SIXTH FRAMEWORK PROGRAMME
MARIE CURIE ACTIONS
HUMAN RESOURCES AND MOBILITY
MARIE CURIE RESEARCH TRAINING NETWORKS (RTN)

COMPUTATIONAL AND COGNITIVE VISION SYSTEMS:
A TRAINING EUROPEAN NETWORK

VISIONTRAIN

Martin Bujňák

Algorithms for Robust Matching in Computer Vision

Contract number: **MRTN-CT-2004-005439**
Project start/end dates: **1/5/2005 – 30/4/2009**
Reporting period: **1/5/2005 – 30/4/2007**

Phd advisor: **Radim Šára** (sara@cmp.felk.cvut.cz)
Host institution: CTU Prague (partner 9), Czech Republic

Contents

1 Introduction: The Phd topic

My work focuses on developing a new robust method for a problem of matching images given a set of regions of interest in each image. This activity falls under Topic 3 and belongs to **WP1: Computational theories and methods for low-level vision, Task 1.3**. We pose the problem as a problem of finding a strict sub-kernel in an oriented graph, where vertices are all possible correspondences, edges capture the structure of constraints and edges orientation represents pairwise comparison between two correspondences based on image similarity.

2 Brief review of the state of the art

The state-of-the-art pipeline for the robust (wide-baseline) matching problem consists of three basic steps:

1. Detecting image points/regions of interest [5],
2. selecting tentative correspondences [6], and finally
3. robust matching.

The RANSAC paradigm [3]) is used to solve the last step by finding the largest set of tentative correspondences that are consistent with the selected geometric model (e.g. epipolar geometry, homography [4], etc).

The success of such methods depends on the discriminability of interest point/region descriptor, the number of required tentative correspondences needed to build geometric model (minimum sample), the ratio of correct and incorrect tentative correspondences, as well as on the selected variant of the RANSAC based algorithm [1]. Optimal solution is guaranteed when the number of iterations in RANSAC main cycle goes to infinity. On the other hand, speed is proportional to the number of iterations and the main cycle terminates when probability that better sample is found is less than a given threshold. This probability is calculated from the estimated number of incorrect correspondences, the size of the minimum sample and the size of the best solution so far. Thus the number of iterations and time complexity of the method cannot be easily expressed in closed form for a given image pair.

3 Approach and methodology

In our work we present a robust deterministic approach to the problem of matching points under a mixture of non-parametric and parametric constraints with unknown parameters. An oriented graph is constructed where vertices are all possible correspondences and edges represent pairwise constraint violations. The orientation of the graph is obtained by image similarity from worse to better node in the edge. In such graph we can describe possible matchings as the set of independent vertex sets.

Note that searching for maximum independent vertex set is an NP [12] complete problem. Thus we pose the matching problem as a problem of finding of not maximum, but 'stable' matching i.e. finding a strict sub-kernel [10] in the oriented graph.

4 Scientific achievements

We have developed a quadratic time complexity algorithm for the class of problems where the parametric constraint has the form of a bijection (like homography).

Comparing to the state of the art pipeline our algorithm enhances the matching process in several ways:

- the algorithm works with all interest points/regions in both images and does not require tentative correspondences to be pre-selected,
- the algorithm works well even when the true correspondences constitute a very small fraction (0.1%) of all high-correlated correspondences,
- sampling is replaced by a deterministic algorithm with quadratic time complexity, where algorithm terminates and returns a unique (single) solution in guaranteed time,
- we have achieved robustness by a deterministic algorithm.

By robustness we understand that the algorithm is capable of rejecting unreliable data or data inconsistent with a model and to select the correct one i.e. the one consistent with a geometrical model.

Our approach to robust matching is based on a radically different notion than standard optimization or robust statistics methods. Robustness is directly addressed via the property of stability, while the loss of optimality is characterized [2]. This approach can be used as alternative method to the matching state of the art pipeline for several different tasks like panorama stitching, aerial images registration, structure from motion computation etc.

5 Results

An algorithm has been developed and implemented for estimating a general homography between two images like in the problem of image stitching into panoramas.

Recall that in our method we have translated the matching problem to the problem of finding of a strict sub-kernel [10] in the oriented graph.

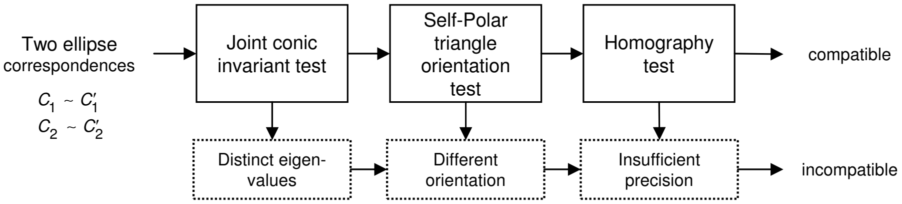
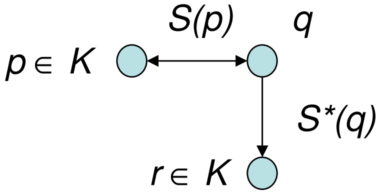


Figure 1: The structure of the geometric test.

Def. 1 *Let $K \subseteq V(G)$. A vertex $p \in K$ is stable with respect to K if every successor of p has a strict successor in $r \in K$.*



where $S(p)$ is a set of successors (bidirectional) of the vertex p and $S^*(q)$ represents a set of strict successors (edge is oriented from vertex q towards p).

Def. 2 *Independent vertex set $K \subseteq V(G)$ is a strict sub-kernel of G if every $p \in K$ is stable with respect to K.*

The graph nodes and edges were build using these components:

- As the image features we have selected ellipses oriented to a local affine frame (LAF) [6]. The ellipses were obtained by a fit to MSER regions [5].
- Uniqueness constraint and joint conic invariants [2] were used for constraint violation tests (edges).
- Edges were oriented by comparing intervals associated with correspondences (graph nodes). The intervals were build using similarity of the images associated with the ellipse and measured by normalized cross-correlation [2].

The edges of our graph are induced by the constraint violations. It means that two graph nodes (correspondences) are connected by an edge when they cannot by in the same solution. Thus two graph nodes must be enough to be able to perform these constraint violation tests. Note that two ellipse correspondences over-constrain a homography which gives a necessary condition for geometric compatibility test.

The geometric test is based on several necessary conditions that are progressively more expensive, see Fig. 1. The first subtest is based on joint conic invariants [7]: Given a pair of conics $\mathbf{C}_1 \leftrightarrow \mathbf{C}_1'$,

Figure 2: An unordered set of images with 3 panoramas on the left, matched images from one of panoramas in the middle and stitched panorama on the right.

$\mathbf{C}_2 \leftrightarrow \mathbf{C}'_2$, corresponding under homography the matrix $\mathbf{C}_2^{-1}\mathbf{C}_1$ has the same eigenvalues as $\mathbf{C}'^{-1}_2\mathbf{C}'_1$. This is used as a preliminary test to eliminate unpromising pairs. The second subtest is based on the eigenvectors of $\mathbf{C}_2^{-1}\mathbf{C}_1$ corresponding to three vertices of the self-polar triangle [7]: If mirroring does not occur and a correspondence is correct then the triangle will preserve its orientation with respect to the vertex order given by the eigenvalues. The last subtest is based on homography $\mathbf{H}$ estimated from the four bitangents of the ellipses $\mathbf{C}_1$, $\mathbf{C}_2$ and the corresponding bitangents between $\mathbf{C}'_1$, $\mathbf{C}'_2$ [9]. The bitangents are ordered by the point of tangency with respect to the first ellipse. We test if the error

$$e = \sum_{i=1}^{4} \|\mathbf{t}'_i - \mathbf{H}\mathbf{t}_i\|^2 + \|\mathbf{t}_i - \mathbf{H}^{-1}\mathbf{t}'_i\|^2 \qquad (1)$$

in pixels between the vertices of the transferred self-polar triangle $\mathbf{H}\mathbf{t}_i$ and their position in the second image $\mathbf{t}'_i$ exceeds a given threshold.

Unlike in RANSAC the algorithm does not need to recover homography in order to work. It can work with necessary geometric constraint instead. Our experiments also showed that joint conic invariants are sufficient to obtain good results.

We have achieved worst-case complexity $O(N^2)$, which is optimal (N^2 is the number of all possible pairwise correspondences). If there are multiple different models explaining the relation between the images and there is no additional information preferring one model over the other then the algorithm cannot decide which of the models is valid and it returns an empty solution.

This work has been submitted to ICCV 2007. Some results are shown in Fig. 2 where the algorithm was used to identify images from different panoramas and to find matches between images. Another example of application shows Fig. 3, where matches between two aerial images were found using our algorithm.

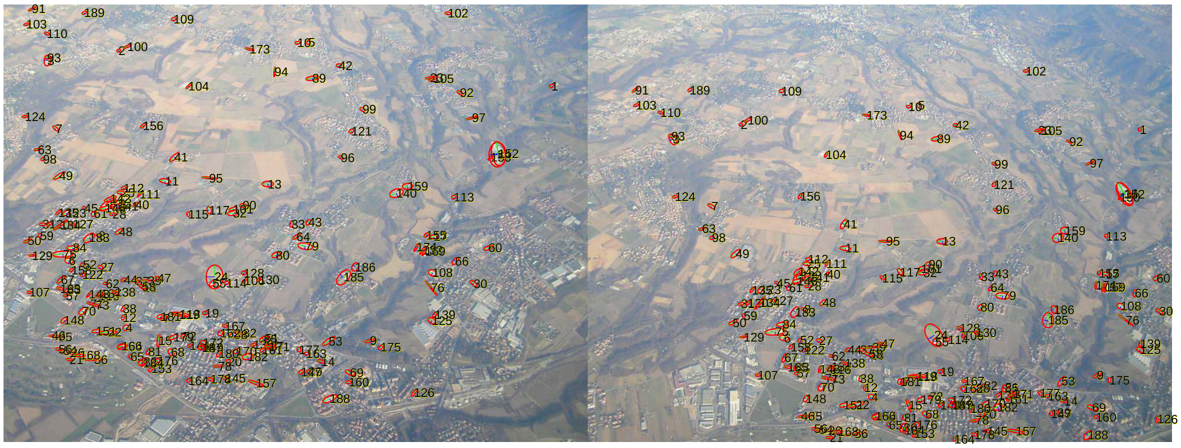


Figure 3: Matched MSER[5] regions of interest in two aerial images.

6 Future work

I plan extending the method to finding multiple geometric models. This requires to investigate a graph sub-structures with particular properties arising from the geometric test. A task is to find an efficient algorithm to detect these sub-structures. We can also try using the proposed method for finding epipolar geometry between two images. There are two promising ways how to achieve that:

1. by using recently developed calibrated 5-point relative pose algorithm [8] or 6-point relative pose and unknown focal length algorithm [11] for the geometric test directly with our algorithm,
2. by searching for multiple homographies (models) in the first step of the algorithm and then match detected homographies using homology constraint [4] as the geometric test in the second step of the algorithm.

Time complexity of our algorithm is optimal with respect to the number of tentative correspondences. On the other hand it seems that it is still possible to reduce the number of geometric tests using appropriate proximity representation. This may open a way to the solution of a fast k-image matching problem.

7 List of publications

We submitted paper to ICCV 2007 [2].

References

[1] 25 years of RANSAC. Workshop in conjunction with CVPR, June 2006.

[2] Martin Bujňák and Radim Šára. A robust graph-based method for the general correspondence problem demonstrated on image stitching. Submitted to ICCV 2007.

[3] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Comm ACM*, 24(6):381–395, 1981.

[4] Richard I. Hartley and Andrew Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, Cambridge, UK, 2000.

[5] J. Matas, O. Chum, M. Urban, and T. Pajdla. Robust wide baseline stereo from maximally stable extremal regions. In *Proc BMVC*, pages 384–393, 2002.

[6] Jiří Matas, Štěpán Obdržálek, and Ondřej Chum. Local affine frames for wide-baseline stereo. In *Proc ICPR*, pages 363–366, 2002.

[7] Joseph L. Mundy and Andrew Zisserman, editors. *Geometric Invariance in Computer Vision*, chapter Geometric Interpretation of Joint Conic Invariants, pages 77–86. The MIT Press, 1992.

[8] D. Nistér. An efficient solution to the five-point relative pose. *PAMI*, 26(6):756–770, June 2004.

[9] C. A. Rothwell, A. Zisserman, and al. Relative motion and pose from arbitrary plane curves. *IVC*, 10(4):250–262, 1992.

[10] Radim Šára. Robust correspondence recognition for computer vision. In Alfredo Rizzi and Maurizio Vichi, editors, *COMPSTAT 2006: Proceedings in Computational Statistics of 17th ERS-IASC Symposium*, pages 119–131, Heidelberg, Germany, August/September 2006. European Regional Section of the International Association for Statistical Computing, Physica-Verlag.

[11] Henrik Stewénius, David Nistér, Fredrik Kahl, and Frederik Schaffalitzky. A minimal solution for relative pose with unknown focal length. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, volume 2, pages 789–794, 2005.

[12] Clifford Stein Thomas H. Cormen, Ronald L. Rivest. *Introduction to Algorithms, Second Edition*. The MIT Press, Cambridge, UK, 2001.