

Balanced Exploration and Exploitation Model Search for Efficient Epipolar Geometry Estimation

Liran Goshen and Ilan Shimshoni, *Member, IEEE*

Abstract—The estimation of the epipolar geometry is especially difficult when the putative correspondences include a low percentage of inlier correspondences and/or a large subset of the inliers is consistent with a degenerate configuration of the epipolar geometry that is totally incorrect. This work presents the Balanced Exploration and Exploitation Model (BEEM) search algorithm, which works very well especially for these difficult scenes. The algorithm handles these two problems in a unified manner. It includes the following main features: 1) balanced use of three search techniques: global random exploration, local exploration near the current best solution, and local exploitation to improve the quality of the model, 2) exploitation of available prior information to accelerate the search process, 3) use of the best found model to guide the search process, escape from degenerate models, and define an efficient stopping criterion, 4) presentation of a simple and efficient method to estimate the epipolar geometry from two scale-invariant feature transform (SIFT) correspondences, and 5) use of the locality-sensitive hashing (LSH) approximate nearest neighbor algorithm for fast putative correspondence generation. The resulting algorithm when tested on real images with or without degenerate configurations gives quality estimations and achieves significant speedups compared to the state-of-the-art algorithms.

Index Terms—Fundamental matrix, robust estimation.

1 INTRODUCTION

THE estimation of the epipolar geometry is an important task in computer vision. The RANdom SAMple Consensus algorithm (RANSAC) [1] has been widely used in computer vision, in particular for recovering the epipolar geometry. The estimation of the epipolar geometry is especially difficult in two cases. The first difficult situation is when the putative correspondences include a low percentage of inliers. The other problem occurs when a large subset of inliers is consistent with a degenerate epipolar geometry.

In the first case, the number of required iterations is usually high. A popular stopping criterion in a RANSAC-like algorithm is

$$I = \frac{\log(1-p)}{\log(1-\alpha^s)} \approx \frac{-\log(1-p)}{\alpha^s}, \quad (1)$$

where s is the size of the random sample, I is the number of iterations, α is the inlier rate, and p is the required probability [1], [2]. For example, for $\alpha = 0.15$, the number of needed iterations for $s = 7$, $s = 3$, and $s = 2$ are $I = 2,695$, 296 , $I = 1,362$, and $I = 202$, respectively, for $p = 0.99$.

- L. Goshen is with the Faculty of Industrial Engineering and Management, Technion-Israel Institute of Technology, 32000 Haifa, Israel.
E-mail: lirang@tx.technion.ac.il.
- I. Shimshoni is with the Department of Management Information Systems, University of Haifa, Haifa 31905, Israel.
E-mail: ishimshoni@mis.haifa.ac.il.

Manuscript received 1 Mar. 2006; revised 30 Nov. 2006; accepted 30 July 2007; published online 29 Aug. 2007.

Recommended for acceptance by J. Oliensis.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0196-0306. Digital Object Identifier no. 10.1109/TPAMI.2007.70768.

Several approaches have been suggested to speed up the RANSAC algorithm. The locally optimized RANSAC (LO-RANSAC) [3] exploits the fact that the model hypothesis from an uncontaminated minimal sample is often sufficiently near to the optimal solution and a local optimization step is carried out only if a new maximum in the size of the support set of the current sample model has occurred. The number of samples that the LO-RANSAC performs achieves good agreement with the theoretical predictions of (1).

In [4], random sampling was replaced by guided sampling. The guidance of the sampling is based on the correlation score of the correspondences. The idea of guided sampling is very promising. However, the correlation score provides only weak evidence to the correctness of the matches. Using this method with a more powerful score can yield more significant speedups. This was achieved in the Progressive Sample Consensus (PROSAC) [5] algorithm, which exploits the similarity between scale-invariant feature transform (SIFT) [6] features. Generally speaking, PROSAC exploits the linear ordering defined on the set of correspondences by the similarity function used in establishing putative correspondences. PROSAC samples are drawn from progressively larger sets of top-ranked correspondences. In our previous work [7], the algorithm generates a set of weak motion models (WMMs). These models approximate the motion of points between the two images using a smaller number of matches and, thus, are computationally cheaper to detect. These WMMs are used to establish probabilities that matches are correct. The RANSAC process uses these probabilities to guide the sampling. WMMs are especially useful when no good prior knowledge is available for this task.

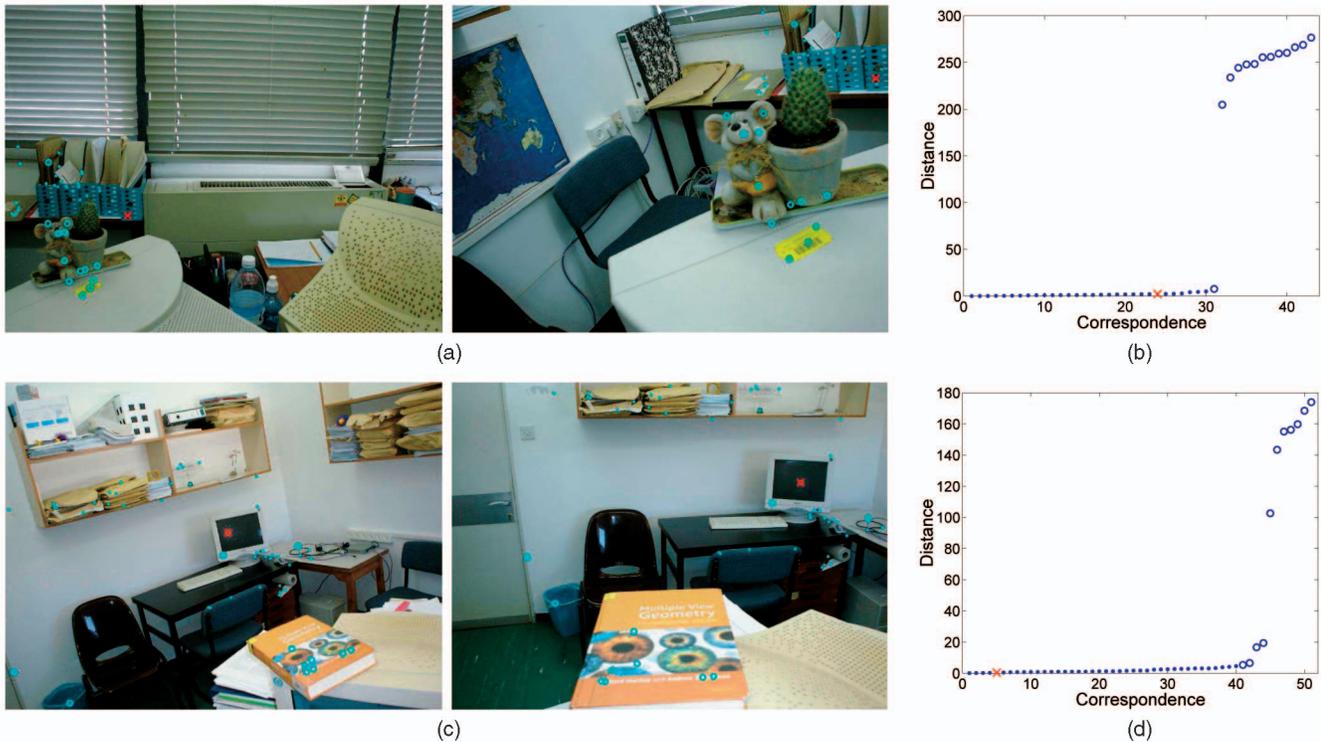


Fig. 1. Image scenes and quality evaluation. The graph shows the distance from the epipolar surface for the inliers (degeneracy inliers are denoted by dots, whereas the nondegeneracy inliers are denoted by circles). The nondegeneracy inliers lie very far from the surface. (a) Flowerpot scene. (b) Result evaluation of the flowerpot scene. (c) Book scene. (d) Result evaluation of the book scene.

Assigning probabilities to putative correspondences was also used to evaluate the score of possible solutions. Domke and Aloimonos [8] used probabilities based on Gabor filters for this purpose.

In [9], [10], the use of three affine region-to-region matches to estimate the epipolar geometry in each RANSAC sample was suggested. To hypothesize a model of the epipolar geometry, a random sample of three region correspondences is drawn. Three region correspondences give nine point correspondences. These are then used to estimate the fundamental matrix F using the linear eight-point algorithm [11]. Under this framework, s in (1) is changed from seven to three, considerably reducing the number of iterations. In [12], which was performed concurrently with our work [13], two pairs of affine matches were used. In that case, it was assumed that some information is available about the internal calibration matrices.

Another approach for dealing with a large number of outliers is to substitute the combinatorial complexity of finding a correct set of matches with a search in the motion parameter space, looking for a set of parameters that is supported by a large set of matches [14]. This approach is most effective when dealing with constrained motion.

The second difficult situation occurs when a large subset of inliers is consistent with a degenerate epipolar geometry. This situation often occurs when the scene includes a degeneracy or close to degenerate configurations. In this case, standard epipolar geometry estimation algorithms often return an epipolar geometry with a high number of inliers that is, however, totally incorrect. The estimation of the fundamental matrix in such situations has been

addressed before. In [15], a RANSAC-based algorithm for robust estimation of epipolar geometry in the possible presence of a dominant scene plane was presented. This algorithm exploits the theorem that if five or more out of the seven correspondences are related by a homography, then there is an epipolar geometry consistent with the septuple, as well as with all correspondences related by the homography. In each iteration, the algorithm selects a sample of seven correspondences. It then detects samples in which at least five correspondences are consistent with a homography. This homography is then used to estimate the epipolar geometry by the plane and parallax algorithm [16].

To illustrate the above difficult situations, consider the following two examples. Fig. 1a shows the flowerpot image scene in which the inlier rate is low, and it includes a dominant degenerate configuration. In this scene, 17 percent of the 252 putative correspondences are inliers and 70 percent of the inliers lie in a small part of the scene that yields a degenerate configuration. A computation of the fundamental matrix based on only inliers from this small space results in a very unstable fundamental matrix. On this scene, RANSAC often fails to find the correct fundamental matrix. Fig. 1a shows a typical result of RANSAC. Dots represent inliers from the degenerate configuration, circles represent inliers that do not belong to the degenerate configuration, and \times represents an outlier that RANSAC detected as an inlier. In this example, RANSAC succeeded in finding all of the inliers that belong to the degenerate configuration, but failed to find any inliers outside it. This is demonstrated in Fig. 1b, which shows the square root of the symmetric epipolar distance of

the inlier pair from the fundamental matrix. The distances of the inliers outside the degenerate configuration are large. Although a large number of inliers were found, the precision of the resulting fundamental matrix is very low. The number of iterations for this scene, according to (1) for $p = 0.99$, is over one million. Fig. 1c shows another example in which the inlier rate is 16.5 percent out of 310 putative correspondences and it includes a dominant-plane degenerate configuration. In this scene, 78 percent of the inliers lie near the plane. Fig. 1d shows a typical result of RANSAC that succeeded in finding part of the inliers that lie near the plane and failed to find any inliers not close to the plane. As a result, the fundamental matrix is totally incorrect, as can be seen in Fig. 1d. The number of iterations required for this scene, according to (1), is again over one million.

In this paper, we propose a novel algorithm for robust estimation of epipolar geometry. The algorithm handles the above two difficult cases in a unified manner. The algorithm can handle not only the planar degeneracy but also scenes that include a variety of degeneracies or close to degenerate configurations.

The balanced exploration and exploitation model (BEEM) search algorithm includes a balanced use of three search techniques borrowed from classical general optimization methods and adapted for use within the RANSAC framework. The first technique is global random exploration, which tests random possible solutions. The second technique is local exploration, which searches for better solutions in the neighborhood of the current best solution, and the last one is local exploitation, which tries to improve the quality of the model by local search methods. Moreover, it exploits available prior information, the distance ratio of the closest to the second-closest neighbors of a SIFT keypoint, to accelerate the search process [6]. The novelty here is to convert each distance ratio assigned to a correspondence into a prior probability that the correspondence is an inlier using empirical nonparametric distributions. We use this probability to guide the sampling process. The algorithm uses the best found model to guide the search process, escape from degenerate models, and define an efficient stopping criterion. This is done by a smart sampling strategy. In addition, we developed a simple and efficient method for global exploration that is able to estimate the epipolar geometry from two SIFT correspondences. The combination of the prior probabilities and the two SIFT estimation methods enables us to find estimations after a very small number of iterations has been tried. This method is only able to provide an initial estimate for the fundamental matrix and needs all of the other components of the system to yield an accurate result.

Considering the system as a whole, the only slow steps left are the generation of the features and their matching. The matching is sped up using the locality-sensitive hashing (LSH) [17] approximate nearest neighbor algorithm. The generation of the SIFT features can be accelerated using the approximation described in [18] or a GPU-based implementation described in [19].

The resulting algorithm, when tested on real images with or without degenerate configurations, gives quality estimations and achieves significant speedups, especially in scenes that include the aforementioned difficult situations.

The paper is organized as follows: In Section 2, the exploration and exploitation search techniques are discussed. Section 3 describes the generation of the prior probability for putative correspondences. Our fast method for global exploration which is able to calculate the fundamental matrix from two SIFT correspondences is presented in Section 4. Section 5 describes a method to estimate the quality of the best found epipolar geometry model. The details of the algorithm are presented in Section 6. Experimental results are shown and discussed in Section 7. The paper concludes in Section 8.

A shorter version of this paper, including some of the results presented here, was presented at the Ninth European Conference on Computer Vision (ECCV '06) [13].

2 EXPLORATION AND EXPLOITATION

Any efficient search algorithm must use two general techniques to find the global maximum: exploration to investigate points in new and unknown regions of the search space and exploitation to make use of knowledge found at points previously visited to help find better points. These two requirements are contradictory and a good search algorithm must strike a balance between them. A purely random search is good at exploration but does no exploitation, whereas a purely hill-climbing method is good at exploitation but does little exploration. Combinations of these two strategies can be quite effective, but it is difficult to know where the best balance lies.

The robust estimation of the fundamental matrix can be thought of as a search process. The search is for the parameters of the fundamental matrix and the set of inliers. Therefore, algorithms that estimate the epipolar geometry can be analyzed according to the way in which they combine the above techniques. The RANSAC algorithm [1] samples, in each iteration, a minimal subset of matches and computes from it a model. This random process is actually an indirect global exploration of the parameter space. In the projection-based M-estimator (pbM) algorithm [20], [21], each exploration iteration is followed by a standard exploitation step. A hill-climbing procedure over the parameter space is performed using a local search algorithm. The LO-RANSAC algorithm [3] makes an exploitation step only when a new good model is found in an exploration iteration. The exploitation step is performed by choosing random samples only from the set of suspected inliers, the model's support set, and computing a fundamental matrix from it. In cases where there exists a degenerate configuration, the exploitation step tends to enlarge the support set, but it includes only inliers belonging to the degeneracy. In our algorithm, we use the LO-RANSAC local optimization step to perform the exploitation stage.

In classical search algorithms such as simulated annealing, a local exploration step exists. There, with a certain probability, a local step in the parameter space is taken which does not improve the quality of the current solution. This step is used to escape from local minima in the parameter space. No similar step exists within the RANSAC family of algorithms. Even if a relatively good model that includes a large number of inliers is found, it is not used

after the exploitation (LO-RANSAC) step has been performed. The algorithm simply returns to random sampling, hoping to find by chance a better model. This problem occurs mainly when the RANSAC process reaches a degenerate set of inliers. We suggest adding an intermediate technique that uses the previous best solution and explores its neighborhood, looking for a better solution whose support set is larger and includes most of the support set of the previous best solution. We use the term neighborhood loosely. When the current solution is supported by a degenerate set, the solution is merely a point on a surface consistent with the support set. The goal of the local exploration step is to find another point on this surface, which can be quite far in the parameter space from the current solution, which is consistent with all of the correct matches. Thus, when we use the term local, we mean it in the support set sense. To achieve this, we need to generate a sample of inliers that includes, in addition to members of the current support set, other correspondences. Once we have a “good” previous solution, it can be assumed that the vast majority of its support set are inliers. Therefore, when choosing a subset for the RANSAC step, we choose most of the subset from the support set and the rest from points that are outside the support set. When such a subset consists only of inliers, the support set of the resulting model tends to break out of the confines of the set of inliers belonging to the degeneracy (the local maximum), yielding a more correct solution. Unlike simulated annealing, in our algorithm, the result of the local exploration step is only used if the size of the support set increases.

When incorporating a local exploration step into the algorithm, several questions have to be addressed. First, local exploration is only effective when the best previous support set nearly includes only inliers. Therefore, it is essential to be able to recognize such sets. Second, depending on the quality of the set, a balance between the application of global exploration, local exploration, and exploitation has to be struck. Finally, how to incorporate available prior information about the quality of each putative correspondence into the general scheme has to be addressed.

The BEEM algorithm includes all of the components described above. Its state diagram is presented in Fig. 2. The algorithm includes the following states and the transitions between them:

- **Prior estimation.** Use prior available information to estimate the probability that a correspondence is an inlier. This probability is used to guide the sampling in the other states.
- **Global exploration.** Sample a minimal subset of correspondences and instantiate the model from the subset. If the size of the support set of the formed model is larger than all of the models that were formed in this state, go to the *exploitation* state; otherwise, go to the *model quality estimation* state.
- **Model quality estimation.** Estimate the quality of the best model found until now based on the size of its support set and the number of iterations that the algorithm has performed until now. Use this quality

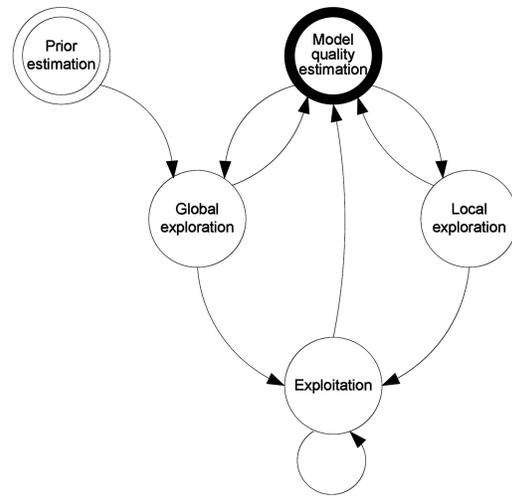


Fig. 2. State diagram of the BEEM search algorithm. The algorithm first assigns probabilities to the putative correspondences and then performs a global exploration step. Depending on the quality of the recovered model, the algorithm performs global or local exploration steps followed by an exploitation step.

estimate to probabilistically choose the next state, *global exploration* or *local exploration*.

- **Local exploration.** Sample a subset of correspondences from the support set of the best model and sample a subset of correspondences from the rest of the correspondences. Instantiate the model from the union of the two subsets. If the size of its support set is larger than all of the models that were previously formed in this state, go to the *exploitation* state; otherwise, go to the *model quality estimation* state.
- **Exploitation.** Iteratively try to improve the last formed model by choosing subsets of matches from the support set and testing their quality. At the end of this process, go to the *model quality estimation* state.

In the following sections, we will describe the main components of the algorithm, which include our methods for prior probability estimation, our fast method for global exploration, the 2-SIFT method, which is used to produce initial solutions to the fundamental matrix estimation, and our method for model quality estimation. The detailed algorithm is given in Section 6.

3 USING PRIOR INFORMATION OF THE MATCH

Each SIFT feature is represented by a descriptor vector whose length is 128. The best candidate match for each SIFT keypoint from the first image is found by identifying the keypoint in the second image whose descriptor is closest to it in a euclidean distance sense. Some features from the first image will have no correct match in the second image. Therefore, it is useful to have the ability to discard them. A global threshold on the distance to the closest feature does not perform well as some descriptors are much more discriminative than others. A more effective measure was suggested in [6] and is obtained by comparing the distance of the closest neighbor to that of the second-closest neighbor. This measure performs well because, for correct matches, the closest neighbor is significantly closer than the closest incorrect match. For false matches,

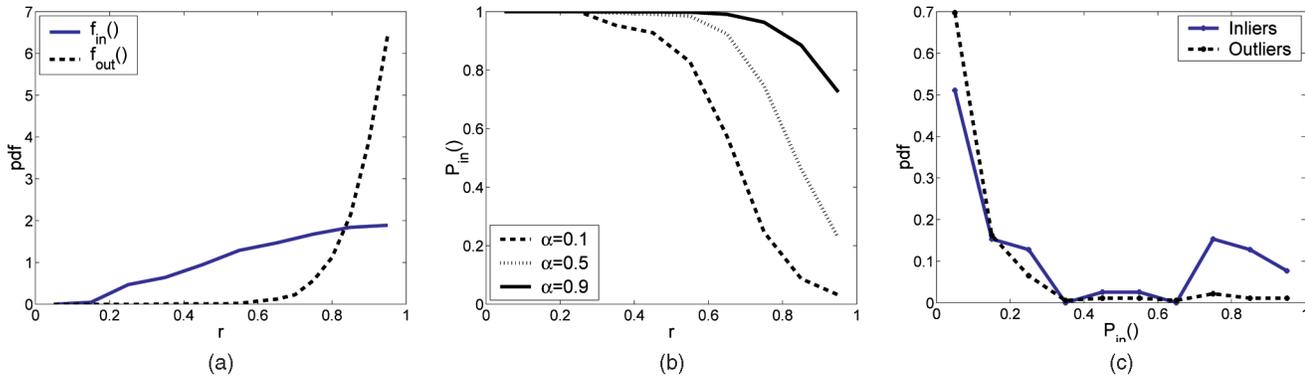


Fig. 3. (a) The empirical distributions of the distance ratio, r , for inlier and outliers were generated based on 20 image pairs. (b) The probability that a correspondence is an inlier as a function of r for several values of the inlier rate, α . (c) The distributions of the estimated probability $P_{in}()$ of the inliers and the outliers for the book scene image pair.

there will likely be a number of other false matches within similar distances due to the high dimensionality of the feature space. We can think of the second-closest match as providing an estimate of the density of the false matches within this region of the feature space. The consequence of this criterion is that repetitive features appearing in the image will also be discarded.

Let r_i be the distance ratio of the closest to the second-closest neighbors of the i th keypoint of the first image. Fig. 3a shows the value of this measure for real image data for inliers and outliers. In [6], it was suggested to reject all matches in which the distance ratio is greater than $r_{thresh} = 0.8$. In our experiments, we also follow this rule. The probabilistic meaning of this is that each correspondence whose score is below this threshold is sampled uniformly. PROSAC exploits this ratio even more and its samples are drawn from progressively larger sets from the set of correspondences ordered by this ratio. This improves the performance of the algorithm. In this work, we make an additional step by giving an empirical probabilistic meaning to this ratio.

The distance ratio can be thought of as a random variable and is modeled as a mixture model:

$$f_r(r_i) = f_{in}(r_i)\alpha + f_{out}(r_i)(1 - \alpha),$$

where

$$f_{in}(r_i) = f(r_i|p_i \leftrightarrow p'_i \text{ inlier}), f_{out}(r_i) = f(r_i|p_i \leftrightarrow p'_i \text{ outlier}),$$

and α is the mixing parameter, which is the probability that any selected correspondence is an inlier. The probability $P_{in}(i)$ that correspondence $p_i \leftrightarrow p'_i$ is an inlier can be calculated using Bayes' rule:

$$P_{in}(i) = \frac{f_{in}(r_i)\alpha}{f_{in}(r_i)\alpha + f_{out}(r_i)(1 - \alpha)}. \quad (2)$$

We estimate this probability in a nonparametric manner. We generate two samples from real images:

- S_{in} , a sample of \tilde{N}_{in} inlier ratio distances and
- S_{out} , a sample of \tilde{N}_{out} outlier ratio distances.

We estimate $f_{in}()$ and $f_{out}()$ using a kernel density estimator over S_{in} and S_{out} , respectively.

We estimate α for a given image pair using the curve fitting of the empirical cumulative distribution function (cdf) of S_{in} , S_{out} , and the set of ratios of the putative correspondences. An empirical cdf over a set of measurements S can be estimated by

$$F(s) = \frac{\sum_{i=1}^{\|S\|} g(s_i, s)}{\|S\|},$$

where

$$g(s_i, s) = \begin{cases} 1, & s_i \leq s \\ 0, & \text{otherwise,} \end{cases}$$

and s_i is the i th element in S .

Let

$$R = \left\{ \rho_j | \rho_j = j \frac{r_{thresh}}{N_R + 1} \right\}_{j=1}^{N_R}$$

be a set of N_R uniformly spaced ratio distances. We obtain a set of the following N_R linear equations:

$$F_r(\rho_j) = F_{in}(\rho_j)\alpha + F_{out}(\rho_j)(1 - \alpha), j = 1, \dots, N_R.$$

These equations are used to estimate α by a least squares technique. Once α has been estimated, $P_{in}()$ can be estimated for all putative correspondences using (2). Fig. 3b shows the probability $P_{in}()$ for several values of α . Fig. 3c shows the distributions of the estimated $P_{in}()$ of the inliers and the outliers for the book scene image pair. As can be seen in the graph, a large portion of the correspondences that received high probabilities are indeed inliers. In this example, the inlier rate of matches with r_{thresh} less than 0.8 is 16.5 percent and the estimated α is 15.7 percent, which is quite accurate.

The estimation of the inlier rate using the prior distributions gives a very good clue about the relation between the two images. If the estimated inlier rate, $\hat{\alpha}$, is close to zero, the two images are probably not related.

Tordoff and Murray [4] use normalized correlation between the regions around putative correspondences as a basis for their probability measure. Comparing their method to ours, several differences are apparent. The evidence we use is more informative, as pointed out by Lowe [6]. The

difference between the two types of evidence is that the correlation score yields an absolute score, whereas the euclidean distance between SIFT features does not in itself indicate the quality of the match. Therefore, the ratio of the distances is used as the basis for the probability estimate. Ratios close to one are considered to be outliers with high probability. Thus, when dealing with repeated structures in the image, the SIFT score is unable to differentiate between this case and outlier matches and discards them. The correlation score, on the other hand, can detect this case but is unable to choose among the different instances of the structure. Therefore, all possible alternatives are assigned similar probabilities, which are all quite low. The result in both cases is similar because, due to the low probabilities, the matches of the repeated structures are rarely chosen. In addition, Tordoff and Murray have to compute the correlation score between all possible matches in order to compute the best match's probability, whereas the SIFT ratio score requires the computation of only two scores.

When comparing our method to PROSAC, we claim that there is a slight disadvantage of not assigning probabilities to the correspondences. When given a set of matches with close probability values, pairs with a slightly higher probability of being correct might be placed much higher in the list and chosen much more often, whereas we will choose all of these pairs with approximately equal probability. When some of these high probability pairs are outliers, the number of iterations needed to find an outlier-free set could increase considerably.

4 GLOBAL EXPLORATION: EPIPOLAR GEOMETRY FROM TWO SIFT CORRESPONDENCES

In [9], [10], using three affine region-to-region matches to estimate the epipolar geometry in each RANSAC sample was suggested. Actually, two regions suffice. Assuming that, for each region-to-region match, there exists a homography that approximates the transformation between the regions, the two homographies can be used to recover the fundamental matrix [22, chapter 13, pp. 337-338]. The fact that the transformation is approximated by a special type of homography such as an affine or even a similarity transformation does not change this fact. Moreover, each transformation can be represented by a set of four pairs of points satisfying the transformation and used as input for the normalized eight-point algorithm, yielding comparable results to the two-homography algorithm. This general principle can be applied to any local region matching method [9], [23], [24], [25], [26], [27].

In our implementation, we chose the SIFT descriptor, which is a very powerful descriptor for image matching. This descriptor is invariant to the similarity transformation, which is not as accurate as the affine transformation or the homography, but, as we will show, worked well in practice. The ability to generate epipolar geometry from two SIFT correspondences instead of seven point correspondences is expected to significantly reduce the runtime according to (1). This ability actually reduces the complexity of the robust estimation of the fundamental matrix to that of a robust estimation of a line from a set of points in space. We

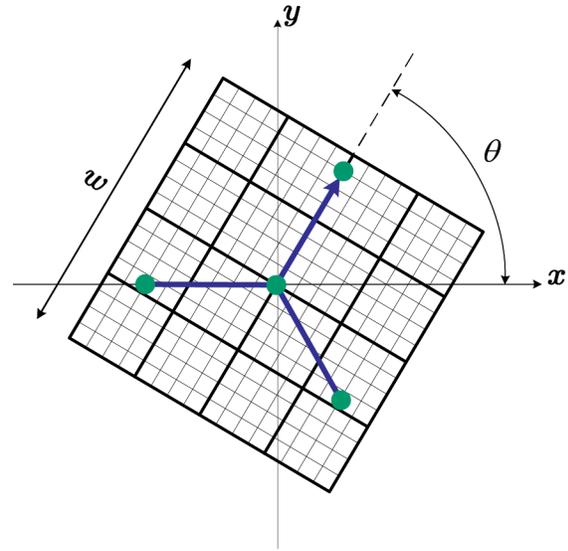


Fig. 4. Illustration of the four-point generation for the SIFT descriptor.

suggest a simple method to estimate the epipolar geometry from two SIFT correspondences. Each SIFT keypoint is characterized by its location $p = (x, y)$, the orientation θ of the dominant gradients, and its scale s . We generate for each SIFT keypoint a set of four points:

$$\begin{aligned} &((x, y), (x + ls \cos(\theta), y + ls \sin(\theta)), \\ &(x + ls \cos(\theta + \frac{2\pi}{3}), y + ls \sin(\theta + \frac{2\pi}{3}), \\ &(x + ls \cos(\theta + \frac{4\pi}{3}), y + ls \sin(\theta + \frac{4\pi}{3})). \end{aligned}$$

We set $l = \frac{7}{8} \frac{w}{2}$, where w is the width of the descriptor window. The configuration of the four points is illustrated in Fig. 4. Thus, the three additional points lie within the descriptor window. A set of two SIFT correspondences gives a set of eight point correspondences. These can be used to estimate the fundamental matrix using the normalized eight-point algorithm [11]. This method is equivalent to finding the fundamental matrix that is consistent with two homographies. The additional points are simply used to represent those homographies. When scoring a hypothesized fundamental matrix, a SIFT correspondence is considered consistent with the hypothesized epipolar geometry only when all coincident four point correspondences, $(p_{s1}, p_{s2}, p_{s3}, p_{s4}) \leftrightarrow (p'_{s1}, p'_{s2}, p'_{s3}, p'_{s4})$, are within their respective error thresholds. The location of the first point in the set is quite accurate, whereas the locations of the last three points are less accurate because they are approximated from the SIFT characteristics. We use the error threshold d for the first point in the set and $d\sqrt{s's}$ for the other three, where s and s' are the SIFT scale parameters of the keypoints of the first and the second SIFT descriptors, respectively, and d is a threshold parameter.

One may wonder how accurate the estimation of the fundamental matrix using the 2-SIFT method is. The 2-SIFT method generates four point correspondences from each SIFT keypoint. These four points are usually quite close to each other and the last three points are estimated less accurately.

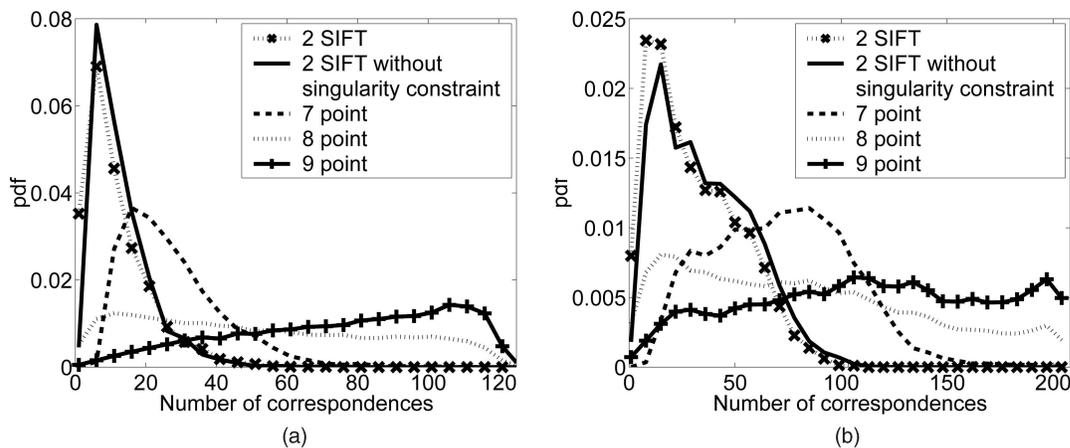


Fig. 5. Algorithm evaluation. For each of the algorithms, 10,000 experiments were run over the inlier correspondences. The number of correspondences supporting the obtained fundamental matrix was recorded and their distribution is shown. (a) Sideways scene. (b) Forward scene.

Therefore, a fundamental matrix that is based on such point correspondences is expected to be less accurate than when the points are accurately estimated and uniformly distributed over the whole image. However, all that is required of this step of the algorithm is to produce a very rough approximation of the fundamental matrix, which will be supported by several additional correct correspondences.

To check the severity of this problem, the estimation quality of the 2-SIFT method was compared to the quality of the seven-point algorithm, normalized eight-point algorithm with eight and nine point correspondences. Two types of real scenes without any dominant degenerate configurations were checked: a scene moving sideways and a scene moving forward. For each scene, the inlier SIFT correspondences were found. For each algorithm, in each scene, 10,000 samples were taken from the inlier correspondences. For each sample, a fundamental matrix was calculated, and the number of correspondences consistent with the model was recorded. The size of the support set of the model quantifies the quality of the model. Fig. 5 shows the results. The horizontal axis gives the size of the support set and the vertical axis represents the distribution of the models that were supported by sets of this size. The results of the 2-SIFT method are less accurate than the 7, 8, and 9-point algorithms, as expected. This can be seen from the graphs as, in many cases, only a small number of inliers support the proposed solution. However, it usually recovers enough supporting inliers to initialize the fundamental matrix estimation process. Clearly, the use of the LO-RANSAC step after the 2-SIFT method is very important to produce a more accurate solution.

To improve the estimation quality, we checked one more method, the 2-SIFT without the singularity constraint (2-SIFT-NSC) method. In this method, the singularity constraint of the fundamental matrix is not enforced. The result is usually an illegal model, but, in the sample step of the algorithm, it is not necessary to work with legal models because the main purpose of the sample step is to detect large amounts of supporting inliers. The results of the 2-SIFT-NSC method, which are also shown in Fig. 5, outperform that of the 2-SIFT method. The reason for this is that the singularity constraint enforcement when applied

in the 8-point algorithm changes the solution in a non-optimal way by projecting the matrix to the closest point on the singularity surface. This is not the optimal solution since the entries of the fundamental matrix do not have equal importance. In addition, the computation of the optimal singular matrix adds to the computational cost. For both reasons, it is better not to apply this step at all. We therefore use the 2-SIFT-NSC method in our algorithm.

The examples shown above deal with motions that do not involve out-of-plane rotation. In these cases, a similarity transformation approximates the local motion well and, therefore, both the SIFT and the 2-SIFT algorithms work well. It is also interesting to check whether the 2-SIFT algorithm will be able to perform in cases where severe foreshortening occurs. This happens when there is a large out-of-plane rotation between the two images. It is well documented that the SIFT feature matching algorithm itself does not work in very high rotation angles. Therefore, the question remains whether the 2-SIFT algorithm will be able to perform at the extreme cases where the SIFT algorithm still works. This might be problematic because the local transformations between the corresponding SIFT features could be far from the similarity transformations assumed by the SIFT algorithm.

To demonstrate the performance of the algorithm in this situation, the algorithm was applied to the two pairs of images shown in Fig. 6. As expected, the fraction of correct matches from the total number of feature pairs is much lower (0.17 and 0.1, respectively) due to the difficulty in matching SIFT features in this case. As in the previous experiment, we plotted the success of the various RANSAC variants in Fig. 7. In these cases also, enough supporting matches were found to enable the BEEM algorithm to start its journey toward the correct solution. In these experiments, the recovered fundamental matrix was quite poor due to the inaccurate SIFT transformations used in its construction. Therefore, enforcing the singularity constraint on it causes a larger deterioration in the solution. This can be clearly seen by comparing the graphs of 2-SIFT to those of 2-SIFT-NSC. The 2-SIFT-NSC method is clearly superior due to the small number of fundamental matrix hypotheses

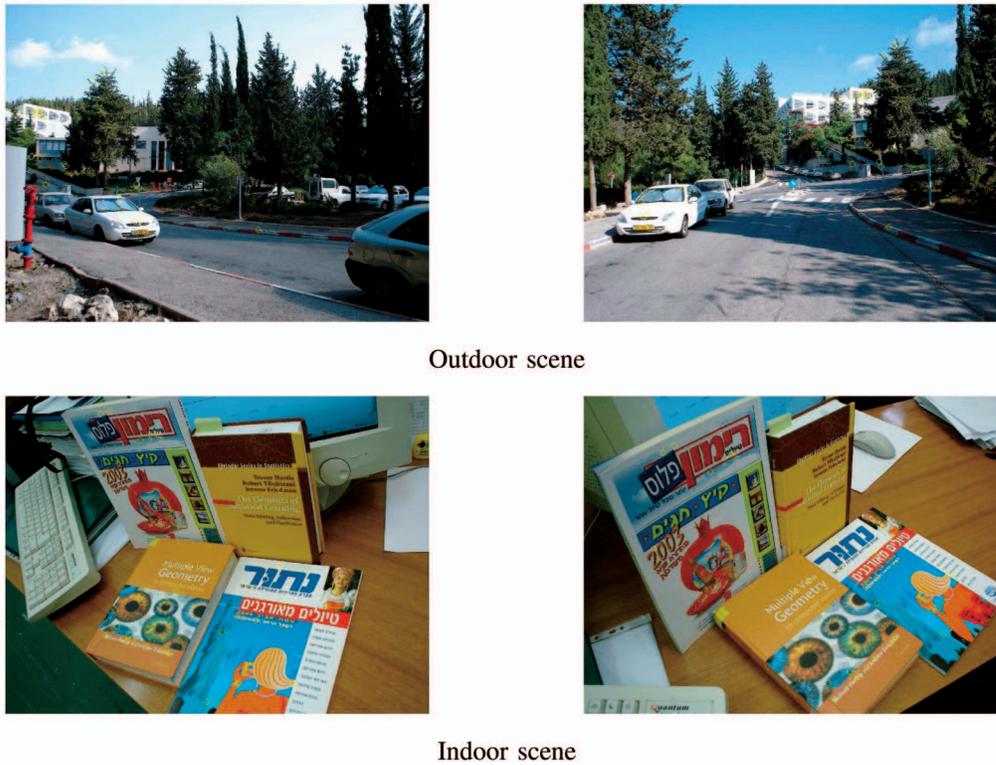


Fig. 6. Scenes with considerable foreshortening.

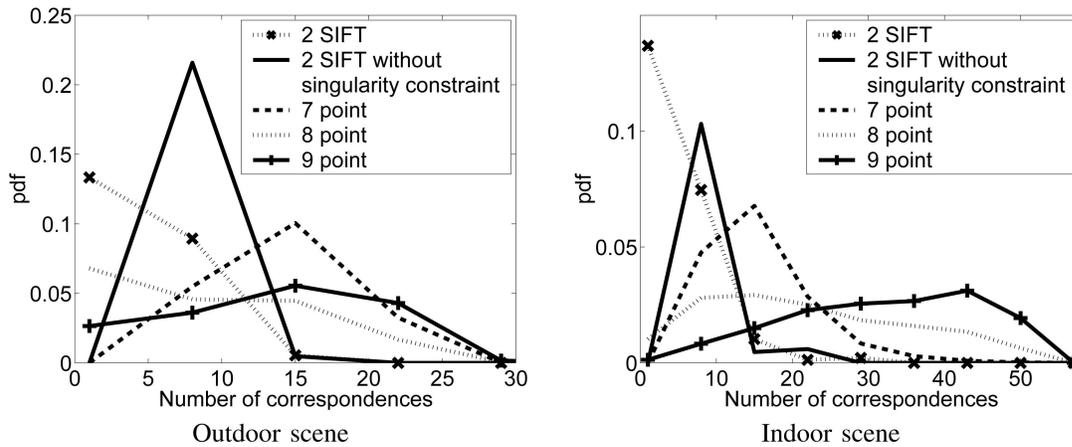


Fig. 7. Scenes with considerable foreshortening.

that were supported by a very small number of correspondences. These experiments demonstrate that, as long as the SIFT process detects correct matches, the 2-SIFT algorithm will be able to exploit them to find an approximate fundamental matrix.

The results presented in this section have demonstrated that the 2-SIFT method generates good results within the general framework of the BEEM algorithm. It cannot be used, however, as a complete method because a fundamental matrix supported by only, say, 10 matches out of a hundred is a poor estimation for the correct solution.

5 BEST FOUND MODEL QUALITY ESTIMATION

In the model quality estimation state, the algorithm estimates the quality of the best found model as an inlier

model, that is, a model in which nearly all of the members of its support set are inliers. When an inlier model is detected, it can help accelerate the search process using the local exploration state, whereas using an outlier model in that state is useless. In such situations, we want to direct the BEEM algorithm to continue to perform global exploration. To achieve this, we have to estimate the probability that the model is supported by outliers that are by chance consistent with it. Let $P_{om}(i/N)$ be the probability that at most i outlier matches support an outlier model from the N putative matches. Let $N_{best} = \max\{N_i\}_{i=1}^I$ be the maximal size of the support set after I iterations achieved by model M_{best} , where N_i is the size of the support set of the i th iteration. Using the above definitions, the probability P_q that M_{best} is not an outlier model is estimated. This is equivalent to the

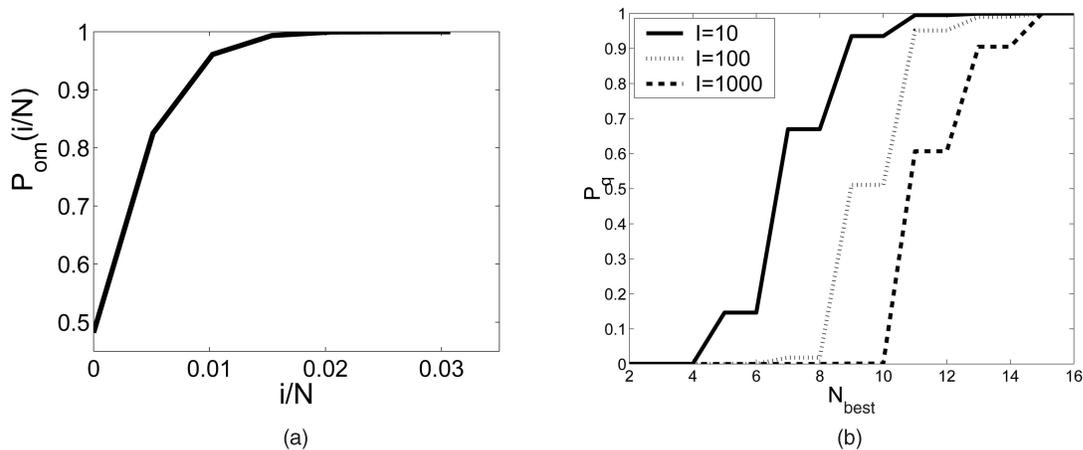


Fig. 8. (a) The cdf $P_{om}(\cdot)$ as a function of the percentage of falsely detected matches i/N from the total number of putative matches. (b) The probability P_q as function of N_{best} for $I = 10$, $I = 100$, and $I = 1,000$, where the number of putative correspondences is set to 400.

probability that, in all of the I iterations, the support set of size N_{best} could not be achieved by an outlier model. Thus,

$$\begin{aligned} P_q &= \prod_{i=1}^I \text{Prob}(N_i < N_{best}) \\ &= \prod_{i=1}^I \text{Prob}(N_i < N_{best}) = (P_{om}((N_{best} - 1)/N))^I. \end{aligned}$$

The BEEM algorithm uses the probability P_q as an estimate to the quality of the best found model. We estimate $P_{om}(\cdot)$ using several unrelated image pairs in a nonparametric manner. We ran the 2-SIFT-NSC algorithm for the above image pairs and recorded the size of the support sets of the outlier models. Fig. 8a shows the cdf $P_{om}(\cdot)$ as a function of the fraction of falsely detected inliers i from the total number of putative matches N . The empirical distribution shows that, when the fraction of detected matches is larger than 0.035, it cannot be a result of a totally incorrect fundamental matrix. As a result, in this case, the algorithm will be directed to perform only local exploration steps. Fig. 8b shows the probability P_q as a function of N_{best} for $I = 10$, $I = 100$, and $I = 1,000$, where the number of putative correspondences is set to 400. Note that, when the number of iterations increases, the “belief” of the algorithm in the correctness of small subsets decreases. As a result, the algorithm tends to do more global exploration.

6 ALGORITHM DETAILS

Up to this point, we have described the principles of the BEEM algorithm. Now, we will put them all together, yielding the complete epipolar geometry estimation algorithm. The algorithm is summarized in Algorithm 1. The details of the algorithm are given as follows:

Fundamental matrix generation. The generation of the fundamental matrix from a given subset S of SIFT correspondences chosen from the set of putative correspondences, C , is done as follows: If $2 \leq |S| < 7$, then we use the normalized eight-point algorithm, where each SIFT correspondence provides four point correspondences, as described in Section 4. If $|S| = 7$, then we use the seven-point algorithm with seven points, one from each SIFT correspondence. If $|S| > 7$, then we use the standard normalized

eight-point algorithm with $|S|$ keypoints provided from the SIFT correspondences.

Exploitation. This state is very similar to the local optimization method described in [3] with a small improvement. In this state, a new sampling procedure is executed. Samples are selected only from the support set S of the previous state. New models are verified against the whole set of putative correspondences. The size of the sample is set to $\min(|S|/2, N_F)$, where N_F is set to 14, as was suggested in [3]. For each fundamental matrix generated from a sample, all of the correspondences in its support set are used to compute a new model using the linear algorithm. This process is repeated until no improvement is achieved. The modification we made to the original LO-RANSAC is that, whenever a larger support set is found, the exploitation process restarts again with it. The algorithm exits this state to the model quality estimation state after I_{LO} iterations without improvement, where I_{LO} is set to 10 in our experiments.

Local exploration. The parameter space close to the best model found so far is searched in this state by choosing a sample of size $\min(|S_{best}|/2, N_F - 1)$ SIFT correspondences from S_{best} and a single SIFT correspondence from $C \setminus S_{best}$. Here again, N_F was set to 14. The fundamental matrix is instantiated from the union of the above subset and the single SIFT correspondence, where the single SIFT correspondence always contributes four point correspondences. This way, the algorithm has a better chance to escape from degenerate configurations.

Once $|S_{best}|$ exceeds $0.035|C|$, according to our empirical model (whose distribution is plotted in Fig. 8a), the model must contain a large number of inliers. As a result, P_q is equal to one. When this happens, the sampling strategy for correspondences from $C \setminus S_{best}$ changes slightly. Each time a new maximum is found, that is, S_{best} is updated, the correspondences in $C \setminus S_{best}$ are sorted in decreasing order according to $P_{in}(\cdot)$. In each iteration, a single SIFT correspondence is chosen from $C \setminus S_{best}$ according to the sorting order and the rest as usual from S_{best} .

Stopping criterion. The BEEM algorithm terminates if, in the last $|C| - |S_{best}|$ exploration samples, the subset S_{best} was not updated and if P_q is equal to one in these samples. This criterion ensures with high confidence that nearly all of the

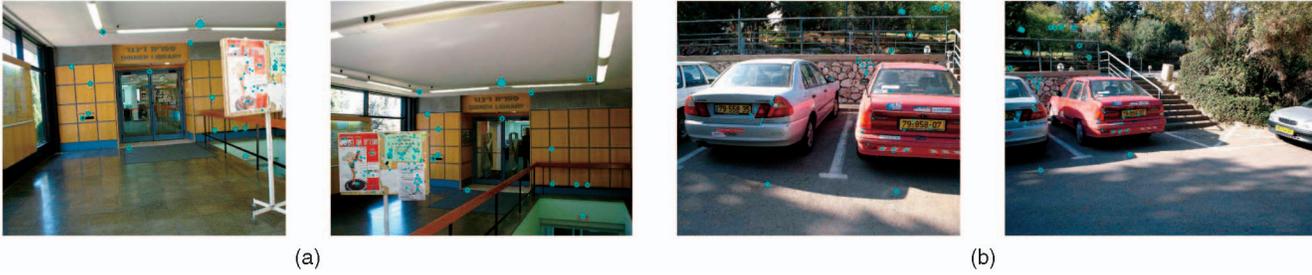


Fig. 9. BEEM experiment image scenes. Degeneracy inliers are denoted by dots, whereas the nondegeneracy inliers are denoted by circles. (a) Board scene. (b) Car scene.

inliers have been detected. This suggested stopping criterion usually terminates much earlier than in the standard approach because, once the algorithm finds a model with an adequate number of inliers, P_q is estimated as one and the algorithm enters the final local exploration iterations. Because the correspondences in $C \setminus S_{best}$ are sorted in decreasing order according to $P_{in}()$, the rest of the inliers are rapidly found. Once S_{best} ceases to change, $|C| - |S_{best}|$ iterations are performed. In the experiments that we have performed, the number of iterations until an adequate number of inliers are found is usually very small thanks to the various components of the BEEM algorithm. As a result, the total number of iterations of the BEEM algorithm is, in practice, slightly higher than the number of outliers in the putative correspondence set. This number is much lower than the bound given by (1).

Algorithm 1: The BEEM algorithm.

begin Prior estimation:

Estimate α and $P_{in}()$ of the set C of putative correspondences.

end

begin Global exploration:

Sample according to $P_{in}()$ a subset of two SIFT correspondences from C ;

Instantiate the fundamental matrix F ;

if the support set S of F is the best found in this state **then**
goto *Exploitation*

else

goto *Model quality estimation*;

end

begin Exploitation:

Execute local optimization with inner RANSAC over S until I_{LO} repetitions without improvement;

if found model with largest support until now **then**
keep its support set in S_{best} ;

end

begin Model quality estimation:

Estimate P_q ;

if the stopping criterion is satisfied **then**
terminate;

Choose with probability P_q to goto *Local exploration*;
otherwise goto *Global exploration*;

end

begin Local exploration:

Sample according to $P_{in}()$ a subset of SIFT correspondences from S_{best} ;

if $P_q < 1$ **then**

sample according to $P_{in}()$ a single SIFT from $C \setminus S_{best}$

else

choose the next SIFT correspondence from $C \setminus S_{best}$;

Instantiate the fundamental matrix F ;

if the support set S of F is the largest found in this state **then**
goto *Exploitation*;

else

goto *Model quality estimation*;

end

7 EXPERIMENTS

7.1 BEEM Algorithm

The proposed algorithm was tested on many image pairs of indoor and outdoor scenes, several of which are presented here. The cases that are presented here are difficult cases in which the inlier rate is low and includes a dominant degeneracy.

For each image, we applied the SIFT method to detect the keypoints. The descriptors of the first image were then stored in an LSH [17] data structure and the descriptors of the second image were used for querying the data structure to find their approximate nearest neighbors to generate putative correspondences. We used the adapted version of the LSH [28] with data-driven partitions. The LSH algorithm is simple for implementation and efficient. For example, the runtime for the generation of the putative correspondences of the book scene was reduced from 25.6 sec using a simple linear search to 0.45 sec using the LSH algorithm on a Pentium 4 CPU 1.70 GHz computer (all of the runtime results in this paper were checked on this computer). The LSH algorithm has been claimed to be faster than other nearest neighbor techniques such as the K-dimensional tree (KD-tree) [17], [29]. This claim was not verified by us for this case.

For illustration reasons, we divided the set of putative correspondences into three sets: outliers, inliers belonging to the degenerate configuration, and the rest of the inliers for which most of them have to be part of the support set in order to generate an accurate fundamental matrix. The images of the scenes are shown in Figs. 1, 9, and 6. Their details are given in Table 1.

For each scene, six algorithms were tested: the BEEM algorithm, LO-RANSAC using samples of two SIFT correspondences to generate fundamental matrices (2SIFT LO-RANSAC), RANSAC using samples of two SIFT correspondences (2SIFT RANSAC), LO-RANSAC using

TABLE 1
The Characteristics of the Tested Scenes

Scene	Degeneracy	N	α	$\hat{\alpha}$	Out.	In.	Deg. In.	Non-Deg. In.
Flowerpot	Small region	252	0.17	0.25	210	42	30	12
Book	Plane	310	0.17	0.16	260	50	44	6
Board	Plane	276	0.27	0.25	201	75	57	18
Cars	Several small regions	272	0.17	0.11	225	47	35	12
Indoor	Plane	310	0.17	0.14	256	54	43	11
Outdoor	None	308	0.1	0.11	277	31		31

For each scene, the table gives the type of degeneracy, the number of correspondences, the inlier rate, the BEEM estimation of the inlier rate, the number of outliers, the number of inliers, the number of inliers belonging to the degeneracy, and the number of inliers not belonging to the degeneracy.

TABLE 2
Results of the Experiments

Scene	Algorithm	Success	Iterations	In.	N.Deg.	Times
Flowerpot	BEEM	100%	(5.0) 213	40.6	11.2	11.1
	2SIFT LO-RANSAC	30%	356	29.8	3.6	10.3
	2SIFT RANSAC	0%	880	16.9	0	21.3
	7pt P-LO-RANSAC	65%	10,000	34.6	7.9	708.9
	7pt LO-RANSAC	15%	10,000	27.2	2.4	704.5
	7pt RANSAC	0%	10,000	19.5	1.2	703.2
Book	BEEM	95%	(6.3) 279	44.1	5.6	12.9
	2SIFT LO-RANSAC	5%	660	27.2	0.6	15.5
	2SIFT RANSAC	0%	2,449	11.2	0.2	46.6
	7pt P-LO-RANSAC	30%	10,000	35.1	1.8	796.3
	7pt LO-RANSAC	0%	10,000	19.9	0.2	789.9
	7pt RANSAC	0%	10,000	16.5	0.5	785.7
Board	BEEM	90%	(1.7) 207	72.4	15.6	11.6
	2SIFT LO-RANSAC	5%	90	57.8	1.9	8.0
	2SIFT RANSAC	0%	1964	31.9	1.0	13.9
	7pt P-LO-RANSAC	15%	10,000	61.3	4.9	761.0
	7pt LO-RANSAC	5%	10,000	57.9	2.1	760.9
	7pt RANSAC	0%	10,000	53.6	1.1	758.2
Car	BEEM	100%	(2.5) 230	44.8	10.9	10.3
	2SIFT LO-RANSAC	30%	533	31.3	5.7	11.2
	2SIFT RANSAC	0%	1,236	14.8	1.0	27.1
	7pt P-LO-RANSAC	70%	10,000	39.2	8.2	701.0
	7pt LO-RANSAC	25%	10,000	27.25	3.9	701.5
	7pt RANSAC	0%	10,000	18.05	2.3	698.4
Indoor	BEEM	100%	(9.7) 272.7	53.7	10.8	14.6
	2SIFT LO-RANSAC	45%	217.1	48.6	5.5	11.7
	2SIFT RANSAC	0%	1,177	22.6	0.8	39.8
	7pt P-LO-RANSAC	55%	10,000	49.6	5.9	846.5
	7pt LO-RANSAC	0%	10,000	15.9	0.9	701.5
	7pt RANSAC	0%	10,000	17.1	0.4	847.3
Outdoor	BEEM	100%	(12.6) 301.4	28.3	N/A	14.7
	2SIFT LO-RANSAC	30%	1,195	19.1	N/A	42.2
	2SIFT RANSAC	0%	2,756	13	N/A	93.2
	7pt P-LO-RANSAC	35%	10,000	19.3	N/A	856.2
	7pt LO-RANSAC	0%	10,000	10.1	N/A	847.6
	7pt RANSAC	0%	10,000	13	N/A	845.7

For each algorithm, the following statistics are presented: the success rate, the number of iterations until the termination of the algorithm, the number of inliers found, the number of inliers outside the degenerate configuration found, and the runtimes. For the BEEM algorithm, the number of global exploration iterations is given in parentheses.

samples of seven point correspondences where the samples were sampled according to the probability $P_{in}(i)$ (7pt P-LO-RANSAC), LO-RANSAC using samples of seven point correspondences (7pt LO-RANSAC), and RANSAC using samples of seven point correspondences (7pt RANSAC). The termination criterion for RANSAC and LO-RANSAC was based on (1) for $p = 0.99$. In cases where the number of iterations exceeded 10,000, the algorithm also terminated. Each algorithm was applied to each image pair 20 times. For each algorithm, the following statistics are presented: the success rate, defined as the percentage of the experiments in which at least 75 percent of the inliers were found and at least 50 percent of the inliers outside the degenerate configuration were found, the

number of iterations until the termination of the algorithm, the number of inliers found, and the number of inliers outside the degenerate configuration found. For the BEEM algorithm, in the iteration column, the average number of global exploration iterations is also given, denoted in parentheses. The runtimes in seconds are given for Matlab implementations. These runtimes are only given for comparative reasons. A C++ implementation could easily speed up the algorithm by an order of magnitude.

The results shown in Table 2 clearly show that the BEEM algorithm outperforms the other algorithms in the way that it deals with degeneracies, almost always detecting most of the inliers outside of the degenerate configuration. The quality of the results as represented by the overall number

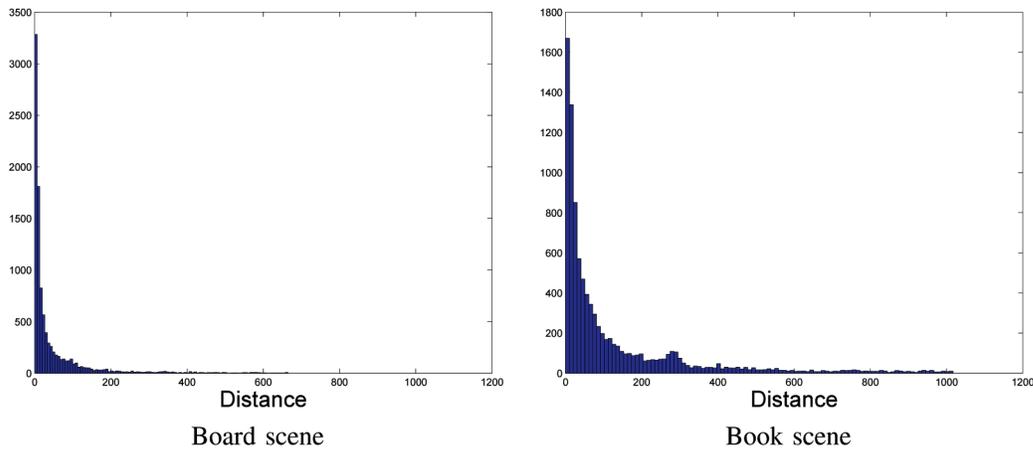


Fig. 10. Distance histograms from computed homographies.

of detected inliers is also much higher. The number of iterations until termination of the algorithm is much lower than for the other algorithms. Finally, the number of global exploration iterations of the BEEM algorithm is very low as a result of the use of the prior information and the 2-SIFT method. As mentioned in the previous section, the number of iterations of the BEEM algorithm is, in practice, slightly higher than the number of outliers in the putative correspondence set. This number is much lower than the number of iterations of the other algorithms.

The results of the other algorithms demonstrate the contribution of each component of the BEEM algorithm to the quality of the detection. Comparing the BEEM algorithm to the 2-SIFT LO-RANSAC, we can see the effects of the local exploration step. This step dramatically increases the success of the algorithm in dealing with degeneracies. This is achieved at no clear additional computational cost. There are challenging cases, such as the outdoor scene whose results are also presented in Table 2, where the local exploration considerably reduces the runtime while improving the result, even though there are no degeneracies in the scene. This is simply an example where the stopping criterion of the BEEM algorithm yields a faster run than the stopping criterion of RANSAC.

When the LO-RANSAC step is removed in the next implementation, the algorithm always fails to detect the degeneracy and requires more iterations. When the 2-SIFT method is replaced by the seven point RANSAC, the complexity dramatically increases and, even when a good solution has been found, the algorithm is not able to stop because the number of iterations has not reached the RANSAC stopping criterion. When the probabilistic sampling is turned off, the success rate is further reduced and the number of recovered inliers decreases. Finally, when comparing the 2-SIFT to the seven point RANSAC, we can see how poorly the 2-SIFT method performs by looking at the number of recovered inliers. This demonstrates that the 2-SIFT method needs the other components of the BEEM algorithm to insure its success. This is because its goal is not to find an accurate fundamental matrix but merely a good starting position that is exploited by the other components.

7.2 Plane Degeneracy

In scenes which contain a dominant plane, algorithms have been proposed to deal with the degeneracy caused by it

[15], [16]. In such cases, the algorithm has to be given a parameter measuring the planarity of the plane. Consider, for example, the two examples presented above, the Board scene and the Book scene. In the first case, an actual plane containing many features is present. In the second scene, the back wall with the shelves is the relatively planar region of the scene. In the following experiment, we compared the planarity of both scenes in the following manner: For both scenes, 10,000 quintuple correct matches from the degenerate plane were sampled and the geometric distance of the fifth point match from the homography computed from the other four matches was calculated. The results are presented in Fig. 10. What can clearly be seen is that the distances in the two cases are very different and, therefore, setting a threshold on the distance determining whether a plane exists or not is required and can vary considerably from scene to scene. Moreover, once the algorithm finds a homography for a nonplanar region, the remaining steps of the algorithm are not guaranteed to succeed.

The BEEM algorithm, on the other hand, does not explicitly model the degeneracy and therefore is not limited to the modeled degeneracy. Therefore, it does not depend on the level of the planarity of the region. It simply detects correct matches that the current solution does not explain. In conclusion, the BEEM algorithm is a nonparametric method, whereas previous methods are model (plane) based and they exploit the model after it has been detected.

8 DISCUSSION

In this paper, we presented the BEEM algorithm for epipolar geometry estimation. It works very well in difficult scenes where the inlier rate is low and/or large subsets of the inlier correspondences are consistent with a degenerate configuration. The BEEM algorithm can replace algorithms from the RANSAC family whenever robust model estimation is needed. The principles of the BEEM algorithm, using prior knowledge, the balanced use of exploration and exploitation within the RANSAC framework, and the generation of approximate (not necessarily legal) models in the RANSAC step, can also be applied in other cases.

The BEEM algorithm can be easily modified to address other estimation problems. Homographies can be robustly estimated from one or two SIFT correspondences. Nister's algorithm [30] for essential matrix estimation can also be

improved under the BEEM framework using two SIFT correspondences instead of five point correspondences, resulting in a faster algorithm. In both cases, the entire BEEM framework is needed in order to improve the results obtained by the 1-2SIFT match algorithm.

The only limitation of the BEEM algorithm is that it relies on correctly matched SIFT features. In cases where the camera underwent considerable out-of-plane rotation, this might not be possible because the local transformation might not be close enough to a similarity transformation. As a result, the SIFT matching process will perform poorly. This problem might be addressed using other types of features that are matched using more accurate transformations such as affine transformations or homographies.

ACKNOWLEDGMENTS

The authors acknowledge the support of Grant 01-99-08430 of the Israeli Space Agency through the Ministry of Science Culture and Sports of Israel.

REFERENCES

- [1] M. Fischler and R. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," *Comm. ACM*, vol. 24, no. 6, pp. 381-395, June 1981.
- [2] P. Torr, "Motion Segmentation and Outlier Detection," PhD thesis, Dept. of Eng. Science, Univ. of Oxford, 1995.
- [3] O. Chum, J. Matas, and J. Kittler, "Locally Optimized RANSAC," *Proc. German Pattern Recognition Symp.*, pp. 236-243, 2003.
- [4] B. Tordoff and D. Murray, "Guided Sampling and Consensus for Motion Estimation," *Proc. Seventh European Conf. Computer Vision*, pp. I: 82-I: 96, 2002.
- [5] O. Chum and J. Matas, "Matching with PROSAC: Progressive Sample Consensus," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. I: 220-I: 226, 2005.
- [6] D. Lowe, "Distinctive Image Features from Scale-Invariant Key-points," *Int'l J. Computer Vision*, vol. 60, no. 2, pp. 91-110, Nov. 2004.
- [7] L. Goshen and I. Shimshoni, "Guided Sampling via Weak Motion Models and Outlier Sample Generation for Epipolar Geometry Estimation," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. I: 1105-I: 1112, 2005.
- [8] J. Domke and Y. Aloimonos, "A Probabilistic Framework for Correspondence and Egomotion," *Proc. Workshop Dynamical Vision*, 2005.
- [9] F. Schaffalitzky and A. Zisserman, "Multi-View Matching for Unordered Image Sets, or 'How Do I Organize My Holiday Snaps?'" *Proc. Seventh European Conf. Computer Vision*, pp. I: 414-I: 431, 2002.
- [10] O. Chum, J. Matas, and S. Obdrzalek, "Enhancing RANSAC by Generalized Model Optimization," *Proc. Sixth Asian Conf. Computer Vision*, pp. II: 812-II: 817, 2004.
- [11] R. Hartley, "In Defense of the Eight-Point Algorithm," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 6, pp. 580-593, June 1997.
- [12] M. Perdoch, J. Matas, and O. Chum, "Epipolar Geometry from Two Correspondences," *Proc., Int'l Conf. Pattern Recognition*, pp. IV: 215-IV: 219, 2006.
- [13] L. Goshen and I. Shimshoni, "Balanced Exploration and Exploitation Model Search for Efficient Epipolar Geometry Estimation," *Proc. Ninth European Conf. Computer Vision*, 2006.
- [14] A. Makadia, C. Geyer, S. Sastry, and K. Daniilidis, "Radon-Based Structure from Motion without Correspondences," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. I: 796-I: 803, 2005.
- [15] O. Chum, T. Werner, and J. Matas, "Two-View Geometry Estimation Unaffected by a Dominant Plane," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. I: 772-I: 779, 2005.
- [16] M. Irani and P. Anandan, "Parallax Geometry of Pairs of Points for 3D Scene Analysis," *Proc. Fourth European Conf. Computer Vision*, pp. I: 17-I: 30, 1996.
- [17] A. Gionis, P. Indyk, and R. Motwani, "Similarity Search in High Dimensions via Hashing," *Proc. 25th Int'l Conf. Very Large Data Bases*, pp. 518-529, 1999.
- [18] M. Grabner, H. Grabner, and H. Bischof, "Fast Approximated Sift," *Proc. Seventh Asian Conf. Computer Vision*, pp. 918-927, 2006.
- [19] S. Sinha, J. Frahm, M. Pollefeys, and Y. Genc, "GPU-Based Video Feature Tracking and Matching," *Proc. Workshop Edge Computing Using New Commodity Architectures*, 2006.
- [20] H. Chen and P. Meer, "Robust Regression with Projection Based M-Estimators," *Proc. Int'l Conf. Computer Vision*, pp. 878-885, 2003.
- [21] S. Rozenfeld and I. Shimshoni, "The Modified pbM-Estimator Method and a Runtime Analysis Technique for the RANSAC Family," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. I: 1113-I: 1120, 2005.
- [22] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge Univ. Press, 2003.
- [23] T. Tuytelaars and L. Van Gool, "Matching Widely Separated Views Based on Affine Invariant Regions," *Int'l J. Computer Vision*, vol. 59, no. 1, pp. 61-85, 2004.
- [24] T. Kadir, A. Zisserman, and M. Brady, "An Affine Invariant Salient Region Detector," *Proc. Eighth European Conf. Computer Vision*, pp. 345-357, 2004.
- [25] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust Wide-Baseline Stereo from Maximally Stable Extremal Regions," *Proc. 13th British Machine Vision Conf.*, pp. 384-393, 2002.
- [26] K. Mikolajczyk and C. Schmid, "An Affine Invariant Interest Point Detector," *Proc. Seventh European Conf. Computer Vision*, pp. I: 128-I: 142, 2002.
- [27] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A Comparison of Affine Region Detectors," *Int'l J. Computer Vision*, vol. 65, no. 12, pp. 43-72, 2005.
- [28] B. Georgescu, I. Shimshoni, and P. Meer, "Mean Shift Based Clustering in High Dimensions: A Texture Classification Example," *Proc. Ninth Int'l Conf. Computer Vision*, pp. 456-463, 2003.
- [29] R. Weber, H. Schek, and S. Blott, "A Quantitative Analysis and Performance Study for Similarity-Search Methods in High-Dimensional Spaces," *Proc. 24th Int'l Conf. Very Large Data Bases*, pp. 194-205, 1998.
- [30] D. Nister, "An Efficient Solution to the Five-Point Relative Pose Problem," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 6, pp. 756-777, June 2004.



Liran Goshen received the BSc and PhD degrees in information system engineering from the Department of Industrial Engineering and Management at the Technion-Israel Institute of Technology, Haifa, in 1996 and 2006, respectively. The PhD studies were done in a framework of direct doctoral track under the guidance of Dr. Ilan Shimshoni and Professor Daniel Keren. From 1996 to 2001, he served in the Israeli Navy. Since 2006, he has been with the Global Research and Advance Development Section at CT Philips Medical Systems as the leader of the Image Processing and Workflow Laboratory. He is the recipient of the Guttwirth fellowship.



Ilan Shimshoni received the BSc degree in mathematics and computer science from the Hebrew University, Jerusalem, in 1984, the MSc degree in computer science from the Weizmann Institute of Science, Rehovot, Israel, in 1989, and the PhD degree in computer science from the University of Illinois at Urbana-Champaign in 1995. He spent three years as a postdoctoral researcher with the Computer Science Department and then several years with the Industrial Engineering Department at the Technion-Israel Institute of Technology. He joined the faculty of the Department of Management Information Systems at Haifa University in 2005 and is currently the chair of that department. He also spent a year on sabbatical at Rutgers University, Piscataway, New Jersey. He has served as a committee member for all of the major conferences in computer vision. His research interests are in the fields of computer vision, robotics, and computer graphics, specializing mainly in the applications of statistical methods in these fields. He is a member of the IEEE.