Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems

Vision-Based Door-Traversal for Autonomous Mobile Robots

C. Eberst
Inst. for Real-Time Computer-Systems
Technische Universität München
D-80333 Munich, Germany
Christof.Eberst@rcs.ei.tum.de

M. Andersson and H. I. Christensen
Centre for Autonomous Systems
Royal Institute of Technology
S-10044 Stockholm, Sweden
{sungam,hic}@nada.kth.se

Abstract

Robust and safe traversal of doors is an important task for autonomous mobile robots. This paper presents a robust and inexpensive approach for recognizing and localizing doors based on monocular grey-level images and minimalistic models. The traversal is achieved by servoing the robot with respect to the door-hypothesis. Robustness against pose errors, scene complexity, and sensing conditions is obtained by combining indexing on significant aggregated features with a multi-view approach that employs consistency over time: pose-hypotheses are filtered via the motion of the robot to reject incorrect ones. The traversal is implemented as a behavior and was evaluated in 150 experiments at different doors under different conditions.

Keywords: robot vision, object recognition, pose estimation, indexing, mobile robots, servoing, behavior-based robotics

1 Introduction

Vision is a powerful modality for recognition of objects and for estimation of object parameters such as pose and size. In a robotics context a number of other sensory modalities like ultra-sound and laser ranging are also available. Most other sensory modalities are slower, provide less information, or are only suitable for use in the context of active sensing strategies. Vision on the other hand provides a tremendous amount of information, which also poses a problem. The imaging process combined with the richness of the potential set of scenes implies that it is often difficult to interpret images and obtain robust estimates. One possible solution to this problem is to use model based information for filtering and interpretation of data. Robot navigation in an in-door environment is a typical example of a scenario where a rich set of information is available. Through introduction of model based information it is possible to transform the stream of video data into a well posed estimation problem through careful use of model based information.

Use of model based information is well known in computer vision. Fennema et al. [8] determined the pose of a platform with respect to the building by projecting CAD models into the image for matching against image line-segments. A similar approach was reported by Christensen et al. [4] for navigation in large scale structures. Use of complete CAD models of buildings requires extensive and accurate modeling of complete structures, which is demanding and often challenging. In terms of recognition and tracking of specific objects model based matching has also been extensively used, as initially suggested by Lowe [13], and later exploited as illustrated in [10, 11, 16].

For operation in a general environment it may not be possible to have explicit metric models for the objects of interest. One example is detection and traversal of doors. For such objects it is possible to provide a constrained model that limits search, yet it is difficult to provide a full metric model. For doors it is possible to specify an approximate height (i.e. height in range 2.0-2.5m) and it is possible to specify a bound on the width (i.e. 0.8-1.2m). Doors can be characterized as a rectangular structure that is up-right. For detailed modeling one can even include articulation in terms of the rotating door-wing [9].

In this paper the problem of using a "simple" polyhedral model of doors for detection and servoing through doors is studied. The problem is divided into several sub-processes: i) detection of potential door structures through line based filtering, ii) verification of hypotheses, and iii) active tracking of a structure over time to facilitate traversal of the door. Through careful analysis of each step it is possible to provide a robust functionality that accommodates navigation in a regular house in the presence of varying illumination, occluding objects, and varying parameters for the doors.

The main contribution of this paper is the demonstration of a simple solution to reliable door-traversal by the integration of simple techniques with the consistency-over-time constraint in a mobile robotic system. Initially the overall system is outlined (Sec. 2), and the methods for detection, verification and filtering are described in detail (Sec. 3). Sec. 4 describes a series of experiments that illustrate the robustness of the presented approach. Finally, conclusions are presented in Sec. 5.

2 The System

2.1 The Robot System

For navigation in an in-door environment a number of techniques using ultra-sonic and laser-ranging techniques are available [3, 5, 7, 12, 15]. These techniques are well suited for navigation in "simple" structured environments. The methods allow easy navigation in hallways and relative open spaces in rooms. In this work the vision based navigation is studied in the context of such a system [1]. The system consists of a Nomad 200 robot equipped with 16 ultra-sonic sonars, a SICK laser scanner and a binocular set of cameras (Fig. 1). The system uses a behavior based approach to control [2]. In this approach different control modules acquire sensory information and generate control hypotheses in terms of a histogram of desired movement commands. The control commands are then fused using a simple voting approach [14].

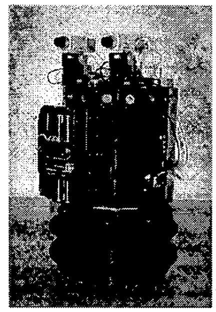

Figure 1: The system.

2.2 The Task

Basic functionalities are available for general navigation in open spaces. For detection and traversal of doorways simple 2D sensing, as provided by sonar and laser-ranging is, however, non-optimal. Vision enables robust detection and tracking of complex structures such as doors. It is here assumed that doors have a passage width of 80-90 cm and a height of 220cm. The robot has a diameter of 65cm. Thus, to ensure a safe passage through doorways, the visual servoing must provide estimates with an error of less than 6cm. Initial estimates of the position of the robot might be up to 40cm off and the orientation might be 15 degrees off (worst case condition). The task for the visual door traversal (VDT) module is to detect the doorway and to provide a continuous estimate of the door position that allows traversal. A problem in this context is that the appearance of the doorway changes significantly as the object is approached. The module must be able to cope with multiple hypotheses and changing appearance of the door. A doorway consists of the door post, and indentation for the door etc., which implies that the geometric structure of the doorway changes with scale. The module must cope with these ambiguities.

2.3 The Conditions

The doors differ in their precise shape and appearance. Closely posed objects, massive walls, persons, and the door-wing itself cause unknown occlusions. The background and the foreground are densely structured and the illumination condition varies strongly, due to windows. The doors show only a weak contrast to the embedding walls (Fig. 2,3,5).

The camera system and the pan-tilt head are not calibrated: Thus, errors in the external parameters, in the internal parameters (e.g. the focal length (6mm), the pixel-size, and the focal-point), and imperfections such as the radial distortion are not compensated. The system is designed to tolerate these errors online, allowing for long-term operation with temporarily changing deterioration of the sensor system.

Other behaviors run in parallel with the door traversal and may pose differing or even conflicting requirements on the robots/sensors pose and orientation. Therefore, the VDT has no complete control and must cope with a suboptimal path and sensor views.

2.4 The Approach

Recognition and localization of the door employ indexing: Extracted, parallel line segments are evaluated according to an indexing-table. This table is built up online for each sensor reading according to the 2D line segments which are predicted for the most current pose estimation. The pose of the object in the image and in 3D-coordinates is directly derived from the entries in the tables, without the interim step of assigning extracted features to predicted ones (see Sec. 3.4).

The approach must tolerate large pose-errors of the robot. The employed recognition, tuned to simplicity, achieves robustness against incorrectly derived pose-hypotheses by filtering these via the motion of the robot: i.e. consistency over time is imperative to the hypotheses.

Robustness of the overall system performance is also achieved by allowing other, complementing behaviors, such as obstacle avoidance or ultra-sonic-based door-traversal, run in parallel with the door traverser behavior.

3 Vision-Based Door-Traversal

3.1 Overview

The door-traversal behavior (VDT) is implemented as a simple state-machine with 5 states.

1. *Initialization:* The VDT takes control over the turret and slows down the robot. The global pose estimate is once requested from the localizer module and the local coordinates from the platform. Since the global pose estimate is permanently corrected by the localizer, all succeeding operations are based on the local coordinates only: Subsequent updates of the pose estimates plus related motion correction by the VDT could lead to collisions with the walls.
2. *Initial search:* The turret is controlled to direct the camera to the initial door hypothesis, according to the current pose estimate of the robot. The images are grabbed and line segments are extracted. Next, a synthetic view of the relevant door is predicted (Sec. 3.2) based on inexpensive wire frame models with visibility ranges [6].

   An indexing table is build up and used for matching and direct pose estimation (see Sec. 3.3). The indexing-keys are predicted/extracted parallel lines in image coordinates, the indexing-values are the distances between these two lines. After each newly integrated pose estimate (see Sec. 3.4), the camera is controlled to focus the center of the object. To cope with high pose errors even close to the door, three separate initial searches for the door are performed by the robot, with the camera directed

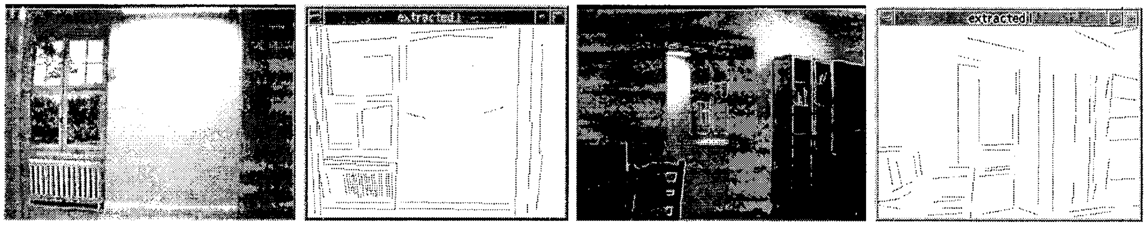

Figure 2: Typical input at short-range (*left*) and long-range (*right*) servoing.

straight, clockwise (cw) and counterclockwise (ccw). The orientation offset (cw and ccw) increases for decreasing distance to the door. The pose estimations from these three searches are fused afterwards.

3. *Servo to traverse point:* According to the permanently updated pose estimate, the robot drives to a point in front of the door that permits for a final pose correction and a safe traversal. During that drive, the VDT permanently directs the camera to the door and uses vision to correct its pose estimates.
4. *Final pose correction:* At the traverse point, servoing terminates and the robot performs a last estimation of its pose and correction of its motion. Since the position uncertainty of the robot is strongly reduced at this point, no cw and ccw scans are required. Furthermore, the pose of the robot aligned to the door's passage assures minimum occlusion of the door-frame.
5. *Blind door traversal:* On the way through the door, the robot can no longer perceive the features of the door frame. It drives blindly. The motion of the robot and the turret are controlled to avoid collisions of the platform or the forklift with the walls.

The next subsections stress the different steps which are employed by the actual behavior: *Scene prediction* (see Sec. 3.2), *matching*, including filtering and indexing (see Section 3.3), *pose estimation* (Sec. 3.4), *fusion of hypotheses*, enforcing consistency over time (see Sec. 3.5), *servoing* (see Sec. 3.6), and the *blind traversal* (Sec. 3.7).

3.2 Scene Prediction

The prediction of the virtual sensor-view is based on the perspective projection of a wire-frame model of the door-frame and the latest pose estimate. The perceptive field of the camera is modeled as a sensor-cone. It is widened the closer the robot is to the object, to compensate for the error of the pose estimates. Self occlusion is taken into account by applying the visibility ranges [6] of each single feature. Imprecisions of this very inexpensive method, the prediction of partially occluded features in their full length, can be neglected since the length is not restrictively applied for matching.

3.3 Matching

Matching must combine a sufficient selectivity with a high robustness against pose errors of the robot and inaccuracies of the map.

Reduction of the complexity: For reducing possible match-candidates, non-fitting extracted line segments are rejected prior to the match. Since large initial errors in the position and orientation of the robot with respect to the door must be tolerated, rejection is based on the less effected criteria described next.

- *Orientation-based preselection:* The orientation of each extracted line segment is compared to the orientation distribution of the predicted line segments, stored in a table. If no equivalent entry exists, the extracted line segment is rejected.
- *Preselection according to region of interest:* The displacement of extracted line segments to the image area, covered by the prediction is used for filtering. Since pose errors of the robot have a strong impact on this aspect, only line segments which are significantly displaced are rejected.
- *Length filtering:* Extracted line segments, shorter than a adaptive minimum are rejected.

Notation: The following abbreviations are used next. In general, all values related to image coordinates are presented with small letters, e.g. (x,y), while capital letters, e.g. (X,Y) refer to world-coordinates. W/w determine the width in 3D/image-coordinates, of: a key/hypothesis (indices key/hyp), or the object/robot/sensor ($Obj/R/S$). The indices $pred/extr$ indicate whether a predicted or extracted value is meant, f, s_p refer to the focal length, the pixel-size. FP_x is focal point (in x).

Indexing: All predicted line segments are tested for being close to parallel, overlapping, and for possessing a minimum slope. In experiments, it turned out to be sufficient to focus on "non-horizontal" line segments, originating in the door frame. Since the camera perceives only a fraction of the door frame when close to the door, this restriction on the vertical features of the door frame allows a uniform approach for wide-range and short-range localization without scanning vertically (tilt). For each pair of these features, an entry in the indexing cell is established. The cell in the indexing table is determined according to the displacement of the two features w_{key}^{pred} in image coordinates (see Fig. 4).

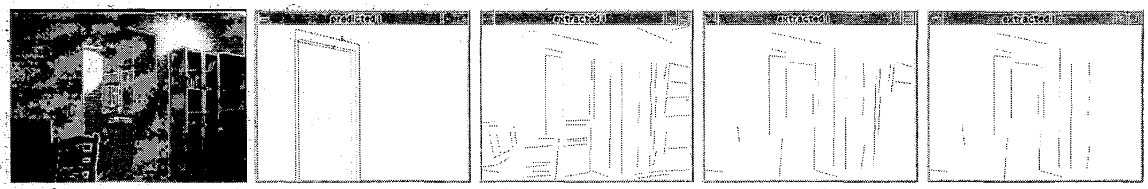

Figure 3: Visualization of the preselection. From left to right: input image, original prediction, extracted line segments, orientation-filtered line segments, pose filtered line segments.

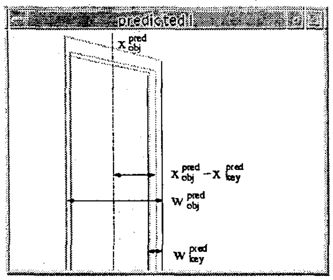

Figure 4: Values of the indexing table for a pair of line segments (bold) derived from the prediction.

For each pair of line-segments, the value of the actual indexing-key w_{key}^{pred} is stored in the indexing table together with their center x_{key}^{pred}, and their distance to the center of the predicted object $(x_{obj}^{pred} - x_{key}^{pred})$. Furthermore, the ratio of the entry's width (w_{key}^{pred}) to the width of the predicted object (w_{obj}^{pred}) is stored. These stored properties allow a rapid retrieval of an estimation of the width and center of the object in image coordinates in the following. Thus, no explicit reference on the predicted line-segments is required.

For each pair of extracted and filtered parallel and overlapping line segments the horizontal displacement of their $(w_{key}^{extr} = 0.5 * |(x_{0_0} + x_{0_1}) - (x_{1_0} + x_{1_1})|)$ and the pose of their center x_{key}^{extr} are calculated. For comparison with the line segment, stored in the indexing table, the estimated minimum and maximum disparity of a corresponding line segment pair are calculated according to $Thr(w_{key}^{pred}) = w_{key}^{extr} \pm \Delta_w$, with Δ_w from equ. 4. c_n reflects noisy line segmentation, ϵ_ϕ,ϵ_{par} are the tolerated error of the orientation and of the pose parallel to the door's passage.

$$\phi = \phi_{Obj} - atan(|\frac{Y_{Obj} - Y_S}{X_{Obj} - X_S}|) \quad (1)$$

$$D_{par} = \sqrt{(Y_{Obj} - Y_S)^2 + (Y_{Obj} - Y_S)^2} * cos(\phi) \quad (2)$$

$$D_{orth} = \sqrt{(Y_{Obj} - Y_S)^2 + (Y_{Obj} - Y_S)^2} * sin(\phi) \quad (3)$$

$$\Delta_w = c_n + w_{key}^{pred} * (\frac{D_{par}}{D_{par} - \epsilon_{par}} - 1 + max(|\frac{sin(\phi + \epsilon_\phi)}{sin(\phi)} - 1|, |\frac{sin(\phi - \epsilon_\phi)}{sin(\phi)} - 1|)) \quad (4)$$

The cells of the indexing table are selected for inspection according to $cell_{min/max} = (int)(n_{cells}(w_{key}^{extr} \pm \Delta_w)/w_{Obj}^{pred})$. To improve the selectivity, all references stored in the indexed cells are tested for being within these boundaries subsequently. To suppress assignments between opposite sides of the door-frames, the remaining references are rejected if the deviation of their centers $x_{key}^{pred} - x_{key}^{extr}$ exceeds 80% of the predicted width of the object w_{Obj}^{pred}.

Reliability of a match: The trustworthiness P of an assignment of a pair of extracted line segments to a pair of predicted line segments, stored in the indexing table, is calculated according to the deviation of their distance and the displacement of their center, both in image coordinates. c_1, c_2, $(0 < c_i < 1.0)$ reflect the significance of the two deviations.

$$P = (1.0 - c_1 * \frac{|w_{key}^{extr} - w_{key}^{pred}|}{\Delta_w}) * (1 - c_2 * \frac{|x_{key}^{extr} - x_{key}^{pred}|}{w_{Obj}^{pred}}); \quad (5)$$

Estimated accuracy of the 2D pose cues: An estimate of the accuracy A of the derived center and width of the object in image coordinates is shown in equ. 6 c_3, $(0 < c_3 < 1.0)$ reflect the significance of the deviation of the reconstructed width in image coordinates from the predicted one.

$$A = \frac{w_{key}^{pred}}{w_{Obj}^{pred}} * (1.0 - c_3 * \frac{|w_{key}^{extr} - w_{key}^{pred}|}{\Delta_w}) \quad (6)$$

3.4 Pose Estimation

Instead of calculating the 3D-pose of the robot with respect to the door for each matching pair of predicted/extracted line segments, the pose x_{hyp} and size w_{hyp} of the object-hypotheses in image coordinates are fused for all consistent hypotheses prior to the 3D pose estimation. For robust pose-estimation, we sacrify a precise determination of the robots orientation ϕ: We freeze the initial orientation estimate of the localizer. Thus, orientation errors of the robot lead to a rotation of the pose estimate of the robot around the center of the door and an imprecise depth estimation. However, this error does not impair directing the camera and does only weakly affect/falsify the orientation in which the robot drives. What we obtain is:

1. a fusion of the pose estimates in 2D (X,Y);
2. simplified equations;
3. a higher robustness: Since the orientation estimation is directly influencing the estimated 3D distance from the robot to the door, small orientation variations of different pose hypotheses would lead to significantly varying 3D poses. This makes fusion more difficult. For a "frozen" orientation, all estimates that originate in a correct assignment are similarly (precisely or imprecisely) located. It must be noted that a robust orientation estimation is rather unlikely since (a) the used model is imprecise for the non uniform doors, (b) the sensor is not calibrated, and (c) the matching is restricted to the vertical features on the door-frame.

Pose in image coordinates: The center and width of a hypothesis are calculated for each pair of predicted/extracted line segments according to equ.7 and 8.

$$w_{hyp} = \frac{w_{key}^{extr}}{w_{key}^{pred}} * w_{Obj}^{pred} \quad (7)$$

$$x_{hyp} = x_{key}^{extr} + \frac{w_{key}^{extr}}{w_{key}^{pred}} * (x_{Obj}^{pred} - x_{key}^{pred}) \quad (8)$$

Hypotheses with similar estimated center x_{hyp} and width w_{hyp} are merged according to a weighted average into the pose x_{fin} and width $w_{(fin)}$. Fused estimations are rejected if their reliability $P(fin)$ is below a certain threshold (with $P(fin) = \sum_{i=1}^{n} P(i)$ for n fused hypotheses). Next, fused hypotheses are rejected if their absolute probability $P(fin)$, or relative probability $P(fin)/max(P(fin))$ is below a certain threshold. $max(P(fin))$ is the maximum probability of all hypotheses after fusion.

Pose in world coordinates: From the fused hypotheses concerning the object-pose in image coordinates, the 3D pose of the sensor with respect to the door is calculated according to the equ. 9 to 10. D_{par} and D_{orth} are the distance parallel and orthogonal to the door's passage. ϕ_S is static.

$$D_{par} = W_{Obj} * (cos(\phi_{Obj} - \phi_S) * f/(w_{fin} * s_p) \quad (9)$$

$$D_{orth} = -D_{par} * ((x_{fin} - FP_x) * s_p/f) \quad (10)$$

The pose of the robot (X_R, Y_R) is obtained by subsequent transformations into the world-coordinate system. It must be noted, that this approach is optimal if mainly one aspect of the object is seen, or (rather) flat objects, such as doors and windows are perceived. Otherwise, the fusion in image coordinates and the subsequent 3D pose calculation should be done separately for different aspects (planes) of the object or for each assignment.

3.5 Fusion and Rejection

All priorly obtained pose estimations are shifted according to the last local motion of the robot, obtained by dead-reckoning. The newly derived pose estimates are fused with old ones according to their weighted average. The usage of old hypotheses dampens the influence of outliers in the pose estimates.

All hypotheses are exposed to a probability-reducing process, which lowers the influence of older measurements, that become more imprecise in the presence of rotational errors. By reducing the probability of non-fused hypotheses stronger than for fused ones, the consistency over time constraint (COT) is enforced.

COT is a very powerful instrument for rejecting incorrect hypotheses caused by false assignments of neighboring structures, such as shelfs (see Fig. 2), or background structures, which show similar vertical image line segments, but differ in their 3D geometry stronger than the individual doors differ from each other. When the robot is moving, hypotheses, resulting from these structures do not evolve at a stable pose. In combination with the usage of aggregated features (here: parallel line segments) for matching, COT enforces robustness and recovery from initial incorrect pose estimates.

3.6 Servoing

The camera is permanently directed to the most current estimated center of the door with respect to the estimated pose of the robot. The wheels are directed to the next goal point in front of the door, or behind the door. The speed of the platform is controlled according to the current state of the traversal and the distance from the robot to the next goal point.

3.7 Final (Blind) Door Traversal

The final drive of the robot must be performed with the turret facing the rear point, behind the door. No rotation is allowed to prevent damages of the forklift of the robot. The actual traversal of the door is therefore performed without the vision system, based on dead-reckoning only. A collision-free door traversal is assured if the VDT is applied in combination with us-based avoidance/traversal behaviors (see Sec. 4).

4 Results

The VDT has been evaluated in 150 experiments at different doors, varying day-time and illumination conditions. Starting points varied for long range and short range traversal, between 1.3m and 4.5m distance to the door. The robustness of the VDT was evaluated with initial pose estimations artificially deteriorated up to 15° and 25% of the distance to the door, exceeding the error of the localizer module by 300% to 1000%.

In addition to common occlusions of the door by the embedding wall, or objects standing between the robot and the door (Fig. 5), the door was hidden temporarily by persons walking through or by in 30 experiments. The door was partially (one frame-side), but permanently artificially occluded in further 20 experiments. The strong correction at Fig. 5, middle column, was caused by the domination of the new hypotheses (with non-occluded door) over the older ones, which has been established with the left door-frame occluded.

Success rate: The VDT failed to execute a safe passage in only one experiment (0.667%), leading to a contact with the bumper of 5mm. In five experiments, the calculated path was closer to the door-frame than an ad-hoc defined safety distance of 6cm (in worst case 3cm).

In 85% of the test-drives, the robot did traverse the doors with a deviation from the center smaller than 5cm. It must be noted, that during the final drive of approximately 1.3m, including corrections of the robots path to compensate for deviations, pose estimations are performed by dead-reckoning only, without vision based pose correction or updates from the global localizer.

Integrated Door-Traversal: The behavior was tested in 30 experiments in combination with a sonar-based door-traverser. The speed and direction-commands where fused by the vehicle controller. Distant to the door, the fused behavior was basically identically to the behavior without US-based door-traverser, besides the capability of avoiding persons and objects blocking the doorway. The fused behavior close to the door was stronger influenced by the US-based door traverser, thereby keeping the robot closer to the center of the free passage, which is not necessarily the center of the door: The door wing shrinks the passable way, if not completely open, as well as objects and persons in or close to the passage.

Processing Time: The VDT, including line segmentation, runs at 5Hz on the on-board Pentium III-450MHz. The frequency increases if the pose hypothesis becomes reliable such that the processed image-region shrinks. To reduce processing time, the slow line-extraction shall be replaced by a rapid one. The VDT is called with 2Hz by the controller.

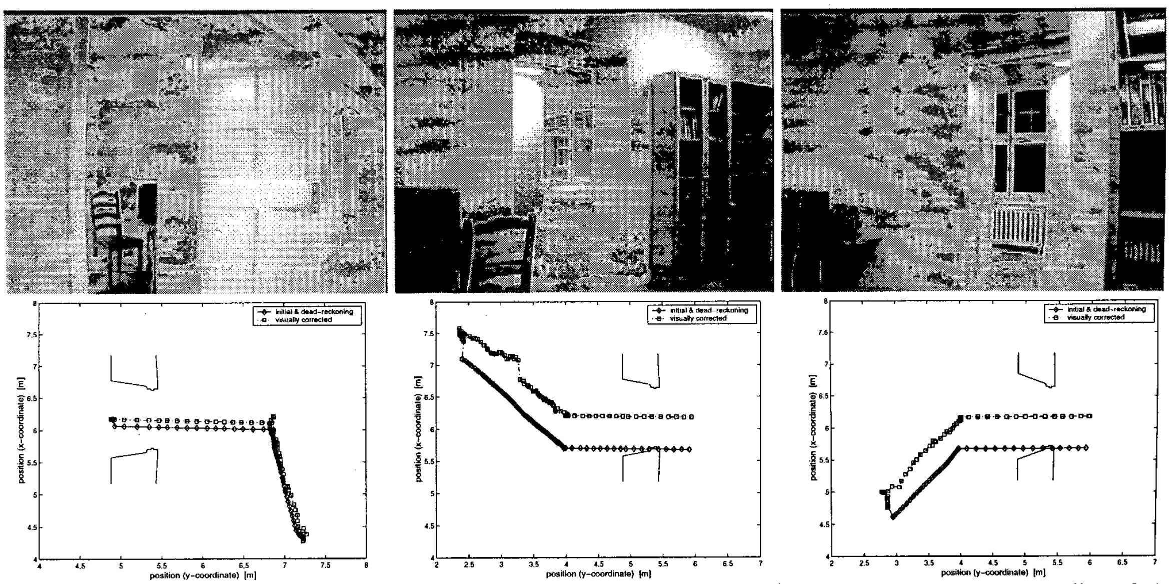

Figure 5: Traversal of the door in the living-room from different starting points (left drive: starting right, middle and right drive: starting left). Top row: Initial scene. Lower row: Plot of the robot's X-Y position (*squares*) with VDT. *Diamonds*: Equivalent pose estimations based on the initial estimation and dead-reckoning. The initial pose estimate was artificially deteriorated for these experiments. The path of the robot was less than 3cm from the center of the passage. The traversal without the pose correction of the video-based door-traverser would have led to collisions with the wall for all cases.

Therefore, more than 600 ms remain per system-cycle (1Hz) for the processing the sonar-based door-traverser, the localizer, and the controller.

5 Conclusion

In this paper we have presented a highly robust vision-based door-traverser behavior (VDT) inexpensive in implementation and computational costs. The recognition and localization of the door is performed via indexing on parallel line segments, without explicit pose estimation per assignment. The approach has been evaluated in 150 experiments and has proven its robustness, separately and in combination with other behaviors.

Robustness is achieved by requiring consistency of pose estimates over time and motion of the robot: incorrect assignments (outliers) are compensated. Recognition and localization as a vision problem is therefore facilitated. The integration into the behavior-based ISR-control system allows for fast application and robust performance.

Acknowledgment

The presented work was supported by the *Swedish Foundation for Strategic Research* and by the *Deutsche Forschungsgemeinschaft* as a part of the research projects on "Telepresence and Teleaction" SFB453 and "Video-based Exploration of Indoor Environments" Fal09/14-1. Acknowledgment to D. Tell, A. Orebäck, P. Jensfelt, O. Wijk, and M. Lindström for their support concerning the sensor-system and the robot.

References

[1] M. Andersson, A. Öreback, M. Lindström, and H.I. Christensen. *Intelligent Sensor Based Robots*, volume 1724 of *Lecture Notes in Artificial Intelligence*, chapter ISR: An Intelligent Service Robot, pp. 291–314. Springer Verlag, Oct. 1999.

[2] R. C. Arkin. *Behaviour Based Robotics*. Intelligent Robots and Autonomous Agents. MIT Press, Boston, MA, 1998.

[3] B. Burks, G. de Saussure, C. Weisbin, J. Jones, and W. Hamel. Autonomous navigation, exploration, and recognition using the HERMIES-IIB robot. *IEEE Expert*, 2(4), Winter 1987.

[4] H. I. Christensen, N.O.S. Kirkeby, S. Kristensen, L. F. Knudsen, and E. Granum. Model-driven vision for in-door navigation. *Robotics and Autonomous Systems*, 12:199–207, 1994.

[5] J. L. Crowley. Navigation for an intelligent mobile robot. *IEEE Trans. on Robotics and Automation*, 1(1):31–41, March 1985.

[6] C. Eberst and J. Sicheneder. Generation of Hypothetical Landmarks Supporting Fast Object Recognition with Autonomous Mobile Robots. In *Proc. IROS'96*, pp. 813–820, Nov. 1996.

[7] A. Elfes. Sonar-based real-world mapping and navigation. RA-3(3):249–265, June 1987.

[8] C. Fennema, A. Hanson, E. Riseman, J.R. Beveridge, and R. Kumar. Model-directed mobile robot navigation. Technical Report COINS TR 90-42, University of Massachusetts, Computer and Information Science, June 1990.

[9] A. Hauck and N. O. Stöffler. A hierarchic world model supporting video-based localisation, exploration and object identification. In *2. ACCV, Singapore, Dec.*, pp. 176–180, 1995.

[10] D. Koller, K. Daniilidis, and H. H. Nagel. Model-based object tracking in monocular image sequences of road traffic scenes. *Int. J. of Computer Vision*, 10(3):257–281, 1993.

[11] H. Kollnig and H.H. Nagel. 3d pose estimation by directly matching polyhedral models to gray value gradients. *I. J. of Computer Vision*, 23(3):282–302, 1997.

[12] J. J. Leonard and H. F. Durrant-Whyte. *Directed Sonar Sensing for Mobile Robot Navigation*. Boston, 1992.

[13] D. Lowe. *Perceptual Organisation and Visual Recognition*. Robotics: Vision, Manipulation and Sensors. Kluwer Academic Publishers, Dordrecht, NL, 1985. ISBN 0-89838-172-X.

[14] P. Pirjanian and H. I Christensen. Behaviour coordination using multiple objective decision making. In *Proc. Conf. on Int. Systems and Advanced Manufacturing*, Pittsburgh, PA, USA, Oct 1997. SPIE.

[15] W.D. Rencken. Autonomous sonar navigationin indoor, unknown and unstructured environments. In *IEEE/RSJ IROS'94*, pp. 127–134, Munich, Sept. 1994.

[16] T. N. Tan, G. D. Sullivan, and K. D. Baker. Model based localisation and recognition of road vehicles. *Int. J. of Computer Vision*, 27(1):5–25, 1998.