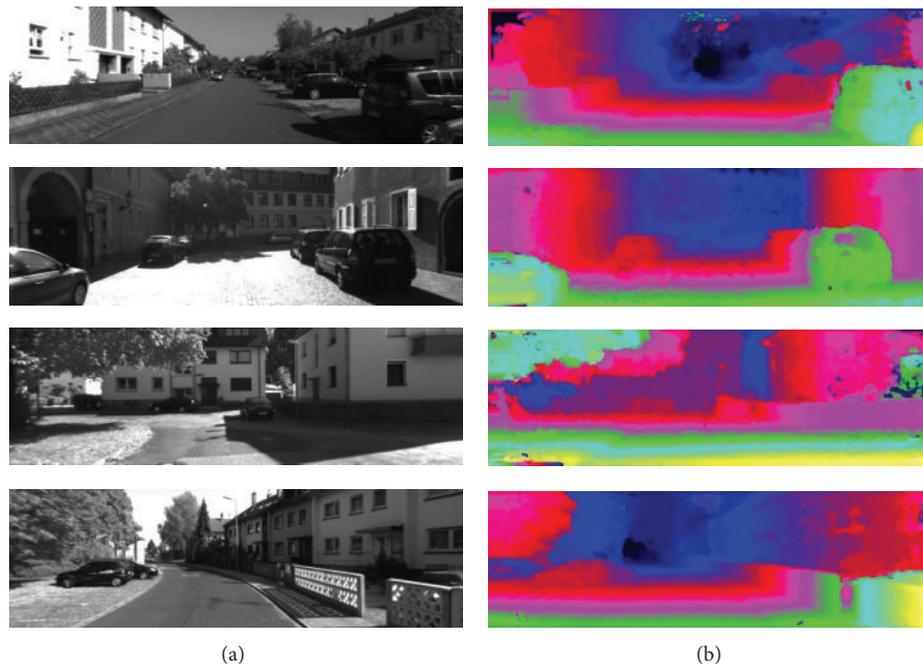


FIGURE 9: Results on urban scenes. (a) Left images. (b) Our method results. Best viewed is in different colors.

TABLE 3: Comparative evaluation results on KITTI test dataset.

	>2 pixels		>3 pixels		>4 pixels		>5 pixels		End-point	
	Out-Noc	Out-All	Out-Noc	Out-All	Out-Noc	Out-All	Out-Noc	Out-All	Avg-Noc	Avg-All
ELAS	10.96%	12.83%	8.24%	9.96%	6.73%	8.24%	5.67%	6.97%	1.4 px	1.6 px
Ours	14.59%	16.08%	9.91%	11.30%	7.32%	8.57%	5.74%	6.87%	1.7 px	1.9 px
GCSF	17.41%	18.73%	12.05%	13.24%	9.22%	10.28%	7.54%	8.49%	1.9 px	2.1 px
GCS	19.03%	20.32%	13.38%	14.54%	10.41%	11.43%	8.64%	9.55%	2.1 px	2.3 px

scenes, short or wide baseline image pairs. Importantly, we introduce a cross ratio restraint model to expand more feature correspondences based on state-of-the-art feature matching.

Our method primarily involves performing point computation in large numbers of segmented regions, which is fit for implementing in GPU and can real-time compute the disparity map of stereo images.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

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, no. 1-3, pp. 7-42, 2002.
- [2] A. F. Bobick and S. S. Intille, "Large occlusion stereo," *International Journal of Computer Vision*, vol. 33, no. 3, pp. 181-200, 1999.
- [3] J. C. Kim, K. M. Lee, B. T. Choi, and S. U. Lee, "A dense stereo matching using two-pass dynamic programming with generalized ground control points," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1075-1082, June 2005.
- [4] J. Sun, N. Zheng, and H. Shum, "Stereo matching using belief propagation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 7, pp. 787-800, 2003.
- [5] H. Sadeghi, P. Moallem, and S. A. Monadjemi, "Feature based dense stereo matching using dynamic programming and color," *International Journal of Computational Intelligence*, vol. 4, no. 3, p. 179, 2008.
- [6] L. Valgaerts, A. Bruhn, M. Mainberger, and J. Weickert, "Dense versus sparse approaches for estimating the fundamental matrix," *International Journal of Computer Vision*, vol. 96, no. 2, pp. 212-234, 2012.
- [7] A. Geiger, M. Roser, and R. Urtasun, "Efficient large-scale stereo matching," in *Proceedings of the 10th Asian Conference on Computer Vision (ACCV '10)*, November 2010.
- [8] B. M. Smith, L. Zhang, and H. Jin, "Stereo matching with non-parametric smoothness priors in feature space," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '09)*, pp. 485-492, Miami, Fla, USA, June 2009.
- [9] L. Tang, H. T. Tsui, and C. K. Wu, "Dense stereo matching based on propagation with a Voronoi diagram," in *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing*, vol. 22, 2002.

- [10] M. Z. Brown, D. Burschka, and G. D. Hager, "Advances in computational stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 8, pp. 993–1008, 2003.
- [11] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and Vision Computing*, vol. 22, no. 10, pp. 761–767, 2004.
- [12] C. Strecha, T. Tuytelaars, and L. van Gool, "Dense matching of multiple wide-baseline views," in *Proceedings of the 9th IEEE International Conference On Computer Vision*, pp. 1194–1201, October 2003.
- [13] Q. Chen and G. Medioni, "Volumetric stereo matching method: application to image-based modeling," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '99)*, pp. 1029–1034, Fort Collins, Colo, USA, June 1999.
- [14] M. Gong and Y. Yang, "Fast stereo matching using reliability-based dynamic programming and consistency constraints," in *Proceedings of the 9th IEEE International Conference on Computer Vision*, pp. 610–617, October 2003.
- [15] M. Lhuillier and L. Quan, "Match propagation for image-based modeling and rendering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 8, pp. 1140–1146, 2002.
- [16] H. Wu, Z. Song, J. Yao, L. Li, and Y. Gu, "Stereo matching based on support points propagation," in *Proceeding of IEEE International Conference on Information Science and Technology*, pp. 23–25, IEEE, Hubei, China, March 2012.
- [17] G. Zeng, S. Paris, L. Quan, and F. Sillion, "Accurate and scalable surface representation and reconstruction from images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 141–158, 2007.
- [18] G. Zeng, S. Paris, L. Quan, and M. Lhuillier, "Surface reconstruction by propagating 3D stereo data in multiple 2D images," in *Proceedings of the European Conference on Computer Vision*, pp. 163–174, 2004.
- [19] J. Cech and R. Sara, "Efficient sampling of disparity space for fast and accurate matching," in *Proceedings of the International Workshop on Benchmarking Automated Calibration, Orientation, and Surface Reconstruction from Images*, 2007.
- [20] L. Wang, H. Jin, and R. Yang, "Search space reduction for MRF stereo," in *Proceedings of the European Conference on Computer Vision*, 2008.
- [21] L. Wang and R. Yang, "Global stereo matching leveraged by sparse ground control points," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3033–3040, June 2011.
- [22] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proceedings of the 4th Alvey Vision Conference*, pp. 147–151, 1988.
- [23] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.
- [24] D. Comaniciu and P. Meer, "Mean shift: a robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [25] A. Geiger, M. Roser, and R. Urtasun, "Urban Scenes Dataset," 2013, http://www.cvlibs.net/datasets/kitti/eval_stereo_flow.php.
- [26] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, "A comparison and evaluation of multi-view stereo reconstruction algorithms," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '06)*, pp. 519–526, June 2006.
- [27] M. Weber, M. Humenberger, and W. Kubinger, "A very fast census-based stereo matching implementation on a graphics processing unit," in *Proceedings of the 12th International Conference on Computer Vision Workshops (ICCV '09)*, pp. 786–793, IEEE, October 2009.
- [28] A. Benschrair, P. Miché, and R. Debric, "Fast and automatic stereo vision matching algorithm based on dynamic programming method," *Pattern Recognition Letters*, vol. 17, no. 5, pp. 457–466, 1996.
- [29] S. Birchfield and C. Tomasi, "Depth discontinuities by pixel-to-pixel stereo," *International Journal of Computer Vision*, vol. 35, no. 3, pp. 269–293, 1999.
- [30] Y. Ohta and T. Kanade, "Stereo by intra- and inter- scanline search using dynamic programming," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, no. 2, pp. 139–154, 1985.
- [31] O. Veksler, "Stereo correspondence by dynamic programming on a tree," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, pp. 384–390, June 2005.
- [32] O. D. Faugeras and R. Keriven, "Complete dense stereovision using level set methods," in *Proceedings of the European Conference on Computer Vision*, June 1998.
- [33] K. N. Kutulakos and S. M. Seitz, "Theory of shape by space carving," *International Journal of Computer Vision*, vol. 38, no. 3, pp. 199–218, 2000.
- [34] L. Alvarez, R. Deriche, J. Sánchez, and J. Weickert, "Dense disparity map estimation respecting image discontinuities: a PDE and scale-space based approach," *Journal of Visual Communication and Image Representation*, vol. 13, no. 1-2, pp. 3–21, 2002.
- [35] C. Strecha, R. Fransens, and L. van Gool, "Combined depth and outlier estimation in multi-view stereo," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 2394–2401, June 2006.
- [36] S. M. Seitz and C. R. Dyer, "Photorealistic scene reconstruction by voxel coloring," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 1067–1073, June 1997.
- [37] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [38] J. Yedidia, W. T. Freeman, and Y. Weiss, "Understanding belief propagation and its generalizations," in *Proceedings of the International Joint Conference on Artificial Intelligence, Distinguished Papers Track*, 2001.
- [39] V. Kolmogorov and R. Zabih, "Multi-camera scene reconstruction via graph cuts," in *Proceedings of the European Conference on Computer Vision*, pp. 82–96, 2002.
- [40] C. Rhemann, A. Hosni, M. Bleyer, C. Rother, and M. Gelautz, "Fast cost-volume filtering for visual correspondence and beyond," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 3017–3024, June 2011.
- [41] P. Pritchett and A. Zisserman, "Wide baseline stereo matching," in *Proceedings of the 6th International Conference on Computer Vision*, pp. 754–760, IEEE, January 1998.
- [42] E. Tola, V. Lepetit, and P. Fua, "A fast local descriptor for dense matching," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '08)*, pp. 1–8, Anchorage, Alaska, USA, June 2008.

- [43] T. Tuytelaars and L. V. Gool, "Wide baseline stereo matching based on local, affinity invariant regions," in *Proceedings of the British Machine Vision Conference*, pp. 412–425, 2000.
- [44] H. Bay, A. Ess, T. Tuytelaars, and L. van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, 2008.
- [45] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1615–1630, 2005.
- [46] S. M. Smith and J. M. Brady, "SUSAN: a new approach to low level image processing," *International Journal of Computer Vision*, vol. 23, no. 1, pp. 45–78, 1997.
- [47] E. Tola, V. Lepetit, and P. Fua, "DAISY: an efficient dense descriptor applied to wide-baseline stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 5, pp. 815–830, 2010.
- [48] R. M. Haralick and L. G. Shapiro, "Image segmentation techniques," *Computer Vision, Graphics, & Image Processing*, vol. 29, no. 1, pp. 100–132, 1985.
- [49] G. Otto and T. Chau, "'Region-growing' algorithm for matching of terrain images," *Image and Vision Computing*, vol. 7, no. 2, pp. 83–94, 1989.
- [50] M. O'Neill and M. Denos, "Practical approach to the stereo matching of urban imagery," *Image and Vision Computing*, vol. 10, no. 2, pp. 89–98, 1992.
- [51] T. Kim and J. Muller, "Automated urban area building extraction from high resolution stereo imagery," *Image and Vision Computing*, vol. 14, no. 2, pp. 115–130, 1996.
- [52] M. Lhuillier and L. Quan, "A quasi-dense approach to surface reconstruction from uncalibrated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 3, pp. 418–433, 2005.
- [53] J. Kannala and S. S. Brandt, "Quasi-dense wide baseline matching using match propagation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, June 2007.
- [54] Z. Megyesi, G. Kós, and D. Chetverikov, "Dense 3D reconstruction from images by normal aided matching," *Machine Graphics and Vision*, vol. 15, no. 1, pp. 3–28, 2006.
- [55] J. Čech, J. Sanchez-Riera, and R. Horaud, "Scene flow estimation by growing correspondence seeds," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 3129–3136, June 2011.
- [56] C. Vogel, S. Roth, and K. Schindler, "Piecewise rigid scene flow," in *Proceedings of the International Conference on Computer Vision*, 2013.
- [57] K. Yamaguchi, D. McAllester, and R. Urtasun, "Robust monocular epipolar flow estimation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2013.
- [58] M. J. Atallah, "Faster image template matching in the sum of the absolute value of differences measure," *IEEE Transactions on Image Processing*, vol. 10, no. 4, pp. 659–663, 2001.
- [59] K. Mikolajczyk, T. Tuytelaars, C. Schmid et al., "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, 2005.
- [60] J. S. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1000–1006, June 1997.
- [61] "Edge Detection and Image Segmentation (EDISON) System," 2014, <http://coehttp://www.rutgers.edu/riul/research/code/EDISON/doc/overview.html>.
- [62] J. Cech and R. Sara, Cech GCS Dataset, 2013, <http://cmp.felk.cvut.cz/~cechj/GCS/>.
- [63] D. Scharstein and R. Szeliski, "Middlebury Stereo Matching Benchmark," 2013, <http://vision.middlebury.edu/stereo/>.
- [64] Q. Yang, L. Wang, and N. Ahuja, "A constant-space belief propagation algorithm for stereo matching," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 1458–1465, San Francisco, Calif, USA, June 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

