MVA2000 IAPR Workshop on Machine Vision Applications, Nov. 28-30, 2000, The University of Tokyo, Japan

7—1 Line Matching through Global Optimization constrained by Projective Invariants

Antonella Branca, Ettore Stella, Arcangelo Distante *
Istituto Elaborazione Segnali ed Immagini - C.N.R.

Abstract

In this work we propose a method for determining correspondences between sets of lines arising when observing a given scene from two different points of views. The method exploits results from the projective geometry and is capable of determining correspondences between images related by any arbitrary projective transformation. The proposed matching approach consists of an optimization process based on satisfaction of global constraints derived from the projective invariance of the cross ratio of any subset of four concurrent coplanar lines.

1 Introduction

Matching two sets of image features extracted from a pair of views of the same scene represents the fundamental step of most computer vision tasks as discrete motion estimation, feature-based stereo, object recognition, topological navigation.

In this work we consider the context of human made environments mainly characterized by planar surfaces, such as walls, floors, ceilings and doors, rich in linear features. In this context straight line segments represent the main features to be considered in any reconstruction process.

Most of the approaches [10] proposed in the past in literature for line matching have been based on the invariance of some geometric attribute that cannot be preserved under any perspective transformation.

Recently, some works have exploited some results of the projective geometry to determine more robust approaches [9], however most of them have concentrated on the epipolar geometry requiring to know a priori some geometric relations as the fundamental matrix or the trifocal tensor.

On the contrary the method we propose, based on some global constraints derived from the projective geometry theory, allows us to determine robust correspondences between line features extracted from two views related by any arbitrary projective transformation, without any a priori information as epipolar geometry, projective camera matrices, corner correspondences. It is only assumed that the scene is locally approximated by a sufficient number of rectangular planar surfaces whose junctions give rise to sets of parallel straight line segments, that are projected in the image plane, under any projective transformation (collineation), into pencils of concurrent lines.

Planarity implies very strong constraints on the kind of correspondences that occur between two images. More precisely, if the tokens that are put into correspondence are produced by coplanar visual features, there exists a linear transformation from the first image coordinates to the second image coordinates. This analytic transformation is a collineation between the two retinal planes considered as projective planes [1], and it is completely specified by a 3 by 3 transformation matrix named homography.

Four concurrent coplanar lines have as their projective invariant the cross-ratio estimated as ratio of ratios of the line parameter distances.

Our aim is to propose a line matching approach consisting of an optimization process constrained by the projective invariance of the cross-ratio of any subset of four concurrent coplanar lines.

Given two views, the two respective sets of lines extracted by connecting adjacent segments extracted from each image using the Canny edge detector are given as input to our algorithm.

Matching of lines is performed by imposing globally the invariance of the cross-ratio of each subset of four lines.

This approach allow us at same time to recover automatically all pencils present in the scene and to match all the related features.

2 Feature Matching

The matching problem involves two sets of features extracted from two different images of a sequence. We assume that the scene is locally approximated by a sufficient number of rectangular planar

*Address: Via Amendola 166/5, 70126 Bari, Italy. E-mail: (branca,stella,distante)@iesi.ba.cnr.it

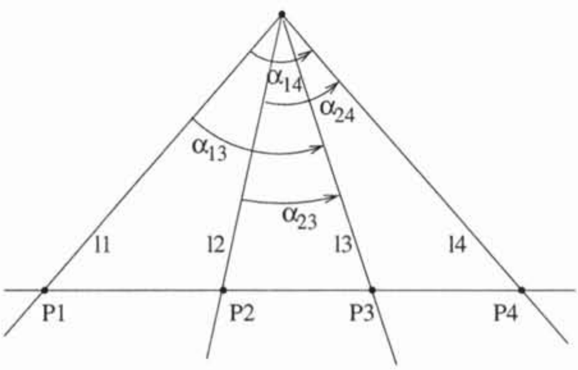

Figure 1: The pencil of four coplanar lines has a cross ratio defined by the angles between lines. Any line intersecting the pencil has the same cross-ratio for the points of intersection of the line with the pencil.

surfaces whose junctions give rise to sets of parallel straight line segments. Parallel lines under a projective transformation meet to a single point giving rise to a pencil of lines. The features to which we are interested will belong to pencils generated by sets of horizontal or vertical coplanar lines present in the scene.

We perform the matching of lines in the parameter space, where all small line segments, obtained by merging sets of neighboring collinear edge strings provided by the Canny edge detector, are represented by a single point.

A pencil of lines is dual with a single line in the parameter space. All lines belonging to a single pencil are represented by collinear points in the parameter space.

Four collinear points $\mathbf{P}_{\mathbf{i}_{i=1,..,4}}$ have as their invariant the cross-ratio defined as ratio of ratios among their distances:

$$cr(\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}, \mathbf{P_4}) = \frac{(X_3 - X_1)(X_4 - X_2)}{(X_3 - X_2)(X_4 - X_1)} \quad (1)$$

where $X_{i_{i=1,..,4}}$ represent the corresponding positions of each point along the line, and $(X_i - X_j)$ is the distance between points $\mathbf{P_i}$ and $\mathbf{P_j}$.

The coordinates X_i of four collinear points are the parameters (a_i, b_i) of four lines $\mathbf{l_i}$ incident to a single point.

Conversely, four concurrent coplanar lines $\mathbf{l_i}$ have as their invariant the cross ratio of their parameters $\mathbf{P_i}$:

$$cr(\mathbf{l_1}, \mathbf{l_2}, \mathbf{l_3}, \mathbf{l_4}) = cr(\mathbf{P_1}, \mathbf{P_2}, \mathbf{P_3}, \mathbf{P_4}) \quad (2)$$

To each subset of four concurrent lines $\mathbf{l_i}$ observed from one view the corresponding lines $\mathbf{r_i}$ in any other view will have the same cross-ratio.

3 The Optimization Model

Most of previous attempts [2, 3, 4, 5, 6, 7], to use the projective invariance of cross-ratio for matching have encountered a lot of difficulties being based on a probabilistic analysis [8]. The performance of probabilistic approaches depends on the choice of rule for deciding whether features have a given cross-ratio and thresholds on the probabilities in the decision rule.

In our work we overcome all such difficulties by exploring the space of all potential matches through an optimization approach.

Exploring the space of potential matches spares us to deciding about local measurements of cross-ratio and give rise to a more robust approach. In order to overcome the exponential complexity, that this global search algorithm could require, we formulate the matching problem through a graph-matching algorithm aiming to estimate a maximum clique into an optimal manner.

We formulate the matching problem using graph matching theory.

The two images to be matched are represented by two *relational graphs*, characterized by

- nodes $(\{n1_i\}, \{n2_j\})$ representing the feature-lines to be matched
- four-order links $(\{L1_{i_1 i_2 i_3 i_4}\}, \{L2_{j_1 j_2 j_3 j_4}\})$ weighted by the cross-ratio $(CR1_{i_1 i_2 i_3 i_4}, CR2_{j_1 j_2 j_3 j_4})$ evaluated on the features associated to the four connecting nodes.

Optimal matches are recovered by searching on the association graph $G = \{\{n_h\}, \{l_{hklm}\}\}$ consisting of nodes $\{n_h\}$ (the candidate matches) connected by four-order links $\{l_{hklmn}\}$ weighted by the cross ratio similarity (3) among the four connected match nodes.

$$C_{hklm} = e^{-(|CR1_{h_i k_i l_i m_i} - CR2_{h_j k_j l_j m_j}|^2)} \quad (3)$$

Theoretically, the association graph G should consists of $N \times M$ nodes $\{n_h\}$, representing candidate matches $\{(\mathbf{P_{h_i}}, \mathbf{P_{h_j}})\}$ and $\binom{N \times M}{4}$ links weighted by compatibility matrix values $\{C_{hklm}\}$. Actually, in order to reduce the number of links involved in the process, we generate a node n_h only if the corresponding features have an high radiometric similarity. We estimate a radiometric measure by averaging the intensity grey values along the neighborhood of each segment of a line.

The nodes representing correct matches among concurrent lines will be mutually compatible according to the weights of the connection links.

The goal is to select all matches mutually compatible according to compatibility matrix C.

The method we have used to do so is based on an iterative nonlinear relaxation labeling process.

A relaxation labeling process takes as input an initial labeling (matching) assignment and iteratively updates it considering the compatibility model until a stable state is reached.

In our problem, to each node n_h is assigned the initial labeling Λ_h representing the degree of confidence of the hypothesis $\mathbf{P}_{\mathbf{h_i}}$ *is matched with* $\mathbf{Q}_{\mathbf{h_j}}$. The relaxation algorithm updates the labeling $\{\Lambda_h\}$ in accordance with the compatibility model, by using a *support function* γ (4) quantifying the degree of agreement of the h-th match with the context.

$$\gamma_h = \sum_{klm} C_{hklm} \Lambda_k \Lambda_l \Lambda_m \qquad (4)$$

The rule of adjustment of the labeling $\{Lambda_h\}$ should increases Λ_h when γ_h is high and decreases it when γ_h is low. This leads to the following updating rule:

$$\Lambda_h = \Lambda_h \gamma_h / \sum_k \Lambda_k \gamma_k \qquad (5)$$

Iteratively all labeling of $\{n_k\}$ nodes are updated until a stable state is reached. In the stable state all $\{n_h\}$ nodes with non-null labeling Λ_h will represent optimal matches. This algorithm posses a Liapunov function. This amounts to stating that each relaxation labeling iteration actually increases the labeling consistency, and the algorithm eventually approaches the nearest consistent solution.

An important property of relaxation labeling process is to provide as output a maximal clique, i.e. a set of nodes mutually compatible (totally connected), so that none other set of nodes mutually compatible including it there exists.

In our context a relaxation labeling process detects a set of concurrent coplanar lines correctly matched.

The whole set of correct matches can then be obtained by applying iteratively the relaxation process on the whole set of potential node-matches pruned of all node-matches selected in the previous iterations.

4 Experimental Results

We have performed a number of tests on real time-varying image sequences acquired in our laboratory with a TV camera (6mm focal length) mounted on a *NOMAD* vehicle. The sequences here reported (fig.1,2) have been acquired while the vehicle was moving toward two doors. The aim of this experiment has been to test the performance of the approach to match correctly a set of coplanar straight lines arising in this context, in order to be able to construct a "door-detection" module useful to an autonomous moving vehicle to perform the self

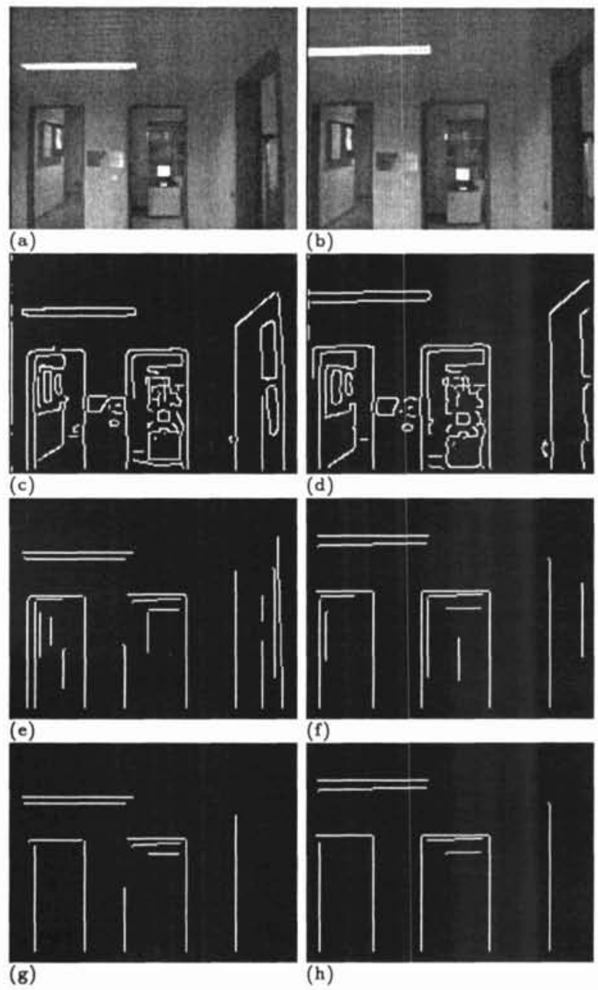

Figure 2: (a)-(b) The two grey level images to be matched, (c)-(d) the images of the edges recovered with the Canny edge detector, (e)-(f) the straight lines selected to be matched (by projection in the parameter space, and verification of collinearity), (g)-(h) the matched lines.

location from the correspondences found between an image model of a "door" and a new image.

5 Conclusions and Future Work

An optimization approach to perform matching of coplanar lines which arise in most human-made indoor environments has been proposed. The method is able to process any set of 3D parallel straight lines when projected in the image plane in concurrent lines generating a pencil. This is not a loss of generality because in indoor environments most of objects or 3D structures carrying on significant information for the scene understanding are rich of this type of straight lines. The approach has the main advantage of don't requiring any a priori information: epipolar geometry, projective camera matrices.

Moreover, though the constraint that is used to detect correct matches is true only for concurrent coplanar lines, it is not required to detect a priori all the appropriate subsets. In fact, our optimization approach detects automatically all subsets of concurrent coplanar line segments correctly matched.

The method is very suitable in the reconstruction of human made environments both indoor (rooms characterized by doors, floor, walls) or outdoor (urban scenes).

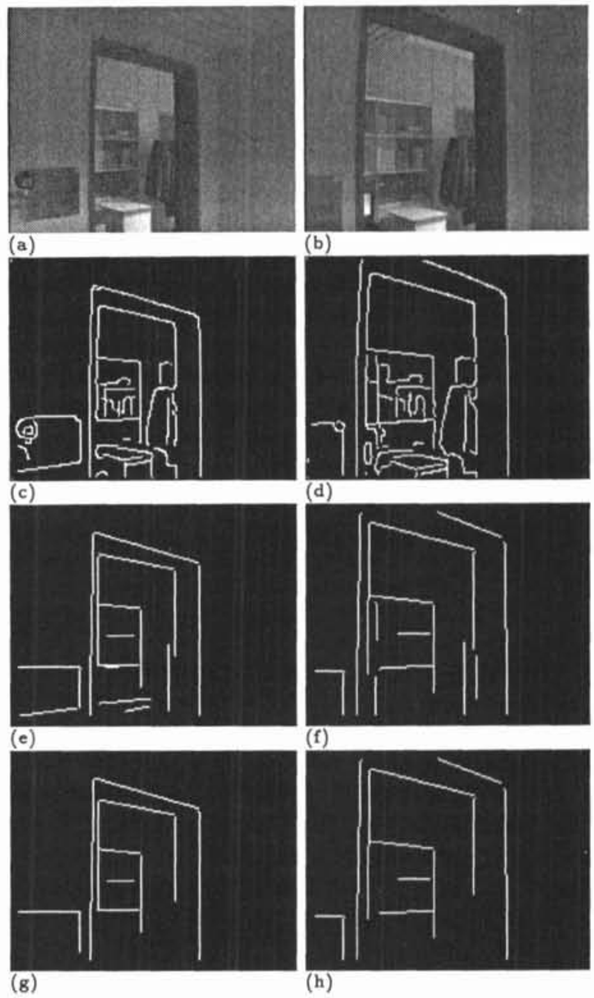

Figure 3: (a)-(b) The two grey level images to be matched, (c)-(d) the images of the edges recovered with the Canny edge detector, (e)-(f) the straight lines selected to be matched (by projection in the parameter space, and verification of collinearity), (g)-(h) the matched lines.

References

[1] J.Semple and G.Kneebone *Algebraic Projective Geometry* Oxford University Press, 1979.

[2] H.Chabbi, M.O.Berger "Using Projective Geometry to Recover Planar Surfaces in Stereovision" *Pattern Recognition* Vol.29, No.4, pp.533-548, 1996.

[3] D.Sinclair, A.Blake "Qualitative Planar Region Detection" *International Journal of Computer Vision* Vol 18 No 1 pp 77-91 (1996).

[4] P.Gurdjos, P.Dalle, S.Castan "Tracking 3D Coplanar Points in the Invariant Perspective Coordinates Plane" *Proc. of ICPR'96*

[5] S.Carlsson "Projectively Invariant Decomposition and recognition of Planar Shapes" *International Journal of Computer Vision* Vol 17 No 2 pp 193-209 1996.

[6] C.A.Rothwell,A.Zisserman,D.A.Forsyth,J.L.Mundy "Planar Object Recognition using Projective Shape Representation" *International Journal of Computer Vision* Vol 16 pp 57-99 1995.

[7] K.Kanatani "Computational Cross Ratio for Computer Vision" *CVGIP: Image Understanding* Vol 60 No 3 pp 371-381 1994.

[8] S.J.Maybank "Probabilistic Analysis of the Application of Cross Ratio to Model Based Vision" *International Journal of Computer Vision* Vol 16 pp 5-33 (1995)

[9] C.Schmid and A.Zisserman "Automatic Line Matching across Views" *CVPR'97* 666-671, 1997.

[10] R.Horaud and T.Skordas "Stereo correspondence through feature grouping and maximal cliques" *PAMI*, 1989.