IMAGE SIGNAL PROCESSING FOR VISUAL DOOR PASSING WITH AN OMNIDIRECTIONAL CAMERA

Luis-Felipe Posada, Thomas Nierobisch, Frank Hoffmann and Torsten Bertram
Chair of Control System Engineering, Technische Universität Dortmund, Dortmund, Germany

Keywords: Omnidirectional vision, Door passing, Robotic behaviour, Visual navigation, Door detection, Visual servoing.

Abstract: This paper proposes a novel framework for vision based door traversal that contributes to the ultimate goal of purely vision based mobile robot navigation. The door detection, door tracking and door traversal is accomplished by processing omnidirectional images. In door detection candidate line segments detected in the image are grouped and matched with prototypical door patterns. In door localisation and tracking a Kalman filter aggregates the visual information with the robots odometry. Door traversal is accomplished by a 2D visual servoing approach. The feasibility and robustness of the scheme are confirmed and validated in several robotic experiments in an office environment.

1 INTRODUCTION

Vision plays an increasingly prominent role in autonomous navigation of mobile robots (DeSouza and Kak, 2002). This development is on the one hand driven by the exponential increase in performance of modern cameras and computers at increasingly economic costs. On the other hand robotic researchers learn to harvest the broad spectrum of robust and effective low-, mid- and high-level vision algorithms developed by the computer vision community over the past two decades for the purpose of robot localisation, map building and navigation. In that context this paper contributes to the development of vision based robot navigation.

Door traversal is a vital skill for autonomous mobile robots operating in indoor environments. Robust and reliable door passing is feasible with laser range scanners as their angular resolution provides sufficient information to distinguish among open doors and other objects such as tables and shelves. However 2D laser range scanner are not suitable for the detection of closed or partially opened doors (Jensfelt, 2001). Equipment of mobile robots with a laser range scanner contributes significantly to the overall cost of the mobile platform.

Door passing that relies on sonar sensors is feasible in some scenarios but is in general not robust and reliable enough to be applicable in all office environments (Budenske and Gini, 1994). In the context of door detection and door passing vision provides an economic albeit reliable alternative to proximity sensors such as sonar or laser range scanner. However, the robust visual detection and localization of doors remains a challenging task despite a number of successful implementations in the past (Eberst et al., 2000; Shi and Samarabandu, 2006; Stoeter et al., 2000; Patel et al., 2002; Murillo et al., 2008).

The authors in (Stoeter et al., 2000) detect doors by means of a monocular camera in conjunction with sonar. The visual door detection is restricted to large views and rests on the assumption of a priori knowledge of the door and corridor dimensions. The final door detection at close range relies on sonar information only.

In (Monasterio et al., 2002) the authors employ a Sobel edge detector combined with dilatation and a filtering operation to detect doors in monocular images. The final door identification employs an artificial neural network to classify the presence or absence of a door based on the sonar data. Initially the robot approaches the door based on the pose estimated from visual information. The final approach and door traversal relies on sonar data. The approach rests on the assumption that the robot initially already faces the door, which excludes more realistic scenarios in which the robot travels along a corridor with parallel orientation to the doors.

The door traversal approach by (Eberst et al., 2000) is robust with respect to individual pose errors, scene complexity and lighting conditions as door hypothesis are filtered and verified for consistency

Felipe Posada L., Nierobisch T., Hoffmann F. and Bertram T. (2009).
IMAGE SIGNAL PROCESSING FOR VISUAL DOOR PASSING WITH AN OMNIDIRECTIONAL CAMERA.
In *Proceedings of the Fourth International Conference on Computer Vision Theory and Applications*, pages 473-480
DOI: 10.5220/0001791404730480

across multiple views. The door detection relies on a binocular pan-tilt camera system whereas our approach uses an omnidirectional camera. An omnidirectional view offers the advantage that an initial scan of the environment for doors executed by rotating the robot base becomes obsolete. In addition the omniview guarantees that the door remains visible throughout the entire door traversal whereas with a conventional perspective camera the door eventually leaves the field of the view such that the final stage of door traversal is performed open loop. The omniview also offers an advantage in scenarios with semi open doors in which the robot still detects the door in the rear view after it has passed the door leaf.

Omnidirectional vision for door traversing has also been investigated by (Patel et al., 2002). However the suggested solution is based on depth information obtained from a laser sensor to guide a mobile platform through a doorway.

To our best knowledge there is still no proposed solution that handles the three problems of door detection, localisation and door traversal in a coherent purely vision based framework. The majority of proposed solutions rely on range sensors in one way or another. Even though the underlying methods for image processing and door frame recognition are standard, our approach is novel as it provides a robust and coherent solution to the entire door detection and navigation problem relying in omnidirectional vision only.

Images contain a large amount of information which necessitates the filtering, extraction and interpretation of those image features that are relevant to the task. Similar to other approaches in the past our door detection scheme relies on a door frame model composed of two vertical door posts in conjunction with a horizontal top segment. The image processing of edges involves edge detection, thinning, gap bridging, pruning and edge linking. Individual edge segments are aggregated into lines by means of line approximation, line segmentation, horizontal and vertical line selection and line merging. Finally the lines extracted from the omnidirectional image are compared with the door frame model which upon successfully matching constitute door hypotheses. These hypotheses are tracked over multiple frames and eventually confirmed. The position of the door relative to robot frame is estimated by a Kalman filter that aggregates the robot motion with the door perception. The vision based door recognition and traversal problem is structured into the three steps: 1) door detection 2) door localisation and tracking and 3) door traversal which are discussed in the three following sections and are illustrated in figure 1. Section 4 reports experimental results of the door traversal in our office environment. The paper concludes with a summary and outlook in section 5.

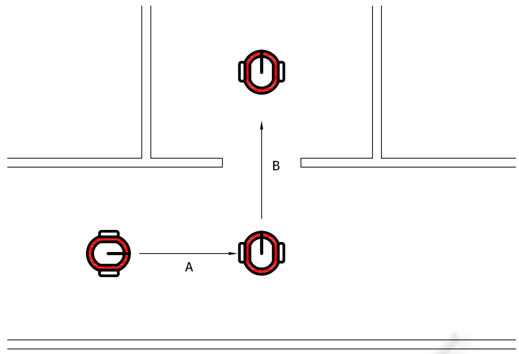


Figure 1: A) Vision based door detection, door localisation and tracking B) Door traversal by visual servoing.

2 DOOR DETECTION

The mobile robot Pioneer 3DX is equipped with an omnidirectional camera which provides a 360° view of the scene. Catadioptric cameras employ a combination of lenses and mirrors. Our camera obeys the single viewpoint property, which is a requirement for the generation of pure perspective images from the sensed images. A formal treatment of catadioptric systems is provided by (Baker and Nayar, 1998; Geyer and Daniilidis, 2001).

The robot navigates through the environment by means of a topological map. The door detection algorithm is designated to detect doors from arbitrary robot view points, including lateral and rear views. The door detection relies on door frame recognition and thus rests on the reasonable assumption that the door frame contrasts with the surrounding background. The map provides no prior information on door locations, however in conjunction with a mission plan it enables the robot to either traverse the left or right of two opposite doors in a corridor.

The door detection is composed into three subsequent fundamental steps: image processing, line processing and door frame recognition. The image processing is executed in the image space; hence, the processed entities are pixels. The line processing is performed in Cartesian coordinates and the entities handled are lines. Finally, the detected lines and their spatial relationship are interpreted to recognize the door. Figure 2 illustrates the visual door detection with pivotal processing steps.

2.1 Image Processing

Histogram equalization is applied to the original image I_O with the objective of obtaining full dynamic range of gray levels. Edges define regions in the image at which the intensity level changes abruptly. Our scheme employs the well known Canny edge detection algorithm (Canny, 1986) not only because it is still the most common edge detector in the vision community, but because is maximizes the signal-to-noise ratio and generates a single response per edge. The Canny edge detector consists of a Gaussian smoothing filter with a kernel of nine pixels and $\sigma = 1.85$, non-maximum suppression and edge thresholding with lower and upper thresholds of 15 and 30.

Morphological image processing is a useful technique for noise removal, image enhancement and image segmentation. The image obtained from morphological processing, edge detection and contrast stretching is shown in figure 2b. The following morphological operations are applied to the binary edge image.

- **Thinning** is a so called hit-or-miss transformation that is similar to erosion and dilation. The structuring element is applied to all pixels, in case the image pixels match the structuring element the corresponding image pixel is set to foreground (1), otherwise it is set to background (0).
- **Gap Bridging** bridges disconnected groups of pixels. Gap bridging is also defined in terms of hit-or-miss transformations and is applied iteratively until it produces no further changes to the image.
- **Pruning** is an operation to complement thinning as it is intended to remove spurs that remain after thinning.
- **Isolated Pixel Removal** eliminates all isolated foreground pixels that are surrounded by background pixels.

The edge detection and morphological processes characterizes individual pixels as edge pixels, but do not consider their connectivity. The edge linking groups these pixels into sets of connected pixels better suited for subsequent processing. We applied edge linking by contour following described in (Kovesi, 2000) as shown in figure 2c.

2.2 Line Processing

The previous operations are performed on pixel level, whereas line processing decomposes a curved contour into a sequence of straight lines. The generated straight line segments are classified into horizontal, vertical and randomly oriented lines, of which the first two are relevant for door detection.

Several techniques are reported in the literature for edge linking and boundary detection such as global processing via Hough transform (Duda and Hart, 1972), RANSAC fitting or local processing by linking pixels analyzing relations in a small neighborhood. Our experience in experiments reveals that the Hough transform fails to detect many meaningful line structures required for door detection.

We employ a split approach to partition the linked contours into straight line segments. A straight line through two points (x_1,y_1) and (x_2,y_2) is represented by:

$$x(y_1 - y_2) + y(x_2 - x_1) + y_2x_1 - y_1x_2 = 0 \quad (1)$$

Contours are represented by ordered lists of connected pixels. In the first step the contour is globally approximated by a straight line connecting the first and final pixel. The pixel (x,y) of maximum orthogonal distance

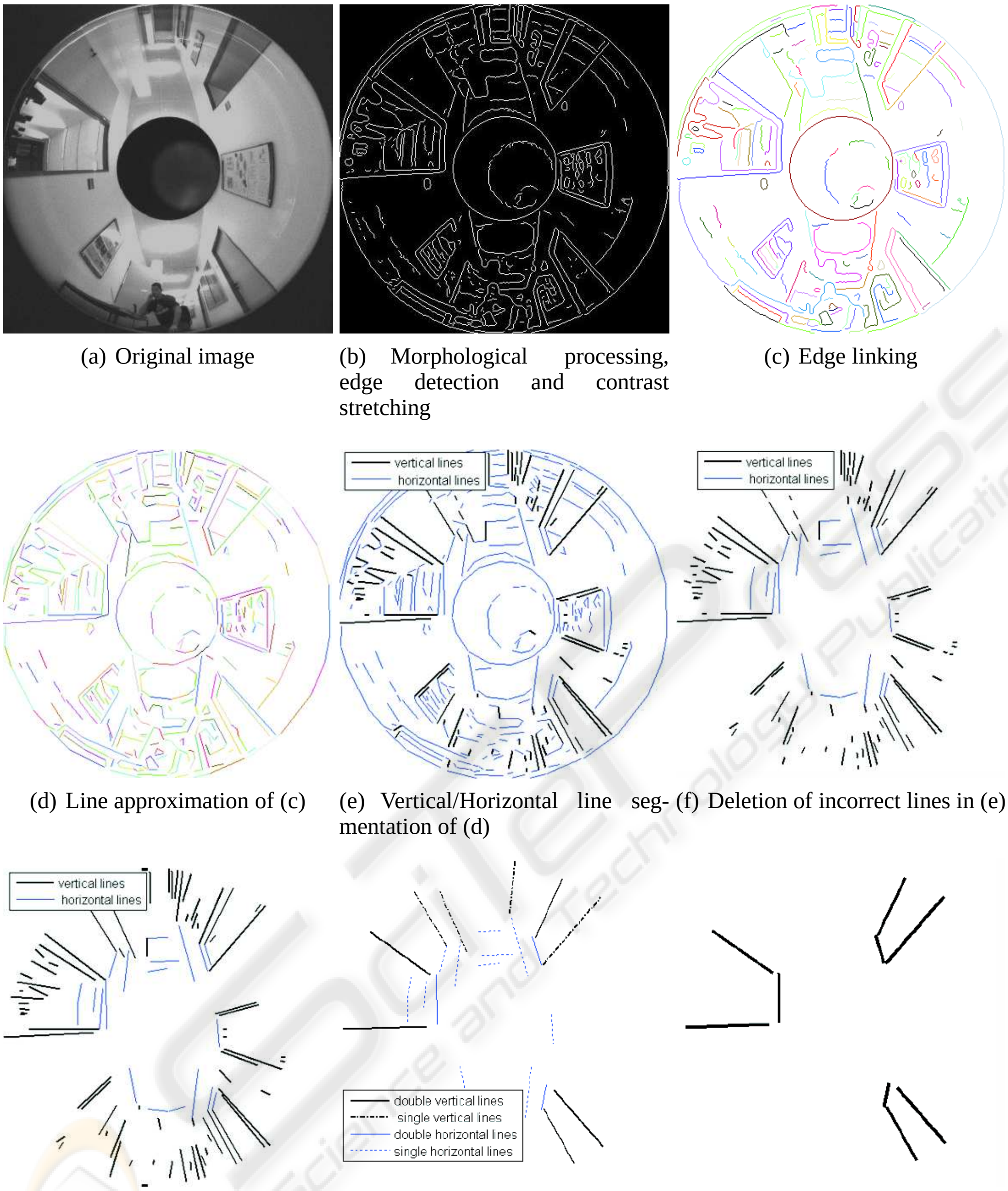
$$d = \left| \frac{x(y_1 - y_2) + y(x_2 - x_1) + y_2x_1 - y_1x_2)}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}} \right| \quad (2)$$

to the line is identified. The original line is split into two straight line segments, in which the maximum distance pixel becomes the end point of the first segment and the starting point of the second segment. This segmentation is recursively repeated until all pixels are located with a distance of less than 2 pixels with respect to their associated straight line. The straight line segmentation results are shown in figure 2d.

The resulting lines are classified into vertical lines, that are potential candidates for door posts, horizontal lines as candidates for the horizontal door case and others not part of door frames.

The lines are classified according to the angle between the line itself and the radial line going through the central principal point. Lines with an angle of 15° are considered vertical and with an angle of 45° horizontal. All other lines are discarded. These thresholds are obtained in several tests with different omnidirectional images containing doors of different position and size. The result of the segmentation is shown in figure 2e.

The upper horizontal part of the door case are expected to emerge in the central region of the omnidirectional image corresponding to higher elevation. In addition the segment is supposed to be of a minimal length. The horizontal candidate lines are selected if their length is between 5 % and 20 % of the image radius and the distance to the center is between 17 %



(a) Original image

(b) Morphological processing, edge detection and contrast stretching

(c) Edge linking

(d) Line approximation of (c)

(e) Vertical/Horizontal line segmentation of (d)

(f) Deletion of incorrect lines in (e)

(g) Merging collinear lines in (f)

(h) Double/Single lines segmentation of (g) and deletion of incorrect vertical lines

(i) Detected door frames in (h)

Figure 2: Overview of image processing steps in door detection.

and 28 % of the image radius. The results of the door frame line selection are shown in figure 2f.

Since the vertical door posts are substantially longer, it is fairly likely that the corresponding line segments are disconnected due to noise, non-uniform illumination and other effects. Therefore it is necessary to merge disconnected vertical lines prior to the detection of vertical door posts. Lines are merged if they are collinear and the gap between the segments is small compared to their overall length. The results of vertical line merging are shown in figure 2g.

Door frames often generate two parallel edges in the image corresponding to the inner and outer edge of the door case. Prior to door detection parallel, nearby edges of similar length are grouped into so called double vertical and horizontal lines. For the purpose of door pose estimation the double line is geometrically represented by the center of both lines.

The presence of double lines is a strong indicator for a door. Eventually all lines are either double vertical lines, single vertical lines, double horizontal lines and single horizontal lines as shown in figure 2h.

2.3 Door Frame Recognition

The final step in the door detection comprehends the matching between plausible combinations of vertical and horizontal lines with multiple potential door frame patterns. In general a door frame is described by two vertical and one horizontal line which endpoints coincide. In practice part of the door frame might be occluded such that two line configurations are also considered. Figure 3 shows the possible combinations of single and double lines that are matched with the lines detected in the image. These door patterns are inspired by the work of (Munoz-Salinas et al., 2004), which defines simple and double door frames. The patterns containing double lines are more meaningful and are therefore matched first. The matching proceeds from the most distinctive pattern a) to less discriminative structure j). In our final implementation only the patterns a) to h) are eventually associated with doors as the remaining patterns i) and j) are ambiguous and tend to produce too many false positives. The results after the door pattern matching are shown in figure 2i.

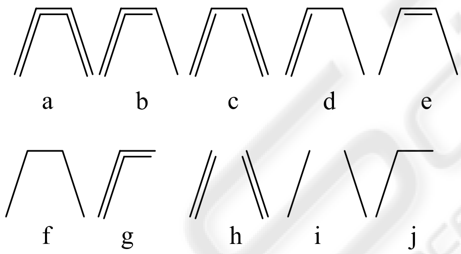


Figure 3: Possible door frame patterns sorted by priority.

3 DOOR LOCALISATION, TRACKING AND TRAVERSAL

3.1 Door Localisation and Tracking

Door localisation estimates the robots current pose (x, y, θ) with respect to the door. In case of monocular cameras the robot pose is usually recovered by triangulation of features from multiple captures taken from different viewpoints. In the literature this localization scheme is known as bearing only localization. The built-in odometer estimates the relative robot motion between consecutive viewpoints. Since both the measurement and the motion are subject to noise and errors, the robot position with respect to the door is estimated with an extended Kalman filter (EKF). The state prediction of the EKF relies on the odometry motion model, which describes the relative robot motion between two consecutive poses by three basic motions: an initial rotation δ_{rot1} followed by a straight motion δ_{trans} and a final rotation δ_{rot2}. The odometry model predicts the relative robot motion between consecutive states (Thrun et al., 2005):

$$\mathbf{x}_t^- = \begin{bmatrix} x_t^- \\ y_t^- \\ \theta_t^- \end{bmatrix} = \begin{bmatrix} x_{t-1}^+ \\ y_{t-1}^+ \\ \theta_{t-1}^+ \end{bmatrix} + \begin{bmatrix} \delta_{trans}\cos(\theta_{t-1} + \delta_{rot1}) \\ \delta_{trans}\sin(\theta_{t-1} + \delta_{rot1}) \\ \delta_{rot1} + \delta_{rot2} \end{bmatrix} \tag{3}$$

in which the superscript (-) denotes the a priori estimate of the process model and the superscript(+) indicates the a posteriori estimate after the correction step.

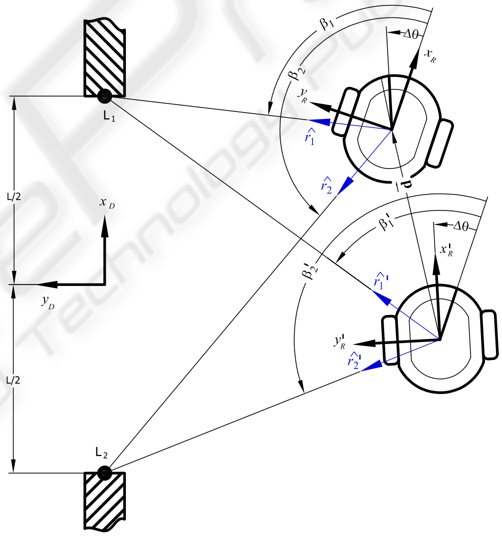


Figure 4: Robot position from bearing.

Figure 4 shows the door coordinate frame $< x_D, y_D >$ located at the center of the door and the two robot coordinate frame at two consecutive poses $< x_R', y_R' >$ and $< x_R, y_R >$.

The location of the door posts L_1, L_2 in robocentric coordinates are represented by

$$\mathbf{r}_{1,2} = k_{1,2} \begin{bmatrix} \cos(\beta_{1,2}) \\ \sin(\beta_{1,2}) \end{bmatrix} \tag{4}$$

The door post L_1, L_2 with respect to the current robot coordinate frame are recovered by triangulation from the two robot poses according to

$$\mathbf{r}_{1,2} = k_{1,2}\hat{\mathbf{r}}_{1,2} = k_{1,2}'\hat{\mathbf{r}}_{1,2}' - \mathbf{d} \tag{5}$$

in which **d** denotes the relative motion between both poses.

Solving equations 4 and 5 for $k_{1,2}$, the door pose with respect to the current robot coordinate frame is given by:

$$\begin{bmatrix} x_d \\ y_d \\ \theta_d \end{bmatrix} = \begin{bmatrix} (k_1 \cos(\beta_1) + k_2 \cos(\beta_2))/2 \\ (k_1 \sin(\beta_1) + k_2 \sin(\beta_2))/2 \\ \arctan\left(\frac{k_1 \sin(\beta_1) - k_2 \sin(\beta_2)}{k_1 \cos(\beta_1) - k_2 \sin(\beta 2)}\right) \end{bmatrix} \tag{6}$$

The robot pose with respect to the door frame predicted from the bearing angles is computed as:

$$\mathbf{z}_t = \begin{bmatrix} -x_d \cos(\theta_d) - y_d \sin(\theta_d) \\ x_d \sin(\theta_d) - y_d \cos(\theta_d) \\ -\theta_d \end{bmatrix} \tag{7}$$

In the correction step of the Kalman filter the posteriori state estimate is obtained by:

$$\mathbf{x}_t^+ = \mathbf{x}_t^- + K_t(\mathbf{z}_t - \mathbf{x}_t^-) \tag{8}$$

in which the Kalman gain K_t depends on the ratio of measurement and process covariance. The measurement error covariance is determined by a prior off-line analysis of door post triangulation accuracy to $\sigma_x^2 = 1$, $\sigma_y^2 = 1$ and $\sigma_\theta^2 = 0.5$. The Kalman filter is initialized based on the first two consecutive measurements of door post bearings.

3.2 Door Traversal

Typically the door is detected in the image for the first time at a separation between robot and door of about two to three meters. The door is tracked continuously by means of the Kalman filter while the robot continues its motion parallel to the corridor. The robot stops once it is located laterally with respect to the door center. At this instance it executes a 90 ° turn towards the door while continuously tracking its relative orientation.

Before initiating the traversal the open door state is verified from a single image and a sonar scan as a failsafe. The region between the door posts is analyzed in terms of its texture. A homogeneous texture indicates a closed door, whereas random texture implies an open door. The sonar scan is merely a failsafe confirmation of the visual classification, no sonar range data is needed for controlling the subsequent traversal. The robot traverses the door at constant velocity by centering itself with respect to the continuously tracked door posts. The visual servoing controls the robots turn rate such both door posts remain equilateral in the omnidirectional view. The Kalman filter is no longer applied as the depth information becomes unreliable at close range and is not needed for guiding the robot through the door.

4 EXPERIMENTAL RESULTS

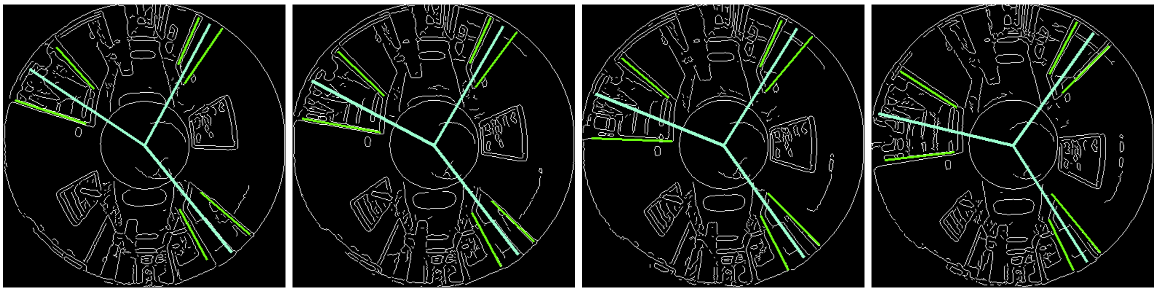
A typical scenario is depicted in Figure 5 consisting of door detection (A), door localisation and tracking (B) and door traversal (C). Similar traversal scenarios with different doors, illumination and directions of approach have been repeated successfully and consistently about hundred times. The door detection algorithm runs at a frame rate of 20 Hz on a standard 2.4 GHz processor and an image size of 388x388 pixels. False positives and false negatives in the door detection mainly occur due to noise, occlusion, illumination effects and confusion with door like objects such as shelves. To render the detection algorithm even more robust, the door frames are tracked over consecutive images during the motion. Initial false positives are eventually rejected in subsequent captures. This validation step is of particular importance for the Kalman filter localization.

The algorithm is tested on 1000 manually labeled images taken from video sequences captured in the office environment of our department. The enviroment is cluttered with objects such as pillars, cabinets or frames that could be confused with doors. The data set contains views taken from the corridor but also from the inside of offices. We assume that our vision based, topological localisation and navigation scheme guides the robot to the vicinity of the door. Thus there is need for our scheme to detect remote or occluded doors.

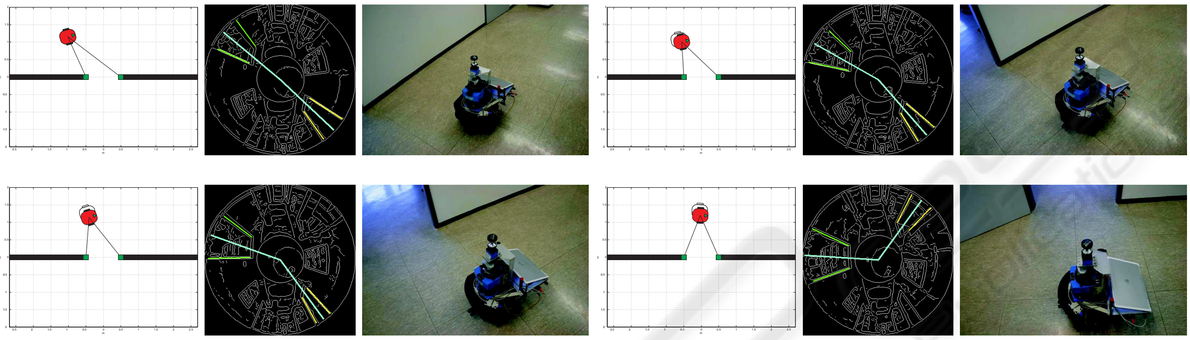
The initial door detection is shown in row (A). False positives in single images for doors amount to 3%, false negatives occur 5% of the time.

Row (B) of figure 5 depicts the state estimate and measurements of the Kalman in terms of robot pose relative to the door (left), the tracked door frames in the omnidirectional view (center) and an external view of the scene (right). During the tracking and approach phase the third door is occluded by a person standing in front of the door. The robot proceeds from its initial position about 1.5m along the corridor away from the door (top left) until is longitudinally aligned with the door center (bottom left) and completion of the 90 ° turn (bottom right).

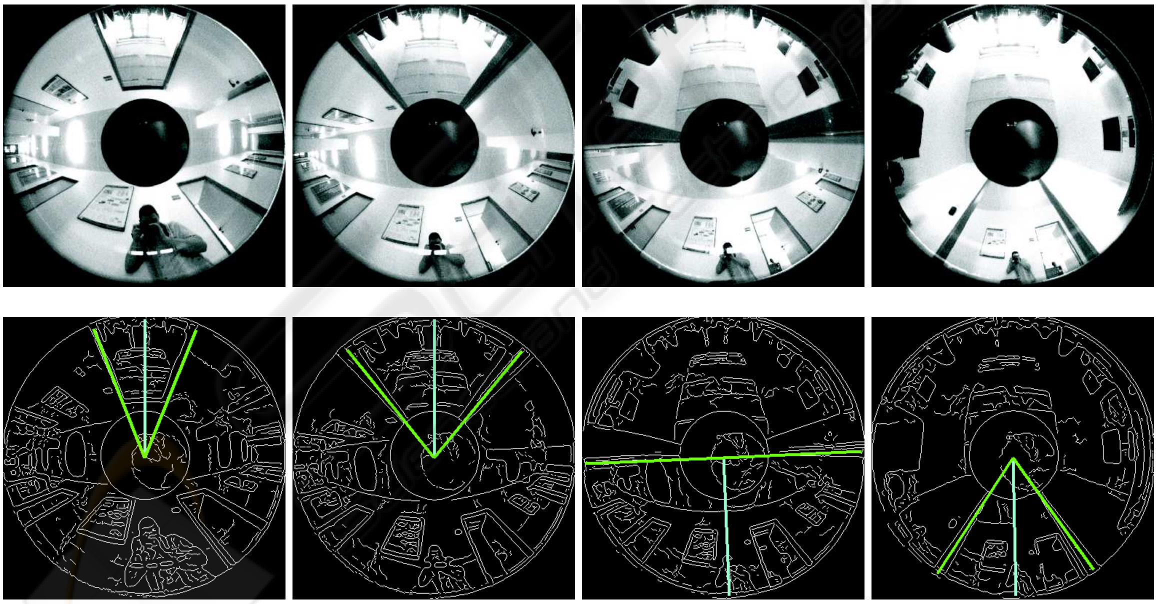
The door traversal stage is depicted in row (C) starting initially heading towards the door (left), approaching the door (center left), passing the door (center right) and after the traversal (right). Notice, that at all times, the detected door posts form an equilateral triangle with the image center indicating accurate alignment of the robot with the door.



A) Vision based door detection



B) Vision based door location and tracking



C) Visual servoing for door traversal

Figure 5: Experimental evaluation of the visual door passing behavior.

5 CONCLUSIONS

This paper introduced a novel framework for vision based door traversal of a mobile robot. The door detection, door tracking and door traversal rely on omnidirectional images only. The door detection is based on the matching of detected line segments with prototypical door patterns. The door localisation and tracking aggregates the visual measurements with the robots odometry information in a Kalman filter. The door traversal follows a 2D visual servoing approach with extracted door posts as image features. The practical usefulness and robustness of the approach are confirmed in several experiments in an office environment. The proposed scheme contributes to our overall objective of achieving purely vision based robot navigation and localisation in indoor environments. It has been successfully integrated and tested with other visual behaviors such as goal point reaching and obstacle avoidance.

REFERENCES

Baker, S. and Nayar, S. (1998). A Theory of Catadioptric Image Formation. In *IEEE Int. Conference on Computer Vision (ICCV)*, pages 35–42.

Budenske, J. and Gini, M. (1994). Why is it so difficult for a robot to pass through a doorway using ultrasonic sensors? In *IEEE Int. Conference on Robotics and Automation*, volume 4, pages 3124 – 3129.

Canny, J. (1986). A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8(6):679–698.

DeSouza, G. and Kak, A. (2002). Vision for mobile robot navigation: A survey. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(2):237–267.

Duda, R. O. and Hart, P. E. (1972). Use of the hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15(1):11–15.

Eberst, C., Andersson, M., and Christensen, H. (2000). Vision-based door-traversal for autonomous mobile robots. In *IEEE/RSJ International Conference on Intelligent Robots and Systems,IROS 2000*, volume 1, pages 620 – 625.

Geyer, C. and Daniilidis, K. (2001). Catadioptric projective geometry. *Int. J. of Computer Vision*, 4(3):223–243.

Jensfelt, P. (2001). *Approaches to Mobile Robot Localization in Indoor Environments*. PhD thesis, KTH.

Kovesi, P. D. (2000). MATLAB and Octave functions for computer vision and image processing. The University of Western Australia. Available from: <http://www.csse.uwa.edu.au/∼pk/research/matlabfns/>.

Monasterio, I., Lazkano, E., Rano, I., and Sierra, B. (2002). Learning to traverse doors using visual information. *Trans. of Mathematics and Computers in Simulation*, vol.60.

Munoz-Salinas, R., Aguirre, E., M., G.-S., and Gonzlez, A. (2004). Door-detection using computer vision and fuzzy logic. *WSEAS Transactions on Systems*, 10(3):3047–3052.

Murillo, A. C., Kosecka, J., Guerrero, J. J., and Sagues, C. (2008). Visual door detection integrating appearance and shape cues. *Robotics and Autonomous Systems*, 56(6):512–521.

Patel, S., Jung, S. H., Ostrowski, J. P., Rao, R., and Taylor, C. J. (2002). Sensor based door navigation for a nonholonomic vehicle. In *IEEE Int. Conference on Robotics and Automation, Washington,DC*, pages 3081–3086.

Shi, W. and Samarabandu, J. (2006). Investigating the performance of corridor and door detection algorithms in different environments. In *Int.Conference on Information and Automation, ICIA 2006*, pages 206–211.

Stoeter, S., Le Mauff, F., and Papanikolopoulos, N. (2000). Real-time door detection in cluttered environments. In *Proc. IEEE Int. Symposium on Intelligent Control*, pages 187 – 192.

Thrun, S., Burgard, W., and Fox, D. (2005). *Probabilistic Robotics (Intelligent Robotics and Autonomous Agents)*. The MIT Press.