

Wide Baseline Feature Matching Using the Cross-Epipolar Ordering Constraint

Xiaoye Lu and Roberto Manduchi

Department of Computer Engineering, University of California, Santa Cruz¹

Abstract – *Robust feature matching across different views of the same scene taken by two cameras with wide baseline and arbitrary rotation is still an open problem. Matching based on appearance alone is unreliable, because a surface point changes its appearance depending on the viewpoint. As a result, this approach may generate "unrealizable" correspondence sets, that is, sets of pairwise correspondences that are not consistent with the epipolar geometry. We propose a novel technique, which is applicable when the epipolar lines in the two images are approximately parallel. Our algorithm only searches in the space of realizable matchings, thereby reducing the likelihood of mismatches as well as the dimension of the search domain. This approach can use any given feature descriptor with an associated distance function, and assumes no knowledge of the intrinsic parameters of the two cameras. The extension to the general case of epipoles in finite position is the object of current research.*

I. INTRODUCTION

Reliable image registration over wide baselines from uncalibrated cameras is an important (and as yet unsolved) problem in computer vision. Whereas in traditional stereo or structure from motion the rigid motion between the cameras may be assumed to be small (which facilitates the feature matching or tracking operation), we are interested in applications where two images of the same scene may be taken from very different points of view, possibly by two different types of cameras, and without knowledge of the intrinsic camera parameters. These conditions should be expected, for example, as a consequence of ad-hoc deployment of networks of miniaturized cameras, which has been proposed for surveillance and monitoring applications.

There is an immense amount of literature covering image registration, and we can only summarize the main results here. If the epipolar geometry is known, stereo matching can be performed along the epipolar lines, generally resulting in dense and reliable measurements. However, in order to derive the epipolar geometry, a bootstrap phase involving full image search is needed. The typical paradigm is to extract a number of feature points from both images, attempt a matching, and then extract a subset of "correct" matches by RANSAC validation [10]. Our paper addresses the first step in this procedure, with the goal to find as many correct matches as possible in the initial phase. Previous work on this

subject includes a classic paper by Deriche *et al.* [5], whereby candidate matches are found within a search window for every feature in one image. A relaxation algorithm is then used to minimize an energy function based on the local structure ("strength") of the matches. Unfortunately, this algorithm does not work well for our purposes, since local structure will change considerably with large baselines due to parallax.

Another approach is based on the definition of robust visual descriptors and on the detection of "distinctive" features, in the hope that the initial matching can rely on visual appearance [2,12,14,18,19,9,16]. For the considered type of applications, scale and rotation invariance is obviously very desirable. A comparison of the performances of some popular descriptors is reported in [17]. Unfortunately, matching based on local appearance alone is not reliable in real-world situations. Lacking structural constraints on the set of correspondences, this approach is haunted by two main problems: ambiguity and appearance change (as a function of the viewpoint). These two issues are in fact related, in the sense that invariant operators that can relieve the dependence of the appearance on nuisance factors (such as rotation, foreshortening) usually increase ambiguity.

In this paper, we introduce an algorithm that allows one to impose realizability constraints on the search space of matchings. This technique can work with any choice of features and feature descriptors. Knowledge of the intrinsic parameters of the two cameras is not required. A strong limitation of our present approach, though, is that it requires that the epipoles of the stereo system are located at infinity (and therefore, that the epipolar lines are parallel in both cameras). This is the case when both image planes are parallel to the baseline. However, the extension to the general case of epipoles in finite positions is possible, and it is the topic of current research. Note that our experiments show that a moderate convergence of the epipolar lines is well tolerated by the algorithm in its present form.

Our approach is based on a structural property that must be satisfied by any realizable matching under the parallel epipolar lines condition. We call this property the "cross-epipolar ordering constraint"; it relies on the natural ordering of the epipolar lines that is preserved in both images. Note that this is a very different concept from the ordering constraint within the same epipolar line that is normally used in stereo matching. The cross-epipolar ordering property constrains the set of candidate

¹ This work was supported by DARPA – MARS 2020 Program, through subcontract 1080049-132209 from Carnegie Mellon University.

matchings; since the correct matching must be realizable, by rejecting non-realizable matchings we decrease the likelihood of misclassification due to ambiguity. In addition, we show that this constrained search can be implemented efficiently. Indeed, exhaustive search within the space of possible matchings has complexity $o(N^6)$, but a coarse-to-fine search scheme that relies on the “smoothness” of the search space can be implemented in $o(N^2 \log N)$ time. Note that, without the cross-epipolar order constraint, exhaustive search for the optimal match has complexity of $o(N!)$.

This paper is organized as follows. Section II describes our algorithm, including the coarse-to-fine search strategy. Section III presents experiments that show the performance and limitations of our approach. Section IV has the conclusions.

II. ALGORITHM DESCRIPTION

We will assume that two images have been taken by two cameras (or possibly the same camera) from different positions. The cameras need not have the same intrinsic parameters; no knowledge of the camera’s parameters is assumed. We do assume that the two fields of view overlap, so that we can expect that some parts of the scene seen from the first camera are visible by the second camera. As mentioned in the Introduction, we also assume that the epipoles of the stereo system are at infinity, and therefore that the epipolar lines within each image are parallel.

Suppose that a number of feature points have been selected in each image by a suitable feature detector. In general, the number N_1 and N_2 of points found in the two images will be different. Our purpose is to find the “correct” correspondences (p_i^1, p_j^2) between points in the two images. To simplify the problem, we will assume that one point in one image can match at most one point in the other image. Note that in practical cases, due to foreshortening and finite scale of the feature operator, one feature in one image may actually correspond to more than one feature in the other image. However, this event is relatively unlikely, and our uniqueness assumption simply means that only one of the two (or more) possible matches will be selected. In graph theory parlance, our correspondence problem is equivalent to finding a *matching* of a *bipartite graph* whose vertices are the disjoint union of the points in the two images [21].

Together with the selection of feature points, we need a criterion for comparing the appearances of two features. A popular approach is to define a descriptor vector for each feature point, which represents salient characteristics of the image in a neighborhood of the point. Together with the descriptor, we need a suitable metric within the feature space. Many types of descriptors and related distance functions have been proposed in the literature

(see for example [20,12]). In general, it is desirable that the descriptor be relatively immune to changes in appearance due to nuisance parameters such as rotation, foreshortening, or change in illumination. At the same time, the descriptor should be rich enough to reduce the risk of ambiguity.

Let us denote the descriptor of a point p_i^k in the k —th image by $v(p_i^k)$, and the distance (in the descriptors’ space) between the i —th point in the first image and the j —th point in the second image by $d(v(p_i^1), v(p_j^2)) \equiv d(i, j)$. We are using the term “distance” rather liberally here – some functions $d(\cdot, \cdot)$ proposed in the literature are not real distances, because they do not satisfy the triangular inequality. For what we are concerned, $d(\cdot, \cdot)$ is a function that yields a positive number, which is equal to 0 when the two descriptors are identical, and hopefully is large when the features look different. The distance $d_{i,j}$ represents the cost associated to an edge of the bipartite graph.

A sensible way to estimate the set of correspondences between the two images is to find the matching that minimizes the corresponding sum of costs (distances). If $N_1 < N_2$, then one may try to match all points in the first image with N_1 points in the second image (*maximum matching* [21]). In general, however, each image will contain some points that have no correspondence in the other image, and therefore maximum matching is not recommended. To avoid trivial solutions (empty matching), it is customary to add a penalty c to the unmatched points. The feature matching problem can thus be formally stated as an overall cost minimization. Let I^k be a set of ordered indices for the k -th image (that is, an ordered subset of $[1, \dots, N_k]$). The notation $I^k(i)$ stands to indicate the i -th element of I^k . Then, the matching problem is equivalent to finding

$$\arg \min_{\substack{I^1, I^2, \Pi \\ |I^1|=|I^2|}} \sum_{i=1}^{|I^1|} d(I^1(i), I^2(\Pi(i))) + c(N_1 + N_2 - 2|I^1|) \quad (1)$$

where $|I^k|$ is the cardinality of I^k , and Π is a permutation function defined on $[1, \dots, |I^1|]$. In other words, the problem is to select two subsets of points (one per image) with the same cardinality, and a permutation matrix that describes the point correspondences. The final matching is represented by the pairs

$$(I^1(i), I^2(\Pi(i))) \text{ for } i = 1, \dots, |I^1|$$

Note that (1) can be generalized to the case of cost c being a function of the features that are removed from the matching.

A naïve approach to the matching problem is by using a greedy algorithm: start by finding the most similar feature pair in the two images, add them to the matching, delete the two features from the original feature lists, and iterate on the remaining features until adding one more match would increase the overall cost. This algorithm has relatively low complexity ($\mathcal{O}(N^3)$) but will fail, in general, to find the optimal solution. Unfortunately, exhaustive search requires $N!$ tests, and therefore is only feasible for very small sets of features.

In fact, complexity is not the only problem with this approach. Consider an ideal noiseless case, meaning that the descriptors of two corresponding points have null distance. In this case, it is clear that the overall cost is minimized by one (or more) matchings that contain the “correct” one. However, the appearance of a feature is often not unique; in these ambiguous cases, it only takes a little noise to generate mismatches. To reduce the risk of mismatches, it would be useful to constrain the search domain to a smaller space that is guaranteed to contain the correct matching. This is indeed possible if only realizable matchings are taken into account. A *realizable* matching is a set of correspondences such that there exist a set of 3-D points and camera parameters that can generate such corresponding image points [10].

There are several methods to verify the consistency of a set of matches. For example, if the epipolar geometry is known, one can check that matches of two points on the same epipolar line in one image are on the same epipolar line in the other image. As another example, if we can find 4 collinear points in one image that match 4 collinear points in the other image, then the cross ratio of such points in both images must be the same. In this work, we use a very simple, non-metric criterion that leads to a simple implementation of the constrained search. This criterion is based on the epipolar ordering constraint, discussed in the next section.

A. The Epipolar Ordering Constraint

We will rely here on the assumption of epipoles at infinity introduced earlier. In this case, the epipolar lines in an image are parallel; let’s denote by ϑ the slope angles of these lines with respect to a given direction (say, the vertical image edge). Each epipolar line is uniquely characterized by the position p of any image point on the line. Suppose now that the slopes ϑ^1 and ϑ^2 in the two images are known. For each image, consider a line orthogonal to the epipolar lines, oriented in an arbitrary (but fixed) direction (see Figure 2). Let \bar{p}_i^k be the (signed) projection of the feature point p_i^k in the k -th image onto this line. It is easy to show that the following statement is true:

Fact 1: A matching is realizable only if there exists a constant $s = \pm 1$ such that, for any couple of matches

$(p_i^1, p_{i_2}^2)$ and $(p_{j_1}^1, p_{j_2}^2)$ the following identity holds:
 $\text{sign}(\bar{p}_{i_1}^1 - \bar{p}_{j_1}^1) = s \cdot \text{sign}(\bar{p}_{i_2}^2 - \bar{p}_{j_2}^2)$.

where the function $\text{sign}(\cdot)$ yields 1 if its argument is positive, -1 if it is negative, and 0 if it is null.

Intuitively, Fact 1 states that as we move steadily across epipolar lines in one image, we should move steadily across epipolar lines in the other image as well (see Figure 1) We call this property the *cross-epipolar ordering constraint*.

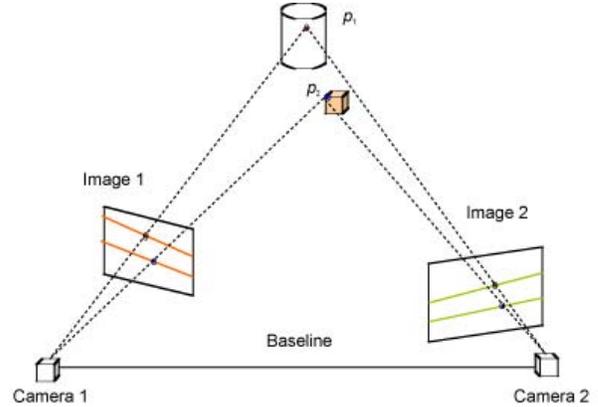


Figure 1: The epipolar geometry of a stereo system.

Let us stress that this constraint is fundamentally different from the ordering constraint *along* an epipolar line that is often used for stereo matching [6]. The latter refers to points within the *same* epipolar line, and applies only when these points are projections from a connected surface. In contrast, the cross-epipolar ordering constraint (i) applies to *any* couple of matches, regardless of whether they are on the same epipolar line, and (ii) it is *always* verified (in the assumption of epipoles at infinity), regardless of whether the viewed surfaces are connected or not.

The angles ϑ^k , together with the orientations of the corresponding lines across, induce an ordering on the feature points of each image (through their projections \bar{p}_i^k). Note that this is only a partial ordering, because two features may lie on the same epipolar line, and therefore may have the same value of \bar{p}_i^k . We will assume hereinafter that this never happens, the algorithm modifications for the general case being trivial.

The benefits of the induced ordering should be clear. For example, consider a simple case whereby each feature point in the first image has a match in the second image, and vice versa (then $N_1 = N_2 = N$). Let I^k be a vector containing the indices of the ordered points in the k -th image. Then, one sees that, for $i = 1, \dots, N$, the correct matching must be either $(I^1(i), I^2(i))$ or

$(I^1(i), I^2(\Pi_{\text{rev}}(i)))$, where Π_{rev} is a permutation operator that simply reverts the order of the list.

More in general, when only a subset of points in each image belong to the correct matching, one should find the sets I^1 and I^2 by minimizing the overall cost (1). Now, however, thanks to the natural ordering in the two sets induced by the angles ϑ^k , the permutation function can only be either Π_{id} (which leaves the order of the list unaltered) or Π_{rev} . For a given permutation function, the sets I^1 and I^2 that minimize (1) can be found with complexity of $\mathcal{O}(N_1 \times N_2)$ using dynamic programming [3]. Note that a similar technique is often used to solve the correspondence problem in intra-epipolar line stereo matching [1,15]. Both cases constrain the data in similar ways, although, as we have discussed, they refer to different geometrical properties.

Unfortunately, in practice, the angles ϑ^k of the epipolar lines in the two images are not known in advance (indeed, the main reason for feature matching here is to determine the epipolar geometry!) Still, the ordering constraint for realizability applies to *any* possible choice of the angles $(\vartheta^1, \vartheta^2)$, which represent the slopes of “candidate” epipolar lines. Thus, one may try to run the algorithm above for all possible values of these angles. As we show in the next section, it turns out that only a finite number of angles need to be considered.

B. Critical Angles

Assume that candidate angles for the epipolar lines in both images have been selected, together with an orientation across lines. As we have seen, these parameters determine a unique ordering of the feature points in both images. Suppose now to change the angle ϑ by a quantity $\Delta\vartheta$ in one image. If the variation $\Delta\vartheta$ is small enough, the ordering is not changed (see Figure 2), and therefore the same matching is found by minimizing (1). Indeed, as $\Delta\vartheta$ is increased, the order of the points changes only at an angle such that two feature points lie on the same epipolar line (see Figure 2). Any such angle will be called a *critical angle*. By computing all

$M_k = \binom{N_k}{2}$ critical angles $\{\vartheta_i^k\}$ for each image, we identify all the possible different orderings of the feature points that generate realizable matchings. Note that M_k is a much smaller number than $N_k!$, the number of permutations that should be checked without this ordering constraint.

Thus, dynamic programming needs only be performed for each pair of critical angles $(\vartheta_i^1, \vartheta_j^2)$, with orderings in the two images point sets induced by any angle in the sectors $]\vartheta_{i-1}^1, \vartheta_i^1[$ and $]\vartheta_{j-1}^2, \vartheta_j^2[$

respectively, with the convention that $\vartheta_0^k \equiv \vartheta_{M_k}^k$. If this is done exhaustively, the overall complexity of matching is $\mathcal{O}(N^6)$. As we show in the next section, this complexity can be greatly reduced by a coarse-to-fine search procedure on the critical angles.

C. Coarse-to-Fine Search

Let $\{\vartheta_1^k, \dots, \vartheta_{M_k}^k\}$ be the collection of critical angles for the k -th image. Assume that the orientation of the line orthogonal to the candidate epipolar lines onto which the position of feature points is projected is “locked” with the lines (see Figure 2). In this way, rather than considering two such orientations per each angle, we can just assume that the epipolar lines are oriented, with slopes spanning the full interval $[0^\circ, 360^\circ]$.

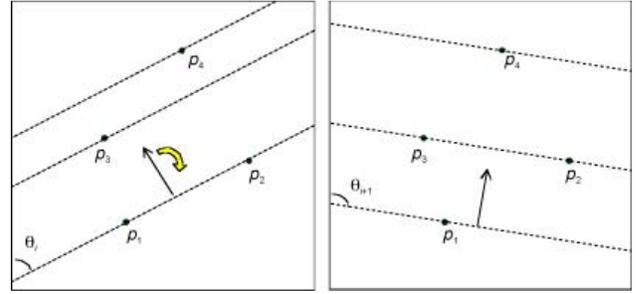


Figure 2: Two consecutive critical angles for the same image.

We will denote by $R(\vartheta_i^1, \vartheta_j^2) \equiv R_{i,j}$ the minimum of function (1) relative to the critical angle pair $(\vartheta_i^1, \vartheta_j^2)$ obtained by dynamic programming. The profile of $R_{i,j}$ for three different image pairs is shown in Figures 5, 8, and 10. The “quincunx” periodicity in the graphs is expected: rotating the slope of the candidate epipolar lines by 180° is equivalent to reversing the order of the points; if this is done in both image, we obtain the same optimal matching. This suggests that search needs only be performed over $2M_1M_2$ orderings.

These figures also highlight the ambiguity still lurking in our approach. Indeed, multiple minima co-exist in the profile of $R_{i,j}$, sometimes at very different locations. Obviously, we cannot hope to completely eliminate ambiguity; however, it is a fact that, by discarding all unrealizable matchings, and therefore restricting our search to a set of matchings that certainly contains the correct one, we also reduce the likelihood of a mismatch due to ambiguity.

Another interesting characteristic of the graphs of $R_{i,j}$ is that a certain degree of “smoothness” is present. Again, this should be expected: point orderings for angles just before and just after a critical angle are identical, except

for two points that are swapped (see Figure 2). Therefore, one may expect the optimal matching (and the resulting overall cost) to be similar in the two cases.

We can exploit this “smoothness” property to implement a coarse-to-fine search mechanism. Our algorithm proceeds as follows. In the first step, only a subset of m uniformly sampled critical angles are considered in both images, and the minimizer $(\vartheta_{i_1}^1, \vartheta_{i_2}^2)$ of R_{i_j} is found over this subset. Then, for the k -th image, only m uniformly sampled angles within the subset of $2M_k/m$ angles around $\vartheta_{i_k}^k$ are considered. The algorithm is iterated, every time considering an interval $m/2$ times smaller than the previous one. Assuming for simplicity that the number of points in the two images is equal ($N_1 = N_2 = N$), it is easy to see that the search over m samples is computed k times with

$$k = \left\lceil \log_{\frac{m}{2}} \frac{2M}{m} \right\rceil + 1 \approx 2 \log_{\frac{m}{2}} N$$

where $M = \binom{N}{2}$ is the number of critical angles. With this algorithm, the complexity is thus reduced to $\mathcal{O}(N^2 \log_{\frac{m}{2}} N)$. We used $m = 8$ in our experiments, with good success.

III. EXPERIMENTS

In this section we present some feature matching experiments using our proposed algorithm. The descriptor vector $v(p_i^k)$ in our experiment was simply a 3-D histogram representing the marginal distribution of color within a 30×30 neighborhood of the chosen feature. This histogram used 10 bins for each color plane. Note that this descriptor is approximately rotation-invariant. More sophisticated descriptors [12] could be used; for these preliminary experiments, however, we felt that a simpler local representation was preferable, because it would help us better analyze the results of our algorithm. Among the several possible distances between histograms proposed in the literature, we selected the Earth Mover’s Distance [22], which gave the best empirical results in our tests. For all our experiments, we compared the results of our algorithm with a greedy algorithm based on the same feature descriptor but not using the cross-epipolar ordering constraint. The greedy algorithm was forced to compute a maximum matching.

In the first experiment we manually selected the feature sets in the two image, while in the other two case presented here, features were selected automatically using the Harris corner detector. The reason for manual feature selection in the first case was to provide a simple and well-controlled case study for analysis.



Figure 3 Matching results using our algorithm. The green asterisks are manually selected feature points. The points with a nearby number have been matched with the point with the same number in the other image. The points are ordered according to the critical angle that minimizes the overall cost (1) under the cross-epipolar ordering constraint.

The two views for the first experiment are shown in Figure 3. 20 points were selected within each image; 17 such points have a correspondence in the other image. In spite of the large baseline, our algorithm detected all 17 matches correctly. The matches are shown in Figure 3, with numbers that reflects the ordering induced by the critical angle that minimizes the cost function. For comparison, we show the actual epipolar lines (computed with a mainstream calibration code on manually selected matches) in Figure 4.



Figure 4 The epipolar lines of the images in Figure 4.

Figure 6 shows the profile of the function R_{ij} , corresponding to the minimum of the cost function at different critical angles. Note that this matching (as well as the matchings in the next two experiments) was computed using the fast coarse-to-fine algorithm of Section II.C. To compute the profile of R_{ij} in Figure 6, we had to exhaustively consider all critical angle pairs. The results of the greedy algorithm not using the cross-epipolar ordering constraint are shown in Figure 6. Only 8 correct matches were found by this technique.

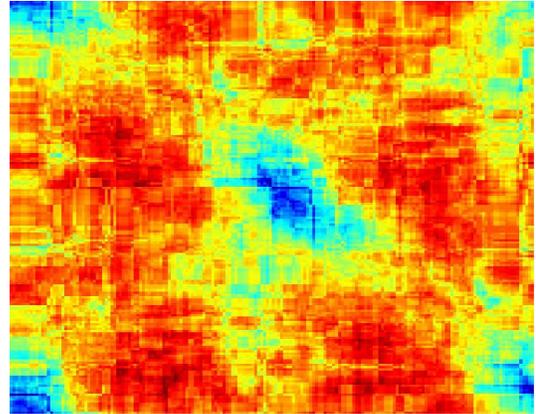


Figure 5: The profile of R_{ij} for the images of Figure 3. Blue color represent small values, red color represents high values.



Figure 6: Matching results with the greedy algorithm without using the cross-epipolar ordering constraint.

In the next two experiments, the features were selected automatically by the Harris corner detector. In the case of

Figure 7, 41 points were found in the first image, and 52 in the second image. By visual analysis, we verified that 20 points in each image had a correspondence in the other image. Our algorithm found 31 correspondences, 19 of which were correct. The greedy algorithm (which, as we said, used the same descriptors and was forced to find a maximum matching), found only 9 correct matches. The blue line shown in the figure corresponds to one candidate epipolar line for the optimal critical angle. It can be verified by inspection that this line is almost parallel to the actual epipolar lines.

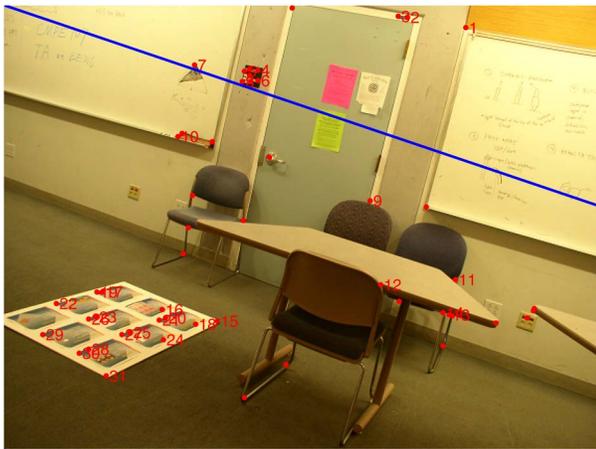
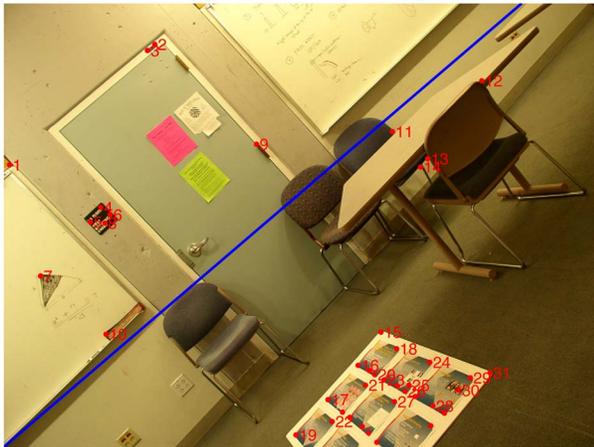


Figure 7: Matching results using our algorithm. The blue line is one candidate epipolar line corresponding to the optimal critical angle.

Figure 9 shows the results for third experiments. 26 points were selected by the corner detector in the first image, and 14 in the second one, with only 8 actual matches. Our algorithm found 9 matches, 6 of which were correct. The greedy algorithm found only 3 correct matches. In spite of the good result, the critical angle pair

that minimized R_{ij} was very different from the slopes of the actual epipolar lines (shown in the right upper corner of the images). This is a consequence of the ambiguity that is visible in the profile of R_{ij} , shown in Figure 10 (note that the graph is not square due to the different number of points, and therefore of critical angles, in the two images).

The CPU time for the direct and for the fast coarse-to-fine implementation of the algorithm on a 1 GHz PC are reported in Table 1. Note that the direct implementation is feasible only for a limited (~20) number of points. In our extensive experiments, the coarse-to-fine implementation gave the same result as the direct implementation in more than 90% of the image pairs considered.

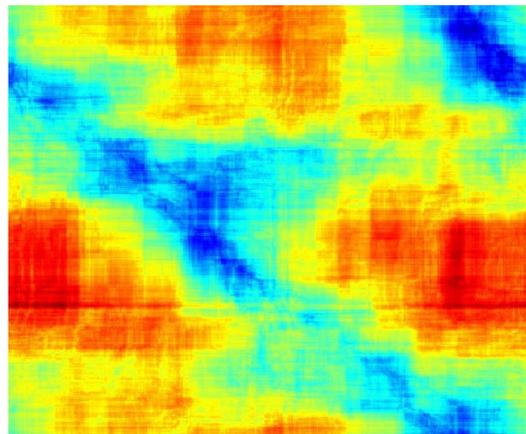


Figure 8: The profile of R_{ij} for the images of Figure 7.

IV. CONCLUSIONS AND FUTURE WORK

We have presented a novel technique for feature matching that constrains the search for the optimal assignment only to those matchings that are realizable. This has the double advantage of reducing the risk of mismatches (because only a subset of possible matchings is searched) as well as the computational cost with respect to exhaustive search over the set of all possible matchings. No prior knowledge of the intrinsic parameters of the cameras or their location is assumed, except that the epipolar lines of the stereo system are assumed to be parallel. Our experiments have shown the good performances of the proposed system as compared to a greedy algorithm that uses the same feature descriptors.

The assumption of epipoles at infinity is the most obvious drawback of our approach. Future work will aim at extending our algorithm to the general case. The main problem with an epipole in finite position is that it is more difficult to find all the orderings (and therefore the realizable matchings) induced by the cross-epipolar constraint.



Figure 9: Matching results using our algorithm. The blue line is one candidate epipolar line corresponding to the optimal critical angle. The insets at the top right corner show the actual epipolar lines.

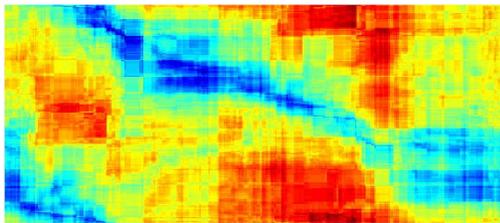


Figure 10: The profile of R_{ij} for the images of Figure 9.

Table 1: CPU times (in seconds) using direct (time 1) and coarse-to-fine (time 2) implementation.

N. points	10	20	30	40	50
CPU time 1	3 s	134 s	1409 s	N/A	N/A
CPU time 2	3 s	4 s	10 s	18 s	34 s

V. REFERENCES

- [1] H.H. Baker and T.O. Binford, "Depth from edge- and intensity-based stereo", *7th Joint Conf. On Artificial Intelligence*, 631-6, 1981.
- [2] A. Baumberg, "Reliable feature matching across widely separated views", *CVPR'00*, 774-81, 2000.
- [3] T.H.Cormen, C.E. Leiserson, R.L. Rivest, Clifford Stein, *Introduction to Algorithms*, P350, MIT Press, 1990.
- [4] F. Dellaert, S. M. Seitz, C. E. Thorpe, S. Thrun: "EM, MCMC, and chain flipping for structure from motion with unknown correspondence", *Machine Learning*, 50:45-71, 2003.
- [5] R. Deriche, Z. Zhang, Q.-T. Luong and O. Faugeras. "Robust recovery of the epipolar geometry for an uncalibrated stereo rig", *ECCV94*, 1:567-76, Stockholm, Sweden, May 1994.
- [6] O. Faugeras, *Three-dimensional computer vision*, The MIT Press, Cambridge, MA, 1996.
- [7] J.E. Hopcroft and R.M Karp, "An $n^{5/2}$ algorithm for maximum matching in bipartite graphs", *SIAM J. Computing*, 2:225-31, 1973.
- [8] L. V. Gool, T. Tuytelaars, V. Ferrari, C. Strecha, J. V. Wyngaerd and M. Vergauwen. "3-D modeling and registration under wide baseline conditions", *Proc. ISPRS Commission III*, Vol. 34, Part 3A, Graz, Sep. 2002, pp. 3-14.
- [9] L.V Gool, T. Moons, and D. Ungureanu, "Affine/photometric invariants for planar intensity patterns", *Proc. ECCV*, 642-51, 1996.
- [10] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2000.
- [11] J. Maciel and J. Costeira, "Towards a global solution of correspondence problem", *IEEE Trans. PAMI*, Feb. 2003.
- [12] K. Mikolajczyk, C. Schmid, "A performance evaluation of local descriptors", *CVPR '03*, Madison, Wisconsin, 2003.
- [13] G.V. Meerbergen, M. Vergauwen, M. Pollefeys, and L.V. Gool, "A hierarchical stereo algorithm using dynamic programming", *IEEE Workshop on Stereo and Multi-Baseline Vision*, 166-74, 2001.
- [14] D.G. Lowe, "Object recognition from local scale-invariant features", *Proc. ICCV*, 1159-57, 1999.
- [15] Y. Ohta and T. Kanade, "Stereo by intra- and inter-scanline search", *IEEE Trans. PAMI*, 7(2):139-54, 1985.
- [16] P. Pritchett and A. Zisserman. Wide baseline stereo matching, *Proc. 6th International Conference on Computer Vision, Bombay, India*, pages 754-760, January 1998.
- [17] J. Puzicha, Y. Rubner, C. Tomasi, and J. M. Buhmann. "Empirical evaluation of dissimilarity measures for color and texture", *ICCV'99*, 2:1165-73, Kerkyra, Greece, 1999.
- [18] C. Schmid, R. Mohr, and C. Baukhage, "Evaluation of interest point detectors", *International Journal of Computer Vision*, 37(2):151-72, 2000.
- [19] C. Schmidt and R. Mohr, "Local grayvalue invariants for image retrieval", *Trans. PAMI*, 19(5):530-4, 1997.
- [20] C. Schmidt, R. Mohr, and C. Baukhage, "Evaluation of interest point detectors", *IJCV*, 37(1):151-72, 2000.
- [21] S.L. Tanimoto, A. Itai, and M. Rodeh, "Some matching problem for bipartite graphs", *Journal of the ACM*, 16(4):517-25, 1978.
- [22] Y. Rubner, C. Tomasi and L. Guibas, "A metric for distributions with applications to image databases", *Proc. ICCV'98*, 1998.