DOOR HANDLE IDENTIFICATION: A THREE-STAGE APPROACH

E. Jauregi, J. M. Martínez-Otzeta, B. Sierra and E. Lazkano [1]

Robotics and Autonomous Systems Group.
http://www.sc.ehu.es/ccwrobot
Computer Science and Artificial Intelligence Department
University of the Basque Country

Abstract: The aim of the work presented here is to develop a door identification subsystem based on door handle recognition. The problem is stated as deciding whether a door handle is present or not in the images taken by the robot while navigating. And it is solved from a supervised classification point of view, combining Machine Learning segmentation with the Hough transform and statistical measures.

Keywords: Machine Learning, Autonomous Mobile Robots, Robot Vision

1. INTRODUCTION

Indoor semi-structured environments are full of corridors that connect different offices and laboratories. It is often necessary to cross the doors to enter a room. Many navigation tasks can be fulfilled by point to point navigation, door identification and door crossing (Li *et al.*, 2004), and endowing the robot with the door identification ability would undoubtedly increase the navigating capabilities of the robot.

Several references can be found that tackle the problem of door identification. Most of them try to identify the door by means of line identification. In (Stoeter and Papanikolopoulos, 2000) a ranger robot is used to launch smaller robots in environmental conditions that can be dangerous for humans. Vision is used to identify doors in corridors and door state (open or close) is evaluated using sonar sensors.

In (Muñoz-Salinas *et al.*, 2005) the authors present a visual door detection system that is based on the Canny edge detector and Hough transform to extract line segments from images. Then, features of those segments are used by a genetically tuned fuzzy system that analyzes the existence of a door.

In (Eberset *et al.*, 2000) a vision-based system for detection and traversal of doors is presented. Door structures are extracted from images using a parallel line based filtering method, and an active tracking of detected door line segments is used to drive the robot through the door.

Another proposal of vision-based door traversing behavior can be found in (Seo *et al.*, 2005). There, the PCA (Principal Component Analysis) pattern finding method is applied to the images obtained from the camera for door recognition.

Stella et al. developed SAURO, a robotic autonomous system oriented to transportation tasks in indoor environments (Stella *et al.*, 1996). SAURO uses a vision-based self-localisation subsystem that allows the robot to determine its

[1] This work has been supported by the Basque Country Government under ETORTEK grant AMIGUNE, and by the MCYT TSI200500390 project.

location to verify that the planned path is correctly followed. It uses an ultrasonic based obstacle avoidance subsystem to detect unknown obstacles along the path and its endowed subsystem uses artificial beacons placed on the doors. If, according to the planned path the robot must cross a door, the door crossing subsystem is activated, and the vision subsystem is used to align the vehicle orientation with the axis of the door. Once the robot is aligned, a straight movement permits the vehicle cross the door safely.

A door identification and crossing approach was also presented in (Monasterio *et al.*, 2002); a neural network based classification method was used for both, the recognition and navigation through the door panels. More recently, in (Lazkano *et al.*, 2006) a Bayesian Network based classifier is used to perform the door crossing task. Doors are assumed to be open, the aperture is identified and the door crossed using sonar sensor information.

The proposal in (Kragic *et al.*, 2002) differs in the sense that doors are located in a map and do not need to be recognized, but rectangular handles are searched for manipulation purposes. The handles are identified using cue integration by consensus: for each pixel the probability of being part of a handle is calculated by combining the gradient and the intensity of every pixel in order to have the degree of membership. The template model is used to obtain the consensus over a region.

Similarly, we propose a method for identifying door handles. But instead of localizing the doors in a map, handles are to be used for door identification. Although most handles in our environment are round in shape, probably, the most appealing method to tackle the handle identification problem would be the SIFT procedure (Lowe, 2004), because it would not impose any assumption about the shape of the object to be recognized. The SIFT method extracts distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. The features are invariant to image scale and rotation, and have shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. But it is known that SIFT suffers from a high computational payload. For on-line applications, each one of the three steps (detection of local extrema, keypoint description computation and keypoint matching) should be computed faster (H. Bay *et al.*, 2006).

Alternatively, we will assume that door handles are circular in shape and hence, the identification is model based. A circular shape detector is used to approximate the region of interest (ROI). Color information is extracted using a supervised classification procedure and afterwards, using some measures about the colors of the surfaces in the surroundings of the ROI, the presence of the handle is confirmed or rejected. Only closed doors are considered.

2. THREE-STAGE DOOR HANDLE IDENTIFICATION

The three-stage process to recognize circular handles proceeds as follows:

(1) Shape detection: this step is designed to approximate the handle area, if any, in the image.
(2) Color segmentation: using the position information of the previous step, the surroundings of the candidate are segmented.
(3) Statistical validation: measures of the segmented image are afterwards needed to classify an image as containing or not a handle.

2.1 Shape detection

As mentioned before, handles are assumed to be circular in shape. Therefore, circle shape detection can be performed to locate handles in the images taken by the robot. Although many circle extraction methods have been developed, probably the most well-known algorithm is the Circle Hough Transform (CHT). Moreover, Hough transform methods have shown to be robust enough to deal with noisy images (Ayala-Ramírez *et al.*, 2006).

The use of the Hough Transform to detect circles was first outlined by (Duda and Hart, 1972) and then, Kimme eta al. (Kimme *et al.*, 1975) gave probably the first known application of the Hough transform to detect circles in real images. In their work, the direction of the gradient at each edge point is used as an additional piece of data which can further constrain the possible parameter values consistent with a given edge point. The center of a circle must lie on the line passing through the edge point along the maximum grey level gradient direction. As a result, instead of incrementing the whole circular cone, only segments of the cone need to be incremented.

Yuen et al. (Yuen *et al.*, 1990) investigated five circle detection methods which are based on variations of the Hough Transform. One of those methods is the *Two stage Hough Transform* and it is implemented in the OpenCV vision library (http://www.intel.com/research/opencv) used in the experiments described later on. This method decomposes the problem into two stages consisting of a 2D Hough Transform to find circle centers, followed by a 1D Hough Transform to determine the radius. Since the center of a circle must lie along the gradient direction of each edge point on the circle, then the common intersection point of these gradients identifies the center of the circle.

The circle finding function can identify a huge number of circles depending on the image background. From the returned list of circles, only the most probable one is considered. However, due to the local navigation strategies of the robot the images will be obtained within the same distance range and therefore, it is possible to know in advance the approximate radius of the handles. In this manner only the identified circumferences with a radius that lies within a known range would be considered as handle candidates. Unfortunately, the existence of a circle on the picture does not guarantee the existence of a handle. Third and fourth images in left column of figure 1 show two false positive considered as handle candidates.

2.2 Color segmentation

The shape extraction step is not reliable enough, neither to confirm, nor to reject, the presence of a handle in an image. Therefore, it needs to be endowed with some complementary processes in order to improve its performance. The approach used here is to use color information around the location of the circle for door recognition. Within the robot environment, a circular handle is always associated to a pladour door. These doors are blue-grey color, presenting different tonalities according to lighting conditions, as the presence or not of electric lamps just in front of the door, or sunlight incoming the area from windows close to the door. The objective of this second step is to segmentate the pladour colored pixels in an image.

To build the classifier we chose *Oc1* (Oblique Classifier 1) (Murthy *et al.*, 1994), a decision tree induction system well suited for applications where the instances have numeric feature values. *Oc1* builds decision trees containing linear combinations of one or more attributes at each internal node. The trees built in this way partition the instance space with both oblique and axis-parallel hyperplanes. Images taken by the robot are represented in RGB color space and thereby, each pixel is a three component vector, each component taking a value that ranges from 0 to 255.

In every classification problem, a training set is required to get a model to be later used when a new instance is presented to the model. To get the training set, we firstly constructed a database of positive instances (those associated to pladour doors), as well as negative instances (those not associated to pladour doors). The size of these databases was about two million pixels, obtained from about sixty images taken from the robot camera in a corridor. From these databases we extracted, in a random manner, 80,000 pixels, 40,000 of them labeled as *pladour* and the remaining 40,000 as *not pladour*. Then, these 80,000 pixels were presented to the *Oc1* tree generation procedure, to get the decision tree used in the segmentation of the images got by the robot camera. The obtained performance after applying 10 fold crossvalidation to this database was 93.61%.

Fig. 1 shows several examples of the segmentation process. The first two rows contain images with handles, although the variant lighting conditions affect the pladour color recognition process and therefore, the segmentation process. Notice for example the upper right corner of the second row images. The bottom rows show images without handles. Original images contain the found circles detected by the Hough transform, in spite of the radius exceeds the preset threshold, for sake of clarity. It can be appreciated that the segmentation of these images does not imply the existence of any handle.

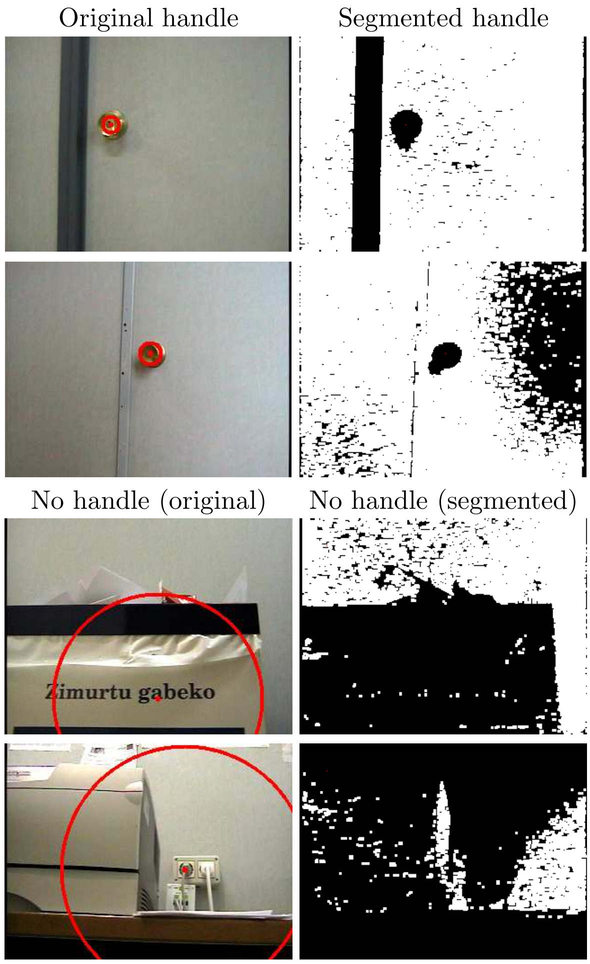

Fig. 1. Segmentation examples

2.3 Statistics for decision taking

So far we are provided with two procedures to recognize the handle and its surroundings:

- Circle detection
- Color segmentation

However, both procedures are prone to errors due to noise, lighting conditions and other objects in the environment (printers, dustbins, wooden doors, tables, and so on). We cannot relay in the accuracy of any of them separately. The aim of this third step is to analyze segmented pixels of the circle and those surrounding the circle in order to definitively confirm or reject the handle candidate. The segmentation process yields a black-and-white image, where white points are those classified as *pladour*, and black ones are those classified as not belonging to *pladour*. To analyze the surroundings of the prospective handle, the coordinates of the center of the circle are passed on from the circle detection procedure to the door recognition module. If the circle of radius r given by the CHT is centered at (x_0, y_0), the following three values are obtained (see Fig. 2):

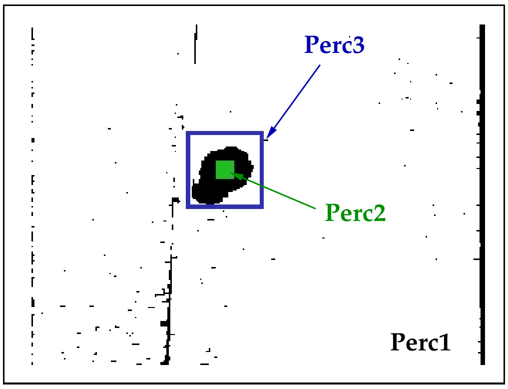

Fig. 2. Zones from which *Perc2* and *Perc3* are computed; *Perc1* is computed over the whole image

- *Perc1*: Percentage of white points (*pladour*) in the whole image. When the robot is close to a pladour door, a huge percentage of the pixels in its visual field are likely to be segmented as white. So, the bigger *Perc1*, the more likely the robot is in front of a pladour door.
- *Perc2*: The pixels around the circle center should belong to the handle, not to the pladour class. This value represents the percentage of black points (*not pladour*) in the 5×5 grid centered at (x_0, y_0). Therefore, the bigger *Perc2*, the more likely the robot is in front of a handle.
- *Perc3*: When the procedure that returns the center of a circle has really recognized a handle, the farther surroundings of that point are expected to be segmented as white, as they do not fall in the handle, but in the pladour door. We define:
  - $S1$: set of points in the squared area centered at (x_0, y_0) and of size length $2 \times r$.
  - $S2$: set of points in the squared area centered at (x_0, y_0) and of size length $2 \times (r+d)$, where d represents the desired shift from circle perimeter.

  Hence, $S = S_2 - S_1$ defines the closest surroundings of the handle and *Perc3* represents the percentage of white points (*pladour*) in S. Again, the bigger *Perc3*, the more likely the robot is in front of a handle.

The experiments were executed for a maximum allowed radius of 20 pixels ($r_{max} = 20$), a shift of 15 pixels from the circle perimeter ($d = 15$), and the confidence level of being in front of a handle should rise above 0.6 to confirm the door ($cl_{TH} = 0.6$). These percentages give us a measure of confidence of being in front of a door and that the circle recognition procedure has really recognized a circular handle. The arithmetic mean of these percentages was defined as the confidence level of being in front of a door (cl_{handle}).

Fig. 3 shows the whole door recognition process graphically.

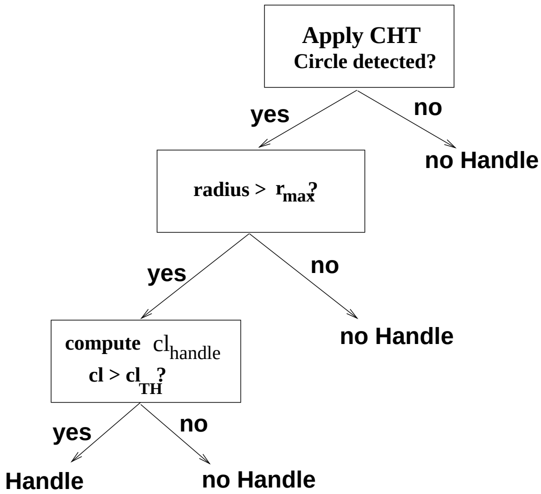

Fig. 3. Door Handle recognition process

3. PERFORMANCE EVALUATION

To test the performance of the proposed approach, first a database of images was built from sequences of images taken in the robot's environments. The obtained images are heterogeneous in the sense that they contain pladour, brick and wooden walls. The database contained almost 2800 images and the positive and negative cases were balanced. In order to evaluate the improvement introduced by the three-stage process, the classification accuracy was computed for the sequence and compared with the accuracy that would be obtained if only shape detection (using CHT) was used. However, accuracy is considered a fairly crude score that does not give much information about the performance of a categorizer.

The F1 measure, also known as F-score or F-measure, is a measure of a test's accuracy and can be viewed as a weighted average of the precision and recall, where an F1 score reaches its best value at 1 and worst score at 0. F1 measure combines both precision and recall into a single metric and favor a balanced performance of the two metrics. Therefore, it is considered more fair to use the F1 measure instead of the common accuracy as the main evaluation metric. Table 1 shows the obtained classification accuracies and the corresponding F1 measure values.

Table 1. Experimental results

Method	perf.	F1
Hough	0.306	0.319
Three stage	0.851	0.685

These results show how both performance measures, the classification accuracy and the F1 value are highly increased using the three stage approach, rising acceptable values to be tested on a real robot.

4. DOOR IDENTIFICATION BEHAVIOR

Tartalo is a PeopleBot robot from MobileRobots. This robot is provided with a Canon VCC5 monocular PTZ vision system, a Sick Laser, several sonars and bumpers and some other sensors. All the computation is carried out in its on-board Pentium III (800MHz). Player-Stage (Gerkey *et al.*, 2003) is used to communicate with the different devices and the software to implement the control architecture is SORGIN (Astigarraga *et al.*, 2003), a specially designed framework that facilitates behavior definition.

To evaluate the robustness of the handle identification system developed, it has been integrated in a behavior-based control architecture that allows the robot to travel across corridors without bumping into obstacles. With the camera pointing perpendicular to the robot heading, the robot processes the images looking for handles. When the robot finds a door, it stops, turns to face the door and simulates the door knocking action by asking for the door to be opened and waiting for someone to open it. If after a certain time the door is still closed, Tartalo turns again to face the corridor and continues looking for a new handle. On the contrary, if someone opens the door the robot detects the opening with its laser and activates a door crossing behavior module that allows it to enter the room.

The door identification behavior has been evaluated in the two different environments (Figures 4 and 6). To avoid false positives, some constraints have been added to the image sequence. On one hand, instead of relying the decision in a single image classification, the robot has to confirm the handle in p consecutive images. On the other hand, to make the CHT results more robust against variable lighting conditions, the deviation of the vertical axis position of the detected circles must be small during the p consecutive images. Figures 5 and 7 plot the door recognition state against time. These data were collected while the robot navigated in the two environments. The dashed line represents the selected value for p, and the yellow colored balls in the figures represent the positive identification of the door handles.

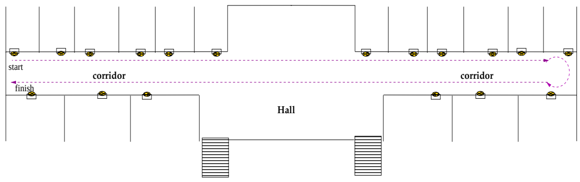

Fig. 4. First environment: 18 handles. The dashed line represents the robot's nominal path

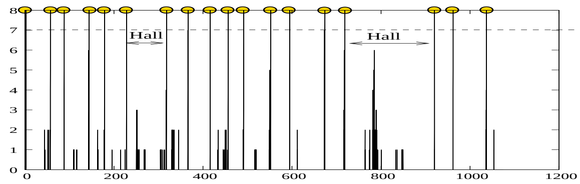

Fig. 5. Environment 1 results: time versus door identification state ($p = 7$)

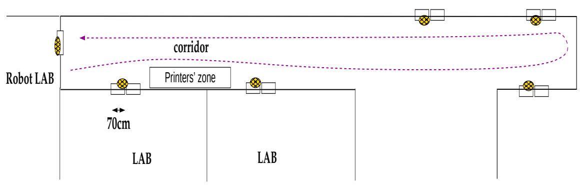

Fig. 6. Second environment: 5 handles. The dashed line represents the robot's nominal path

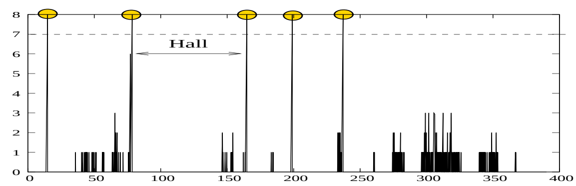

Fig. 7. Environment 2 results: time versus door identification state ($p = 7$)

These runs show a good performance of the robot behavior, successfully recognizing the 18 consecutive handles of the first environment and the five handles of the second one. During the many times the robot run, we found that the tendency of giving false positives was insignificant. However, we got a few false negatives when the lighting conditions were extreme, as for example, the dead end areas of large corridors with no artificial lighting and no windows around. The dark conditions of these handles made difficult the segmentation process to correctly identify the pladour color, and hence, made the 3-stage process fail when the door should have been identified.

5. CONCLUSIONS AND FUTURE WORKS

This paper presents the approach developed for the automatic recognition of the presence of a door handle in the environment of the robot. A three step classifier has been developed to perform the desired classification, and experiments performed in a real robot-environment system show good recognition of the handle.

As future work, the disadvantages of CHT should be overcomed (high computational and storage requirements) trying other circle detection methods (Rad *et al.*, 2003) or even stochastic techniques as proposed in (Ayala-Ramírez *et al.*, 2006).

Also, different kinds of handles in different environments (wooden doors, rectangular door handles, etc.) need to be identified. The generalization capabilities of the procedure –shape extraction, segmentation and probabilistic validation – need to be analyzed.

REFERENCES

Astigarraga, A., E. Lazkano, I. Rañó, B. Sierra and I. Zarautz (2003). SORGIN : a software framework for behavior control implementation. In: *CSCS14*. Vol. 1. pp. 243–248.

Ayala-Ramírez, V., C.H. Garcia-Capulin, A. Perez-Garcia and R. R. Sanchez-Yanez (2006). Circle detection on images using genetic algorithms. *Pattern Recognition Letters* **27**(6), 652–657.

Duda, R. and P. E. Hart (1972). Use of Hough transform to detect lines and curves in pictures. *Communications of the ACM* **15**(1), 11–15.

Eberset, C., M. Andersson and H. I. Christensen (2000). Vision-based door-traversal for autonomous mobile robots. In: *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*. pp. 620–625.

Gerkey, B. P., R. T. Vaughan and A. Howard (2003). The Player/Stage project: tools for multi-robot and distributed sensor systems. In: *Proc. of the International Conference on Advanced Robotics (ICAR)*. pp. 317–323.

H. Bay, H., T. Tuytelaars and L. Vam Gool (2006). SURF: Speeded up robust features. In: *Proceedings of the 9th European Conference on Computer Vision*.

Kimme, C., D. Ballard and J. Sklansky (1975). Finding circles by array accumulators. *Communications of the ACM* **18**(2), 120–122.

Kragic, D., L. Petersson and H. I. Christensen (2002). Visually guided manipulation tasks. *Robotics and Autonomous Systems* **40**(2-3), 193–203.

Lazkano, E., B. Sierra, A. Astigarraga and J. M. Martínez-Otzeta (2006). On the use of bayesian networks to develop behavior for mobile robots. *Robotics and Autonomous Systems*.

Li, W., H. I. Christensen and A. Orebäck (2004). An architecture for indoor navigation. In: *Proceedings of the IEEE International Conference on Robotics and Automation*. pp. 1783–1788.

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60**(2), 91–100.

Monasterio, I., E. Lazkano, I. Rañó and B. Sierra (2002). Learning to traverse doors using visual information. *Mathematics and Computers in Simulation* **60**, 347–356.

Muñoz-Salinas, R., E. Aguirre and M. García-Silvente (2005). Detection of doors using a genetic visual fuzzy system for mobile robots. Technical report. University of Granada.

Murthy, S. K., S. Kasif and S. Salzberg (1994). A system for induction of oblique decision trees. *Journal of Artificial Intelligence Research* **2**, 1–33. ftp://blaze.cs.jhu.edu/pub/oc1.

Rad, A. Ajdari, K. Faez and N. Qaragozlou (2003). Fast circle detection using gradient pair vectors. In: *VIIth digital image computing: techniques and applications* (S. Ourselin C. Sun, H. Talbot and T. Adriaansen, Eds.). pp. 879–887.

Seo, M. W., Y. J. Kim and M. T. Lim (2005). *LNAI*. Chap. Door Traversing for a Vision Based Mobile Robot using PCA, pp. 525–531. Springer-Verlag.

Stella, E., G. Cirelli, G. Attolico and A. Distante (1996). Sauro: An autonomous mobile vehicle for indoor environment. In: *MELECON*. Vol. 2. pp. 1145–1150.

Stoeter, S. and L. M. Papanikolopoulos (2000). Real-time door detection in cluttered environments. *Proceedings of the IEEE International Symposioum on Intelligent Control* pp. 187–192.

Yuen, H. K., J. Princen, J. Illingworth and J. Kittler (1990). Comparative study of Hough transform methods for circle finding. *Image and Vision Computing* **8**(1), 71–77.