Feature Matching in Stereo Images Encouraging Uniform Spatial Distribution

Xiao Tan[a,b], Changming Sun[b], Xavier Sirault[c], Robert Furbank[c], Tuan D. Pham[d]

[a] *The University of New South Wales, Canberra, ACT 2600, Australia.*
[b] *CSIRO Computational Informatics, Locked Bag 17, North Ryde, NSW 1670, Australia.*
[c] *CSIRO Plant Industry, Clunies Ross Street, Canberra, ACT 2601, Australia.*
[d] *Aizu Research Cluster for Medical Engineering and Informatics, The University of Aizu, Fukushima 965-8580, Japan.*

Abstract

Finding feature correspondences between a pair of stereo images is a key step in computer vision for 3D reconstruction and object recognition. In practice, a larger number of correct correspondences and a higher percentage of correct matches are beneficial. Previous researches show that the spatial distribution of correspondences are also very important especially for fundamental matrix estimation. So far, no existing feature matching method considers the spatial distribution of correspondences. In our research, we develop a new algorithm to find good correspondences in all the three aspects mentioned, i.e., larger number of correspondences, higher percentage of correct correspondences, and better spatial distribution of correspondences. Our method consists of two processes: an adaptive disparity smoothing filter to remove false correspondences based on the disparities of neighboring correspondences and a matching exploration algorithm to find more correspondences according to the spatial distribution of correspondences so that the correspondences are as uniformly distributed as possible in the images. To find correspondences correctly and efficiently, we incorporate the cheirality constraint under an epipole polar transformation together with the epipolar constraint to predict the potential location of matching point. Experiments demonstrate that our method performs very well on both the number of correct correspondences and the percentage of correct correspondences; and the obtained correspondences are also well distributed over the image space.

Keywords: 3D/stereo scene analysis, Stereo image processing, Feature

correspondences.

1. Introduction

A stereo image pair contains two images of the same scene or object captured from different viewpoints. The two images are related to each other by the epipolar geometry where the correspondence of a point in one image lies on its corresponding epipolar line in the other image [1, 2]. This geometric information can be represented by a singular matrix which is known as the fundamental matrix [3, 4]. When the intrinsic camera parameters are known, an essential matrix can be obtained which is related to the fundamental matrix. The epipolar geometry limits the search of correspondences between two images to 1D search; that is, the correspondences are restricted to lie on the corresponding epipolar lines which are defined by the fundamental matrix. Estimating the fundamental matrix as well as the corresponding epipolar lines are studied in [5, 6, 7]. For detailed descriptions and the comparison on a variety of techniques on fundamental matrix estimation, readers can refer to the reviews in [8, 9, 10]. Finding correspondences through the recovered epipolar geometry helps reduce the correspondence search dimension and helps find the correspondences correctly as well. Similar ideas are exploited in [11, 1, 12]; however, they do not consider if a 3D point lies in front of the image plane or behind the image plane when a pair of 2D matching points are back projected into a 3D coordinate. The relationship between 3D points and the image plane is named "cheirality" by Hartley in [13]. A back projected 3D point of an image correspondence has a real geometry or from a real scene if and only if it lies in front of both cameras. Following the terminology in [13], we call a set of correspondences a strong realization with respect to a given fundamental matrix $\mathbf{F}$ if a configuration exists so that all triangulated 3D points lie in front of both cameras. Conversely, a weak realization is a set of correspondences which follow the epipolar geometry defined by $\mathbf{F}$ but not all of them correspond to real 3D points. Cheirality is found to be very useful in many studies such as camera calibration [14], pose estimation [15], and camera motion estimation [16], since it enforces the match to have a real geometry. Suppose that we find three candidate matching points q_1, q_2 and q_3 for a feature point p, Fig. 1 shows the back projecting results of these three correspondences under a fixed epipolar geometry. Point M lying in front of both cameras has a real geometry. Point L, the back projecting result from

correspondence (p, q_3), does not have a real geometry as it lies behind the first camera (π_1). Point N, the back projecting result from correspondence (p, q_2), does not have a real geometry since it lies behind both cameras. Let p_∞ be the point at infinity along the line of sight of p, to make sure that the resulting 3D point have a real geometry, the search range is constrained along the epipolar line between e' and the reprojection of p_∞ on π_2. However, the reprojection location of p_∞ is not fixed and can not be estimated without knowing the intrinsic parameters of the cameras. For weak calibrated cases where the intrinsic parameters of cameras are unknown, the constraint from the reprojection of the infinity point during the matching process is not considered.

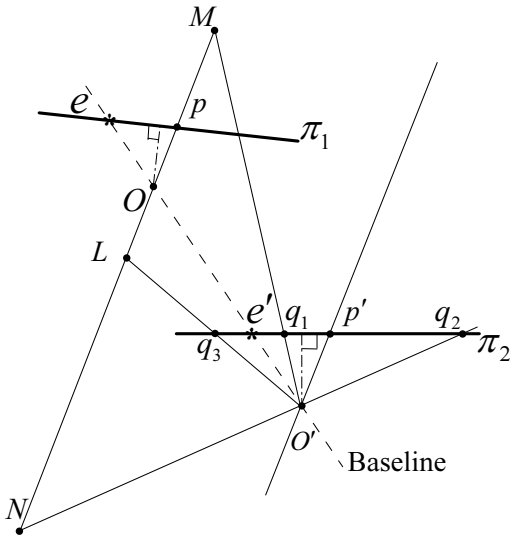


Figure 1: Illustration of triangulated 3D points from point correspondences with and without a real geometry. (p, q_1) is the correspondence with a real geometry; (p, q_2) and (p, q_3) are correspondences without a real geometry. O and O' are the camera centers; e and e' are the epipoles in the two images; π_1 and π_2 are two principal planes; line $O'p'$ is parallel to the line spanned by O and p. Thus, p' is the reprojection of p at infinity along line Op. The triangulated 3D point has a real geometry only when the matching point of p lies between e' and p'.

In this paper, the homogenous coordinate of a feature point is denoted by the bold font notation of that feature point. For example, the homogenous coordinate of feature point p is denoted by $\mathbf{p}$. The corresponding epipolar line of point p is parameterized by $\mathbf{Fp}$ under the homogenous coordinate

system. Let (p_1, q_1) and (p_2, q_2) be two pairs of point correspondences lying on a pair of corresponding epipolar lines denoted by (l, l'). e and e' are the epipoles which separate the two epipolar lines into the upper half and the lower half (see Fig. 2). In a strong realization, the point on one half of l can only correspond to the point on the corresponding half of l' [17]. In other words, suppose that both p_1 and p_2 have real geometries, if p_1 and p_2 lie on the same side of e, q_1 and q_2 will lie on the same side of e'; if p_1 and p_2 lie on different sides of e, q_1 and q_2 will lie on different sides of e'.

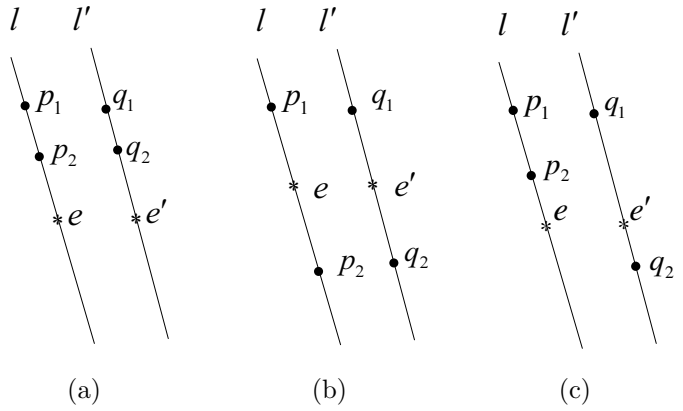


Figure 2: Illustration of corresponding half epipolar lines. (a) Strong realization of two correspondences lying on the same sides of the epipole. (b) Strong realization of two correspondences lying on different sides of the epipoles. (c) Weak realization of two correspondences.

As pointed out in [18], the correspondence between half epipolar lines is also helpful in finding correct correspondences. However, there is no feature matching algorithm which considers the cheirality property so far.

As to feature correspondence, two main steps are usually involved in most existing methods. In the first step, the feature points or keypoints are detected and a specified descriptor is obtained for each feature point within a local region. In the second step, for a feature point in one image, its matching point is found by comparing the similarity between its descriptor with those of its matching candidates in the other image. Many advanced feature point detectors have been proposed in the literature, such as the scale invariant feature transform (SIFT) [19] and the speeded up robust features (SURF) [20]. A simple method for finding feature correspondences is proposed in [19] where the ratio between the best match and the second best

match of each feature point is used to decide whether a matching pair is correct or not. This method uses no extra constraint and is easy to implement. Its disadvantage is that it is sensitive to the selection of the predefined threshold. That is to say if the threshold is high, many correct correspondences may be removed; if the threshold is low, many false correspondences may be kept in the matching result. Some studies have been carried out to solve this problem. In [21], a pairwise constraint is used to build an adjacent matrix and the correspondences are obtained by finding the principal eigenvector. In [22], false correspondences are rejected according to a geometric constraint. A global solution to a sparse correspondence problems is proposed in [12] where the problem of finding correspondences is formulated as an integer optimization problem and then solved by concave programming on a convex set. In [23], the feature correspondence problem is solved together with object recognition by an agglomerative correspondence clustering process. In [24], a triangle constraint is used to find more correspondences and detect false correspondences. The triangle constraint holds for images that are related by a homography, but does not hold for images under general perspective transformation. All the methods described above focus on improving the percentage of correct correspondences or finding more matching points; however none of them considers the spatial distribution of the correspondences. Studies in [1, 9] found that if correspondences only lie in small parts of stereo images, the estimation of the epipolar geometry is highly unstable. A further study [25] showed that uniformly distributed correspondences are very important, especially for the epipolar geometry estimation between a pair of stereo images. As for stereo images, a good set of feature correspondences therefore should have three characteristics: (1) The total number of correct correspondences is large; (2) They are matched correctly; and (3) The correspondences are distributed as uniformly as possible in the images.

The wide baseline images may only contain a small amount of corresponding parts and no correspondence exists in other parts, causing the spatial distribution of correspondences not uniform. Therefore, the uniform spatial distribution is not very important in wide baseline cases. However, for short baseline cases, the corresponding parts cover most parts of images and uniform distribution should be encouraged. By saying encouraging uniform spatial distribution we do not mean to enforce a constant spatial distribution, since the spatial distribution is related with the characteristic of the image and the adopted feature detector, and a constant spatial distribution is usually

unachievable. However, "encouraging uniform spatial distribution" means to attempt to find more correspondences in places where no correspondences or fewer correspondences exist and to find fewer correspondences in places where many correspondences exist. The method described in this paper finds feature correspondences in images without calibration. On the one hand, the epipolar geometry is useful in finding feature correspondences. On the other hand, the epipolar geometry estimation benefits from reliable and well distributed correspondences. Therefore, it is natural to carry out these two processes in a cooperative manner. Comparing with the state-of-the-art methods, our method performs well in the total number of correspondences, matching accuracy and spatial distribution. Moreover, we are able to obtain a highly accurate fundamental matrix as the byproduct from our method.

2. Cheirality Constraint under Epipole Polar Transformation

Suppose we have already calculated the fundamental matrix $\mathbf{F}$ associating with the two views. The strong realization requires that the correspondences are between half epipolar lines which are defined in $\mathbf{F}$ [17]. To determine this correspondence, we carry out a coordinate transformation called "epipole polar transformation". A similar idea is used in [26] for image rectification. In our method, we only use polar parametrization and do not carry out the image rectification. The epipole polar transformation maps a point from its Cartesian coordinate to the polar coordinate around the epipole. In practice, the epipole is usually not at infinity. If the epipole is at infinity, we can perform a projective transformation to the image so that the epipole in the transformed image is a finite point. Then epipole polar transformation can be applied to the transformed image. A discussion on this projective transformation is given in the Appendix. The epipole polar transformation first translates the epipoles of both the reference and the source images to the origin of the image using the following two matrices respectively

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & -e_x \\ 0 & 1 & -e_y \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{T}' = \begin{bmatrix} 1 & 0 & -e'_x \\ 0 & 1 & -e'_y \\ 0 & 0 & 1 \end{bmatrix} \tag{1}$$

where (e_x, e_y) and $\left(e'_x, e'_y\right)$ are the (x, y) coordinates of the epipole in the reference and the source images respectively. After the translation, the co-

ordinates of the two matching points p and q become

$$\begin{bmatrix} x'_p \\ y'_p \\ 1 \end{bmatrix} = \mathbf{T}\mathbf{p}, \quad \begin{bmatrix} x'_q \\ y'_q \\ 1 \end{bmatrix} = \mathbf{T}'\mathbf{q} \tag{2}$$

where $\left(x'_p, y'_p\right)$ and $\left(x'_q, y'_q\right)$ are the coordinates of p and q under the coordinate system where the origin is at the epipole of the image. Then, a polar transformation is performed on two images respectively such that

$$\begin{cases} x'_p = r\cos\theta \\ y'_p = r\sin\theta \end{cases}, \quad \begin{cases} x'_q = r'\cos\theta' \\ y'_q = r'\sin\theta' \end{cases} \tag{3}$$

where (θ, r) and (θ', r') are the polar coordinates of p and q after carrying out epipole polar transformation. Denote the fundamental matrix after the translation by $\mathbf{F_e}$, we have

$$\begin{aligned} \mathbf{F_e} &= \left(\mathbf{T}^{-1}\right)^{\mathrm{T}} \mathbf{F} \mathbf{T}'^{-\mathbf{1}} \\ &= \begin{bmatrix} 1 & 0 & e_x \\ 0 & 1 & e_y \\ 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}} \mathbf{F} \begin{bmatrix} 1 & 0 & e'_x \\ 0 & 1 & e'_y \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 & 0 \\ f_4 & f_5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned} \tag{4}$$

As the elements in the last row and the last column of $\mathbf{F_e}$ are all 0, we have

$$\begin{bmatrix} \cos\theta \\ \sin\theta \\ 1 \end{bmatrix}^{\mathrm{T}} \mathbf{F_e} \begin{bmatrix} \cos\theta' \\ \sin\theta' \\ 1 \end{bmatrix} = 0 \tag{5}$$

It can be proved that (see Appendix) a set of strongly realizable correspondences, (p_n, q_n), should satisfy

$$\exists\, \theta_c, \ \ \forall n : \operatorname{sgn}\big(\sin\left(\theta_n\right)\big) = \operatorname{sgn}\big(\sin\left(\theta'_n + \theta_c\right)\big) \tag{6}$$

This is a necessary condition of a strong realization (not sufficient) enforcing the correspondence between half epipolar lines. As $\theta_n = 0$ and $\theta'_n + \theta_c = 0$ hold in condition (6), θ_c can be obtained by solving the function of θ_c given by (5) where $\theta = 0$ and $\theta' + \theta_c = 0$. Two solutions satisfy this function. To decide which solution is correct, we use a set of point correspondences which are the inliers in the fundamental matrix estimation step as will be

described in Section 3.1. As most of these correspondences have real geometries (some false correspondences may not have real geometries), we choose the angle that can make more correspondences satisfy condition (6). For a given θ_c, the cheirality constraint enforces that all correspondences should satisfy condition (6).

3. Reliable Feature Correspondences

Given some feature points in the two images, we firstly select some seed correspondences or reliable feature correspondences. In this step we do not attempt to find all the correct correspondences in the image pair. We only extract some reliable correspondences which will be used as seed correspondences. The obtained disparities from these correspondences are used to guide finding more correspondences in the next step.

3.1. Candidate Correspondences Exploration

For each feature point in the reference image, we find its matching point among all feature points in the source image by applying the "winner-take-all" (WTA) strategy based on a similarity score which is measured by the Euclidean distance between two descriptors. Different feature points in the reference image may have the same matching point in the source image. Therefore, if several feature points in the reference image have the same matching point in the source image, only the correspondence which has the smallest feature distance will be considered and all other correspondences will be discarded based on the uniqueness constraint. Then a left-right consistency check is carried out. Feature correspondences passing the left-right consistency check constitute S, the set of candidate correspondences. A fundamental matrix $\mathbf{F}$ is then estimated from S by using the normalized 8-points algorithm [7] with least-median-of-squares (LMedS) [9] method.

After obtaining $\mathbf{F}$, we identify false correspondences by checking whether a feature correspondences is consistent with $\mathbf{F}$. To this end, we measure the symmetric epipolar distance [9] of a feature correspondence (p, q) with respect to $\mathbf{F}$. More specifically, we first calculate two corresponding epipolar lines by

$$\mathbf{F}\mathbf{q} = (l_1, l_2, l_3)^{\mathrm{T}} \tag{7}$$

and

$$\mathbf{F}^{\mathrm{T}}\mathbf{p} = (l'_1, l'_2, l'_3)^{\mathrm{T}} \tag{8}$$

where l_3 and l'_3 are normalized to 1. Then the symmetric epipolar distance is given by

$$\mathrm{dis}\,(p,q) = w\left|\mathbf{p}^{\mathrm{T}}\mathbf{F}\mathbf{q}\right| \tag{9}$$

where $w = \left(\frac{1}{l_1^2+l_2^2} + \frac{1}{l_1'^2+l_2'^2}\right)^{1/2}$. We check each correspondence in S whether its symmetric epipolar distance is smaller than a predefined value, ε

$$\mathrm{dis}\,(p,q) \leq \varepsilon \tag{10}$$

The parameter ε is the permitted maximum error of (p,q) measured in the pixel domain. As discussed in [6], the best value of parameter ε is related to the characteristics of the image, the feature detector and the matching method used. The robust estimation method as proposed in [6] can be used to find this best value. However, we found that for the SIFT feature and the WTA based matching as used in our method, ε can be empirically set to a fixed value 5 for all images and the final results are fairly constant when ε varies from 1 to 9. We use $\mathbf{F}$ and the correspondences satisfying (10) and adopt the method as described Section 2 to determine θ_c. Then we rule out the correspondences not satisfying (6) from S.

3.2. Adaptive Disparity Smoothing Filter

The polar disparity between a feature point p and its correspondence q is defined as follows

$$d\,(p) = r - r' \tag{11}$$

where r and r' are the radius in the polar coordinates of p and q as defined in Section 2. To filter out false correspondences, we employ a 3D smoothness constraint which states that any 3D points whose corresponding 2D points in images are close to each other should be close to each other in 3D space. This constraint is verified in many works and is also widely used in stereo matching algorithms [27, 28, 29]. 3D smoothness constraint can be encoded by the polar disparity smoothness in a neighborhood of a feature point. In the rest of this paper, the polar disparity is called disparity for short.

Consider a candidate match, (p_i, q_j), where i and j are the indices of feature points in the two images. According to the smoothness constraint in the disparity domain, when the match (p_i, q_j) is a correct match, the disparity between p_i and q_j should be similar to the disparities of its neighboring correspondences. On the other hand, when its disparity is different from those of its neighboring correspondences, it is more likely that this correspondence is a false match.

Correspondences in S links a set of feature points P in the reference image to a set of feature points Q in the source image. For a feature point $p \in P$, its neighboring feature points is denoted by $N(p)$. The number of neighboring feature points can be set to be the fixed proportion of total feature points as adopted in [23]. In our method, we set this number to be a fixed value 10 and find that the performance is quite similar with using the method in [23]. In our adaptive disparity smoothing filtering (ADSF) method, we decide whether the correspondence associated with p is correct or not by the disparity similarity between p and the pixels in $N(p)$. To this end, for each feature point p_k $(k = 1, 2, ..., 10)$ in $N(p)$, we first calculate its weight for a given p

$$w_k = \frac{\exp\left(-\frac{\|p,p_k\|}{\alpha}\right)}{\sum_{p_i \in N(p)} \exp\left(-\frac{\|p,p_i\|}{\alpha}\right)} \tag{12}$$

where $\|p, p_i\|$ is the spatial distance between p and p_i, and α is the normalization parameter.

Then, the feature points in $N(p)$ are sorted into an ascending order by their disparities, such that

$$\forall p_i, p_j \in N(p) \quad i < j \Rightarrow d(p_i) \leq d(p_j) \tag{13}$$

After that, we calculate the weighted median disparity of a feature point p by

$$d_{\text{wm}}(p) = d(p_{\text{wm}}) \tag{14}$$

where p_{wm} is a feature point in $N(p)$ given by

$$p_{\text{wm}} = \arg\min_{p_r} \left| \sum_{i=1}^{r} w_i - 0.5 \right| \tag{15}$$

Given $d_{\text{wm}}(p)$, we obtained a subset of the neighboring feature points whose disparities are similar to $d_{\text{wm}}(p)$ by

$$N_s(p) = \left(p_r \in N(p) \,\middle|\, |d(p_r) - d_{\text{wm}}(p)| < \beta\right) \tag{16}$$

where β is a threshold for removing false correspondences with large disparity errors and correspondences on different objects (see Fig. 3(a)). Finally, we decide whether the correspondence of p is correct or not by checking

$$\frac{|d(p) - d_{\text{wm}}(p)|}{\text{std}(N_s(p))} < \gamma \tag{17}$$

where std $(N_s(p))$ is the standard deviation of disparities in $N_s(p)$ and γ is a parameter controlling the permitted maximum disparity difference with respect to the standard deviation. The correspondence of p will be regarded as a correct correspondence if (17) is satisfied; otherwise it is regarded as a false correspondence.

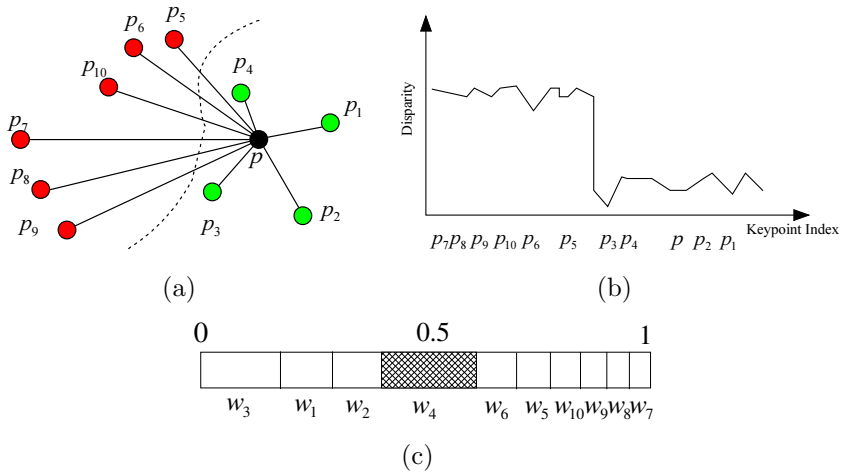


Figure 3: Illustration of ADSF. (a) Feature point p and the feature points in $N(p)$, $p_{\text{wm}} = p_4$; the dashline indicates an object boundary; the green dots are the feature points in $N_s(p)$; the red dots are not included in $N_s(p)$; (b) The disparities of feature points in (a); (c) The weights of feature points which are ranked in ascending order with respect to their disparities (the weight of p_{wm} is denoted by the gray region).

Actually, condition (17) is an amended version of using Mahalanobis distance to measure the similarity. The mean value as used in the Mahalanobis distance is sensitive to false correspondences with large disparity errors. Instead of using the mean value, we therefore use the weighted median disparity as defined in (14), a robust estimation against the values of the disparities of outliers. The weighted median disparity is similar to the median value of the disparities, but it favors the disparities of the feature points which are closer to the feature point p in the spatial domain. As Fig. 3 shows, although the median of $N(p)$ is $d(p_6)$, the disparity $d(p_4)$ obtained from the weighted median disparity is a more accurate estimate for $d(p)$. Actually, the median value can be obtained by setting w_r in (12) equal for all $p_i \in N(p)$.

Moreover, the standard deviation over set $N(p)$ used in the Mahalanobis distance is also sensitive to false correspondences with large disparity errors and correspondences on different objects. When using the standard deviation

over all correspondences in $N(p)$, the performance will be affected. Therefore, in this study, we use the standard deviation over a subset of $N(p)$ as defined in (16). Since (16) only selects the neighboring feature points whose disparities are similar with each other, the subset is less likely to contain false correspondences.

The parameter α in (12) is set to the mean distance between two neighboring feature points so that the weights in (12) are invariant to image scaling

$$\alpha = \frac{\sum_{p \in P} \sum_{p_i \in N(p)} \|p, p_i\|}{\sum_{p \in P} \#N(p)} \tag{18}$$

where $\#N(p)$ is the number of elements in $N(p)$. The average feature points density is defined as

$$D_{avg} = \frac{\#S}{W \times H} \tag{19}$$

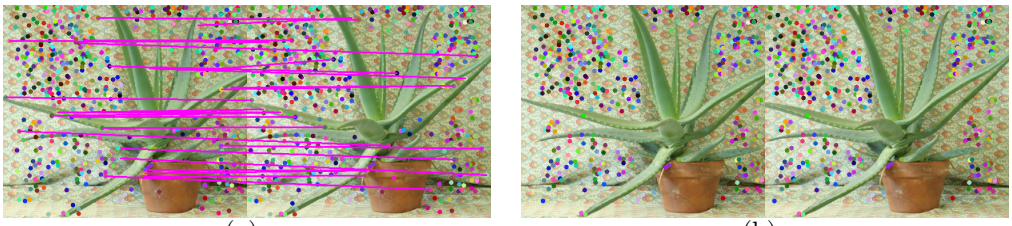
where W and H are the width and height of the image; $\#S$ is the total number of correspondences in S. Generally, a low density of feature points in the image indicates the feature points are far from each other, and the disparity difference between neighboring feature points may be large. Conversely, a high feature point density indicates that the disparities of neighboring feature points are similar to each other. Therefore, when D_{avg} is small, a large value β should be adopted, while a small β is for large D_{avg}. To this end, β is defined as

$$\beta = \frac{w_\beta}{D_{\text{avg}}} \tag{20}$$

where w_β is a parameter controlling β against D_{avg}. ADSF will permit a larger disparity difference in the neighboring feature points by setting w_β to a larger value. Fig. 4 shows an example of ADSF.

4. Finding More Correspondences Considering Spatial Distribution

In this step, we try to find more correspondences considering the spatial distribution of currently obtained correspondences. The correspondences are expected to be distributed as uniformly as possible in the image. Notice that we do not enforce but encourage a uniform spatial distribution. Because at some regions it is hard or even impossible to find correspondences reliably,



(a) (b)

Figure 4: Performance of the ADSF. Correspondences are denoted with the same color in images. Only false correspondences are denoted by lines. (a) Correspondences before performing ADSF. (b) Results by ADSF.

e.g., at textureless regions. We eliminate unreliable correspondences in such regions and make the correspondences distributed uniformly in regions where many correspondences exists.

After ADSF, we have obtained a set of seed correspondences which are considered to be correct. We denote the set of seed correspondences by S_{-}. Let P_{-} and Q_{-} be two sets of feature points in the reference and source images respectively relating to S_{-}. Correspondences to be found in this step is between the pixels in the complement of P_{-} and the pixels in the complement of Q_{-} (the two complements are denoted by $\overline{P}_{-}$ and $\overline{Q}_{-}$ respectively).

For efficiently finding correspondences, we predict the region of potential matching point of a given feature point, $p' \in \overline{P}_{-}$. In [30], the epipolar constraint is found to be useful for finding correspondences. In our work, we found that using the epipolar constraint, the cheirality constraint, and the disparity smoothness constraint together is more powerful. The epipolar constraint and the cheirality constraint require the location of q', the potential matching point of p', should satisfy (6) and (10). Let $d_N(p')$ be the set of disparities of the 10 nearest neighboring feature points (denoted by $N(p')$) of p' in P_{-}. According to the disparity smoothness constraint, the disparity between p' and its matching point should be close to the disparity in $N(p')$. Thus, the matching candidates of p' is defined as the feature points whose disparities with p' are within a specified range,

$$\min\big(d_N(p')\big) - \kappa \leq d(p') \leq \max\big(d_N(p')\big) + \kappa \tag{21}$$

where κ is the permitted disparity shift against the minimum and maximum disparities of $d_N(p')$. According to the disparity smoothness constraint in (17), κ is given by

$$\kappa = \gamma \cdot \mathrm{std}\big(N(p')\big) \tag{22}$$

where $\mathrm{std}\big(N\left(p'\right)\big)$ is the standard deviation of the disparities in $N\left(p'\right)$; γ is a parameter controlling the permitted maximum disparity difference with respect to the standard deviation, which is the same as the one used in Equation (17) The region of possible matching points in the source image for a pixel p' in the reference image is shown in Fig. 5. As there may be no feature point satisfying the requirements of Eqs. (6), (10), and (21), p' may have no matching points. If the feature distance between p' and q' is lower than a threshold $\tau\left(p', q'\right)$, the possible correspondence will be regarded as a correct correspondence. Following the uniqueness constraint, if more than one feature point in the reference image has the same matching feature point in the source image, only the one with the smallest feature distance will be selected.

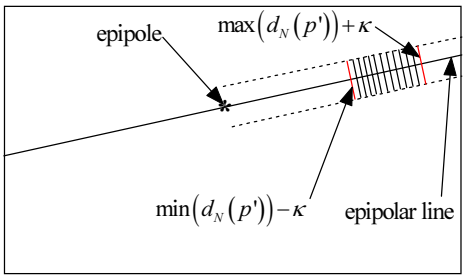


Figure 5: Illustration of the region of possible matching points in the source image for pixel p' in the reference image. Dash lines are the boundaries defined by the epipolar constraint and cheirality constraint; Red line are the boundaries defined by the disparity smoothness constraint.

To make correspondences to be distributed as uniformly as possible in images, a higher threshold $\tau\left(p', q'\right)$ is preferred where fewer correspondences exist, so that more new correspondences can be found. For a correspondence (p', q'), the number of seed correspondences within an $L \times L$ region around p' and q' are denoted by $\mathrm{num}\left(p'\right)$ and $\mathrm{num}\left(q'\right)$ respectively. L is given by $\sqrt{\frac{H \cdot W}{N_{\mathrm{C}}}}$, where H and W are the height and width of the image; N_{C} is the total number of seed correspondences in S_{-}. The threshold $\tau\left(p', q'\right)$ is then given by

$$\tau\left(p', q'\right) = \tau_r \left(1 - \frac{\mathrm{num}\left(p'\right) \cdot \mathrm{num}\left(q'\right)}{\max_{S_+}\big(\mathrm{num}\left(p'\right) \cdot \mathrm{num}\left(q'\right)\big)}\right) \tag{23}$$

where S_+ is the set of all possible correspondences; and τ_r is a predefined threshold. As $\tau\left(p', q'\right)$ is smaller than τ_r, any correspondence whose feature

distance is larger than τ_r will be rejected regardless of the spatial distribution. From this point, τ_r is the threshold for eliminating unreliable correspondences as mentioned at the beginning of this section. As the SIFT feature is a normalized feature descriptor, τ_r ranges from 0 to 1. $\tau\left(p', q'\right)$ becomes larger at places where the correspondence density is low and becomes smaller at places where the correspondence density is high.

Some false correspondences may exist in the result of the correspondence finding step described early in this section. To eliminate any false correspondences, we carry out the ADSF on the union of the set of correspondences obtained in the correspondence finding step and the set S_{-}. We keep performing the correspondences finding and ADSF iteratively until no more new correspondences can be found. We found this iterative method performs much better than just carrying out feature matching once. During the iterations, the correct correspondences may be checked many times by ADSF. In order to reduce the computational cost, we do not check a correspondence further if it is regarded as a correct match for three times by ADSF. For a feature point p' in the reference image, several candidate matching feature points usually exist in the region of possible matching points in the source image. The matching score of the correct correspondence may not always be the best among those of the candidate correspondences. Suppose that p' has more than one candidate matching points, q'_i, which are ranked by their matching scores. In the first round of matching, q'_1 is chosen. If the match (p', q'_1) is rejected by ADSF in this round of matching, q'_2 will be chosen in the next round, and the process may be repeated to go through all the candidates.

5. Fundamental Matrix Re-estimation

After finding more and uniformly distributed correspondences, the **F** matrix is re-estimated using the obtained correspondences to guide the next round of matching process. We found that when the re-estimated **F** matrix is close to the previously estimated **F** matrix, the obtained correspondences are nearly identical to those obtained in the previous round of matching. To estimate the similarity between two sequential estimated **F** matrices, we adopt the fundamental matrix comparison method as described in [9], which measures the similarity using the Sampson distance and calculates an error

called Sampson error in this paper. If this Sampson error is smaller than a threshold τ_s, the matching process will stop. Otherwise, the matching process stops when N_i, the maximum number of iteration, is reached. Setting the maximum number of iteration to a large value and setting τ_s to a small value would require more iteration before the matching process stops. In our experiments, we set $\tau_s = 1.0$ and $N_i = 4$ and found this setting works well for the testing images.

6. Algorithm Steps

Here we summarize the methods described in previous sections and give the steps of our algorithm.

1. Detect feature points and calculate their local descriptors.
2. Obtain candidate correspondences S using the method as described in Section 3.1.
3. Calculate the fundamental matrix $\mathbf{F}$ using S.
4. Remove false correspondences in S using $\mathbf{F}$ based on (10).
5. Remove false correspondences in S according to the cheirality constraint using $\mathbf{F}$.
   (a) Calculate the epipoles of two images.
   (b) Calculate the new image coordinate for each feature point using the translation given in (1).
   (c) Perform polar transformation and obtain the polar coordinates, (θ, r), for each feature point.
   (d) Calculate θ_c by using the method as described in Section 2.
   (e) Remove false correspondences in S using the obtained θ_c based on (10).
6. Carry out ADSF on the set of candidate correspondences (S) and obtain S_- as described in Section 3.2.
7. Find more correspondences using the method as described in Section 4 and add them into S_-.
8. If no correspondence is found then go to step 10; otherwise set S_- to be S, and go back to step 3.
9. Re-estimate the $\mathbf{F}$ matrix. If the maximum number of iteration is reached or the re-estimated $\mathbf{F}$ is close to that in the previous iteration, go to step 10; otherwise go back to step 4.
10. Output correspondence results.

7. Experimental Results

The proposed method is suitable for all fixed scene applications where two images are captured from different view points. The two images are therefore related by a fundamental matrix. When the images are captured from a single fixed point of view, they are related by a homography rather than a fundamental matrix. In such a case, the method described in this paper can still be used. However the set of equations used to estimate the fundamental matrix is underdetermined and its solution is not uniquely defined. Therefore, we choose the solution which has the minimum norm to be the fundamental matrix. Three performance measurements are used in our experiments, i.e., the percentage of correct correspondences, the total number of correct correspondences, and the spatial distribution of correspondences. We compared our method with the latest developed algorithms: the triangle constraint measurement algorithm (TCM) [24], the agglomerative correspondence clustering method (ACC) [23], and the widely used ratio of distances measurement (RDM) method [19].

7.1. Performance Evaluation

In the first, we do not consider the fundamental matrix re-estimation which will be discussed in the following section. Thus, we have three parameters, i.e., w_β, γ, and τ_r. We tune these parameters using the Oxford dataset [31] with the ground truth. Correspondences in the ground truth are generated from the provided homography.

For an obtained correspondence (p, q), a corresponding region pair is defined by the two 3×3 regions around p and q respectively. For each pixel within the corresponding region in one image, we check if its ground truth matching point lies within its corresponding region in the other image. If the ground truth matching points of all pixels in the corresponding region pair are all outside of the corresponding region pair, the match (p, q) will be regarded as a false match; otherwise it will be regarded as a correct match. This measurement is robust against small location errors of feature points at depth boundaries. Fig. 6 shows the percentage of correct correspondences obtained by our method when varying parameters w_β and γ. The performance w.r.t. parameter τ_r are shown in Fig. 7.

We set $w_\beta = 0.2$, $\gamma = 2$, and $\tau_r = 0.3$ empirically and carry out the comparison between the our method and other methods using the Middlebury datasets [32]. Fig. 8 shows the reference images of the datasets from the Middleburry website. The ground truth correspondences can be obtained

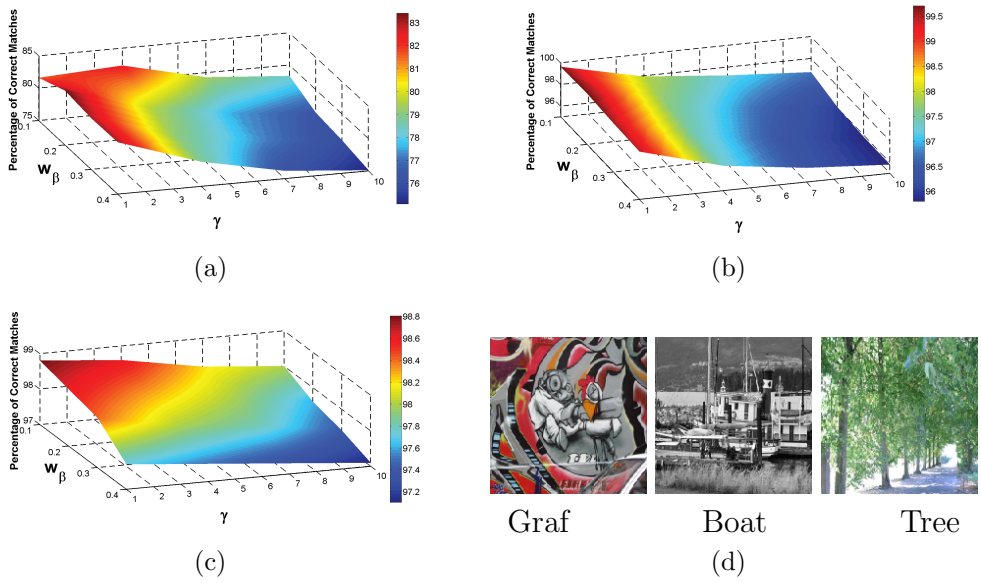
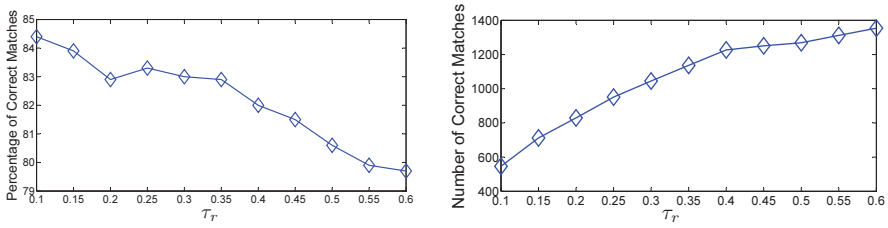


Figure 6: Matching performance when varying w_β and γ by fixing $\tau_r = 0.3$. (a) The performance on the image "Graf". (b) The performance on the image "Boat". (c) The performance on the image "Tree". (d) The images used in the parameter estimation step.

from the provided disparity maps. In the comparisons, the best parameters in the competing methods are used as described in [24, 23, 19]. The performance evaluation is carried out on the percentage of correct correspondences, the total number of correct correspondences, and the spatial distribution of the correspondences. The spatial distribution of correspondences is measured based on the standard deviation of the point density [25], which is low when the result is good. In our paper, we adopt the SIFT feature detector and descriptor in the our method. For all other methods, we adopt their default feature detectors and descriptors as described in their original studies. As RDM and TCM also use the SIFT feature detector and descriptor, we run the experiments on the same feature points for RDM, TCM and our method. We also carried out further experiments using the same feature points from the SIFT detector and descriptor as adopted in the other three methods for the ACC method. The corresponding results are denoted as ACC(SIFT). In these experiments, we rotate the source image of each image pair with a random angles to make the experiments more general. The angle is the same for all matching algorithms for fairness. Fortunately, as the features used in all these algorithms are robust to the rotation, the rotation angel does not



(a) (b)

Figure 7: Performance of our method on the "Graf" dataset with respect to parameter τ_r by setting w_β and γ to constants: $w_\beta = 0.2$, and $\gamma = 2$. (a) The percentage of correct correspondences with respect to τ_r. (b) The number of correct correspondences with respect to τ_r.

impact the results from any of these algorithms.

Table 1 and Table 2 give the performance comparison results between the proposed uniform distribution matching (UDM) method and the competing algorithms. Compared with the other three algorithms, our method greatly improves results with respect to the spatial distribution of the correspondences. The number of correspondences obtained from our method is much more than those from other methods. The performance of our method with respect to percentage of correct correspondences is comparable with that of ACC and is much better than those of ACC(SIFT), RDM and TCM.

To check the effectiveness of using the cheirality constraint on matching, we compared the results of our algorithm using the cheirality constraint and the results without using it. According to [2], when the epipoles in the two images are both outside of the images, any set of correspondences between two image points which satisfy the requirement of a weak realization will also meet the requirement of a strong realization condition, i.e., condition (6). There is no need for checking the performance on images in this case. We therefore choose some images where the epipole obtained from initial estimation lies within the image to verify the effectiveness of using the cheirality constraint. Table 4 shows the results when using the cheirality constraint and the results without using it. The comparison on the results shows that the matching correctness is improved when using the cheirality constraint.

The runtime of the algorithm relates to the image size and the number of feature points. In the experiments, the runtime of our algorithm on the "Tsukuba" image pair with 384×288 pixels and 336 features is 6.0 seconds on a platform with 2.0 GHz CPU and 4 GB RAM without using any SSSE

(a) teddy (b) cones (c) cloth (d) laundry

(e) art (f) baby (g) venus (h) dolls

(i) aloe (j) tsukuba (k) bowling (l) books

Figure 8: The images used to evaluate the matching performance (only the reference image of the image pairs are shown).

Table 1: The comparison on the percentage of correct correspondences (the best performance is colored in red).

	Images												
Method	teddy	cones	tsukuba	venus	art	cloth	laundry	dolls	aloe	books	bowling	baby	averaging
RDM	95.6	95.9	92.0	90.1	85.2	97.2	56.2	94.2	87.5	74.3	94.2	94.0	88.03
TCM	89.2	81.6	91.8	98.5	76.4	94.7	75.0	89.5	90.6	91.1	79.3	90.5	87.35
ACC	83.4	96.2	94.1	94.5	76.5	97.1	88.0	95.5	92.5	92.4	93.0	94.4	91.47
ACC(SIFT)	87.6	94.6	92.0	96.5	79.0	98.1	73.2	92.2	91.2	87.6	91.6	92.5	89.7
UDM	93.6	96.2	97.8	98.1	85.5	99.0	83.0	94.3	97.6	89.1	90.5	94.7	93.28

Table 2: The comparison on the total number of correct correspondences (the best performance is colored in red).

	Images												
Methods	teddy	cones	tsukuba	venus	art	cloth	laundry	dolls	aloe	books	bowling	baby	averaging
RDM	201	312	297	282	166	623	191	299	452	176	139	203	278
TCM	112	122	123	62	154	377	213	248	451	113	46	179	183
ACC	171	208	132	138	74	238	127	189	217	145	45	94	148
ACC(SIFT)	297	378	415	277	216	773	204	332	563	248	157	225	340
UDM	315	460	457	340	292	994	447	564	815	401	198	304	466

Table 3: The comparison on the standard deviation of point density (the best performance is colored in red).

	Images												
Methods	teddy	cones	tsukuba	venus	art	cloth	laundry	dolls	aloe	books	bowling	baby	averaging
RDM	1.41	1.42	1.44	1.38	1.56	1.29	1.41	1.22	1.14	1.39	2.06	1.54	1.44
TCM	2.11	1.51	1.49	1.57	1.95	1.54	2.44	2.74	1.50	1.65	1.88	1.68	1.84
ACC	1.52	1.36	1.38	1.52	1.64	1.28	1.48	1.20	1.23	1.73	1.45	1.67	1.46
ACC(SIFT)	1.39	1.38	1.41	1.47	1.72	1.33	1.54	1.36	1.40	1.55	1.75	1.51	1.48
UDM	1.31	1.32	1.35	1.31	1.46	1.23	1.23	1.10	1.30	1.27	1.69	1.41	1.33

Table 4: The comparison between using and without using the cheirality constraint.

	Matching Accuracy	
Images	Without Using the Cheirality Constraint	Using the Cheirality Constraint
graf	81.4%	83.6%
boat	97.2%	98.7%
art	82.3%	85.5%
cloth	95.3%	99.0%

instruction sets or GPU acceleration. The runtime of other algorithms on the same image pair and the same platform are 1.7 seconds for RDM, 8.6 seconds for TCM, 9.4 seconds for ACC and 11.6 for ACC(SIFT). The ACC, TCM, ACC(SIFT) algorithms are implemented in MATLAB with C++ codes and the other two are implemented in C++.

Some more experiments are carried out on other real images including different types of objects and the viewpoint varying from close views to distant views. The depth of objects in some images varies smoothly (first two pairs). But some of the images contain objects with large depth jumps (last two pairs). The reference image (b) of the first image pair in Fig. 9 is rotated by 180 degrees. The parameters used here are the same as before, i.e., $\tau_r = 0.3$, $w_\beta = 0.2$, and $\gamma = 2$. The results are also very convincing (see Fig. 9).

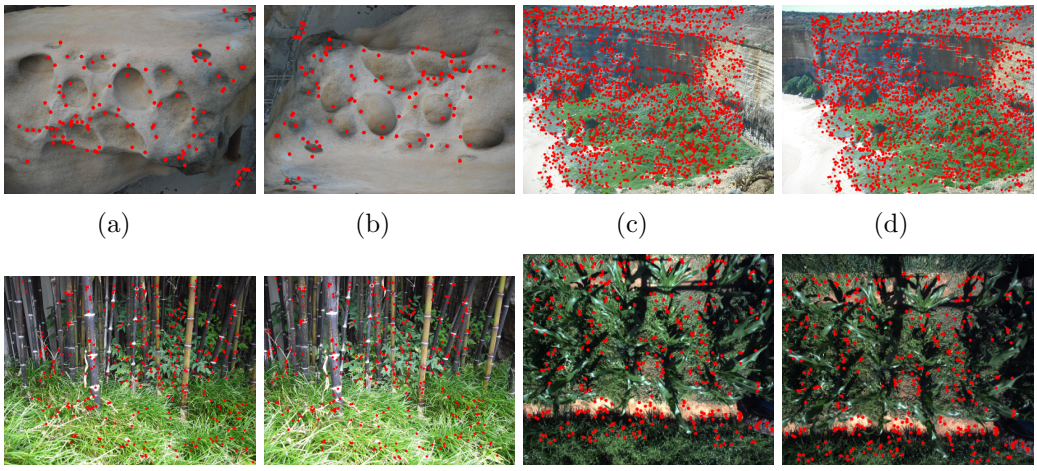


Figure 9: The matching results of other real images captured from different view points (red points are the matched feature points in image pairs).

7.2. Fundamental Matrix Re-estimation

As we use the fundamental matrix to guide the matching process. Our algorithm would be affected by an incorrectly estimated fundamental matrix. The reason is that the correspondence of a feature point may not locate inside the region of possible matching points predicted by an incorrect epipolar geometry. Our remedy for this problem is setting ε to a large value and

decreasing it gradually as the iteration goes. Using this method, our algorithm is robust to incorrectly estimated initial fundamental matrix. To show this, we run our algorithm with an arbitrarily picked matrices (satisfying rank 2 constraint) as the initial fundamental matrix, such as, $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and compare the estimate fundamental matrix with the groundtruth. Fig. 10 shows the results.

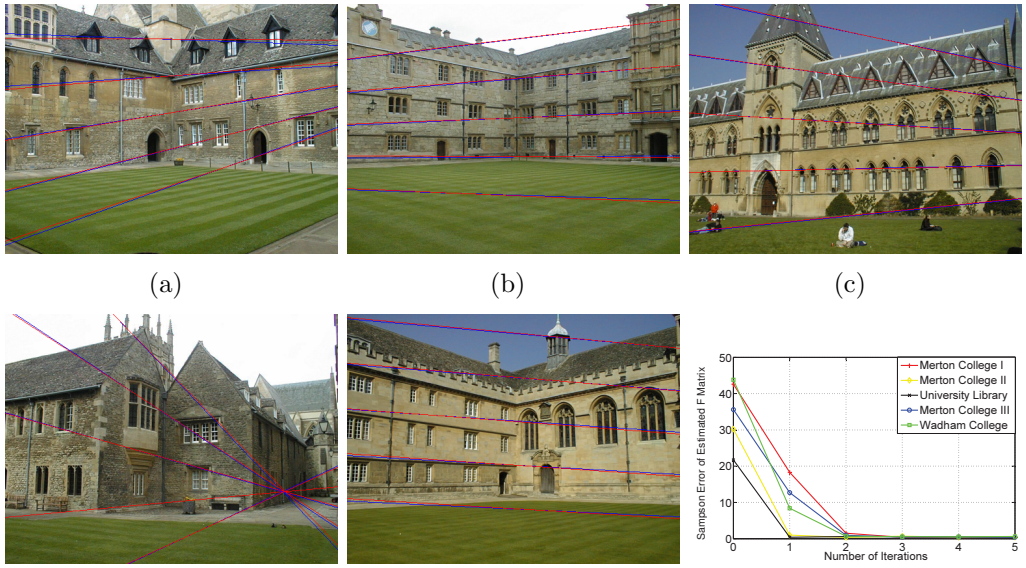


Figure 10: Results of the fundamental matrix estimation. (a)-(e): Red lines are epipolar lines obtained from the groundtruth fundamental matrix. Blue lines are epipolar lines obtained from the estimated fundamental matrix. As the estimated fundamental matrix is close to the groundtruth fundamental matrix, it is not easy to see the difference between the red and blue lines. (f): The Sampson error of the estimated fundamental matrix compared with the groundtruth fundamental matrix. Images are from VGG website [31] based on works [33, 34].

In practice, we found that the fundamental matrix estimated from initial correspondences by mutual check is usually good enough for our method to find almost all possible correspondences in one run. That is the reason why our method provides good results (see Table 1 to Table 2) even without

using iteration. However, the re-estimated fundamental matrix is always much better than that from initial matching points.

7.3. Discussion

Our algorithm carries out feature correspondence and fundamental matrix estimation in a joint fashion. One common problem in the fundamental matrix estimation methods is that when the correspondences are crowed in one part of the image and leave other parts of the image empty, the estimate fundamental matrix would overfitting this large group of crowded correspondences. Using this set of correspondences for fundamental matrix estimation is dangerous. Our algorithm finds a large number of features with high accuracy and more important features are encouraged to be distributed cover the image as much as possible. This enables us to provide a correct fundamental matrix (see Fig. 11). In Fig. 11 the same fundamental matrix estimation method (normalized 8 points algorithm with LMedS[9]) and rectification method [35] are applied to two sets of correspondences. One set of correspondences is obtained by using the growing scheme in Section 4 for encouraging uniform distribution, the other is obtained without using it. Thanks to the uniform distribution scheme, better rectification results are obtained. However, this benefit does not come without a cost. In the place where no matching point exists, the uniform distribution constraint may sometimes may introduce incorrect correspondences. Although most of these incorrect correspondences would be filtered out by ADSF, it is not suitable for wide baseline matching where the scene changes a lot.

8. Conclusions

In this paper, we present a new method for finding feature correspondences between two images. We propose an epipole polar transformation method to extend the cheirality constraint into finding feature correspondences. The proposed ADSF algorithm efficiently removes false correspondences using the smoothness constraint on the polar disparity difference of neighboring correspondences. More correspondences are then found based on the epipolar constraint and the cheirality constraint considering the spatial distribution of the correspondences. An iterative method integrating the finding of more correct correspondences and the removal of false correspondences is used. Experiments are carried out using real images with ground truth matching information. Compared with the state-of-the-art methods, our method performs very well on all these three aspects: percentage

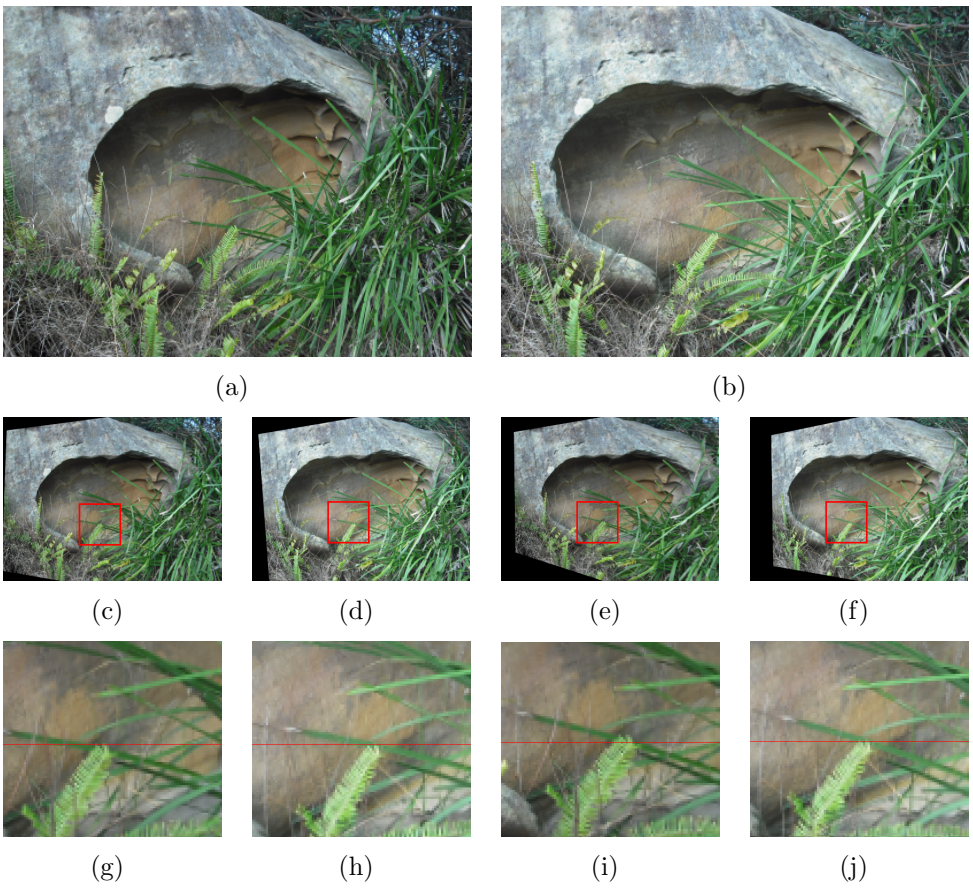


Figure 11: Comparison between the rectification results by encouraging uniform distribution and not encouraging uniform distribution. When encouraging uniform distribution in the matching process, the corresponding pixels are correctly aligned along horizontal lines. (a) and (b): Stereo images. (c) and (d) Results without considering the distribution. (e) and (f) Results when considering the distribution. (g) and (h): Close view of (c) and (d). (i) and (j): Close view of (e) and (f).

of correct correspondences, number of correspondences, and correspondence distribution. The excellent performance of the proposed method on other real images also proves the efficiency of the method.

9. Web Demo

A web demo of the algorithm proposed in this paper is available at:
`http://vision-cdc.csiro.au/f_match/`

Appendix

Projective Transformation When the Epipole Is an Infinite Point.
Suppose that the epipole is an infinite point, $\begin{bmatrix} m & n & 0 \end{bmatrix}^{\mathrm{T}}$, where $m^2+n^2 = 1$. We can apply the following projective transformation to the image, so that the epipole can be projected onto a finite point, $\begin{bmatrix} \frac{mk}{2} & \frac{nk}{2} & 1 \end{bmatrix}^{\mathrm{T}}$

$$\mathbf{T}_\infty = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ (mk)^{-1} & (nk)^{-1} & 1 \end{bmatrix} \tag{24}$$

where k is a large positive value comparing with the image size. For exmaple, when the image size is about 1000×1000 pixels, k can be set to 10^6. We choose a large positive value for two reasons. First, a large k makes the projective transformation $\mathbf{T}_\infty$ to be quasi-affine with respect to image size, which will not split the image. The transformation $\mathbf{T}_\infty$ maps the line $L_\infty : nx + my + mnk = 0$ to infinity. When k is large enough, line L_∞ will be out of the image. According to [35], the transformation is quasi-affine with respect to all pixels in the image. Second, noticing that when k is large, $\mathbf{T}_\infty$ is close to the identity matrix, or when k is large, the difference between the original image and the transformed image is small.

Proof of Equation (6).
According to the Theorem 2 in [17], a strong realization for a set of correspondences, (p_n, q_n), should satisfy

$$\begin{aligned} &\exists \mathbf{F}' \in R^{3\times 3}, \mathrm{rank}\ \mathbf{F}' = 2\,,\ \exists \mathbf{e} \in R^3, \mathbf{e} \neq \mathbf{0}, \mathbf{e}^{\mathrm{T}}\mathbf{F}' = \mathbf{0}, \\ &\forall n\ \exists w_n > 0 : w_n \mathbf{e} \times \mathbf{p}_n = \mathbf{F}'\mathbf{q}_n \end{aligned} \tag{25}$$

We now prove that a set of weak realized correspondences satisfies (25), only when it satisfies condition (6).

As the origin of the new coordinate shifts to the epipole under the epipole polar transformation, we will have two other conditions: $\mathbf{e}^{\mathrm{T}}\mathbf{F}' = \mathbf{0}$ and $\mathbf{F}'\mathbf{e} = \mathbf{0}$, where $\mathbf{e} = [0, 0, 1]^{\mathrm{T}}$ and $\mathbf{F}'\mathbf{e} = \mathbf{0}$. Therefore, we have

$$\mathbf{F}' = \begin{bmatrix} f_1' & f_2' & 0 \\ f_4' & f_5' & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{26}$$

The equation $w_n\mathbf{e} \times \mathbf{p}_n = \mathbf{F}'\mathbf{q}_n$ multiplied by $\mathbf{p}_n^{\mathrm{T}}$ from the left is

$$\mathbf{p}_n^{\mathrm{T}}\mathbf{F}'\mathbf{q}_n = w_n\mathbf{p}_n^{\mathrm{T}}\mathbf{e} \times \mathbf{p}_n = 0$$

Since all $\lambda\mathbf{F}'$ ($\lambda \neq 0$) satisfy this equation, the sign of $\mathbf{F}'$ is lost in $\mathbf{F}_{\mathbf{e}}$ ($\mathbf{F}_{\mathbf{e}} = \lambda\mathbf{F}'$). Substituting the coordinates of a matching pair into $w_n\mathbf{e} \times \mathbf{p}_n = \mathbf{F}'\mathbf{q}_n$ ($\mathbf{e} = [0, 0, 1]^{\mathrm{T}}$), we have

$$\begin{aligned} \begin{bmatrix} -\sin\theta_n \\ \cos\theta_n \\ 0 \end{bmatrix} &= k \begin{bmatrix} f_1' & f_2' & 0 \\ f_4' & f_5' & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \sin\theta_n' \\ \cos\theta_n' \\ \frac{1}{r'} \end{bmatrix} \\ &= k \begin{bmatrix} f_1' & f_2' & 0 \\ f_4' & f_5' & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{R}_{\theta_{\mathbf{c}}}^{-\mathbf{1}} \begin{bmatrix} \cos(\theta_n' + \theta_c) \\ \sin(\theta_n' + \theta_c) \\ \frac{1}{r'} \end{bmatrix} \end{aligned} \tag{27}$$

where $k = \frac{r'}{rw_n} > 0$, $\mathbf{R}_{\theta_{\mathbf{c}}}^{-\mathbf{1}}$ is the inverse of the rotation matrix with the angle θ_c, $\mathbf{R}_{\theta_{\mathbf{c}}}$, given by

$$\mathbf{R}_{\theta_{\mathbf{c}}} = \begin{bmatrix} \cos\theta_c & -\sin\theta_c & 0 \\ \sin\theta_c & \cos\theta_c & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{28}$$

Let θ_c be the angle satisfying $\cos\theta_c = \frac{-f_2'}{\sqrt{{f_1'}^2 + {f_2'}^2}}$ and $\sin\theta_c = \frac{-f_1'}{\sqrt{{f_1'}^2 + {f_2'}^2}}$. Thus, (27) can be written as

$$\begin{bmatrix} -\sin\theta_n \\ \cos\theta_n \\ 0 \end{bmatrix} = k \begin{bmatrix} 0 & a & 0 \\ b & c & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \cos(\theta_n' + \theta_c) \\ \sin(\theta_n' + \theta_c) \\ \frac{1}{r'} \end{bmatrix} \tag{29}$$

where $a = -\sqrt{{f_1'}^2 + {f_2'}^2} < 0$ ($\sqrt{{f_1'}^2 + {f_2'}^2} \neq 0$ follows from $\mathrm{rank}(\mathbf{F}') = 2$). Further, because

$$\sin\theta_n = -ka\sin(\theta_n' + \theta_c) \tag{30}$$

and $k > 0$, condition (6) is proved.

Acknowledgments

We thank Mohammad Hosseini for his comments on this paper. Tan was partially supported by the China Scholarship Council. Sun was partially supported by the CSIRO's Transformational Biology Capability Platform. A preliminary version of the work appeared in [36].

References

[1] Z. Zhang, R. Deriche, O. Faugeras, Q. Luong, A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry, Artificial Intelligence 78 (1-2) (1995) 87–119.

[2] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, 2003.

[3] O. Faugeras, B. Mourrain, On the geometry and algebra of the point and line correspondences between N images, in: IEEE International Conference on Computer Vision, 1995, pp. 951–956.

[4] Q. Luong, O. Faugeras, The fundamental matrix: Theory, algorithms, and stability analysis, International Journal of Computer Vision 17 (1) (1996) 43–75.

[5] R. Hartley, R. Gupta, Computing matched-epipolar projections, in: IEEE Conference on Computer Vision and Pattern Recognition, 1993, pp. 549–555.

[6] P. Torr, D. Murray, The development and comparison of robust methods for estimating the fundamental matrix, International Journal of Computer Vision 24 (3) (1997) 271–300.

[7] R. Hartley, In defense of the eight-point algorithm, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (6) (1997) 580–593.

[8] T. Huang, A. Netravali, Motion and structure from feature correspondences: A review, Proceedings of the IEEE 82 (2) (1994) 252–268.

[9] Z. Zhang, Determining the epipolar geometry and its uncertainty: A review, International Journal of Computer Vision 27 (2) (1998) 161–195.

[10] X. Armangué, J. Salvi, Overall view regarding fundamental matrix estimation, Image and Vision Computing 21 (2) (2003) 205–220.

[11] E. Nishimura, G. Xu, S. Tsuji, Motion segmentation and correspondence using epipolar constraint, in: Asian Conference on Computer Vision, 1993, pp. 199–204.

[12] J. Maciel, J. Costeira, A global solution to sparse correspondence problems, IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (2) (2003) 187–199.

[13] R. I. Hartley, Chirality, International Journal of Computer Vision 26 (1) (1998) 41–61.

[14] D. Nistér, Calibration with robust use of cheirality by quasi-affine reconstruction of the set of camera projection centres, in: IEEE International Conference on Computer Vision, Vol. 2, 2001, pp. 116–123.

[15] W. Xu, J. Mulligan, Robust relative pose estimation with integrated cheirality constraint, in: International Conference on Pattern Recognition, 2008, pp. 1–4.

[16] D. Nister, An efficient solution to the five-point relative pose problem, in: IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, 2003, pp. II–195.

[17] T. Werner, T. Pajdla, Cheirality in epipolar geometry, in: IEEE International Conference on Computer Vision, Vol. 1, 2001, pp. 548–553.

[18] T. Werner, T. Pajdla, Oriented matching constraints, in: BMVC, 2001, pp. 1–10.

[19] D. Lowe, Object recognition from local scale-invariant features, in: IEEE International Conference on Computer Vision, Vol. 2, 1999, pp. 1150–1157.

[20] H. Bay, T. Tuytelaars, L. Van Gool, SURF: Speeded up robust features, European Conference on Computer Vision (2006) 404–417.

[21] M. Leordeanu, M. Hebert, A spectral technique for correspondence problems using pairwise constraints, in: IEEE International Conference on Computer Vision, Vol. 2, 2005, pp. 1482–1489.

[22] M. Brown, D. Lowe, Recognising panoramas, in: IEEE International Conference on Computer Vision, Vol. 2, 2003, pp. 1218–1225.

[23] M. Cho, J. Lee, K. Lee, Feature correspondence and deformable object matching via agglomerative correspondence clustering, in: IEEE International Conference on Computer Vision, 2009, pp. 1280–1287.

[24] X. Guo, X. Cao, Good match exploration using triangle constraint, Pattern Recognition Letters 33 (7) (2012) 872–881.

[25] J. Seo, H. Hong, C. Jho, M. Choi, Two quantitative measures of inlier distributions for precise fundamental matrix estimation, Pattern Recognition Letters 25 (6) (2004) 733–741.

[26] M. Pollefeys, R. Koch, L. Van Gool, A simple and efficient rectification method for general motion, in: International Conference on Computer Vision, Vol. 1, IEEE, 1999, pp. 496–501.

[27] J. Sun, N.-N. Zheng, H.-Y. Shum, Stereo matching using belief propagation, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25 (7) (2003) 787–800.

[28] V. Kolmogorov, R. Zabih, Computing visual correspondence with occlusions using graph cuts, in: IEEE International Conference on Computer Vision, Vol. 2, IEEE, 2001, pp. 508–515.

[29] H. Hirschmuller, D. Scharstein, Evaluation of cost functions for stereo matching, in: IEEE Conference on Computer Vision and Pattern Recognition, IEEE, 2007, pp. 1–8.

[30] R. Hartley, A. Zisserman, Multiple view geometry in computer vision, Vol. 2, Cambridge Univ Press, 2000.

[31] Oxford Vgg `http://www.robots.ox.ac.uk/~vgg/`.

[32] Middlebury Stereo Website `vision.middlebury.edu/stereo/`.

[33] T. Werner, A. Zisserman, Model selection for automated architectural reconstruction from multiple views, in: Proceedings of the British Machine Vision Conference, 2002, pp. 53–62.

[34] T. Werner, A. Zisserman, New techniques for automated architectural reconstruction from photographs, in: ECCV, 2002, pp. 541–555.

[35] R. I. Hartley, Theory and practice of projective rectification, International Journal of Computer Vision 35 (2) (1999) 115–127.

[36] X. Tan, C. Sun, X. Sirault, R. Furbank, T. D. Pham, Feature correspondence with even distribution, in: International Conference on Digital Image Computing: Techniques and Applications, 2012, pp. 1–7.