

Intelligent overhead sensor for sliding doors: a stereo based method for augmented efficiency

Luca Bombini, Alberto Broggi, Michele Buzzoni, and Paolo Medici

VisLab – Dipartimento di Ingegneria dell'Informazione
Università degli Studi di Parma
{bombini,broggi,buzzoni,medici}@vislab.it
<http://www.vislab.it>

Abstract. This paper describes a method to detect and extract pedestrians trajectories in proximity of a sliding door access in order to automatically open the doors: if a pedestrian walks towards the door, the system opens the door. On the other hand if the pedestrian trajectory is parallel to the door, the system does not open. The sensor is able to self-adjust according to changes in weather conditions and environment. The robustness of this system is provided by a new method for disparity image extraction.

The rationale behind this work is that the device developed in this paper avoids unwanted openings in order to decrease needs for maintenance, and increase building efficiency in terms of temperature (i.e. heating and air conditioning). The algorithm has been tested in real conditions to measure its capabilities and estimate its performance.

Keywords: safety sensor, sliding doors, obstacle detection, pedestrian detection, trajectory planning, stereo vision

1 Introduction

Current access control systems for automatic door control require a sensor able to detect a moving object or a pedestrian crossing the gate. This approach does not take into account the trajectory of pedestrians, and therefore can not estimate movements. As an example, if a pedestrian crosses the area in front of the door but does not want to cross the gate, the control access board detects his/her presence and anyway opens the door. In this case, the system is not efficient, since it leads to a waste of energy in terms of electricity, air conditioning, or heating and decrease the system lifetime with unnecessary open/close actions.

The sensor presented in this paper solves the problem of unwanted openings exploiting a stereo vision system based on a 3D reconstruction: CMOS cameras with *fish-eye* optics placed over the sliding doors are used to acquire stereo images.

A very general setup has been studied to fit different kinds of automatic door accesses. Another feature of this system is that it can be integrated or adapted to a wide range of existing doors and customer requirements.

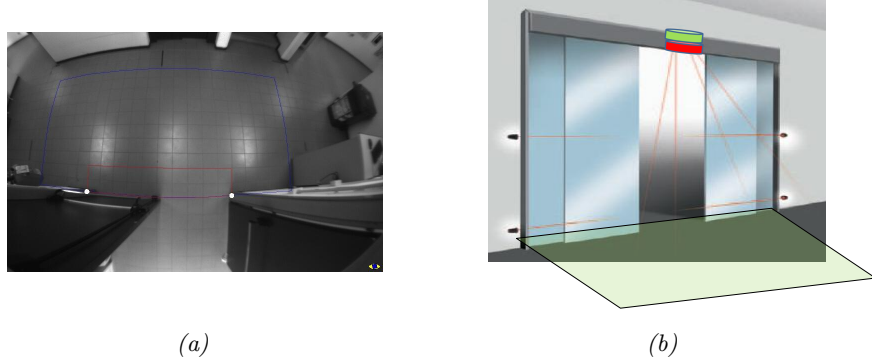


Fig. 1. (a) An image acquired by a single camera installed on an sliding doors access. (b) The cameras system (green) is placed on the door access and can replace usual sensor (red). The area monitored by the new vision-based system (green plane) is larger than the one covered by infrared sensors (red line) (b) and therefore allows the detection of pedestrian trajectories.

A pair of low-cost cameras are installed on the top of the door, overlooking the crossing area (fig. 1.a). The cameras are connected to an embedded computer that processes images using a stereo vision based algorithm. The computer is connected to the access control board and therefore can open or close the door as required.

This paper is organized as follows: in Section 2 the algorithm structured is sketched; Section 3 shows a the new method for the computation of the disparity image; results, final considerations, and ideas for future developments are then presented in Section 4.

2 The algorithm

This section presents the overview and implementation of the algorithm used in the system. Figure 2 shows the flowchart of the main algorithm steps.

Given two stereo images, the direction of objects moving in the scene is detected and analyzed to trigger the doors opening.

The algorithm does not require calibration: the IPM images (Inverse Perspective Mapping) [1, 2] of the scene are generated by a Look Up Table (LUT) computed off-line (see Section 3).

The images obtained so far are used to create the reference background thanks to a standard background algorithm generation using a circular buffer of the last N images [4]. The two background images are compared against input images performing a difference: binarized images (masks) are then generated after thresholding and dilating the difference images.

A Sobel filter is used to reduce the changing light conditions problem and to produce an edges map that is used to generate Disparity Space Image (DSI) [3].

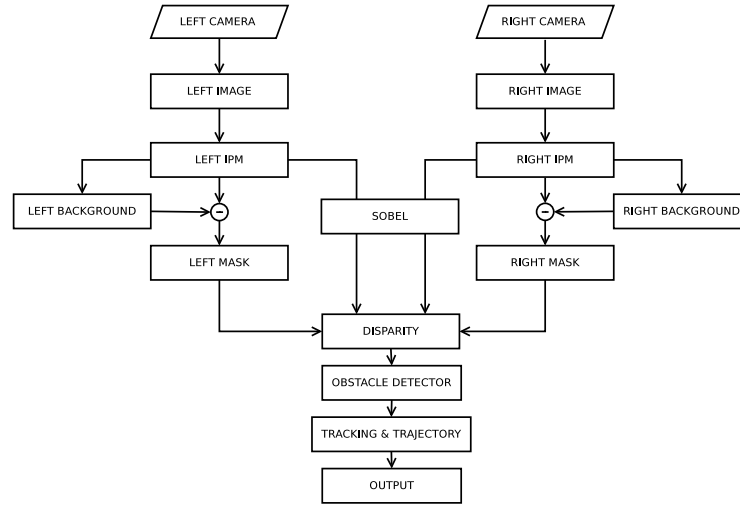


Fig. 2. Flowchart of the proposed algorithm.

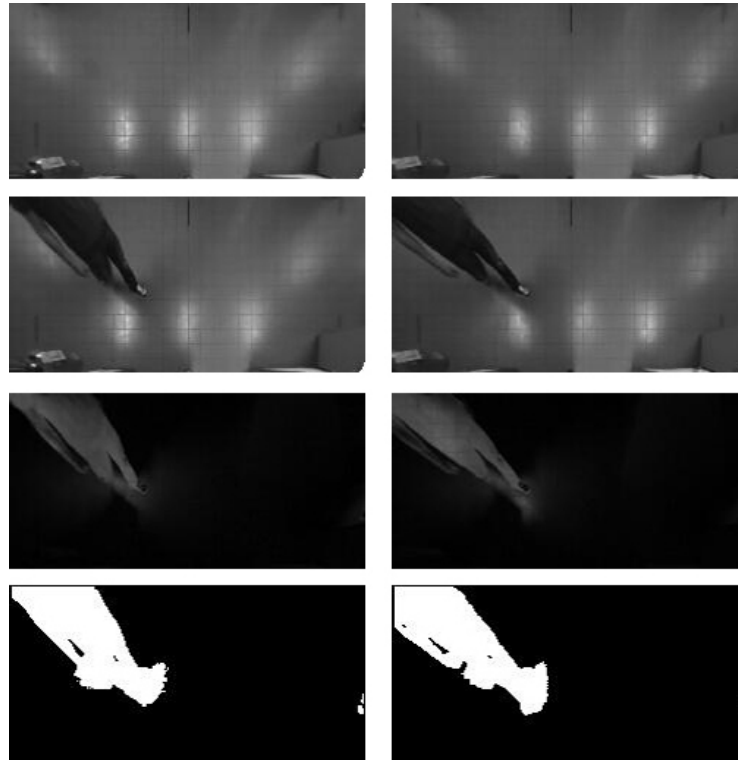


Fig. 3. From top: background, current, difference, and binarized images for left and right cameras.

Thanks to the preliminary use of the IPM, the resulting DSI has to be interpreted using a specific approach (see Section 3)

The DSI image is filtered using the two masks previously calculated, therefore considering only the points of the first mask and validating only the points that match in the other mask. This approach reduces false positives and errors in DSI computation, such as the ones due to repetitive patterns and reflections.

The obtained images are used by the obstacle detector to find each obstacle in the scene. In fact, the DSI image is clustered using the contiguity of points in the image and their location in the world as constraints (see Section 3). The result of this step is a list of blobs that belong to the obstacles in the framed scene (fig. 4.a). Each blob allows the computation of the 3D transformation related to the corresponding obstacle. And, specifically, the size and position with respect to the ground. The obstacles with a small size and an height lower than 20 cm are discarded. The detected size and position of the remaining blobs are then used to track obstacles. The trajectory of each obstacle is computed through a linear regression of the past positions (fig. 4.b). The system triggers the door opening (fig. 4.c) if the trajectory is aimed towards the door for two successive frames or if the object is close to the door.

For safety reasons a small proximity area near the door is considered as well; the door will be opened if any obstacle is detected inside this area, independently from its direction of movement.

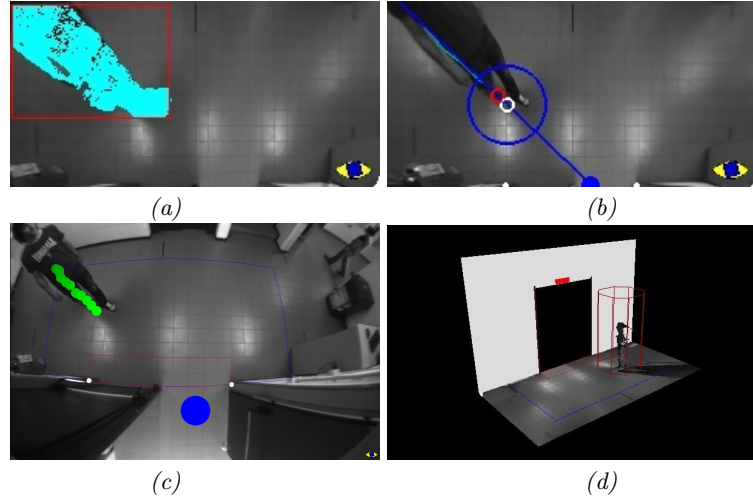


Fig. 4. (a) A cluster DSI Blob, (b) the computed trajectory crosses the doorstep (delimited by the two white markers), (c) output, (d) 3D-reconstruction

3 Disparity and DSI from stereo IPM transformation

In this section a novel method to compute the disparity on IPM images is presented. Initially, the formulas to generate the IPM images are analyzed, then the concept of disparity on these images to create the DSI will be introduced.

3.1 Inverse Perspective Mapping

The Inverse Perspective Mapping (IPM) is a geometric transformation that allows to remove the perspective effect from images acquired by a camera [1, 2]. It resamples the source image mapping each pixel in a different position and creating a new image (IPM image) which represents an bird-eye view of the observed scene.

Unfortunately is not possible to correctly remap all pixels of the source image but only those that belong to a known plane (usually the ground, $Z = 0$). The points that do not belong to this plane are remapped in an incorrect mode that depends on their position in the world and the position of the cameras.

An off-line LUT is used for the generation of the IPM, that can be defined from the intrinsic and extrinsic parameters of the vision system or, as in this case, through a manual mapping of points of a calibration grid and a subsequent interpolation. The LUT allows to remove the perspective effect and the distortion of the source image all together [5–7].

A point not belonging to the $Z = 0$ plane is remapped on the IPM image in a specific way that can be understood by studying the process of generation of the IPM. The IPM image is the result of two perspective transformations: a direct one and a reverse one. The world points are mapped into a perspective image through a perspective transformation and then in an IPM image assuming all pixels belonging to the ground.

The equation of a perspective transformation can be expressed in three-dimensional homogeneous coordinates:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{A}\mathbf{R} \left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \right) . \quad (1)$$

where (X, Y, Z) are the coordinates of the points in the world, (x_0, y_0, z_0) are the coordinates of the camera position in the world, and (u, v) are the coordinates in the perspective image. \mathbf{A} , and \mathbf{R} are the matrices (3×3) of the intrinsic and rotation parameters, respectively.

The IPM transformation can be defined as:

$$\left(\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \right) = \mathbf{R}^{-1}\mathbf{A}^{-1} \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} . \quad (2)$$

With Z' the height from ground of the point to remap. Without additional information and assuming $Z' = 0$, equations (1) and (2) lead to:

$$\left(\begin{bmatrix} X' \\ Y' \\ 0 \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \right) = \overbrace{\mathbf{R}^{-1} \mathbf{A}^{-1} \mathbf{A} \mathbf{R}}^I \left(\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} - \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \right) . \quad (3)$$

which, since the homogeneous coordinates are defined up to a proportionality factor λ , gives the final expression:

$$\begin{bmatrix} X' - x_0 \\ Y' - y_0 \\ -z_0 \end{bmatrix} = \begin{bmatrix} X - x_0 \\ Y - y_0 \\ Z - z_0 \end{bmatrix} \lambda . \quad (4)$$

λ can be obtained thanks to the $Z' = 0$ assumption:

$$\lambda = \frac{-z_0}{Z - z_0} . \quad (5)$$

and, can be substituted in eq.(4), allowing X', Y' computation:

$$X' = -z_0 \frac{X - x_0}{(Z - z_0)} + x_0 \quad Y' = -z_0 \frac{Y - y_0}{(Z - z_0)} + y_0 . \quad (6)$$

It can be noticed that equations (6) do not depend on the lens distortion, intrinsic parameters and orientation of the cameras, but only on their position in the world. Therefore, the IPM can be easily computed using a LUT and not requiring a complex calibration step.

From these results, it is evident that the same point in the world (X, Y, Z) is remapped on different IPM coordinates (X', Y') depending on the camera position in the world. Only if the world point belongs to the ground, $Z = 0$, the equations are reduced at $X' = X$ and $Y' = Y$.

Equations (6) can be merged together, to obtain the important result:

$$\frac{X' - x_0}{X - x_0} = \frac{Y' - y_0}{Y - y_0} . \quad (7)$$

A vertical obstacle features constant X and Y , therefore equation (7) can be written as:

$$X' - x_0 = m'(Y' - y_0) \quad (8)$$

where the line slope m' is defined as:

$$m' = \frac{X - x_0}{Y - y_0} . \quad (9)$$

All the world points that belong to a vertical obstacle are then remapped into IPM image on the line connecting the pin-hole and the base position of the obstacle. This line depends on the coordinates (X, Y) of the world point and on camera location (x_0, y_0, z_0) . Since the dependence from Z has been removed, eq. (7), the z_0 coordinate does not influence the line equation on which the vertical obstacles are mapped, but only their size in the IPM image.

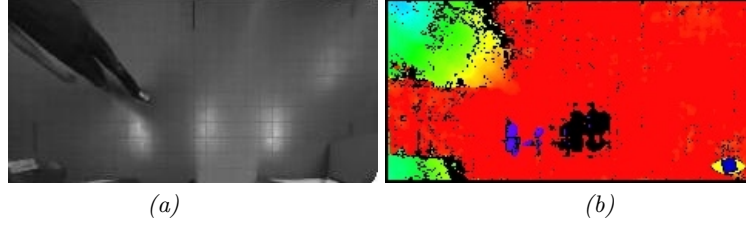


Fig. 5. (a) Right IPM images, (b) DSI.

3.2 The disparity and the 3D reconstruction on IPM images

Let us consider a system with two stereo cameras aligned on a line parallel to Y -axis, with the same height h and with baseline b . The right camera coordinates are $(x_0, -b_r, h)$, while the left ones are (x_0, b_l, h) .

Using equations (6), the IPM coordinates of a generic world point $P(X, Y, Z)$ can be computed. The point is then remapped applying the IPM to the images captured by the cameras:

$$X'_r = X'_l = -h \frac{X - x_0}{Z - h} + x_0 . \quad (10)$$

$$Y'_r = -h \frac{Y + b_r}{Z - h} - b_r \quad Y'_l = -h \frac{Y - b_l}{Z - h} + b_l \quad (11)$$

Solving the coordinates Y'_r as a function of the coordinate Y'_l :

$$Y'_r = Y'_l - b \left(\frac{h}{Z - h} + 1 \right) = Y'_l + b \frac{Z}{h - Z} . \quad (12)$$

and given that the coordinates X'_r and X'_l are the same for the two images, ($X'_r = X'_l$), the concept of disparity is defined as:

The disparity of a world point with coordinate (X, Y, Z) projected onto two IPM images is equal to:

$$\Delta = Y'_r - Y'_l = b \frac{Z}{h - Z} . \quad (13)$$

This definition of disparity leads to the following considerations:

- Since in the IPM image only points of the world with a height less than the height of the camera, $Z < h$, can be remapped, the disparity is always positive.
- The disparity of points does not depend on coordinates (X, Y) of the points in the world, but only on their coordinate Z . This implies that all points placed at the same height Z in the world have the same disparity Δ in the IPM image. The ground, having no height ($Z = 0$), has zero disparity $\Delta = 0$. In addition, the disparity of features has an upper limit of h .

- The disparity proportionally depends on the distance between the two cameras (baseline).
- The disparity position does not depend on the x_0 coordinate of the cameras.

The knowledge of the disparity of an IPM image allows a three-dimensional reconstruction of the scene. The formulas to estimate the world coordinates of a IPM point derive from the previous definition of disparity; in fact, the Z coordinate can be directly obtained from eq.(13):

$$Z = h \left(1 - \frac{b}{\Delta + b} \right) = h \frac{\Delta}{\Delta + b} . \quad (14)$$

with h being the height from ground and b the cameras baseline.

Replacing this value for Z in eq.(6), it can be obtained:

$$X'_r = -h \frac{X - x_0}{\left(h \frac{\Delta}{\Delta + b} - h \right)} + x_0 \quad Y'_r = -h \frac{Y + b_r}{\left(h \frac{\Delta}{\Delta + b} - h \right)} - b_r . \quad (15)$$

Therefore allowing to compute the world coordinates (X, Y) as function of disparity:

$$X = X'_r \left(\frac{b}{\Delta + b} \right) + x_0 \left(\frac{\Delta}{\Delta + b} \right) \quad Y = Y'_r \left(\frac{b}{\Delta + b} \right) - b_r \left(\frac{\Delta}{\Delta + b} \right) . \quad (16)$$

3.3 DSI on IPM images

The Disparity Space Image computation is traditionally used on standard stereo images, but it can be efficiently used for IPM images too (see fig. 5.b). In fact the algorithm for the implementation of DSI [3] requires to search for each pixel of the right image the corresponding pixel on the left image (homologous pixel). In IPM images, the homologous pixels are searched for in the same row in each images. In fact, equation 13, shows that the coordinates X' of a world point only depends on its height and is the same in the two images, therefore easing DSI computation.

In addition it can further noticed for DSI on IPM images that:

- The disparity of pixels is always positive.
- The ground has zero disparity.
- The disparity of vertical object increases with its height, featuring a gradient along the line defined by eq. (7).
- Given that the same obstacle is differently deformed by the IPM depending on camera position, the search of homologous points has to cope with differences when computing the matches.
- If the best correlation between two homologous point is under a given threshold, the DSI value is undefined, indicating that the pixel has not been found in the other image.
- The size of the IPM images is a parameter that highly influences the computation time of the DSI. More precisely the time needed for DSI computations is cubically proportional to the IPM image size.

4 Results and conclusions

The system was tested in real conditions to measure its capabilities and evaluate its performance, using as computing architecture a SmallPC industrial computer with a processor Intel Core 2 Duo 2.5GHz and 2G DDR DRAM. The system was developed using Linux operating system and C++ programming language. Two cameras were installed over a public area door acquiring many image sequences. The format video was set to 752×480 pixel at 15 Hz, but it was reduced to 240×120 pixel in IPM images to reduce the computational load. The images have been analyzed in laboratory and the algorithm has been tested in a lot of situations.

The system has proven to work correctly and reach good results in different situations (fig. 6). The most critical situation that affects the system performance is the presence of an excessive number of obstacles in the scene. In such a case, the objects trajectory is not correctly detected, due to difficulties in tracking all objects, and therefore generating some false positives leading to unwanted openings. Anyway, the system always assures the door opening when a person really wants to cross the gate thanks to the implementation of the proximity area attention (fig. 6.c).

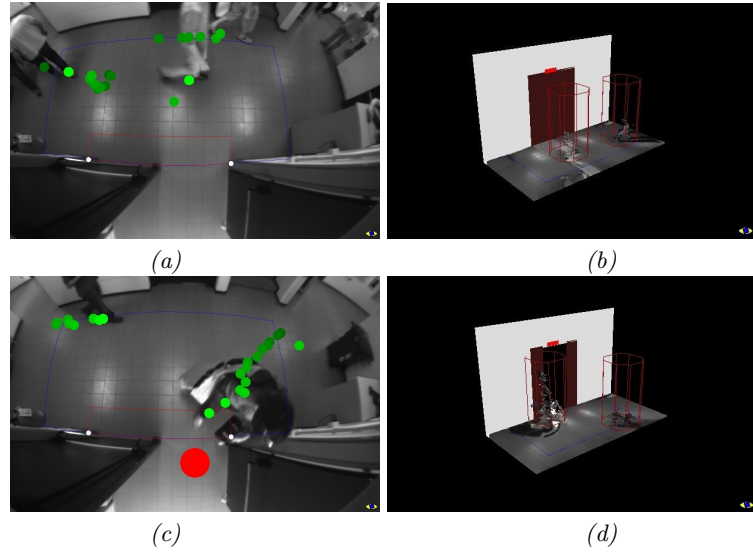


Fig. 6. (a) Algorithm output in a multi-pedestrian scenario, (b) 3D-reconstruction, (c) a pedestrian detected in proximity area, (d) 3D-reconstruction

The stereo system presented in this paper has a great advantage with respects to classic disparity obstacle detectors, in fact it uses a calibration based on a LUT. The LUT for the removal of the perspective effect and camera distortion

is computed with a simple grid positioned in the area of interest during the first installation. This area is easily tunable acting on software parameters and therefore enabling the use of the system for different doors and environments. An automatic calibration application is currently under development: it will ease the installation and reduce the installation time, making easier the work of the installers.

A new installation and testing in a shopping center is scheduled to compare and tune the system performance against a classic infrared approach. An extension of the system for monitoring bus doors also for safety purposes and not only for optimization is currently under development.

A patent [10] was filed to cover this approach that might turn into a product.

5 Acknowledgements

The work has been supported by the Industria2015 AUTOBUS Italian project.

References

1. Bertozzi, M., Broggi, A., and Fascioli, A.: Stereo Inverse Perspective Mapping: Theory and Applications Image and Vision Computing Journal, vol. 8(16):pp. 585-590, 1998.
2. Bertozzi, M., Broggi, A., and Fascioli, A.: An extension to the Inverse Perspective Mapping to handle non-flat roads. In Procs. IEEE Intelligent Vehicles Symposium'98, pp. 305-310. Stuttgart, Germany, ott. 1998.
3. Felisa, M., and Zani, P.: Incremental Disparity Space Image computation for automotive applications, In Procs. IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems, St.Louis, Missouri, USA, October 2009.
4. Benezeth, Y., Jodoin, P.M., Emile, B., Laurent, H., and Rosenberger, C.: Review and evaluation of commonly-implemented background subtraction algorithms, ICPR 2008 19th International Conference on Pattern Recognition, Dec 2008.
5. Claus, D., and Fitzgibbon, A. W.: A rational function lens distortion model for general cameras. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 05), vol. 1, pp. 213-219. San Diego, Calif, USA, giu. 2005.
6. Devernay, F., and Faugeras, O.: Straight lines have to be straight. Machine Vision and Applications, vol. 13(1):pp. 14-24, ago. 2001.
7. Tsai, R. Y.: A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses. IEEE Journal of Robotics and Automation, pp. 323-334, ago. 1987.
8. Jahne, B.: Digital Image Processing. Springer, Berlin, Germany, Fifty ed., set. 2002.
9. Pratt, W.K.: Digital Image Processing. Addison-Wesley, Milano, Italy, third ed., mag. 2001.
10. Bombini, L., Buzzoni, M., Felisa, M., and Medici, P.: Sistema per il Controllo di Porte Automatiche, March 2010, CCIAA di Milano, Patent application MI2010A000460.