

# SIMPLE MONOCULAR DOOR DETECTION AND TRACKING

*Rafiq Sekkal\*, François Pasteau\*, Marie Babel\*, Baptiste Brun\*, Ivan Leplumey\*\**

INSA, IRISA

\*Lagadic Team, \*\*Intuidoc Team  
Rennes, France

## ABSTRACT

When considering an indoor navigation without using any prior knowledge of the environment, relevant landmark extraction remains an open issue for robot localization and navigation. In this paper, we consider indoor navigation along corridors. In such environments, when considering monocular cameras, doors can be seen as important landmarks. In this context, we present a new framework for door detection and tracking which exploits geometrical features of corridors. Since real-time properties are required for navigation purposes, designing solutions with a low computational complexity remains a relevant issue. The proposed algorithm relies on visual features such as lines and vanishing points that are further combined to discriminate the floor and wall planes and then to recognize doors within the image sequences. Detected doors are used to initialize a dedicated edge-based 2D door tracker. Experiments show that the framework is able to detect 82% of doors on our dataset while respecting real time constraints.

**Index Terms**— Door detection, visual features, vanishing point, line extraction, 2D tracking.

## 1. INTRODUCTION

The navigation within unknown indoor environment using electrical wheelchair remains a challenging task. Indeed, in an dynamic environment, disabled people have to be more vigilant while steering the wheelchair, thus inducing fatigability that affects their autonomy. Besides, depending on the disabilities, navigation accuracy can be deeply altered. In particular going through doors and taking an elevator in a secure way and without collisions remains a relevant issue.

In this context, the system has to precisely determine in real-time the location of doors within corridors, in order to further propose a related path planning. Since doors provide stable and meaningful structure, we first have to design an automatic door detection system able to initialize the tracking process.

Recent works have been focused on door detection for indoor navigation. To this aim, different type of sensors (laser

range, camera, etc.) have been considered. Among state-of-the-art solutions, the more efficient ones (up to 90% of recognized doors) are typically based on laser range finder [1, 2]. However, the prohibitive cost of such systems prevents from spreading them onto electrical wheelchairs. Visual solutions based on appearance and shape cues have also been proposed [3] : this solution requires a training step. [4, 5] jointly use a variety of visual features including color, texture and edge intensity that are combined in an AdaBoost classifier to identify doors in the scene. However, since indoor environment such as corridors are typically textureless, miss-detections commonly appear, and only close doors can be detected.

The proposed solution is thus based on features that remain independent of door/wall color or texture. Vanishing points of the scene are first extracted. Since the robot evolves in a corridor, only the vanishing point that corresponds to the end of the corridor should be detected. An original time-consistency framework is then implemented, in order to avoid incoherent behavior between frames. Then, lines related to floor/wall boundaries are estimated. This information is then used to detect 3D planes associated to walls. Naturally, since doors are located in walls, the door location research process is limited to these planes, thus accelerating the overall process.

In addition, the proposed framework should be able to track the detected doors over frames by estimating their position within each frame of the video. Different tracking techniques have been proposed in the literature: they mainly differ in geometric features that are used to describe the tracked objects, e.g. contours [6, 7], straight lines or segments [8], etc. Since doors are characterized by straight lines, we used the 2D Moving Edges (ME) tracker [8, 9] while implementing an oriented gradient mask. The ME algorithm performs a one dimensional research on the normal of the contour at each sampled point and achieves real time performances. This process is then able to retrieve geometrical structures.

This paper is organized as follows. Section 2 presents the proposed door detection framework. Then, section 3 introduces the tracking process approach developed for this system based on the local gradient information computed during the navigation task. Finally, in section 4, experiments and results

are shown to validate our approach.

## 2. GENERAL FRAMEWORK DOOR DETECTION

Since our goal is to recognize doors in corridors environment, we first have to determinate related geometrical features. Indeed, doors can be represented by a rectangular shape where width is given by the distance between its two doorposts. Doors located in the same corridor classically share the same characteristics. In particular, door shape should respect the same height/width ratio. The challenge is then to estimate this width in a corridor taking into account the distance of the door from the camera which strongly affects the door width. If we assume that the camera optical axis is always horizontal, this implies that doorposts are nearly verticals.

The general door detection framework is divided in two main steps. First we estimate the wall/floor boundary. To this aim, lines are extracted and further cross-checked with temporally consistent vanishing point estimation to detect 3D planes associated to walls. Then rectangular shapes that respect a door-like height/width ratio are extracted within wall planes. Ratio is empirically fixed according to handicap compliant standard of door shapes.

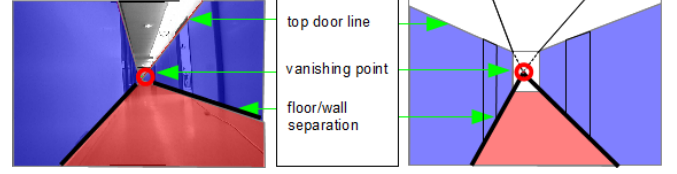
### 2.1. Features Selection

#### 2.1.1. Line extraction

Line detection is the first step of the proposed detection framework. We used the LSD (Line segment Detection) algorithm that is based on local gradient orientations in the image [10]. However, the main issue of this algorithm is the cutting of one line when gradient change the direction. To add robustness to this scheme, a merging line process is applied to group lines that can be considered as single line. To this aim, for each couple of lines, if they share the same slope and their extremities are closed enough to each others, they are merged to form a unique line. The merging process is only applied on the vertical lines, in order to recover the best the doorposts for further rectangular shape detection.

#### 2.1.2. Vanishing Point Estimation

Estimation of vanishing point remains an open issue, especially when aiming at designing an accurate as well as a low complex solution. Vanishing points correspond to locations where a significant number of intersecting features are observed. Classically, vanishing point estimation processes are based on gaussian sphere projection framework [11, 12]. In the proposed approach, for each new frame we compute the current vanishing point  $vp_t$  in the current frame  $I_t$ . Lines segments are classified into two categories: vertical lines that define infinite vanishing points and non vertical lines. Only these latter are used for the gaussian sphere projection, thus speeding up the process.



**Fig. 1.** General corridor geometrical structure: floor and wall planes automatically defined

In order to ensure a temporal consistency of the vanishing point, we apply a low-pass filter along the sequence. A factor  $\alpha$  is used and empirically tuned to ensure a smooth variation of the resulting estimated vanishing point  $vp_t$  at the current frame  $I_t$  according to the previous ones. To this aim, a temporal filtering process is used, following

$$\begin{cases} \widehat{vp}_0 &= vp_0 \\ \widehat{vp}_t &= \alpha \times \widehat{vp}_{t-1} + (1 - \alpha) \times vp_t \end{cases}$$

where  $0 < \alpha < 1$ .

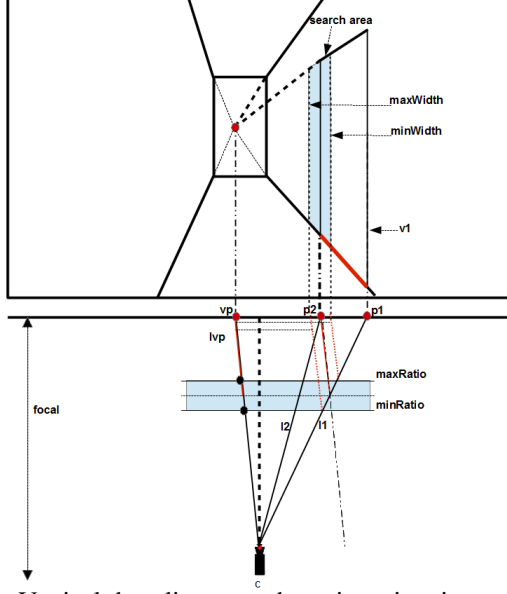
### 2.2. Wall/floor line boundary detection

In the literature, different techniques have been proposed for wall/floor boundary detection. In [13] corners corresponding to the intersection of a vertical line and the floor plane are used to define wall/floor features. In [14], a dynamical Bayesian network model is applied on each column of the picture to estimate the floor boundary. In order to detect the wall floor boundary, we first look for a set of non vertical lines as wall/floor boundary. Each wall contains a set of vertical lines that correspond to different elements (doorposts, closet, boards, etc.). Wall/floor boundaries are characterized by the fact that the lowest extremities of doorposts intersect this particular line. As a consequence, wall/floor boundaries correspond to vanishing lines that cross the most vertical lines bottom extremities. To minimize false positive detection, a maximal distance between the vertical line extremity and the vanishing line is then defined. By applying the same principle on the top extremities, we can easily detect the vanishing line of top of doors on the same wall. Wall plane is then immediately obtained (Figure 1).

### 2.3. Door recognition

Once the 3D geometrical structure has been detected, door search area is reduced to the wall limited by door height. Last step consists of finding rectangular shapes able to represent doors. Using two parameters (distance to the vanishing point and the camera orientation in the corridor), we estimate door width while considering standard size of doors. Thus, the idea is to find a pair of vertical lines (two doorposts) that fits the door model.

We then compute the door width relatively to the representation depicted on Figure 2, that relies on a classical



**Fig. 2.** Vertical door lines search: ratio estimation according to the vanishing point

pinhole camera model. For each vertical line  $v1$ , we then define a reduced search area bounded by  $widthMax$  and  $widthMin$  that are respectively computed using a  $ratioMax$  and  $ratioMin$  according to the formula (1) where we expect to find another vertical line that corresponds to the second doorpost.

$$widthMin = \frac{dist(P_1, vp)}{dist(vp, c)} \times \frac{length(vl)}{ratioMin} \quad (1)$$

with:

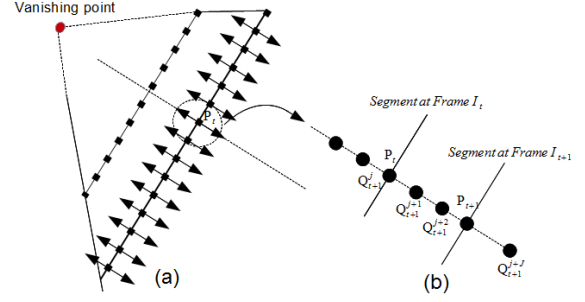
$dist(a, b)$  the distance between two points  $a, b$ .  
 $length(l)$  the length of a line segment  $l$ .

As soon as a candidate is found, the idea is to check if this structure can be labelled as a door. To this aim, a gradient based filter is applied on the lines that remain close to the two top extremities of the vertical lines. If a horizontal line is then observed, the structure is said to be a door.

### 3. DOOR TRACKING

The door tracker consists of a 2D edge tracker that is only applied on doorposts. It allows local tracking of edge points. Our algorithm is inspired from the *Moving Edge* (ME) algorithm [8, 9, 15]. Contours are first sampled at a regular interval. At each sample point, a one dimensional search is performed on the normal of the contour for each corresponding edges as depicted on Figure 3.

Actually, for each sample point  $P_t$  in the image  $I_t$ , the corresponding point  $P_{t+1}$  in the next image  $I_{t+1}$  is expected to



**Fig. 3.** 2D Edge Tracking: point position using the oriented mask filtering

be on the normal axis since the tracker deals only with doorposts which are always vertical. Estimated positions are then characterized by the set of points  $\{Q_{t+1}^j, j \in [-J, J]\}$  located on either side of  $P_t$ . Then, for each position  $Q_{t+1}^j$ , an oriented and derivative convolution mask is applied in order to compute the local gradient  $G_t^j$ . The mask  $M$  is a  $n \times m$  matrix where  $n < m$ . This oriented mask is optimal towards the contour orientation. The new position is giving by

$$P_{t+1} = Q_{t+1}^{j*} = \underset{j \in [-j, j]}{argmax} G_{t+1}^j \quad (2)$$

with

$$G_{t+1}^j = |I_{t+1} * M + I_t * M| \quad (3)$$

that corresponds to the square root of a log-likelihood ratio.

### 4. RESULTS

To validate our approach, seven test sequences have been realized with a single camera embedded onto the wheelchair. Sequences are representative of indoor navigation throughout corridors.

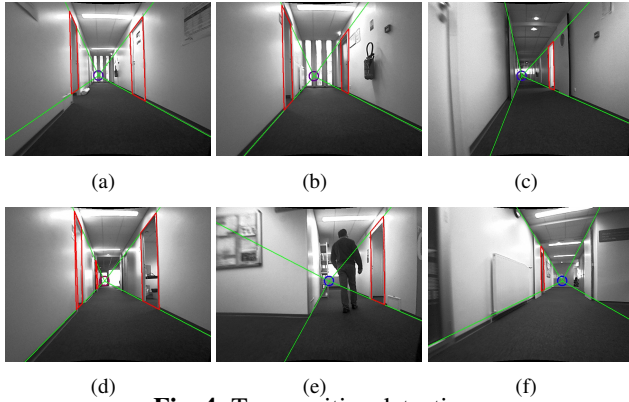
First experiments aim at evaluating the lone detection framework performances.

Fig. 4 shows the obtained detected door shapes with different experimental conditions. As a whole, most of situations are correctly analyzed leading to automatic door detection. Fig. 4(a), 4(b), 4(c) and 4(d) show true positive multiple detection of either open or closed doors at different distances and for different illumination conditions. Moreover, in the presence of moving obstacles, corridor crossings (Fig. 4(e)) or other rectangular shapes (Fig. 4(f)), detection is not affected.

The proposed algorithm will fail when particularly hard illumination conditions occur leading to low local gradient levels or bad vanishing point observation. In these cases, slightly moving forward the wheelchair is sufficient enough to overcome this issue while remaining acceptable in terms of use case. Indeed, we observed that an average of 8 images are required to recover the missing door, then corresponding of less than half a second.

	images	doors	good detection	wrong detection
total	3414	6356	5212	459

**Table 1.** Performance of door detection and tracking algorithm on sequences taken from different corridors in the Inria building



**Fig. 4.** True positive detections

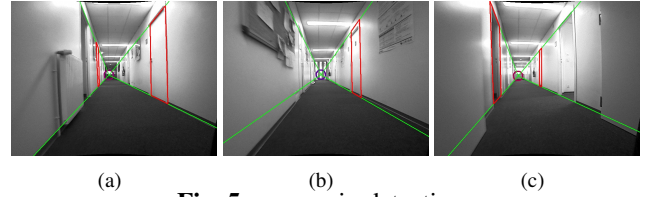
Tracking results are illustrated by Fig. 6. As soon as a door is detected, the 2D edge-based tracker described in section 3 is applied. The two doorposts are then tracked, even if they are partially masked. Additional results can be found on <sup>1</sup>. Experiments then show the ability of our global scheme to detect and track doors within corridors. As a whole, most of situations are correctly analyzed leading to automatic door detection. As shown in Tab. 1, our dataset contains 6356 doors that appear along the sequences: the proposed detection algorithm was able to detect over 82% of them.

## 5. CONCLUSION

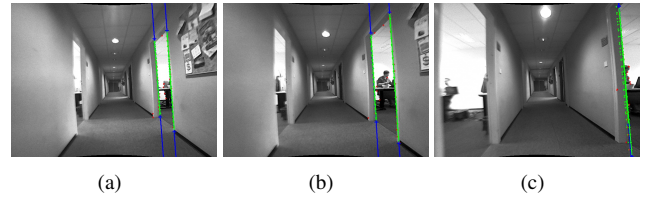
In this paper, detection algorithm dedicated to automatic door recognition has been designed. This scheme consists of the automatic initialization step of a complete tracking process that is intended to be embedded onto a robotic wheelchair. Based on a single camera, this approach is shown to be simple and efficient relatively to the state-of-the-art solutions. In particular, closed or open doors are easily detected. Thanks to a simple geometrical model of doors, the joint use of the LSD algorithm and a temporally consistent vanishing point estimation process provides reliable results, by eliminating most of false detections. In particular, most of rectangular shapes other than doors are not detected, thus leading to a robust framework.

Coupled with a dedicated real-time 2D edge-based tracker, the process provides a robust framework able to locate and track along the sequence the doors within a given corridor. This approach remains relevant even if major parts of doors are occluded.

<sup>1</sup><http://www.irisa.fr/lagadic/team/rafiq.sekkal-eng.html>



**Fig. 5.** some missdetections



**Fig. 6.** Door tracking: only doorposts are tracked using 2D Edge tracker. (a) frame 1, (b) frame 144 and (c) frame 250

This framework is however mainly dependant of the vanishing point estimation process, so that the geometrical model of doors can be efficiently extracted. In some cases, the environment is not favourable, leading to non detection situations. Nevertheless, the miss of a door can be compensated as soon as the wheelchair is moving. This solution remains realistic in our use case since the resulting frame rate is sufficient enough.

## 6. REFERENCES

- [1] D. Anguelov, D. Koller, E. Parker, and S. Thrun, "Detecting and modeling doors with mobile robots," in *IEEE International Conference on Robotics and Automation*, 2004, vol. 4, pp. 3777 – 3784.
- [2] J. Hensler, M. Blaich, and O. Bittel, "Real-Time Door Detection Based on AdaBoost Learning Algorithm," in *Research and Education in Robotics - EUROBOT 2009, Communications in Computer and Information Science*, Vol. 82, 2010, p. 61.
- [3] A. C. Murillo, J. Kořecká, J. J. Guerrero, and C. Sagüés, "Visual door detection integrating appearance and shape cues," *Robot. Auton. Syst.*, vol. 56, no. 6, pp. 512–521, June 2008.
- [4] Z. Chen and S.T. Birchfield, "Visual detection of lintel-occluded doors from a single image," in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, june 2008, pp. 1 –8.
- [5] Zhichao Chen, Yinxiao Li, and Stanley T. Birchfield, "Visual detection of lintel-occluded doors by integrating multiple cues using a data-driven markov chain monte carlo process," *Robotics and Autonomous Systems*, vol. 59, no. 11, pp. 966 – 976, 2011.
- [6] M.-O. Berger, "How to track efficiently piecewise curved contours with a view to reconstructing 3d objects.," in *In Int. Conf on Pattern Recognition, ICPR94*, 1994, pp. 32–36.
- [7] N. Vaswani, Y. Rathi, A. Yezzi, and A. Tannenbaum, "Deform pf-mt: Particle filter with mode tracker for tracking nonaffine contour deformations," *IEEE Transactions on Image Processing*, vol. 19, no. 4, pp. 841 –857, april 2010.

- [8] S. Boukir, P. Bouthemy, F. Chaumette, and D. Juvin, "A local method for contour matching and its parallel implementation," *Machine Vision and Applications*, vol. 10, pp. 321–330, 1998.
- [9] P. Bouthemy, "A maximum likelihood framework for determining moving edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 5, pp. 499–511, may 1989.
- [10] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, G. Randall, "LSD: a Line Segment Detector," *Image Processing On Line*, 2012.
- [11] C. Rother, "A new approach for vanishing point detection in architectural environments," in *In Proc. 11th British Machine Vision Conference*, 2000, pp. 382–391.
- [12] K. Boulanger, K. Bouatouch, and S. Pattanaik, "ATIP: A Tool for 3D Navigation inside a Single Image with Automatic Camera Calibration," in *EG UK conference*, 2006.
- [13] K. Ok, D.-N. Ta, and F. Dellaert, "Vistas and wall-floor intersection features - enabling autonomous flight in man-made environments," in *Workshop on Visual Control of Mobile Robots*, 2012.
- [14] E. Delage, Honglak Lee, and A.Y. Ng, "A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2006, vol. 2, pp. 2418 – 2428.
- [15] E. Marchand and F. Chaumette, "Feature tracking for visual servoing purposes," *Robotics and Autonomous Systems*, vol. 52, no. 1, pp. 53–70, July 2005.