A Graph-Based Approach to Corner Matching Using Mutual Information as a Local Similarity Measure

Manolis I. A. Lourakis, Antonis A. Argyros and Kostas Marias
Institute of Computer Science, Foundation for Research and Technology - Hellas (FORTH)
Vassilika Vouton, P.O. Box 1385, GR 711 10, Heraklion, Crete, GREECE
{lourakis, argyros, kmarias}@ics.forth.gr http://www.ics.forth.gr/cvrl/

Abstract

Corner matching constitutes a fundamental vision problem that serves as a building block of several important applications. The common approach to dealing with this problem starts by ranking potential matches according to their affinity, which is assessed with the aid of window-based intensity similarity measures. Then, actual matches are established by optimizing global criteria involving all potential matches. This paper puts forward a novel approach for solving the corner matching problem that uses mutual information as a window similarity measure, combined with graph matching techniques for determining a matching of corners that is globally optimal. Experimental results illustrate the effectiveness of the approach.

1. Introduction

Image corners, that is points of localized image structure which are formed at the boundaries of different brightness image regions, constitute one of the most widely used types of image features in computer vision. Determining the correspondence between two sets of corners extracted from a pair of views imaging the same scene, is a prerequisite to a broad range of vision tasks, including, among others, discrete motion estimation, feature-based stereo, object recognition and localization, image registration and camera self-calibration. The correspondence (or matching) problem, can be defined as that of identifying features in each set having distinct counterparts in the other set.

Most approaches in the literature for solving the correspondence problem implicitly assume that the employed images have a short baseline. In other words, it is assumed that the camera displacement and the change in camera orientation between images are small. Therefore, pixels lying within small rectangular image windows (templates) centered on corresponding corners should have similar intensities and their disparities could be well approximated by the same 2D translation. Considering two sets of corners detected in the images to be matched, the affinity of potential corner matches that satisfy a maximum disparity constraint is quantified with the aid of local, window-based similarity measures. Typically, the latter include metrics such as standard intensity cross-correlation, normalized (i.e. zero mean) cross-correlation and sum of squared intensity differences. Following the evaluation of the similarities corresponding to potential pair matches, actual matches are determined using various schemes that involve more global criteria. For example, Horaud and Skordas [5] identify maximal cliques in a relational graph, Zhang [13] employs relaxation labeling, Pilu [9] relies upon the SVD of a proximity matrix, Smith et al [11] propose a “winner takes all” strategy, Jung and Lacroix [6] estimate an approximate affine transformation among matches and Maciel and Costeira [8] resort to concave programming. In contrast to the above approaches, when the corners to be matched originate from images acquired from considerably different viewpoints, the assumption of the same disparity and photometric characteristics for all pixels in the neighborhoods of corresponding corners is no longer valid. In such cases, changes in illumination and perspective foreshortening effects should be more accurately accounted for. One of the first approaches along this line was that of Baumberg [1], who employs affine texture invariants which explicitly take into account linear transformations of the image data. Purely geometric approaches exploiting scene constraints have also been proposed [7].

This paper proposes a novel method for solving the short baseline corner matching problem that combines an information theoretic local similarity measure and graph matching techniques. More precisely, the strength of potential matches is assessed by employing mutual information, a random variables similarity measure that is well-established in the medical image registration domain. Then, corner matching is achieved through the solution of a maximum weight maximum cardinality flow problem on a graph whose edges are weighted by the mutual information scores. Using a state of the art graph algorithm, the associated flow problem can be solved very efficiently. To the best of our knowledge, the work reported here is the first to employ mutual information as a window similarity measure for corner matching. The rest of the paper is organized as follows. Section 2 introduces mutual information and provides a brief discussion regarding its use for comparing image regions. Section 3 defines the maximum weight maximum cardinality graph matching problem. Section 4 builds upon the two previous ones to describe the proposed corner matching technique. Experimental results are provided in section 5 and section 6 concludes the paper.

2. Mutual Information

Mutual information (MI) is an information theoretic similarity measure assessing the dependence of one random variable on another. MI has been extensively used in the area of multimodal medical image registration yielding very good results; a recent relevant survey can be found in [10]. MI is defined with the aid of Shannon entropies, without the need for knowing the exact functional form of the random variables to be compared. Specifically, let $p(a)$ and $p(b)$ be the probability distribution functions pertaining to two random variables A and B respectively. The mutual information I corresponding to the two random variables is defined as

$$I(A,\ B) = H(A)\ +\ H(B)\ -\ H(A,\ B), \qquad (1)$$

where $H(A)$ and $H(B)$ are the marginal entropies derived from the probability distribution functions corresponding to A and B, i.e.

$$H(A) = -\sum_a p(a)\log_2 p(a)\ ,\ H(B) = -\sum_b p(b)\log_2 p(b).$$

Moreover, $H(A,\ B)$ is the joint entropy derived from the joint distribution function $p(a,\ b)$ as

$$H(A,\ B) = -\sum_a\sum_b p(a,\ b)\log_2 p(a,\ b).$$

Intuitively, maximizing MI is equivalent to minimizing the joint entropy relative to the marginal ones. MI can thus be thought of as a measure of how well one random variable explains the other, i.e. a measure of the amount of information A contains about B and vice versa. When comparing digital images using MI, the intensities of pixels within image regions are treated as random variables whose probability density functions are either approximated using discrete gray level histograms or estimated using Parzen window techniques. MI is maximized at the optimal alignment of image regions, for which the amount of information they contain about each other is maximal. Rather than being restricted to comparing image regions related with a single 2D translation, MI can be adapted to handle regions undergoing more complex (e.g. affine) 2D transformations [10].

A known problem with MI is that it is computed on overlapping parts of images and is therefore sensitive to both the amount and content of overlap. To overcome this, Studholme et al [12] have proposed a normalized measure of MI that is less sensitive to changes in overlap:

$$NMI(A,\ B) = \frac{H(A)\ +\ H(B)}{H(A,\ B)}. \qquad (2)$$

3. The Maximum Weight Maximum Cardinality Graph Matching Problem

Owing to the fact that the task of corner matching will be formulated in section 4 as a maximum weight maximum cardinality graph matching problem, this section introduces the latter and discusses how can a publicly available software package be extended in order to deal with it.

Corner matching can be thought of as an application of the linear assignment problem. In the classical assignment problem, the goal is to find an optimal, one-to-one assignment of agents to tasks which ensures that all tasks are completed. The objective might be to minimize the total time to complete a set of tasks, to maximize profit, or to minimize the cost of the assignments. The assignment problem is a particular instance of a broader class of combinatorial optimization problems that can be reduced to a maximum weight maximum cardinality bipartite graph matching problem. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is called bipartite if the set of vertices $\mathcal{V}$ can be divided into two disjoint subsets such that no edge from $\mathcal{E}$ connects vertices in the same set. A matching in a graph $\mathcal{G}$ is a subset $\mathcal{M}$ of $\mathcal{E}$ such that no two edges share a common vertex. A maximum cardinality matching is a matching with a maximum number of edges. If the edges of the graph have an associated weight, then a maximum weight matching is a matching such that the sum of the weights of the edges in the matching is maximum. A maximum weight maximum cardinality matching is thus a maximum cardinality matching with maximum weight [2].

Since a maximum weight maximum cardinality matching problem can in turn be cast as a maximum weight flow problem, the notions of flow networks and flows are introduced next. A flow network is a directed graph $\mathcal{N}$ having two distinct nodes s (the "source") and t (the "sink"). A non-negative function $\mathrm{cap}()$ specifies for each edge e of $\mathcal{N}$ its capacity $\mathrm{cap}(e)$, that is the maximum amount of flow that can pass from e. A flow in a network is a real function defined on edges and satisfying the capacity constraint, i.e. the flow over an edge must not exceed its capacity, and the flow conservation constraint, i.e., the flow out of s must be the same as the flow into t. In a maximum weight flow problem, the goal is to convey a given amount of flow from s through the network to t, in such a way that the sum of the weights of the employed edges is maximum. To solve the maximum weight maximum cardinality matching problem for a bipartite graph $\mathcal{G}$, a corresponding flow network $\mathcal{G}' = (\mathcal{V}', \mathcal{E}')$ is defined as follows: Let s and t be new vertices not in $\mathcal{V}$ and $\mathcal{V}' = \mathcal{V} \cup \{s, t\}$. If the vertices of $\mathcal{G}$ are partitioned as $\mathcal{V} = \mathcal{L} \cup \mathcal{R}$, the edges $\mathcal{E}'$ of $\mathcal{G}'$ are given by $\mathcal{E}^{'} = \{(s,u) : u \in \mathcal{L}\} \cup \{(u,v) : u \in \mathcal{L},\ v \in \mathcal{R} \text{ and } (u,v) \in \mathcal{E}\} \cup \{(v,t) : v \in \mathcal{R}\}$. Each edge in $\mathcal{E}^{'}$ is assigned unit capacity. Edges that originate from the source or terminate at the sink are assigned zero weight. The remaining edges are assigned the weight of their corresponding edge in the bipartite graph. With the previous definitions plus a supply and demand of n flow units at s and t respectively, a maximum weight flow through $\mathcal{G}^{'}$ corresponds to a maximum weight matching for $\mathcal{G}$ [2]. The cardinality of this matching is equal to n, which is the amount of flow offered to $\mathcal{G}^{'}$ at its source.

A maximum weight flow problem can be easily transformed to a minimum weight (i.e. minimum cost) one by negating edge weights. A high performance, freely available code for the minimum cost flow problem is CS2, employing a scaling push-relabel algorithm and developed in C by Goldberg [3]. Apart from providing a solution to the minimum cost flow problem, CS2 is also capable of providing negative answers to cases that are unsolvable due to the flow supply at s exceeding the network's capacity. To cope with the maximum weight maximum cardinality matching problem, this last feature of CS2 can be exploited to extend it appropriately. Notice that if the maximum weight flow problem of cardinality m is unsolvable, the same holds true for all such problems having cardinality $m^{'} > m$. Hence, the maximum cardinality of a solution to the maximum weight flow problem can be determined by embedding CS2 in a binary search framework, as follows. Let U be an upper bound for the maximum cardinality of a solution to the maximum weight flow problem[1]. Since a trivial lower bound for the sought cardinality is 0, if the maximum weight flow problem of cardinality $\frac{U}{2}$ is solvable, then the maximum cardinality must be in the range $[\frac{U}{2}+1,\ U]$, otherwise it has to lie within $[0,\ \frac{U}{2}-1]$. CS2 is thus applied recursively using a flow supply equal to the midpoint of the current capacity range. The process stops when the current capacity lower bound exceeds or becomes equal to the upper one.

4. The Proposed Corner Matching Method

Conventional applications of MI for medical image registration aim at estimating a single, global 2D transformation relating the images to be registered. Therefore, they are based on estimating MI for large patches or even for the whole image frames. However, when dealing with corner matching, there exists no single 2D transformation relating all corresponding corners between the two images. Consequently, in this case MI needs to be estimated locally using small image windows centered on image corners.

The proposed method starts by detecting corners in both images. In our implementation, the Harris corner detector [4] has been employed. Most uses of matched corners do not benefit from large numbers of corners that are concentrated in the same image region. Aiming at making the detected corners more evenly distributed in images, the latter are divided into tiles using a 5×5 grid and the N most prominent corners are detected in each tile. Since the size of image windows used for computing MI is small, the corresponding discrete histograms for 8 bit grayscale images are likely to be sparse, having little statistical meaning. To avoid this contingency, the graylevels of window pixels are requantized using fewer than 256 intensity bins (64 in our implementation). In certain cases, it is also beneficial to smooth image windows prior to requantization using a 3×3 Gaussian filter. After the image window preprocessing operations have been completed, potential matches are determined and their affinity is assessed. More specifically, for each corner from the first image, all second image corners that are such that their underlying disparity is less than one third of the smallest image dimension, qualify as its potential matches. The similarity score of each pair of potential corner matches is computed from the corresponding image windows as their normalized MI defined by Eq. (2). The similarity score for potential corner matches exceeding this disparity threshold is taken to be equal to zero. Additionally, the similarity score of all pairs of potential matches whose NMI score is lower than a certain threshold is reset to zero. This is due to the fact that mismatched corners are expected to have lower NMI compared to correct ones, and therefore their number can be reduced by thresholding the NMI scores. Apart from reducing the mismatched corners, thresholding also reduces the number of potential matches, thus reducing the size of the whole matching problem.

The two sets of corners detected in the images define the vertices of a bipartite graph whose edges correspond to potential match pairs having a nonzero similarity score. Edge weights are determined from the NMI scores pertaining to the corresponding potential matching pairs. Evidently, a solution of the maximum weight maximum cardinality problem for this graph gives rise to a solution of the corner matching problem that maximizes both the total number of matched corners and the global similarity of the selected matches. To solve this optimization problem, the procedure outlined in section 3 is employed to transform the constructed bipartite graph into a flow network. The supply of incoming flow to this network, which determines the maximum cardinality of the corresponding matching problem, is determined using CS2 coupled with the previously explained binary search scheme. When dealing with applications where a certain, fixed number of matched corners m is desired, corner matching amounts to solving only one instance of the maximum weight graph matching problem. Thus, corner matching can be achieved by a variant of the proposed method which involves a single execution of CS2 on the flow graph with an input flow of m units. It is also

1 U can, for example, be selected as the smallest of the cardinalities of the two corner sets to be matched.

Figure 1. The first two images from the Arenberg castle sequence (courtesy U. Leuven/VISICS).

worth pointing out that the procedure just described for determining pair matches using graph matching techniques can be applied regardless of the exact similarity measure used for ranking potential pair matches.

5. Experimental Results

Due to space limitations, this section reports experiments conducted with a single image pair only. Qualitatively similar results were obtained when using other image pairs. Corners have been detected with subpixel accuracy using the Harris operator. Using bilinear interpolation, subpixel intensity information has also been computed for the 19×19 windows used for comparing the detected corners. To reduce the computational overhead for calculating NMI, lookup tables have been employed to precompute the logarithms involved in Eq. (2) for all possible discrete probability values. The experiment reported here was performed with the aid of the first two images of the well-known "castle" sequence, shown in Fig. 1. Around 600 corners were detected in each image and, after eliminating potential matches with a NMI score less than 1.15, 391 matches were established by the proposed method. Those matches were employed to robustly estimate the fundamental matrix corresponding to the two images, using the algorithm described in [14]. The root median square (RMedS) distance of points in the second image from their corresponding epipolar lines determined by the estimated fundamental matrix was 0.36 pixels, indicating that the majority of matched corners were indeed correct ones. Although space limitations prevent us from presenting quantitative timing results, we have found experimentaly that using CS2 for determining actual matches requires a few hundred milliseconds, thus being substantially faster than our optimized implementation of the relaxation labeling technique of [13]. Next, the performance of NMI as a window similarity measure was compared to that of two other measures, namely normalized cross-correlation (NCC) and sum of squared differences (SSD). To achieve this, NMI, NCC and SSD where in turn used to assess the similarity of potential matches and then the graph matching technique of section 4 was employed to determine the actual matches for each measure; no other parameters were changed among runs. Furthermore, to avoid fiddling with dissimilar similarity thresholds and at the same time ensure that the same number of corner matches was determined for each measure, the fixed cardinality variant of the proposed graph matching technique with a supply of 300 was employed. Thus, the solutions to three graph matching problems provided the 300 best matches according to each similarity measure. Then, the matches common to all three runs were determined and were found to be 247. The performance of each similarity metric is measured using the fraction of mismatches it generates. To avoid the tedious procedure of manually classifying matches into correct and erroneous ones, the 247 common matches were used to estimate the underlying fundamental matrix with [14]. Then, this estimate was used to label as mismatches the matches whose distance from their corresponding epipolar lines in the second image was larger than one pixel. NMI was found to produce 37 mismatches, i.e. 12.3% of the total matches, NCC 36 (12%) and SSD 31 (10.3%). These results indicate that NMI performs very similar to NCC/SSD and agree with the findings of [11] which reports that SSD outperforms NCC. However, before drawing definitive conclusions regarding the performance of NMI relative to NCC/SSD, further comparative experiments should be conducted.

6. Conclusions

This paper has presented a novel approach to the problem of corner matching. This approach is based on the use of mutual information and has demonstrated the feasibility of using the former as an intensity window similarity measure. Additionally, this paper has explained the reduction of the corner matching problem into a maximum weight maximum cardinality problem for which a solution can be obtained in a rigorous and efficient manner. Experimental results have demonstrated the effectiveness of the approach.

References

[1] A. Baumberg. Reliable Feature Matching Across Widely Separated Views. In *Proc. of CVPR'00*, pages 774–781, 2000.

[2] T. Cormen, C. Leiserson, and R. Rivest. *Introduction to Algorithms*. 1990.

[3] A. Goldberg. An Efficient Implementation of a Scaling Minimum-Cost Flow Algorithm. *J. Algorithms*, 22:1–29, Jan. 1997. Implementation at `http://www.avglab.com/andrew/soft.html`.

[4] C. Harris and M. Stephens. A Combined Corner and Edge Detector. In *Proc. of the 4th Alvey vision conference*, pages 147–151, UK, Sep. 1988.

[5] R. Horaud and T. Skordas. Stereo Correspondence Through Feature Grouping and Maximal Cliques. *IEEE Trans. on PAMI*, 11(11):1168–1180, Nov. 1989.

[6] I. Jung and S. Lacroix. A Robust Interest Points Matching Algorithm. In *Proceedings of ICCV'01*, volume 2, pages 538–543, 2001.

[7] M. Lourakis, S. Tzurbakis, A. Argyros, and S. Orphanoudakis. Feature Transfer and Matching in Disparate Stereo Views Through the Use of Plane Homographies. *IEEE Trans. on PAMI*, 25(2):271–276, Feb. 2003.

[8] J. Maciel and J. Costeira. A Global Solution to Sparse Correspondence Problems. *IEEE Trans. on PAMI*, 25(2):187–199, Feb. 2003.

[9] M. Pilu. A Direct Method for Stereo Correspondence Based on Singular Value Decomposition. In *Proc. of CVPR'97*, pages 261–266, 1997.

[10] J. Pluim, J. Maintz, and M. Viergever. Mutual Information Based Registration of Medical Images: A Survey. *IEEE Trans. on MI*, 22(8):986–1004, Aug.2003.

[11] P. Smith, D. Sinclair, R. Cipolla, and K. Wood. Effective Corner Matching. In *Proc. of BMVC'98*, volume 2, pages 545–556, Southampton, UK, Sep. 1998.

[12] C. Studholme, D. Hill, and D. Hawkes. An Overlap Invariant Entropy Measure of 3D Medical Image Alignment. *Pattern Recognition*, 32(1):71–86, 1999.

[13] Z. Zhang. A New and Efficient Iterative Approach to Image Matching. In *Proc. of ICPR'94*, pages 563–565, 1994.

[14] Z. Zhang, R. Deriche, O. Faugeras, and Q.-T. Luong. A Robust Technique for Matching Two Uncalibrated Images Through the Recovery of the Unknown Epipolar Geometry. *AI Journal*, 78:87–119, 1995. Also as INRIA RR-2273.