Tracking Line Segment without Knowledge of Camera Motion

S. Ait Kaci Azzou
LRSD laboratory,
computer science department,
University Ferhat Abbas Setif,
Algeria,

M.C Baba Hnini
Computer Science Department
University of Biskra, Algeria

ABSTRACT

In this paper, a new method for line segments matching for indoor reconstruction will be presented. The problem of line segment tracker is addressed, which does not require any knowledge about the motion of the camera nor the structure of the observed scene. The slopes of lines segment as feature to track are used in order to deal with the instability of the endpoint extraction. The proposed method relies solely on the geometry layout of the slope of the segment and not on photometric or color profiles and does not need any knowledge about the parameters of the camera or its rotation. The obtained results showed the feasibility of the method.

General Terms

Computer vision

Keywords

Tracking, line segment, Matching.

1. INTRODUCTION

The development of algorithms for tracking objects in image sequences is a major problem for many applications related to computer vision, robotics, etc.. A reliable extraction and tracking of spatio-temporal visual information is indeed a key to the success or failure of such applications.

Two main methods for tracking exist: those based on the contour and those based on the texture of the object. The first approach is essential to track the primitives in the image space or 3D geometric primitives like (points, lines, circles ...). The contour of the object, the projection of the contours of a 3D object, etc. The latter uses a correlation criterion related to the information given by the grayscale pattern of the object or other information presented in this pattern (color...).

Tracking method based edge depends on the strong spatial gradients defining the edge of the object or some geometric primitives present in a pattern (points, lines, distances ...). With regard to tracking in the space of the image (2D tracking), this approach describes the object to follow using geometric primitives as special points [1,2], angles, contours [3,4], line segments [5, 6] or ellipses [7], etc..

However, just look around us to deduce that the line segment is a very expressive primitive in our environment. Line segments in an image represent significant information about the structure of the perceived scene: it is basic primitives on which many applications can be built (recognition of structured objects [8], or 3D reconstruction structured scenes in robotics [9, 10]). Segments were indeed several interesting properties: they are naturally perfectly straight lines in the scene; they can be used to estimate various geometric transformations, and are invariant under change of perspective. But this property is only exploitable if the segments extraction algorithms are robust and stable: classical methods are fragile, they are very sensitive to image noise, lighting conditions and change views.

A common way to obtain line segments is to link gradient maximum points [11] in order to assume a contour and apply one of the methods summarized in [12,13,14].

The tracking methods of the segment, developed, are based either on the Kalman filtering [15], where the assumption of "short baselines" is issued, either by using the epipolar line bundles to constrain to constrain good matches [16]. This method is fast and accurate but requires knowledge of the epipolar geometry thus requires camera calibration which makes it unsuitable for camera calibration.

In [17], a segment matching method with an application to wide baseline stereo is presented. Here, feature points (SIFT, HOG, ...) are extracted and used as anchor points. The actual segments pairing are performed by first, grouping putative matches using their color profiles. Then, anchor-point/segment sideness consistency is exploited to sort out the matches. Obviously, if no or few point features were detected and paired, the segment matching will rely solely on color profiles which are known to be unstable.

The proposed approach in this paper, uses line segments rather than points as landmarks, since there are some advantages in using line segments. Images of artificial environments with little texture contain many line segments, whereas few point features can be localized in such a scene. Moreover, line segment detection is more reliable than point detection. Line segment matching is also more robust than point matching with respect to partial occlusion and viewpoint changes.

In this paper, a new method is presented for segments tracking. The proposed method is suitable in an indoor environment.

The main problem when dealing with line segments is the instability of the endpoint extraction along the line. To avoid this instability, the proposed tracker uses as primitive the

slope of the line segment because the slopes are more robust to noise than the endpoint of line segments.

One of the highlights, of the method is that using uniquely the mathematical relation between the slopes, on image sequences. The method does not require any knowledge of the calibration of the camera or its motion..

2. POSITION OF THE PROBLEM

In this work, horizontal rotational camera movement is assumed. The proposed geometrical model is illustrated by figure 1. After each rotation of the camera, a new image is acquired and noted IM^i, where **i** refers to the position number of the camera. In each image some of the line segments or interest points, defining segment, are extracted.

$IM^i = \{S_j^i, j=1..n_i\}$ where n_i is the number of segment lines located in IM^i .

B. Geometrical model for the motion of camera rotation

The geometrical model of image formation is defined by (see figure 1).

- Oi is the impact point of the image IMi defined as the intersection of the optical axis with the image plane.

- L is the center of projection of the camera, OiL represent the focal length f.

- (Oi Ui Vi) is the internal reference associated to IMi.

- The theoretical external reference (OXYZ) is defined so as : OX is parallel to OiUi , OY parallel to the optical axis,

The axis of rotation OZ is parallel to OiVi.

- The camera is fixed so that the rotational axis passes through the optical axis. Due to the uncertainty of the mechanics, the OZ axis is close to the impact point whose position is at the distance d between O and L.

- ex, ez define the dimensions of the pixel.

None of the defined parameters is known. The aim, is to develop a mathematical relation that allows the computation of the matching parameters without knowledge of the angle of camera rotation. This relation must be independent of the camera model and will use only the coordinates of image points in the different images. The figure 2, shows that when the camera is rotated around the origin O, the projection center L and image plan IMi are also rotated with the same angle.
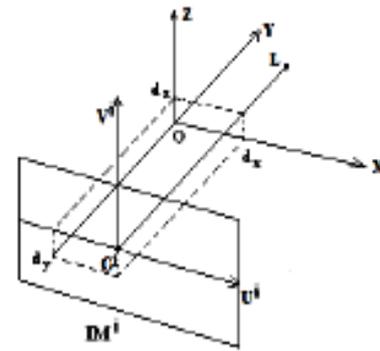
Fig. 1. Geometrical model for image formation
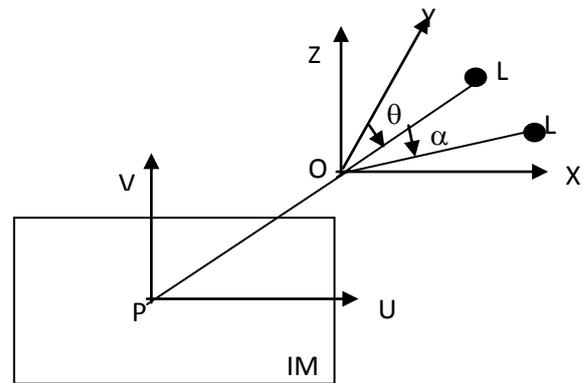
Fig. 2. The geometrical model of camera rotating

3. TRACKING METHOD

3.1 Equation of line segment in the initial image

In the general case, where the camera is rotated by an angle α, any point $M_i(x_i, y_i, z_i)$ of the 3D space is projected into m_i^1, where its projective coordinates on the image plane IM^1 are [18]:

$$u_1 = (ex).f. \frac{-\cos(\alpha).x + \sin(\alpha).y}{\sin(\alpha).x + \cos(\alpha).y + d - f} \quad (1)$$

$$v_1 = (ez).f. \frac{-z}{\sin(\alpha).x + \cos(\alpha).y + d - f} \quad (2)$$

As the reference (OXYZ) refers to the initial position of the camera, the angle α may be considered as equal to zero. Equations 1 and 2 will serve to determine the new coordinates of points after two rotations of the camera.

Let

S_i be a line segment in the 3D scene

S_i^1 be the image of S_i , on the image IM^1

v= $a_{i,l}$.u + $b_{i,1}$, be the equation of S_i^1 relatively to (O^{1}U^{1}V^1)

The coordinates of m_i^1 image on IM^1: of any point M$_i$(x$_i$, y$_i$, z$_i$) of the segment S_i are :

$$u\left(m_i^1\right) = ex.f \cdot \frac{X_{i,1}}{Y_{i,1}+d-f} \qquad (3)$$

$$v\left(m_i^1\right) = ez.f \cdot \frac{-Z_{i,1}}{Y_{i,1}+d-f} \qquad \textbf{(4)}$$

Where

$$X_{i,1} = -x'_i.\cos(\alpha) + y'_i.\sin(\alpha)$$

$$Y_{i,1} = x'_i.\sin(\alpha) + y'_i.\cos(\alpha)$$

$$Z_{i,1} = z'_i.$$

As $v\left(m_i^1\right) = a_{i,1}.u\left(m_i^1\right) + b_{i,1}$, $Z_{i,1}$ can be written :

$$z_{i,1} = a_{i,1} \cdot \frac{ex}{ez} \cdot X_{i,1} - \frac{b_{i,1}}{ez.f}\left(Y_{i,1}+d-f\right) \quad (5)$$

3.2 Images of line segments in new images after camera rotation

After a second rotation of the camera with an angle β, the segment S_i will be projected on IM^2 as S_i^2 .

Let v= $a_{i,2}$.u + $b_{i,2}$ be the equation of S_i^2 .

Following the same steps described in 3.1 :

$$z_{i,2} = a_{i,2} \cdot \frac{ex}{ez} \cdot X_{i,2} - \frac{b_{i,2}}{ez.f}\left(Y_{i,2}+d-f\right) \qquad (6)$$

Where

$$X_{i,2} = -x'_i.\cos(\alpha+\beta) + y'_i.\sin(\alpha+\beta)$$

$$Y_{i,2} = x'_i.\sin(\alpha+\beta) + y'_i.\cos(\alpha+\beta)$$

$$Z_{i,2} = Z_{i,1}$$

After little transformation X$_{i,1}$ and Y$_{i,1}$, become

$$X_{i,1} = X_{i,2}.\cos(\beta) - Y_{i,2}.\sin(\beta)$$

$$Y_{i,1} = X_{i,2}.\sin(\beta) + Y_{i,2}.\cos(\beta)$$

The substitution of X$_{i,1}$, and Y$_{i,1}$ in (5) by their expressions gives :

$$z_{i,1}(ez.f) = X_{i,2}.C_1 - C_2 \qquad (7)$$

Where

$$C_1 = a_{i,1}.(ex.f)\cos(\beta) + b_{i,1}\sin(\beta)$$
$$C_2 = Y_{i,2}\left(a_{i,1}.(ex.f)\sin(\beta) + b_{i,1}\cos(\beta)\right) + b_{i,1}(d-f) \qquad (8)$$

Dividing equation (7) by (Y$_{i,2}$ + d$_y$ - f), and after transformations it becomes :

$$\frac{a_{i,2} - a_{i,1}.\cos(\beta)}{\sin(\beta)} = \frac{b_{i,1}}{f.ex} \qquad (9)$$

(5)

3.3 Computation of the tracking parameters

(6)

For another rotation movement of the camera with angle θ,
a new equation is obtained for (7)segments S_i and S_j where v= $a_{i,3}$.u + $b_{i,3}$ is the equation of the segment S_i^3 in image IM^3.

Equation (9) is also valid for the same segment S_i after the rotation θ :

$$\frac{a_{i,3} - a_{i,1}.\cos(\theta)}{\sin(\theta)} = \frac{b_{i,1}}{f.ex} \qquad (10)$$

(11) is obtained after dividing equation (9) by (10) :

$$\frac{a_{i,2} - a_{i,1}.\cos(\beta)}{a_{i,3} - a_{i,1}.\cos(\theta)} = \frac{\sin(\beta)}{\sin(\theta)} = \sigma \qquad (11)$$

σ is constant value. (9)

For any segments taken from images 1,2 and 3 if they are matched they will give the same value of σ.

(10)

(11)

(12)

3.4 Principle of Algorithm

The main idea of the method is to compute the values of (σ) using three triplets of line segments $\left(S_{i1,}^1 S_{j1,}^1 S_{k1}^1\right)$, $\left(S_{i2,}^2 S_{j2,}^2 S_{k2}^2\right)$ and $\left(S_{i3,}^3 S_{j3,}^3 S_{k3}^3\right)$ chosen respectively from the first, second and the third image.

In the case where the three triplets are matched, solution of equations (11) by means of the slopes of these segment lines gives the same values of (σ).Let

- $\left(S_{i1,}^1 S_{j1,}^1 S_{k1}^1\right)$ be a selected randomly from the first image

- $\left(S_{i2,}^2 S_{j2,}^2 S_{k2}^2\right)$ and $\left(S_{i3,}^3 S_{j3,}^3 S_{k3}^3\right)$ be two triplets chosen randomly from the second and the third image so as they not match the first triplet.

In this case, the resolution of equations (11), by means of the slopes of these segment lines, gives empirical values of (σ_v).

Let A be the event corresponding to the same values of (σ_v) computed using other three triplets. As there are n(n-1)(n-2) possible values generated from all possible combinations, the probability of this event is equal to

$$\Pr ob(A) = \frac{1}{n x (n-1) x (n-2) - 1}$$

Let B be the event corresponding to another three triplets producing the same values of (σ_v). As the event B is independent of the event A, the probability that the two events occur one after the other is equal to :

Prob(B∩A)=Prob(B)xProb(A).

The repeat of this reasoning over all triplets, give the conclusion that the probability that the same values of (σ_v) will be computed m times decreases towards zero when m increases.

This result constitutes our main idea. It is inspired from the basic principle of the Hough Transform [19], generating all possible values of (σ_v). The values (σ_v) computed from the maximum number of three triplets constitute the solution of our problem, and these triplets of line segments constitute the result of the matching images.

3.5 Algorithm

Let H be a function computing value of (σ) using a set of three triplets of line segments.

$$H\left(\left(S_{i1,}^1 S_{j1,}^1 S_{k1}^1\right)\left(S_{i2,}^2 S_{j2,}^2 S_{k2}^2\right)\left(S_{i3,}^3 S_{j3,}^3 S_{k3}^3\right)\right)=\left(\sigma_i\right)$$

Where

(σ_i) is the ith value of (σ)

BEGIN

- REPEAT

- Select randomly three triplets

$$\left(S_{i1,}^1 S_{j1,}^1 S_{k1}^1\right)$$

$$\left(S_{i2,}^2 S_{j2,}^2 S_{k2}^2\right), \left(S_{i3,}^3 S_{j3,}^3 S_{k3}^3\right)$$

- Compute the slopes

- $(a_{i1,1,} a_{j1,1}, a_{k1,1}) for \left(S_{i1,}^1 S_{j1,}^1 S_{k1}^1\right)$

- $(a_{i2,2,} a_{j2,2}, a_{k2,2}) for \left(S_{i2,}^2 S_{j2,}^2 S_{k2}^2\right)$

- $(a_{i3,3,} a_{j3,3}, a_{k3,3}) For \left(S_{i3,}^3 S_{j3,}^3 S_{k3}^3\right)$

- Compute the value of σ_v , by resolving the equation (11)

$$card\left(H^{-1}\left(\sigma_v\right)\right)= card\left(H^{-1}\left(\sigma_v\right)\right)+1$$

- Until all triplets are selected

- Select (σ_s) where:

$$card\left(H^{-1}\left(\sigma_s\right)\right)= Max\left(card\left(H^{-1}\left(\sigma_i\right)\right)\right)$$

Selected value (σs) corresponds to the set of three triplets of line segments which are matched lines.

End.

4. EXPERIMENTAL RESULTS

4.1 Data of experiments

Two kinds of data are used :

- simulated data serving for the study of the impact of many factors on the uncertainty estimation of rotation angles with the orientation of line segments, the amounts of angles and noise.
- Images of an interior 3D scene acquired by a camera performing pan rotation.

To do this, a mechanical device for the movement of the camera, are used. This included a fixed support and a guidance and motor reducer drive device to move the camera in a horizontal rotational manner (see figure 3), where the rotation axis passes through the center of the camera. A set of images was taken by the camera after two rotations. For each position, an image of the scene was acquired. Figures 4, 5 and 6 illustrate an example of three acquired images.

Fig.3.The mechanical support for camera rotation

4.2 Uncertainty in the estimation of segments slopes

The influences of the noise over the uncertainty of the estimation of the tangencies of segments angles are studied. Instead of adding a Gaussian noise with a mean and standard deviation σ to image points, displaced the extreme points of line segments into new positions such that the difference between the new and old slopes is maximal. If m_l, m_2 are the two extremities, their new positions are m'_1, m'_2; or m''_1, m''_2 so as $m'_1m'_2$; and $m''_1m''_2$ are the tangencies to the discs $Dl(m_l,r)$, $D2(m_2,r)$ (see figure 7).

In order to study what are the more robust categories of slopes to noise, the value of 3 pixels for the ray r are used, which allowed the application of algorithm on new slopes that are noisier than in real images.
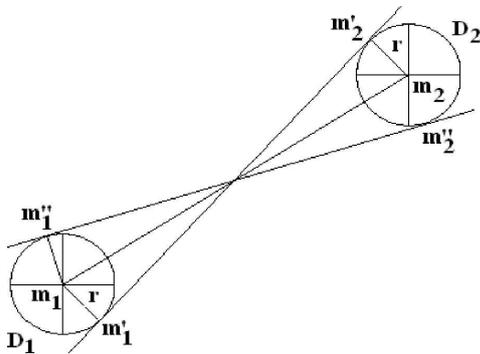
Fig.7. Noisy line segment

Histograms in figure 8 and 9 illustrate the results obtained for different categories of slopes.

It is clear that the considerable noise decreases the accuracy in the estimation of the slopes.

However, we can conclude that the slopes of long segment are more robust to noise that the short one. And with the same length of line segments, the horizontal one is more robust than the other categories.
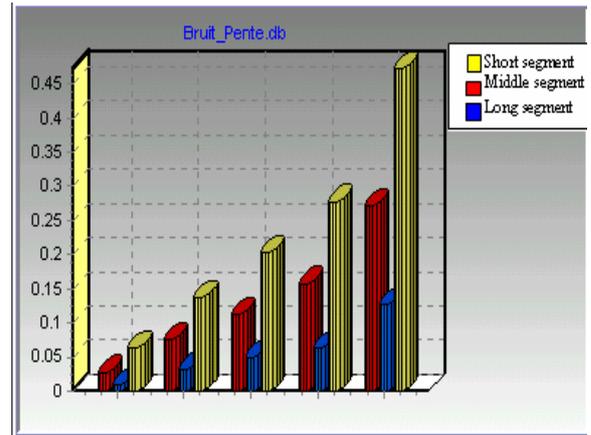
Fig.8 Histogram noise/slope
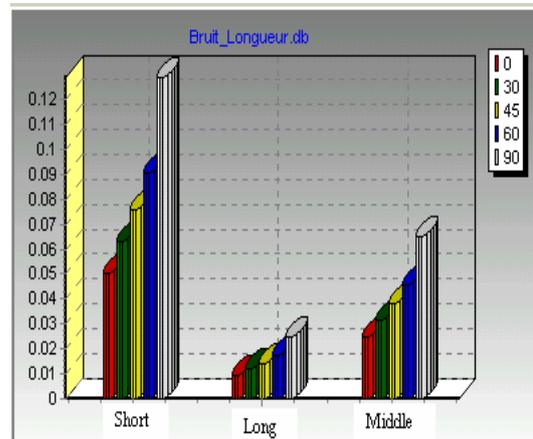
Fig.9. Histogram noise/length

4.3. Tracking line segment for indoor image

Contour points of the three images illustrated in figure 4,5 and 6 are extracted using the Canny-Deriche detector [20] and approximated by a set straight line segment (see figure 10). The algorithm of section 3.5 is applied to the estimation matching line segments.

In order to increase the accuracy of computed values of (σ), all combinations of segments are used, for which the absolute slopes appertain to horizontal or oblique segment. If n is the number of matched lines for each one of the three images, the number of values of (σ) is C^3_n values.

The correct matches between contour lines of the three images is given by the computed value of (σ) with a high score. Figure 11 illustrates line segments matched with the same color.

Fig 4. Initial image   **Fig 5. Second image**   **Fig 6. Final image**
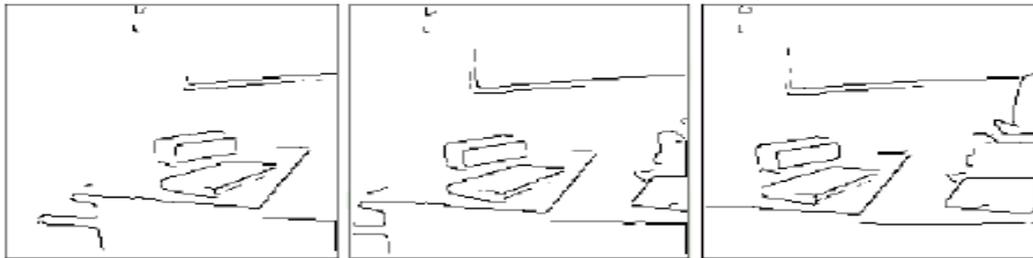
Fig.10. Extracted contours
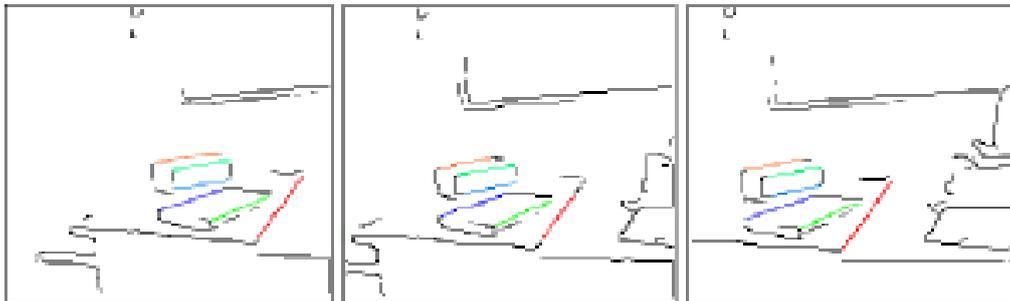
Fig. 11. Same color indicates the segments in correspondance

The average values of (σ) computed are (0,497237), while the real values are (0,491961).

The result appears very interesting because the estimated value is very near of real one.

4.4. Other Examples of experimental results

a) Sequence of 3 real images (puce)

Images are segmented by Canny-Deriche method
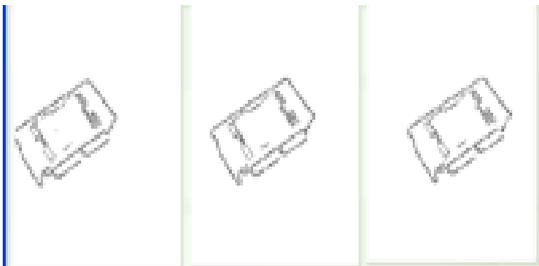
The value of (σ) which has the big vote gives the line segments which are in correspondence.
The matching segments are as follow
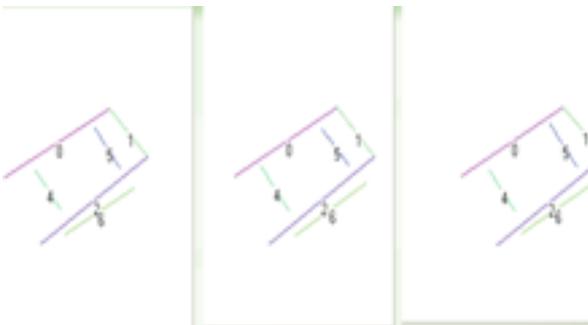
The segments with same colors are in correspondence.

B) Sequence of 3 real images (chair)

Images are segmented by Canny-Deriche method
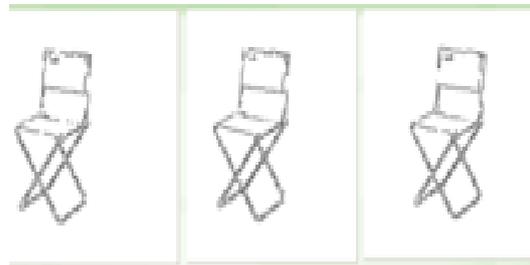
The value of Sigma which has the big vote gives the line segments which are in correspondence.
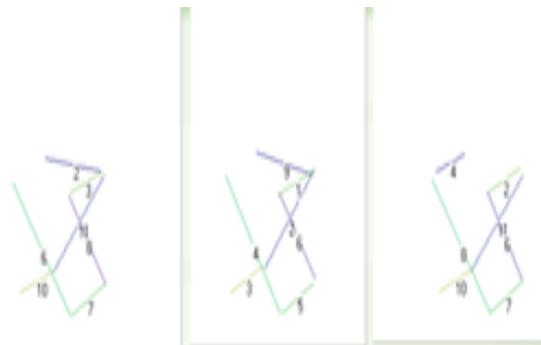
The segments with same colors are in correspondence. The Number on the segment indicate the order of its extraction on image

Example: the segment 0, 11, 0 represent the same segment and are matching.

5. CONCLUSION

In this paper a novel method is presented to match line segments in an indoor environment.

Our contribution in this work is :

*   The proposition of a method which resolves the problem of matching the segment features without any knowledge about camera's motion or its parameters comparing to others witch necessitate the calibration of the cameras.

*   The proposed algorithm is based uniquely on the simple mathematical formulation exiting between the slopes of line segments in successive images where the others suppose the camera motion are known or necessitate a set of initial matching features.

*   The method is robust to noise because the slope is more robust than the endpoint of the segment.

*   The method is completely intensity-blind as no photometric information was used to achieve the matching.

The tracker algorithm tested on synthetic data generated by computer gives expected results on noisy and not noisy data.

Applied also to real images segmented by Canny-Deriche's method, the proposed algorithm gives suitable results. Furthermore, the experiment shows that the algorithm gives better results for the oblique segments ($< = 45$ °) and for longer segments with regard to the short ones.

The proposed method depends strongly on the robustness of the algorithm of the extraction of line seegment

6. REFERENCES

[1] B.D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI'81*, pages 674–679, 1981

[2] J. Shi and C. Tomasi. Good features to track. In *CVPR'94*, pages 593–600, Seattle, Washington, June 1994.

[3] M.-O. Berger. How to track efficiently piecewise curved contours with a view to reconstructing 3D objects. In *ICPR'94*, pages 32–36, Jerusalem, October 1994.

[4] A. Blake and M. Isard. *Active Contours*. Springer Verlag, April 1998.

[5] S. Boukir, P. Bouthemy, F. Chaumette, and D. Juvin. A local method for contour matching and its parallel implementation. *MVA'98*, 10(5/6):321–330, April,1998.

[6] G. Hager and K. Toyama. The XVision system: A general-purpose substrate for portable real-time vision applications. *CVIU'98*, 69(1):23–37, January 1998.

[7] M. Vincze. Robust tracking of ellipses at frame rate. *Pattern Recognition*, 34(2):487 – 498, February 2001

[8] P. David and D. DeMenthon. Object recognition in high clutter images using line features. In Tenth IEEE International Conference on Computer Vision, 2005.,volume 2, pages 1581–1588, 2005.

[9] T. Lemaire and S. Lacroix. Monocular-vision based SLAM using line segments. In IEEE International Conference on Robotics and Automation, Roma (Italy), April 2007.

[10] P. Smith, I. Reid, and A. Davison. Real-time monocular SLAM with straight lines. In British Machine Vision Conference, Edinburgh (UK), Sep. 2006.

[11] Canny, J., A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. vol.8(6), pp. 679-714, 1986.

[12] V. Nguyen, S. G¨achter, A. Martinelli, N. Tomatis, and R. Siegwart. A comparison of line extraction algorithms using 2d range data for indoor mobile robotics. Auton. Robots, 23(2):97–111, 2007.

[13] L. Chen and al. A simple tracking approach for line extraction. In Proceedings of the International Conference on Wavelet Analysis and Pattern Recognition, 2007.

[14] J-M Morel R. Grompone von Gioi, J. Jakubowicz and G. Randall. Lsd : A fast line segment detector with a false detection control. IEEE Transaction on Pattern Analysis and Machine Intelligence, Dec. 2008.

[15] R. Deriche and O. D. Faugeras. Tracking line segments. In O. D. Faugeras, editor, *ECCV*, volume 427 of *Lecture Notes in Computer Science*, pages 259–268. Springer, 1990.

[16] C. Schmid and A. Zisserman. Automatic line matching across views. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 666–671, 1997.

[17] H. Bay, V. Ferrari, and L. Van Gool. Wide-baseline stereo matching with line segments. In CVPR '05: Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1, pages 329–336, Washington, DC, USA, 2005. IEEE Computer Society

[18] Faugeras, O., Luong, Q.T., Papadopoulo, T., The Geometry of Multiple Images. MIT Press, Cumberland (2000)

[19] Shapiro, L. & Stockman, G., Computer Vision, Prentice-Hall, Inc. 2001

[20] Shapiro, L. & Stockman, G., Computer Vision, Prentice-Hall, Inc. 2001

[21] Canny, J., A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. vol.8(6), pp. 679-714, 1986.

[22] Ding, W. and Marchionini, G. 1997 A Study on Video Browsing Strategies. Technical Report. University of Maryland at College Park.

[23] Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems

[24] Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[25] Sannella, M. J. 1994 Constraint Satisfaction and Debugging for Interactive User Interfaces. Doctoral Thesis. UMI Order Number: UMI Order No. GAX95-09398., University of Washington.

[26] Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[27] Brown, L. D., Hua, H., and Gao, C. 2003. A widget framework for augmented interaction in SCAPE.

[28] Y.T. Yu, M.F. Lau, "A comparison of MC/DC, MUMCUT and several other coverage criteria for logical decisions", Journal of Systems and Software, 2005, in press.

[29] Spector, A. Z. 1989. Achieving application requirements. In Distributed Systems, S. Mullender