Feature Transfer and Matching in Disparate Stereo Views through the use of Plane Homographies

M. Lourakis, S. Tzurbakis, A. Argyros, S. Orphanoudakis

Computer Vision and Robotics Lab (CVRL)
Institute of Computer Science (ICS)
Foundation for Research and Technology – Hellas (FORTH)
Heraklion, Crete, Greece

April 2002
CMP, Prague

Presentation outline

- Problem statement
- Overview
- Method description
- Experimental results
- Discussion

Problem statement

"Given some corresponding features in two stereo images, match them with features extracted from a second stereo pair that is captured at a distant location"

Overview

- Matching is a fundamental problem in Computer Vision, appearing in different forms in tasks such as:
  - Discrete motion estimation
  - Feature-based stereo
  - Object recognition
  - Image registration
  - Camera self-calibration
  - Image based rendering

- Fundamental difficulty: Feature matching should rely upon the affinity of photometric or geometric properties between different views of a scene.

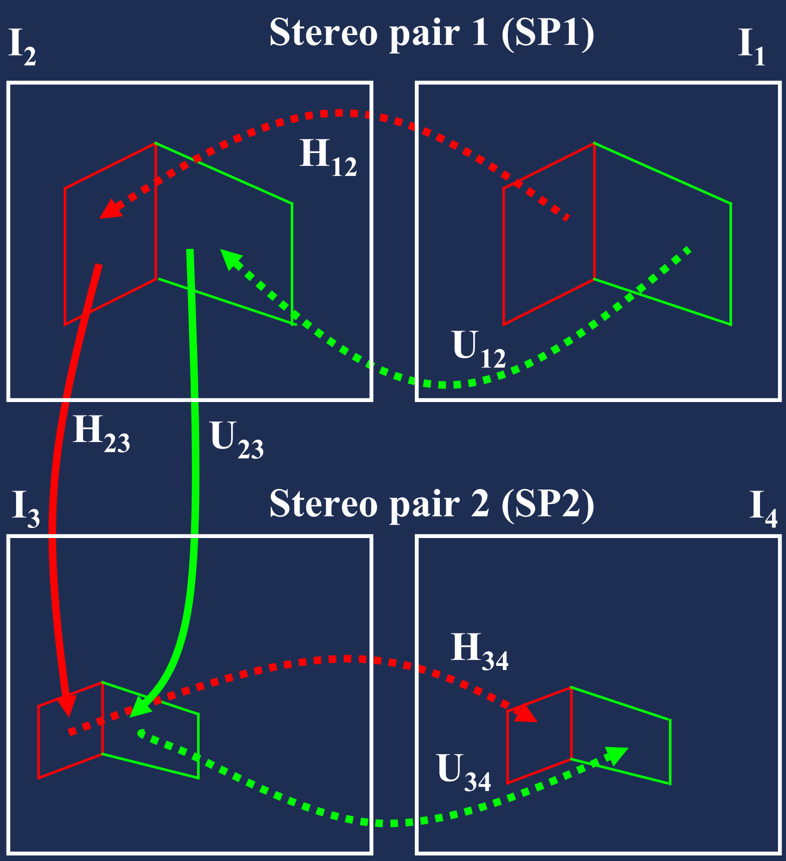
Existing approaches

- Adopt a semi-automatic approach and assume a priori knowledge of geometric constraints that are satisfied by the different views, e.g:
  - Georgis, Petrou, Kittler (IVC ‘98),
  - Schmid and Zisserman (CVPR ‘97)
  - Pritchett and Zisserman (ICCV ‘98),
  - Faugeras and Robert (IJCV ‘96)
- Exploit quantities that remain unchanged under general perspective projection and can be directly computed from the images (i.e. projective invariants). Due to the lack of general-case view invariants, make assumptions regarding the structure of the viewed scene , e.g:
  - Meer, Lenz, Ramakrishna (IJCV ‘98)
  - Lourakis, Halkidis, Orphanoudakis (IVC ‘00)

Our approach

- Assume that the viewed scene contains at least two planes
- Use existing approaches (Lourakis, Halkidis, Orphanoudakis, IVC ‘00) to match planes between disparate views.
- Match points and lines not belonging to the two planes using geometric constraints defined by the matched planes.

Our approach



SP1 → **(A)** Compute feature correspondences → **(B)** Segm. dominant planes and compute H_{12}, U_{12}

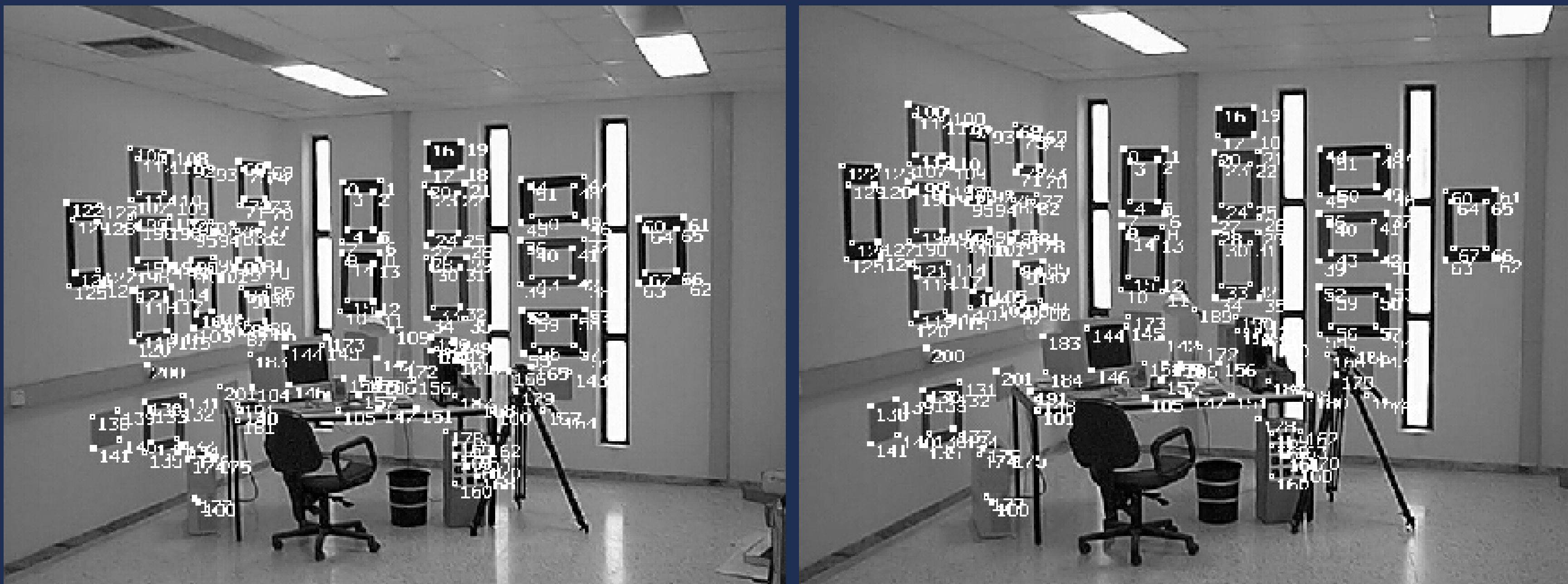
SP2 → **(A)** Compute feature correspondences → **(B)** Segm. dominant planes and compute H_{34}, U_{34}

(C) Match planes between stereo pairs and compute homographies H_{23} and U_{23}

(D) Match features between I_2 and I_3, using reprojection based on H_{23}, U_{23} and geometric constraints

(A) Computing Intra-stereo feature correspondences: points

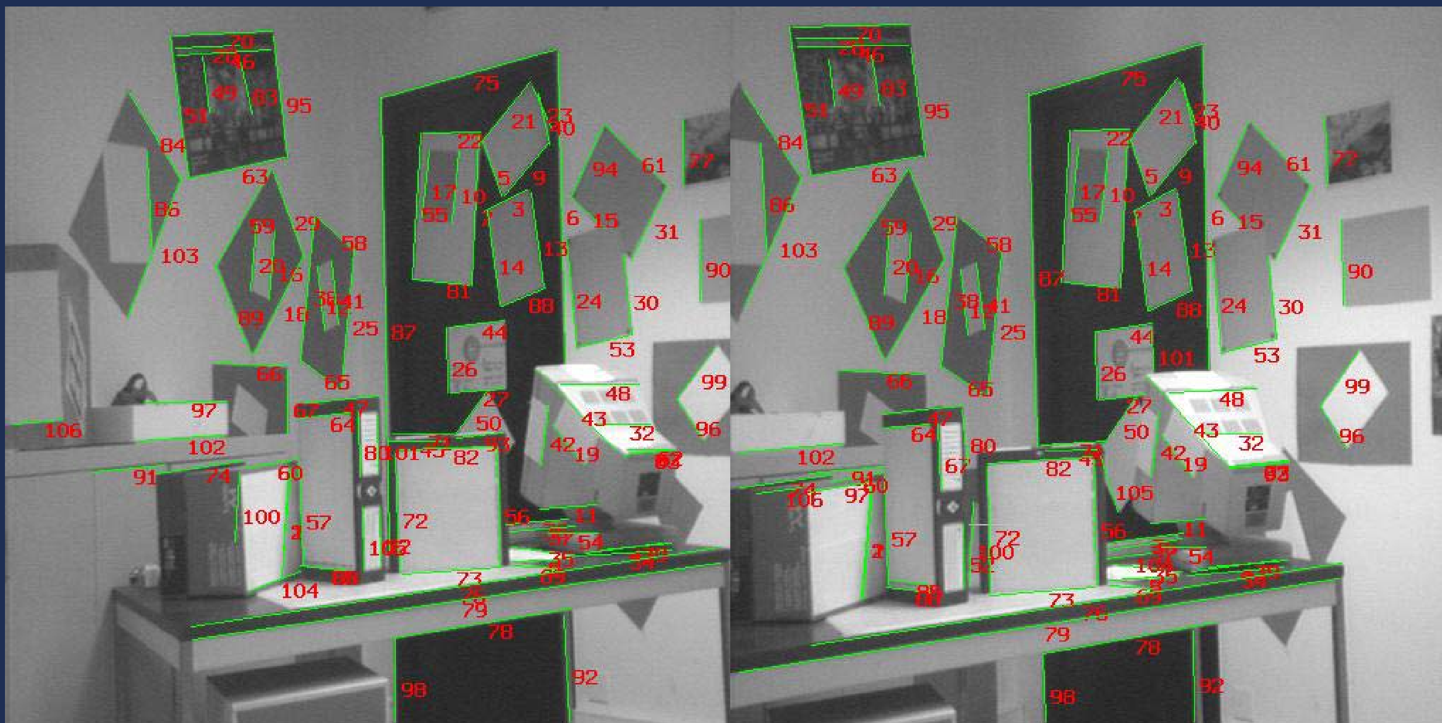
- Point correspondences based on [1]



[1] 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*, vol. 78, pp. 87–119, 1995.

(A) Computing Intra-stereo feature correspondences: lines

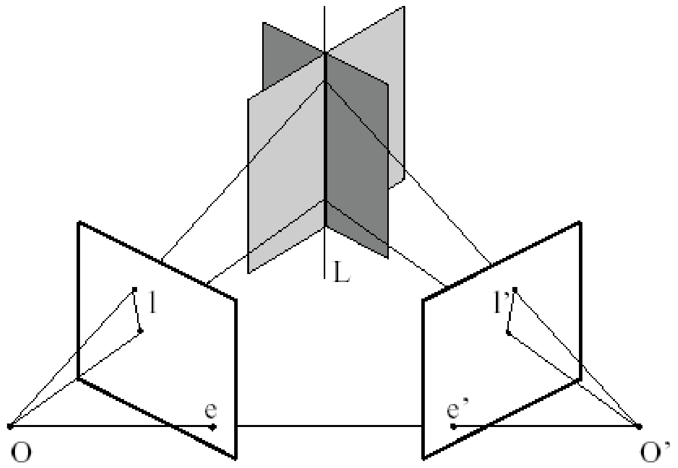
- Line correspondences based on [2]



[2] M.I.A. Lourakis, "Establishing Straight Line Correspondence," Tech. Rep. 208, ICS/FORTH, Aug. 1997, available at ftp://ftp.ics.forth.gr/tech-reports/1997.

(B) Segmentation of planar surfaces

A plane can be defined by a point and a line



$$\mathbf{H}(\mu) = [\mathbf{l}']_\times \mathbf{F} + \mu \mathbf{e}' \mathbf{l}^T, \quad \mu \in R$$

$$([\mathbf{l}']_\times \mathbf{F}\mathbf{p}) \cdot \mathbf{q} + \mu(\mathbf{e}' \mathbf{l}^T \mathbf{p}) \cdot \mathbf{q} = 0.$$

- The homographies of the planes defined by all possible pairs of lines and points are computed.

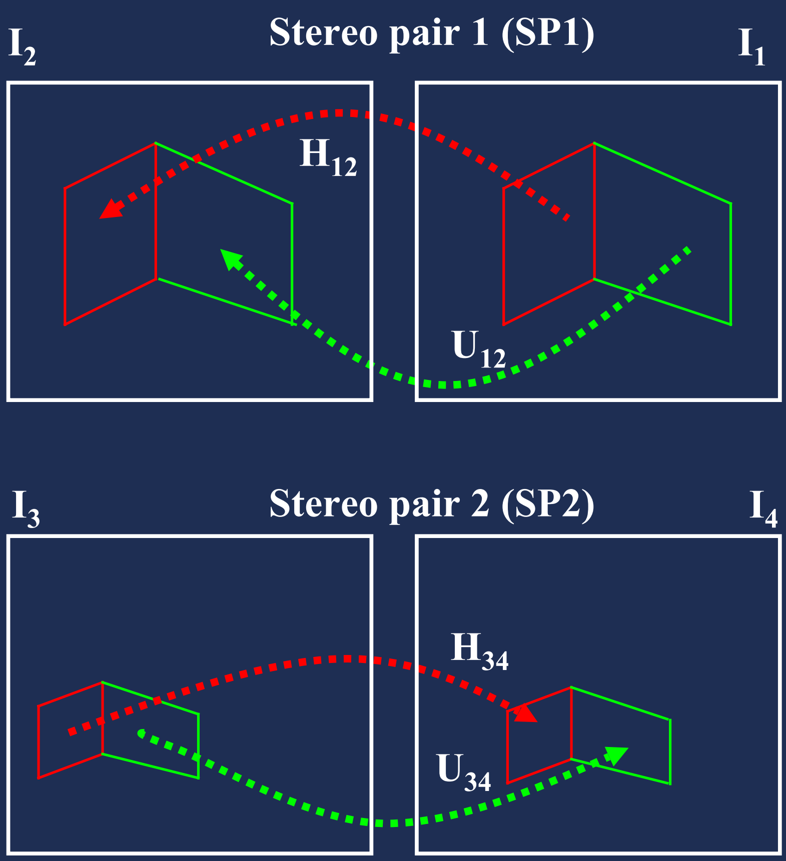
(B) Segmentation of planar surfaces

- Each of the induced homographies is used to predict the location of every feature in one image to the other image.
- Each feature votes in favor of the homography that best predicted its location in the other image. In addition, this feature is assumed to belong to the plane defined by the homography in question.
- The two planes that receive the largest and second largest number of votes are identified as the two most prominent ones.
- The homographies are re-estimated using robust regression, based on the full set of features belonging to them, based on [3].
- Using the homography estimates as initial solutions the two homographies are refined using a non-linear method, based on [4].

[3] M.I.A. Lourakis, S.T. Halkidis, and S.C. Orphanoudakis, "Matching Disparate Views of Planar Surfaces Using Projective Invariants," *Image and Vision Computing*, vol. 18, pp. 673–683, 2000.

[4] Q.-T. Luong and O.D. Faugeras, "Determining the Fundamental Matrix with Planes: Instability and New Algorithms," in *Proc. of CVPR'93*, 1993, pp. 489–494.

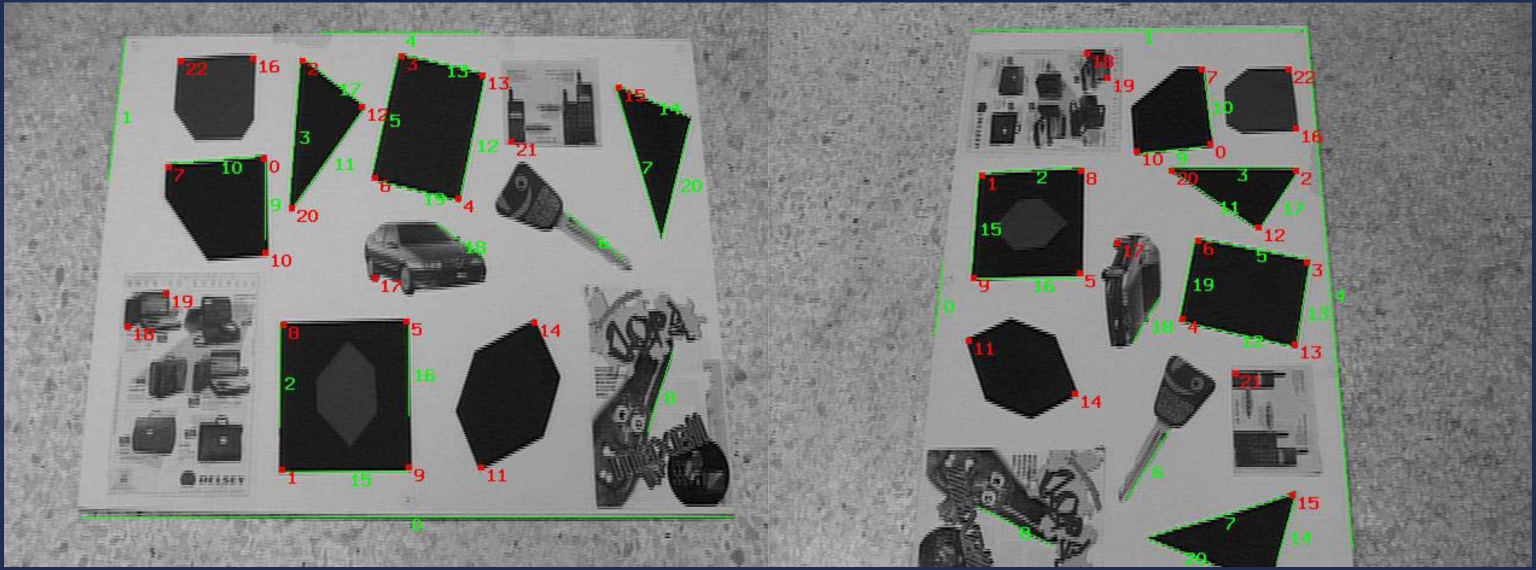
Our approach



SP1	SP2
(A) Compute feature correspondences	**(A)** Compute feature correspondences
(B) Segm. dominant planes and compute H_{12}, U_{12}	**(B)** Segm. dominant planes and compute H_{34}, U_{34}

(C) Matching planes between disparate views

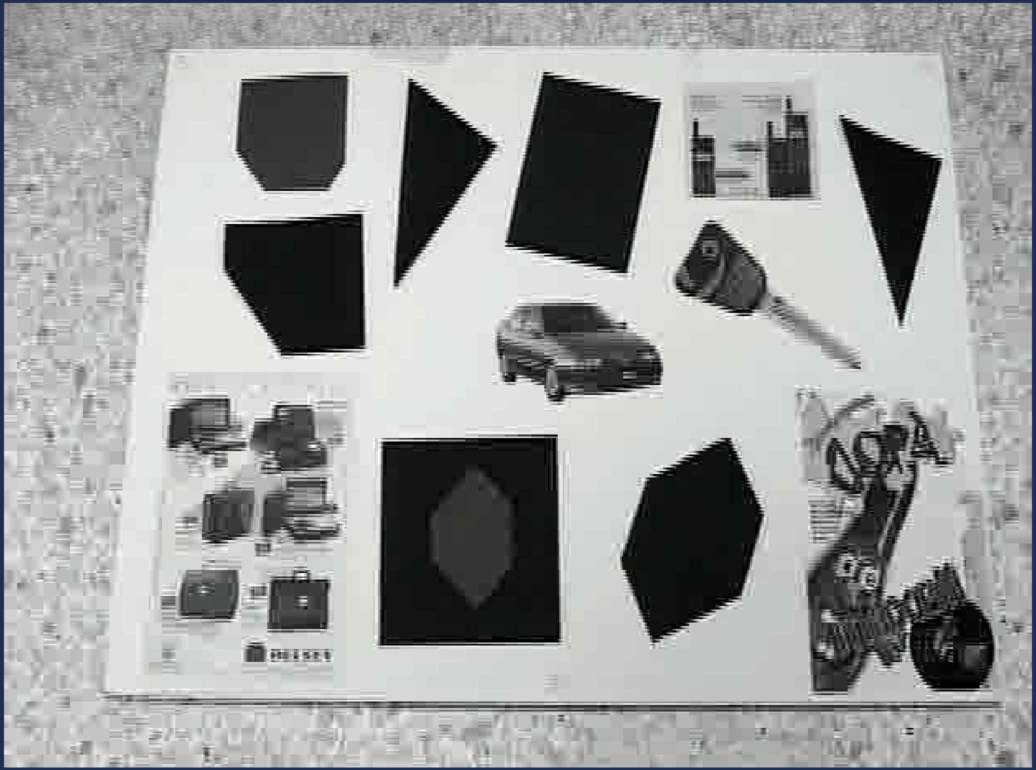
- Employ a randomized search scheme, guided by geometric constraints derived using the two-line-two-point projective invariant [3]



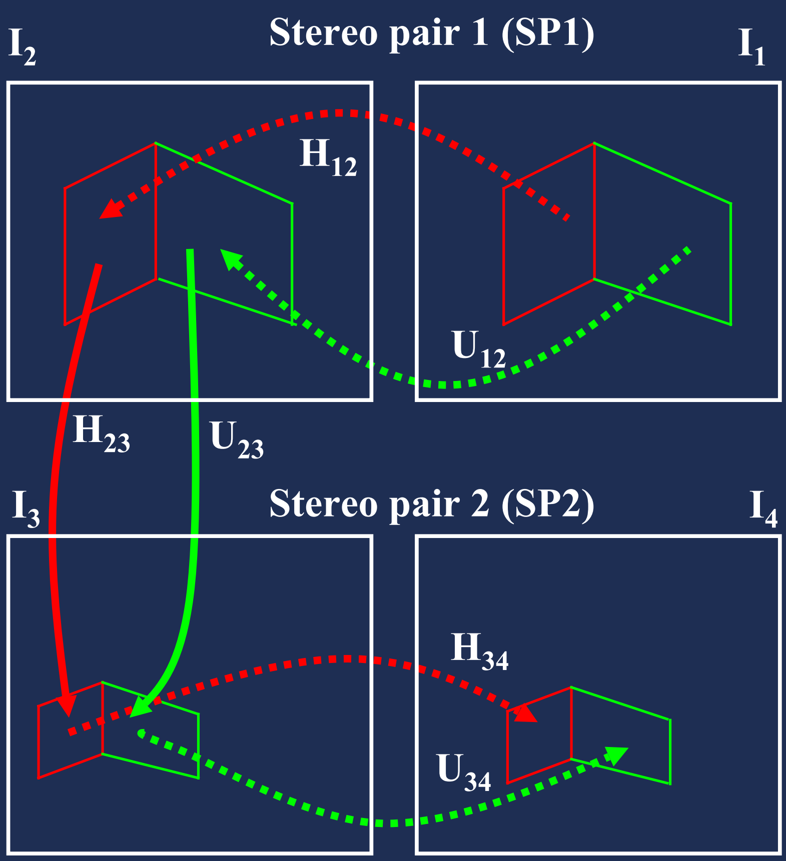
[3] M.I.A. Lourakis, S.T. Halkidis, and S.C. Orphanoudakis, "Matching Disparate Views of Planar Surfaces Using Projective Invariants," *Image and Vision Computing*, vol. 18, pp. 673–683, 2000.

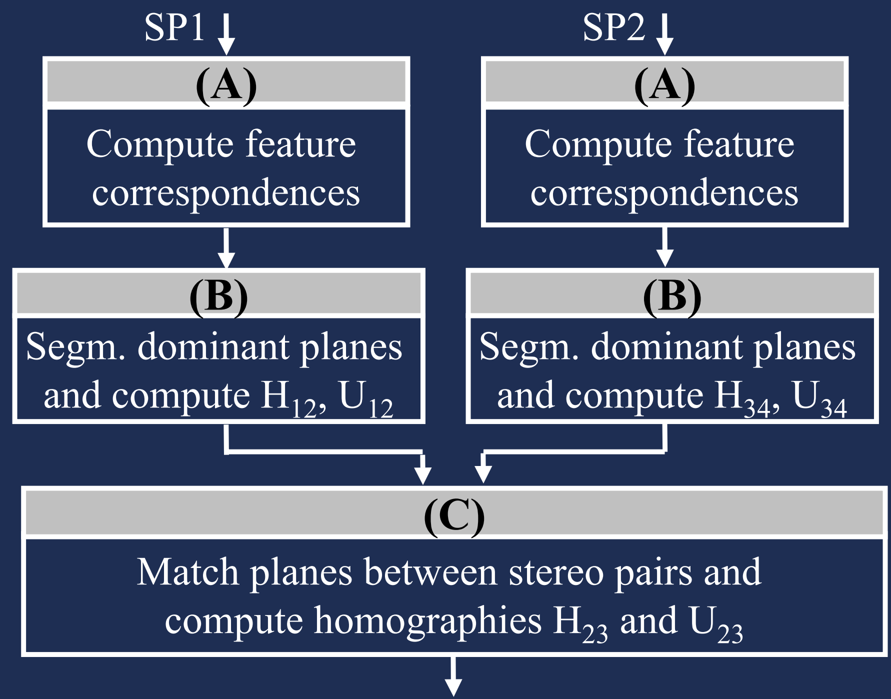
(C) Matching planes between disparate views

- Result of homography-based warping of second image onto the first:

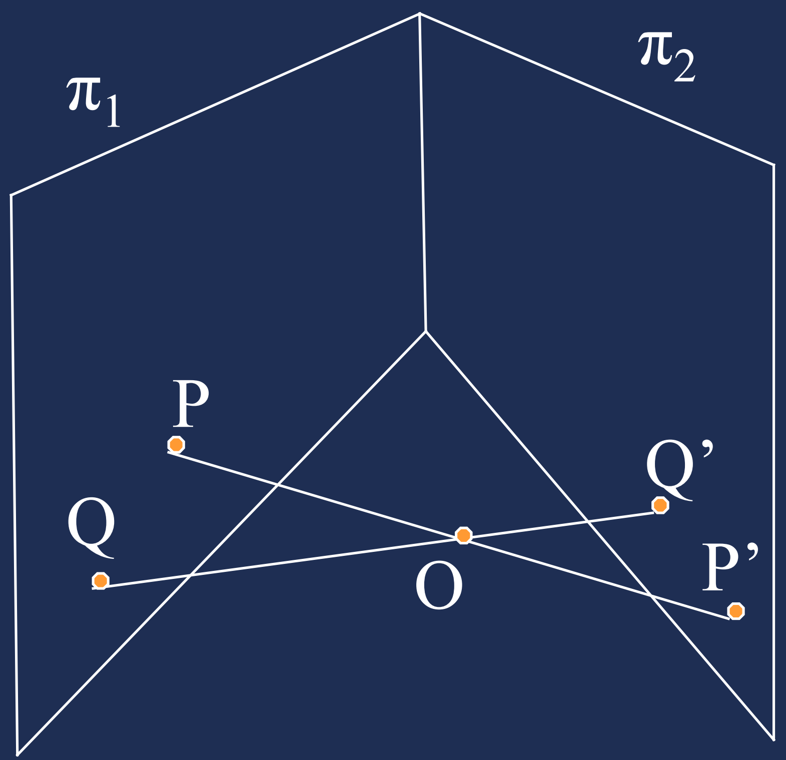


Our approach





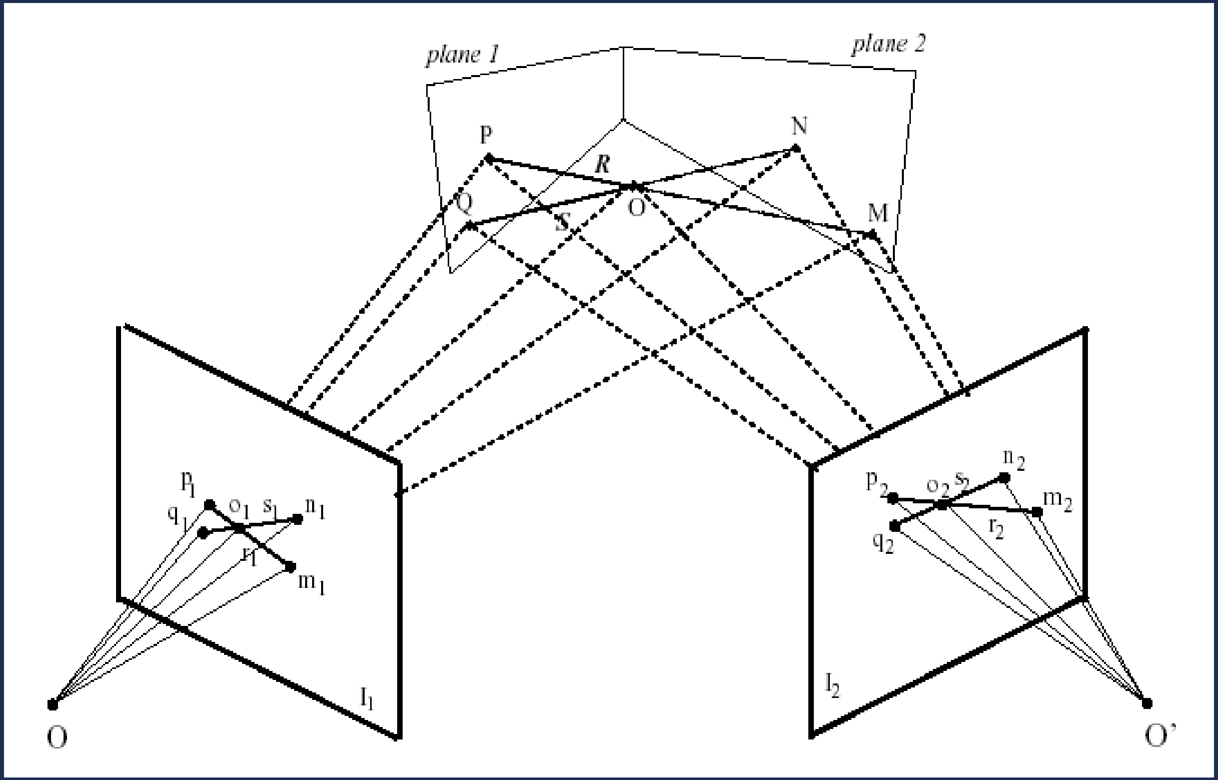
(D) Match features not belonging to the planes, between stereo pairs



Given two planes π_1 and π_2, a point O not belonging in either planes can be represented as the intersection of lines with end points on the planes.

Thus, if we know how planes and their points are transferred in distant views, we may transfer all 3D points as well.

(D) Match features not belonging to the planes, between stereo pairs

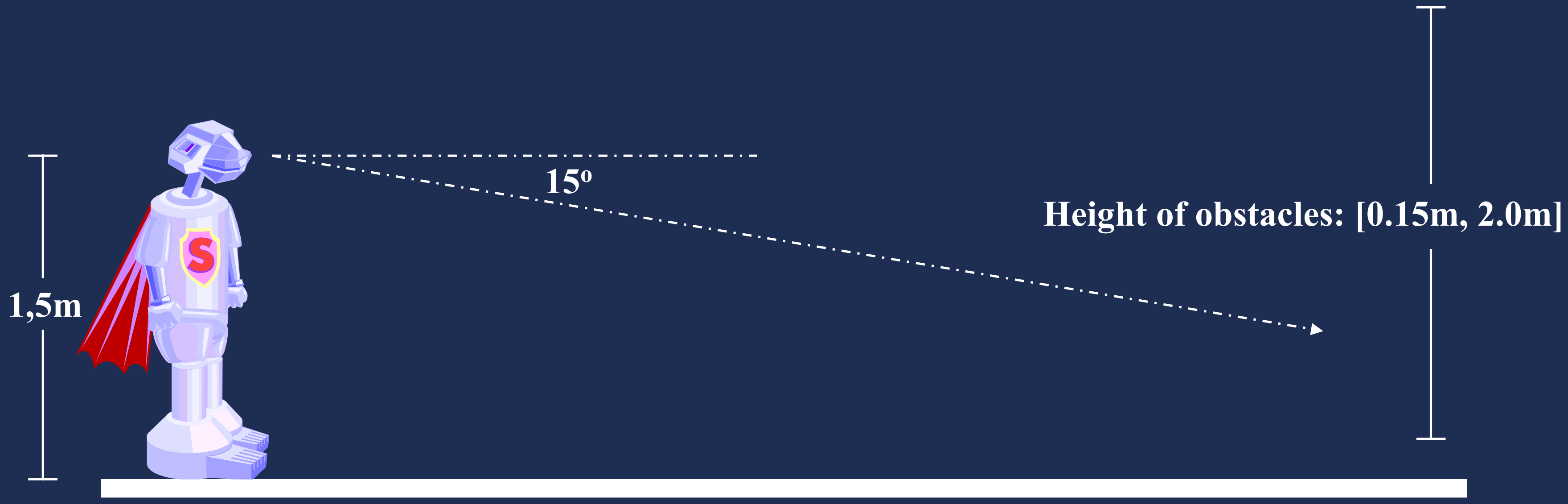


- For the case of lines, it suffices to transfer their end-points
- For robustness, more points on the line can also be transferred

(D) Match features not belonging to the planes, between stereo pairs

- MATCHING
  - After transferring features between stereo-pairs, matching is performed by considering proximity between transferred and extracted features
  - Feature proximity is quantified by the Euclidean distance between the normalized homogeneous vectors representing transferred and extracted features.

Experimental results: on the accuracy of plane segmentation

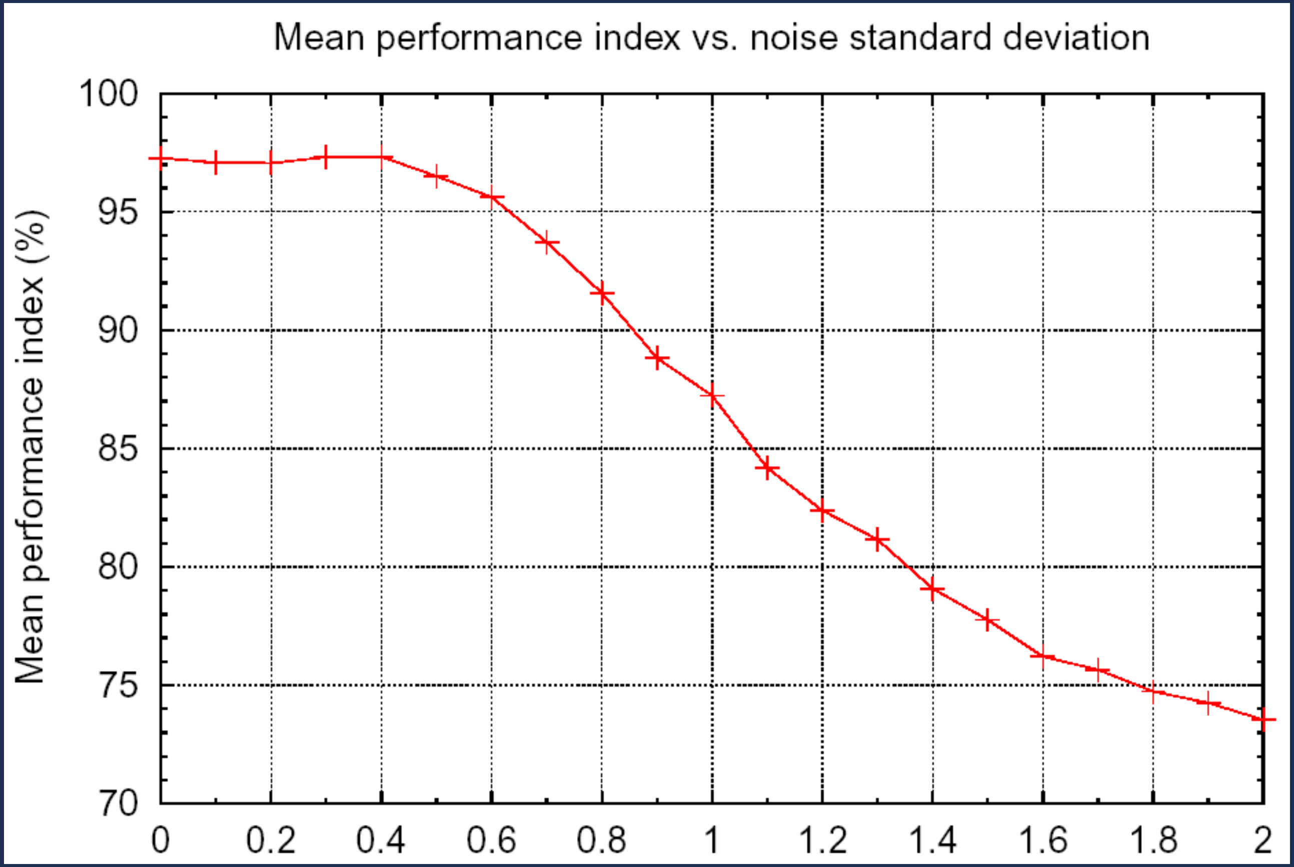


Experimental results: on the accuracy of plane segmentation

Simulated scenario related to an obstacle detection task

- Camera overlooking a planar floor at a height of 1.5 meters, with its optical axis at an angle of 15^{o} with respect to the horizon.
- Simulated retina: 750§750 pixels, focal length: 700 pixels.
- Camera is moving rigidly with 3D translational velocity of (0.100, 0.181, 0.676) meters/frame
- Synthetic scene consists of 300 3D points and 60 3D lines
- 100 3D points and 20 3D lines are assumed to lie on the simulated 3D floor.
- The heights of the 3D points not lying on the floor are assumed to be uniformly distributed in the range of [0.15, 2.0] meters.
- The standard deviation of the Gaussian noise added to retinal projections was varied between 0.0 and 2.0 pixels in steps of 0.2.

Experimental results: on the accuracy of plane segmentation

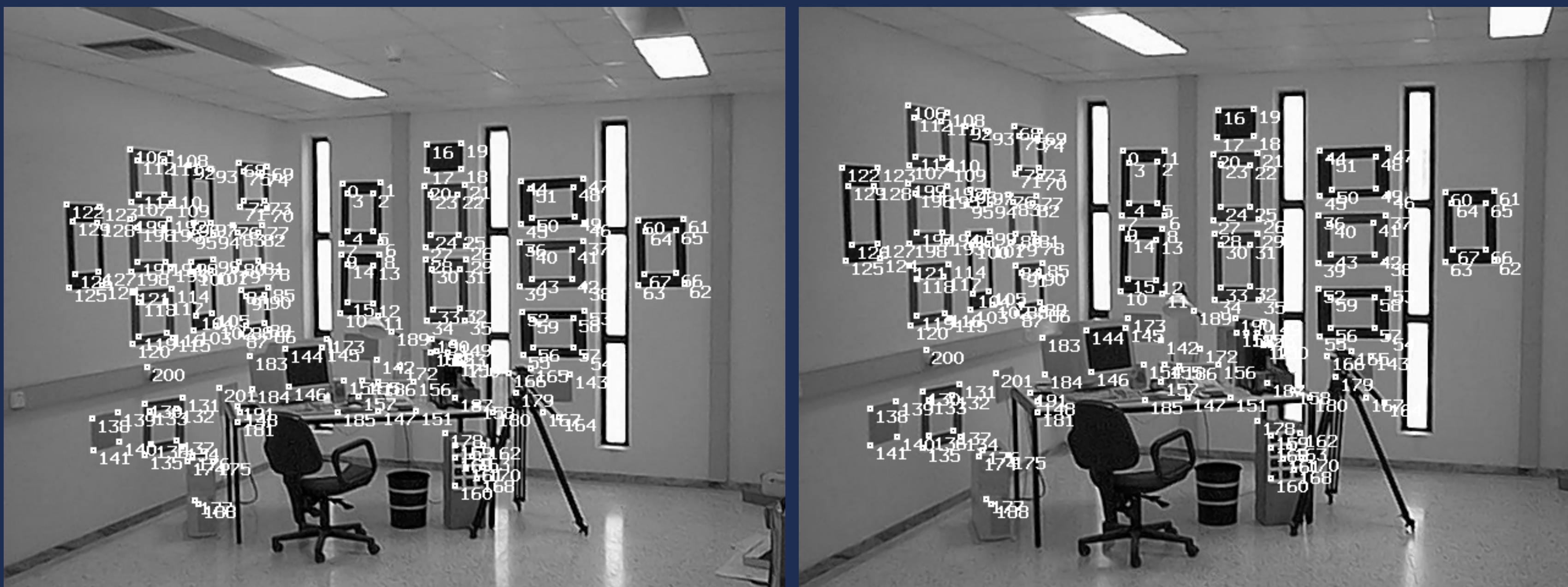


Experimental results (indoor scene)

Input images

Experimental results (indoor scene)

Intra stereo-pair matches

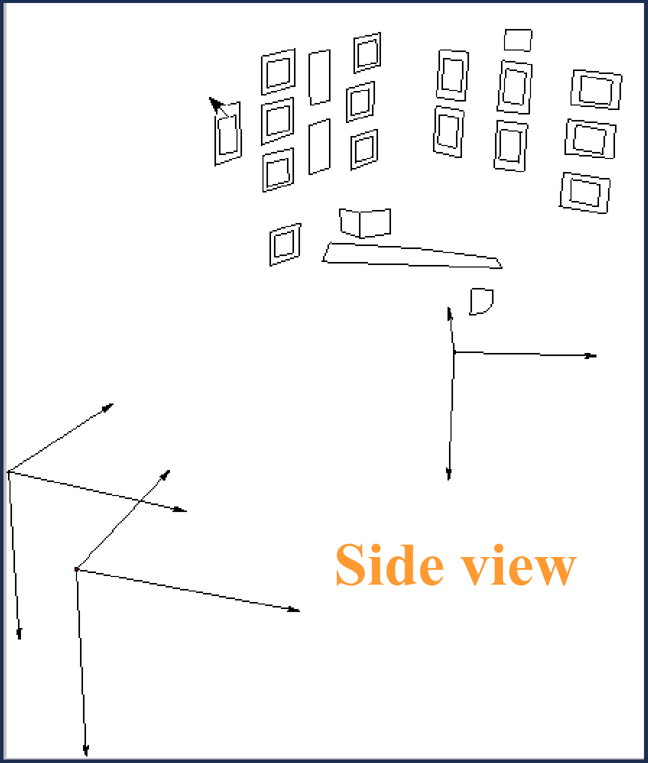


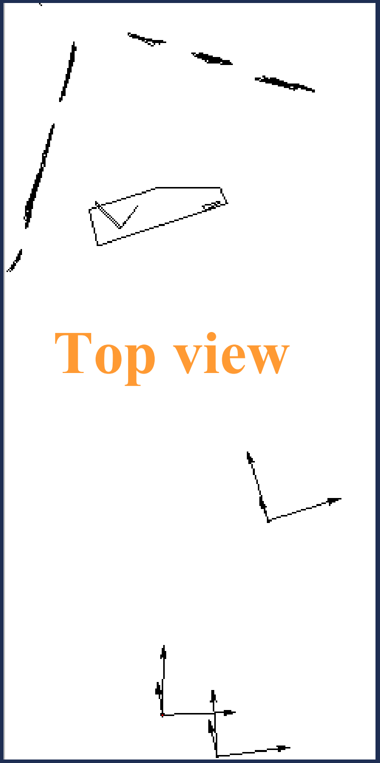
Experimental results (indoor scene)

Inter-stereo pair matches and planes defined

Experimental results (indoor scene)

3D Reconstruction





Experimental results (outdoor scene)

Input images

Experimental results (outdoor scene)

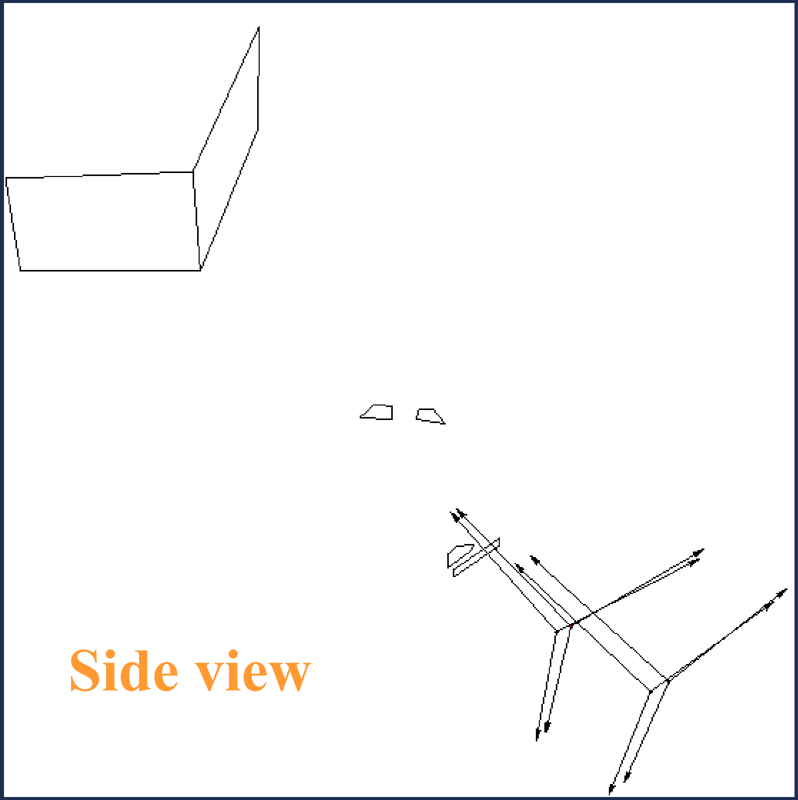
Intra stereo-pair matches

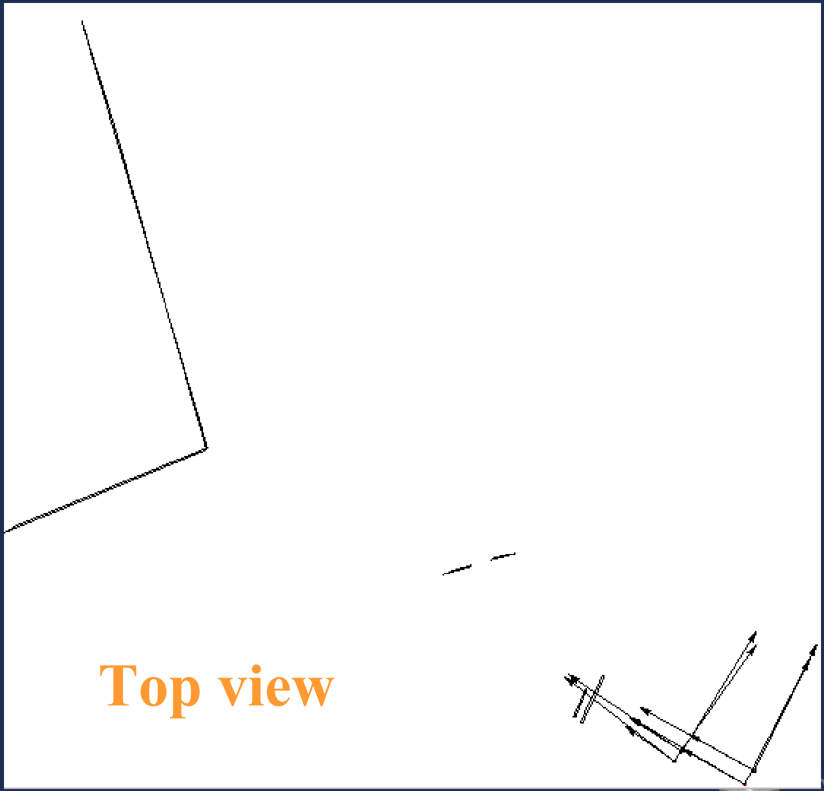
Experimental results (outdoor scene)

Inter-stereo pair matches and planes defined

Experimental results (outdoor scene)

3D Reconstruction





Sample quantitative results

Average reprojection error (in pixels)

	Proposed method	Faugeras, Robert (IJCV '96)
Indoor scene	1.59	2.95
Outdoor Scene	1.79	6.15

Discussion

- Benefits of the approach:
  - Fully automatic method
  - Capability to handle matching features in images that have been captured from significantly different viewpoints
  - Exploits constraints arising from the structure of the scene which are valid regardless of the viewpoints of images
  - No camera calibration required
  - Multiple constraints are defined for each feature thereby increasing robustness by over-determining the solution
  - Points and lines are handled in a unified manner

- Disadvantages of the approach:
  - Assumption regarding the existence of at least two planes in the viewed scene