Segment-Based Structure from an Imprecisely Located Moving Camera

J.M. Martínez
josemari@sun1.cps.unizar.es
University of Zaragoza
Zaragoza. Spain. Europe

Z. Zhang
zzhang@sophia.inria.fr
INRIA
Sophia-Antipolis. France. Europe

L. Montano
lmontano@mcps.unizar.es
University of Zaragoza
Zaragoza. Spain. Europe

Abstract

A probabilistic geometric model for 2D image line segments is first presented. We then propose a method, using 2D segments, to accumulate evidence along an image sequence of a polyhedral scene. The proposed method degrades gracefully with camera location noise. The matched 2D segments are fused with an extended Kalman filter to reconstruct the scene structure.

The correspondences are established using sequential data association. The matching function encodes the consistency between the 2D segments and the 3D segment computed from their fusion. The computation of a 3D segment from two 2D segments is overconstrained, and so using our model, the after-fusion consistency test can reject false matches even from two images, allowing the pruning of false matching hypotheses at early stages.

Two examples are provided. The first determines a scene structure when the camera location is known precisely; the structure is then compared with that obtained by trinocular stereo. The proposed method improves 3D segment orientation computation and reduces the number of spurious segments. A second experiment demonstrates the graceful degradation of the performance, especially in orientation, with respect to the camera location noise.

1 Introduction

This paper is devoted to the reconstructing a polyhedral static scene from images taken by a mobile camera whose location is known up to an uncertainty level. We compute the matching of the segments along a sequence, as well as an estimation of both the 3D scene segments location and their covariances. We use as input data the following information:

- The segments extracted from the images (not the correspondences), and their corresponding covariance matrices. These covariances depend on the feature extraction process, and on the intrinsic parameters of the camera calibration.
- The camera uncertain location estimate, and its location covariance matrix.

We propose an image segment model, which we will call *2D segment*. Classical image segments representations [4, 5] consider image segments as 2D entities in the image plane. The 2D segment we propose is a 3D entity composed of the supporting line projection plane, plus the midpoint projection ray. Similarly the consistency between a projection and a 3D segment is not tested on the image, but considering a statistical distance between two tridimensional entities: the *2D segment* and the 3D segment. A similar model to match and reconstruct polyhedral scenes was proposed in [8], although they proposed to use points and lines, but not an specific treatment for line segments. The segment covariance is tuned as proportional to image segment length, inspired in [14], although they use it for 3D segments, while we use it for 2D segments.

We represent in a unique feature the segment supporting line and its midpoint,which is constrained to rest on that line. The observation of a 3D segment imposes that the 3D segment supporting line is detected as the 2D segment supporting line, and that the 3D segment midpoint is detected as the 2D segment midpoint. The different observations of an scene segment are fused with an Extended and Iterated Kalman Filter (EIFK). The covariance along the 2D segment direction is set proportional to its length; so, the constraint that the 3D midpoint images as the 2D midpoint is considered, along the 2D segment direction, with a low weight.

The correspondences are computed using a tracking for several targets technique, the Split-Track filter, [2, 7, 13]. We exploit the fact that the scene is static and that the camera motion is fairly well known. The consistency between corresponding 2D segments and the 3D segment computed by their fusion is determined with a batch test [11] (instead of the classical recursive test). Due to non-linearities, this test performs better, especially at early stages. It should be noticed, that due to the image segment model, a 3D segment location from two images is an overconstrained problem, which allows to reject false matches with two images, controlling the Split-Track hypothe-

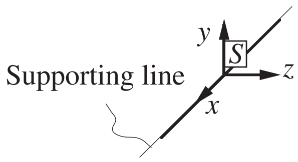

Figure 1: A 3D segment and its reference S.

ses explosion. Besides, using the score of the previous test, a unique track is selected among all the tracks which include the same image segment.

Two experiments are presented. First we compare the results obtained from our technique with the results from a trinocular stereo [1]. A second experiment shows the graceful degradation with the camera location uncertainty, situation in which a stereo based on epipolar constraint would not perform well [10]. .

Section 2 contents a brief introduction to the probabilistic model used for geometric information. Section 3 presents the 2D segment and Sect. 4 the 3D segment. Section 5 describes the measurement equation between a *2D segment* and a 3D segment. In Sect.6 the matching algorithm and the acceptance tests are described. Section 7 shows the experimental results and Sect. 8 the conclusions.

2 Geometric information

The SPmodel is a probabilistic model that associates a reference G to every geometric element $\mathcal{G}$ [12]. Its location is given by the transformation t_{WG} relative to a base reference W. To represent this transformation, a *location vector* $\mathbf{x}_{WG}$, composed of three Cartesian coordinates and three Roll-Pitch-Yaw angles is used:

$$\mathbf{x}_{WG} = (x, y, z, \psi, \theta, \phi)^T$$

$$t_{WG} = \mathrm{Trans}(x, y, z) \cdot \mathrm{Rot}(z, \phi) \cdot \mathrm{Rot}(y, \theta) \cdot \mathrm{Rot}(x, \psi) \quad (1)$$

The estimation of the location of an element is denoted by $\hat{\mathbf{x}}_{WG}$, and the estimation error is represented *locally by a differential location vector* $\mathbf{d}_G$ relative to the reference attached to the element. Thus, the true location of the element is:

$$\mathbf{x}_{WG} = \hat{\mathbf{x}}_{WG} \oplus \mathbf{d}_G$$

where $\oplus$ represents the composition of location vectors (the inversion is represented with $\ominus$). The SPmodel also exploits the concept of *symmetries* of a geometric element, defined as the set $\mathcal{S}_G$ of transformations that preserve the element. For example (see Fig. 1), the symmetries of a 3D line segment is the set of continuous rotations around the its supporting line.

To account for the continuous motion symmetries, we assign in $\mathbf{d}_G$ a null value to the degrees of freedom corresponding to them, because they do not represent an effective location error. The *perturbation vector*, $\mathbf{p}_G$, is formed by the non null elements of $\mathbf{d}_G$:

$$\mathbf{d}_G = B_G^T \mathbf{p}_G \quad ; \quad \mathbf{p}_G = B_G \mathbf{d}_G$$

Where B_G is a selection matrix. e.g. for a 3D line segment, B_G, selects rows 1,2,3,5 and 6.

Based on these ideas, the SPmodel represents the information about the location of a geometric element $\mathcal{G}$ by a self-binding matrix and the triplet:

$$\mathbf{L}_{WG} = (\hat{\mathbf{x}}_{WG}, \hat{\mathbf{p}}_G, C_G)$$

$$\mathbf{x}_{WG} = \hat{\mathbf{x}}_{WG} \oplus B_G^T \mathbf{p}_G; \hat{\mathbf{p}}_G = E[\mathbf{p}_G]; C_G = Cov(\mathbf{p}_G)$$

Transformation $\hat{\mathbf{x}}_{WG}$ is an estimation taken as base for perturbations, $\hat{\mathbf{p}}_G$ is the estimated value of the perturbation vector, and C_G its covariance matrix. When $\hat{\mathbf{p}}_G = 0$, we say that the estimation is *centered*.

3 The 2D segment

To represent the observation of a 3D segment with a camera, we define a geometric element called *2D segment*. Despite its name it is a 3D entity. The reference to locate a 2D segment, D, is defined as (Fig. 2):

- origin attached to the camera optical center.
- z axis normal to the plane defined by the camera optical center, and also to the projection of the 3D segment on the image plane.
- $-y$ axis pointing to the center of the projection of the 3D segment in the image plane.

Thus, a 2D segment includes the following information:

- A projection plane where the 3D segment supporting line rests on.
- A projection ray, where the center of the 3D segment approximately rests on. It is defined by the origin of D and its y axis. The center of a segment is not invariant under projection, but in any case, the covariances will be defined in such a way that this constraint tolerates deviations along the segment comparable to $\approx 30\%$ of the segment length.
- The point of view, which is the origin of D.

As was said in Sect. 2, to define a geometric entity location, the elements of the triplet $\mathbf{L}_{WD} = (\hat{\mathbf{x}}_{WD}, \hat{\mathbf{p}}_D, C_D)$ should be computed. We suppose the location estimate is unbiased, so $\hat{\mathbf{p}}_D = 0$. The 2D segment has one symmetry, translation along y because it preserves the projecting elements. Because of that $\hat{\mathbf{p}}_D$ has dimension 5. The rest of this section is devoted to compute $\hat{\mathbf{x}}_{WD}$ and C_D.

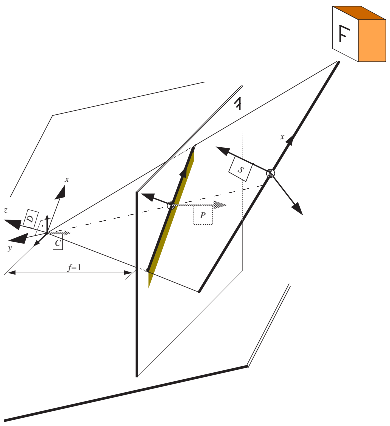

Figure 2: Relationship between the 3D segment S; its projection in the normalized image plane, P; the camera optical center, C, and the 2D segment, D.

3.1 Computing $\hat{\mathbf{x}}_{WD}$

The 2D segment location estimate is computed as:

$$\begin{aligned}\hat{\mathbf{x}}_{WD} &= \hat{\mathbf{x}}_{WC} \oplus \hat{\mathbf{x}}_{CD} \\ \hat{\mathbf{x}}_{CD} &= (0,0,0,\psi_{CD},\theta_{CD},\phi_{CD})^T\end{aligned}$$

where $\hat{\mathbf{x}}_{WC}$ is the camera location estimate in the world reference, read from the robot sensors. $\hat{\mathbf{x}}_{CD}$ is computed using geometric relations from x_{CP}, y_{CP} and ϕ_P, the location of the segment in the image (Fig. 3).

3.2 Computing C_D

The 2D segment location noise, has two independent sources of noise: the camera location noise and the image noise. So, we approximate the 2D segment location noise covariance, C_D (5×5), as the addition of two terms:

$$C_D = J_{DC} C_C J_{DC}^T + J_{DP} N C'_P N^T J_{DP}^T \qquad (2)$$

where C_C is the covariance matrix for the camera location, C'_P is the covariance for the line segment location in the unnormalized image plane. J_{DC} is computed by selecting rows and columns from the jacobian [11] of the relative location vector $\hat{\mathbf{x}}_{DC}$. J_{DP} is computed from $\hat{\mathbf{x}}_{DP}$ values. N converts the covariance from the unnormalized to the normalized camera. C'_P is computed as:

$$C'_P = \operatorname{diag}\left(\sigma_{xP}^2, \sigma_{yP}^2, \sigma_{\phi P}^2\right)$$

- σ_{xP} is proportional to the image segment length, n (in pixels). (see Fig. 3):
  $$\sigma_{xP} = \kappa n \qquad (3)$$
  It mimics, for image segments, the model proposed for scene segments in [14]. The experimental value for κ is 0.14.
- σ_{yP}^2 and $\sigma_{\phi P}^2$ are computed from the covariances of the image segment extreme points, σ_p^2. Some correlation between the extreme points is considered, due to systematic calibration error.

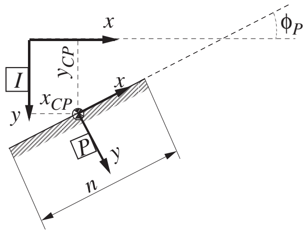

Figure 3: P is the image of a segment, in the normalized image plane, I. The x direction of P is coherent with the segment gray level gradient direction. n is the segment length in pixels.

4 The 3D segment

The method uses a statistical data association technique for correspondences computing, so, an EIFK is used for every scene 3D segment to accumulate evidence from its observations. The state of each EIFK is the 3D segment location, $\mathbf{x}_{WS_{k|k}}$, k being the number of fused observations. As the scene is considered static, the prediction state equation is trivial:

$$\mathbf{x}_{WS_{k+1|k}} = \mathbf{x}_{WS_{k|k}} \qquad (4)$$

and we drop the $|k$, denoting both of them with $\mathbf{x}_{WS_k}$

The 3D segment location is defined by a reference, S, whose x axis is aligned with the segment supporting line; it is attached to a fixed point of the segment (an approximation to the midpoint). It should be noticed that the length is not explicitly considered in the state vector. The 3D segment has only a symmetry because only the d.o.f. rotation around x is not a true location error (see Fig 1). The location covariance matrix, C_S, is 5×5. Thus, a 3D segment is represented by the triplet: $\mathbf{L}_{WS_k} = (\hat{\mathbf{x}}_{WS_k}, \hat{\mathbf{p}}_{S_k}, C_{S_k})$

The fused segment length is determined from the projections of all the segment images on the fused supporting line. For noisy images, the length is computed from the intersection of the projections, and for noise-free images from the union.

5 The measurement equation

Both the EIFK fusion of the 2D segments to determine a 3D segment location, and the matching along the sequence are based on the same measurement equation. The SPmodel measurement equation is based on the relative location, $\mathbf{x}_{DS}$, between the measurement reference, D, and the state reference, S, (see Fig. 2). Theoretically, if a 2D segment (D) is the projection of a 3D segment (S), to pass from D to S, the x translation, the z translation, and the y rotation, should all be 0. Expressing these constraints as an implicit measurement equation:

$$\begin{aligned}\mathbf{r}_{DS} = \mathrm{f}\left(\mathbf{x}_{DS}\left(\hat{\mathbf{x}}_{DS}, \mathbf{p}_S, \mathbf{p}_D\right)\right) = \\ (x_{DS}, z_{DS}, \theta_{DS})^T = 0 \qquad (5)\end{aligned}$$

$$\mathbf{x}_{DS} = \ominus\mathbf{x}_{WD} \oplus \mathbf{x}_{WS} = \\ = \left(x_{DS}, y_{DS}, z_{DS}, \psi_{DS}, \theta_{DS}, \phi_{DS}\right)^T$$

$\mathbf{p}_S$ is an incremental correction for 3D segment location. $\mathbf{p}_D$ is the unbiased measurement noise, it is the perturbation vector of the measurement, i.e., the 2D segment. Linearizing the equation (5) for EIFK, we obtain:

$$\mathbf{r}_{DS} \approx H\mathbf{p}_S + G\mathbf{p}_D \tag{6}$$

$$H = \left.\frac{\partial \mathrm{f}}{\partial \mathbf{p}_S}\right|_{(\hat{\mathbf{x}}_{DS}, \mathbf{p}_S = \mathbf{p}_D = 0)} \quad G = \left.\frac{\partial \mathrm{f}}{\partial \mathbf{p}_D}\right|_{(\hat{\mathbf{x}}_{DS}, \mathbf{p}_S = \mathbf{p}_D = 0)}$$

The proposed measurement equation (5) considers:

- The collinearity in the image of a 3D segment and its projection $\left(\ z_{DS} = 0 \quad \theta_{DS} = 0\ \right)$
- The overlapping in the image of the 3D segment and its projection ($x_{DS} = 0$).

Equation (5) enforces that the center of a segment is invariant under projection, which is not strictly true. This is alleviated because image segment covariance along its direction is proportional to its length, so errors compatible with σ^2_{xP} (3) are allowed.

6 Matching

We propose a *Split-Track* filter [2] to compute the correspondences. For our system we exploit both the scene staticity and the camera motion knowledge. This section is devoted to computation of matches for the "next image".

Corresponding 2D segments are gathered in a track. All 2D segments in a track, at time k, are fused with a EIFK to compute the state: the 3D segment location $\mathbf{L}_{WS_k}$. As the scene is static, it is also the state prediction for next step (see Sect 4).

Prediction test Using the state prediction and the camera location estimate, all the 2D segments in image $k+1$ which falls in the $\mathbf{L}_{WS_k}$ validation region are considered. If there is only one validated 2D segment, it is added to the track; if more than one, a new track is created for each additional validated 2D segment. If no 2D segment is validated, the track is eliminated.

To validate the 2D segment we use a test based on a statistical distance derived from the linearized measurement equation (6):

$$\begin{aligned} \mathbf{r}^T_{DS} C^{-1}_{\mathbf{r}_{DS}} \mathbf{r}_{DS} &\leq \chi^2_{3,\alpha} \\ C_{\mathbf{r}_{DS}} &= HC_S H^T + GC_D G^T \end{aligned} \tag{7}$$

where $\chi^2_{3,\alpha}$ is a chi-square for 3 d.o.f with α confidence level.

After-fusion test The prediction test normally accepts several matches, which implies an explosion in the hypothesized tracks tree. To prune out this tree, after including a new 2D segment in a track, it is tested if all the gathered 2D segments in the track are consistent with 3D segment resulting from their fusion.

$$\sum_{k=1}^{n} \mathbf{r}^T_{D_k S_n} \left(G_k C_{D_k} G_k^T\right)^{-1} \mathbf{r}_{D_k S_n} \leq \chi^2_{3n-5,\alpha} \tag{8}$$

Unlike the usual recursive test [2], the previous test is a batch test proposed in [11]. All the distances are computed with respect to S_n the 3D segment estimate with n images. Theoretically both test are equivalent for a linear system; however our problem is non-linear, then the batch test overperforms because all the linearizations are made around the last estimate.

Unicity test Finally, the score of test (8) is used to select a unique track for each multiple matched image segment. This test is applied after processing several images. As can be seen in the experiments, this test is very useful to reject false matchings.

Track initiation Besides knowing how to process the next image it is necessary an initial guess for the first image. For every 2D segment detected in the first image, D_1, a 3D segment, S_1, is considered as initial guess. S_1 is contained in the 2D segment, and located in the middle of the working space defined by:

$$\begin{aligned} \hat{\mathbf{x}}_{D_1 S_1} &= (0, y, 0, 0, 0, \phi) \\ y &\in [y_{\min}, y_{\max}] \quad \phi \in [-\pi/2, \pi/2] \end{aligned}$$

The covariance for the d.o.f constraint by the 2D segment are computed from the 2D segment covariance, however the the covariance for depth, y axis, and rotation inside the 2D segment, ϕ are guessed from:

$$\sigma_y = \frac{y_{\max} - y_{\min}}{2 \times 1.96} \quad \sigma_\phi = \frac{\pi}{2 \times 1.96}$$

considering the working space as the 95% acceptance region for a χ^2 test.

7 Experimental Results

A first experiment computes a scene structure from the camera location known with stereo calibration precision; the performance of the method is compared with that of a trinocular stereo [1]. A second experiment shows how the method performance degrades gracefully with camera location uncertainty.

The experiments process the images of a trinocular rig. The segments are extracted with the Burns's method [3]. The covariance matrix of camera location uncertainty in the experiments is a diagonal matrix:

$$C_C = \mathrm{diag}\left(\sigma^2_P, \sigma^2_P, \sigma^2_P, \sigma^2_O, \sigma^2_O, \sigma^2_O\right) \tag{9}$$

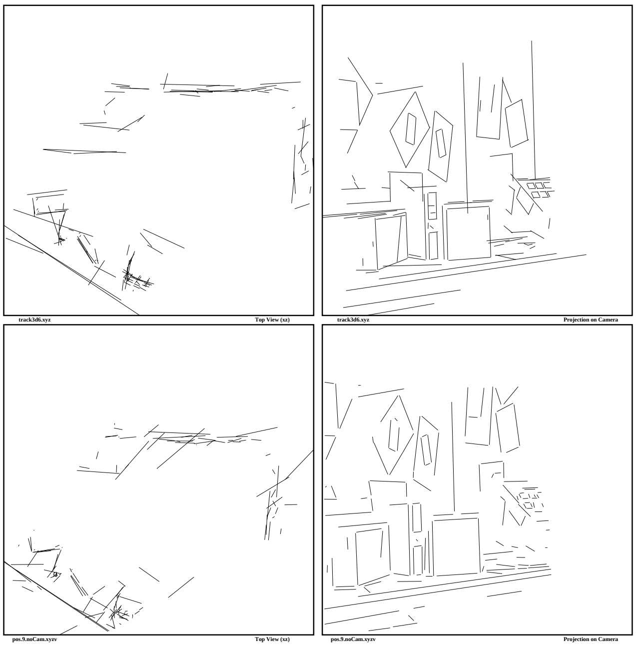

Figure 4: Scene reconstruction in top and camera view. Bottom figures for the trinocular stereo, top figures for the proposed method.

where the covariance in position, σ_P^2, and the covariance in orientation, σ_O^2, will be defined for each experiment.

7.1 Comparison with a Trinocular Stereo

The camera location used is that of the stereo calibration. In order to take profit from the camera location precision, its location covariances (9) are set low:

$$\sigma_P = 0.005\text{mm.} \qquad \sigma_O = 0.0035\text{dg.}$$

The factor κ (3) for covariance along the segment direction is set experimentally to 0.14. The α for both consistency tests, (7) and (8) is set to 0.05. The length of the 3D segments are computed as the union of the projected lengths. Only segments longer than 10 pixel are considered. Figure 4 shows the reconstructed scenes by both trinocular stereo and the proposed method.

The proposed method produces a better scene reconstruction; it is nearly spurious-free. It performs well in cluttered areas: the monitor, especially its lattice; in areas where there are several nearly parallel segments, e.g. middle left in the camera view. Theoretically, all the segments belonging a wall should be seen on a line in top view; deviations from this line shows the imperfections in reconstruction. The proposed method does nearly not produce any segment non-parallel to these ideal lines.

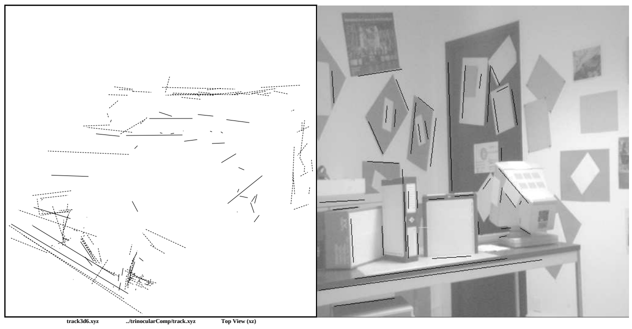

Figure 5: Top and camera view the for a typical sample when camera orientation noise is 20% the β value. Dashed segments are the noise-free reconstruction.

7.2 Graceful degradation with noise

It is analyzed with a Monte Carlo simulation. The camera location is perturbed according to:

$$\mathbf{x}_{WC} = \hat{\mathbf{x}}_{WC} \oplus \mathbf{d}_C$$

where $\hat{\mathbf{x}}_{WC}$ is the camera location estimate taken from the trinocular rig calibration; $\mathbf{d}_C$ is a normally distributed perturbation vector, whose covariance matrix is (9). The number of samples for the simulation is 50. In these noisy experiments, only the segments longer than 20 pixels are considered. The length of the 3D segment is computed as the intersection of the projected lengths. The factor κ (3) is set to 0.14.

Table 1, shows camera location noise for two simulations. The noise is defined by its standard deviation values, σ_P and σ_O and by the relative value of the orientation noise $\frac{2\sigma_O}{\beta}$. The reference value β is the typical angle of the projection rays of a point detected by two cameras; it indicates the effect of an angle error when computing the depth of a point from two views. From the typical scene depth, 3175 mm., and the typical camera baseline, 300 mm., $\beta = 5.4$dg..

As a result, the table shows the mean number of hypotheses at each step: prediction with one image (pt1), after-fusion with 2 images (af2), prediction with 2 images (pt2), after-fusion with 3 images (af3), and unicity test (ut). The number of hypotheses is reduced with the after-fusion test, what allows to control the complexity of the algorithm. It also shows the ability of the unicity test to reject false matches. The last column shows the correct matches among the detected ones.

Figure 5 shows a typical reconstruction when orientation noise is 20% the β value. Top view shows how errors in position are bigger than in orientation.

8 Conclusions

The proposed image segment model combines the properties of a point, its midpoint, with those of its supporting line. All the geometric information con-

$2\sigma_P$	$2\sigma_O$	$\frac{2\sigma_O}{\beta}$		mean number of hypotheses					ok
mm	dg	%	α	pt1	af2	pt2	af3	ut	%
1.0	0.54	10	0.05	903	167	149	115	69	87
2.0	1.09	20	0.5	1096	59	52	35	25	61

Table 1: Monte-Carlo simulation results. Hypotheses number after prediction test, "pt"; after after-fusion, "af"; after unicity, "ut". The digit stands for the number of images considered. "ok" stands for the true matches among the detected ones.

cerning the segment is managed together with the SP-model, which weights together the "point properties" and the "line properties" of a segment depending on the relative location of the image segments, on the image noise covariance and on the camera location covariance. The improvement in the segments orientation is due to the effect of the line properties.

In general, the computed orientation of the scene segments improves that of the classical stereo based on the epipolar constraint. The number of spurious is also reduced. When the camera location noise increases, the epipolar constraint cannot be any longer applied. However, the proposed system performance degrades gracefully and it mainly affects the position of the segments, rather than their orientation. This is because the segment representation can profit from the fact that the orientation computing is not so sensitive to camera orientation noise as the position computing is.

The after-fusion consistency test cuts down the hypotheses tree. It can reject some tracks using only two images, because, unlike the fusion of two image lines, the fusion of two 2D segments is an *overconstrained problem*. Thus, the proposed test can prune out the hypotheses tree at early stages. The unicity test based only on the score of the after fusion test, has a good spurious rejection. It should be noticed, that we only consider the consistency of each track to compute correspondences, even to select a unique track when a image segment is multiple matched. The trinocular stereo [1] uses a global neighborhood test to select the unique match.

In order to improve reconstruction from a unprecisely located camera our future work goal is twofold. On one hand, to use the proposed segment representation to compute both the structure and the camera motion. On the other hand to develop active sensing techniques for the image segment model in order to determine a point of view for the next image, that could reduce the structure uncertainty, while taking profit from the available information for correspondences [9].

It should be noted that the proposed system only manages *geometrical information*; the performance of the system can be increased using additional non-geometrical parameters, such as the image segment bright parameters [6].

Acknowledgments

This work was developed while José María Martínez was visiting the INRIA Lab. at Sophia-Antipolis. The authors wish to express their gratitude to the members of the ROBOVIS Project at INRIA Lab. for their assistance. This work was partially supported by CICYT-TAP-94-0390 and by the Aragonese CAI-CONAI (Programa Europa IT-10-94).

References

[1] N. Ayache. *Artificial Vision for Mobile Robots: Stereo Vision and Multisensory Perception.* MIT Press, Cambridge, MA, 1991.

[2] Y. Bar-Shalom and T.E. Fortmann. *Tracking and Data Association*, volume 179 of *Mathematics in Science and Engineering.* Academic Press, INC., San Diego, 1988.

[3] J.B. Burns, A.R. Hanson, and E.M. Riseman. Extracting straight lines. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(4):425–455, 1986.

[4] J.L. Crowley and P. Stelmaszyk. Measurement and integration of 3-D structures by tracking edge lines. In *First European Conference on Computer Vision*, pages 269–280, Antibes, France, 1990.

[5] R. Deriche and O. Faugeras. Tracking line segments. In *First European Conference on Computer Vision*, pages 259–268, Antibes, France, 1990.

[6] J.J. Guerrero and J.M Martínez. Determination of corresponding segments by tracking both geometrical and brightness information. In *7th. Int. Conf. on Advanced Robotics*, pages 703–709, Spain, September 1995.

[7] Cox J.C. A review of statistical data association techniques for motion correspondence. *Int. Journal of Computer Vision*, 10(1):53–66, 1993.

[8] Jezouin J.L. and Ayache N. 3d structure from a monocular sequence of images. In *3th. Int. Conf. on Computer Vision*, pages 441–445, Osaka, 1990.

[9] J. M. Martínez and L. Montano. A camera motion strategy to loacalize uncertain 3d lines. In *1993 IEEE Inter. Conf. on Systems, Man and Cybernetics*, pages 517–522, Le Touquet-France, October 1993.

[10] Loung Q.T. and Faugeras O.D. Active head movements to solve stereo correspondence. In Neumann B., editor, *ECAI 92. 10th European Conf. on Artificial Intelligence.* John Wiley & Sons, Ltd., 1992.

[11] J.D. Tardós. *Integración Multisensorial para Reconocimiento y Localización de Objetos en Robótica.* PhD thesis, Dpto. Inge. Eléctrica e Informática, University of Zaragoza, Spain, Febrero 1991.

[12] J.D. Tardós. Representing partial and uncertain sensorial information using the theory of symmetries. In *IEEE Int. Conf. on Robotics and Automation*, pages 1799–1804, Nice, France, May 1992.

[13] Zhang Z. and Faugeras O. Three-dimensional motion computation and object segmentation in a long secuence of stereo frames. *Int. Journal of Computer Vision*, 7(3):211–241, 1992.

[14] Z. Zhang and O. Faugeras. A 3d world model builder with a mobile robot. *Int. Journal of Robotics Research*, 11(4):269–284, August 1992.