DECSAI
Departament of
Computer Sciences and Artificial Intelligent

Detection of doors using a genetic visual fuzzy system for mobile robots

DECSAI-SI-2005-02

Authors:

Rafael Muñoz-Salinas

Eugenio Aguirre

Miguel García-Silvente

1.1 abstract

Doors are common objects in indoor environments and its detection can be used in robotics tasks such as map-building, navigation and positioning. In this work we present a new approach to door-detection in indoor environments using computer vision. Doors are sought in grey-level images detecting the segments that form their doorframes. The Hough Transform is used in order to extract the segments in the image after applying the Canny edge detector. Features like size, direction, or distance between segments are used by a fuzzy system to analyze if the relationship between them reveals the existence of a door. The system has been tuned using a genetic algorithm to achieve the maximum performance in detecting the doors of our environment. For that purpose, a large database of images containing doors seen from different angles and distances has been created. The method has shown to be able to detect typical doors under strong perspective deformations and it is fast enough to be used for real-time applications in a mobile robot.

1.2 Introduction

Environment perception is one of the most challenging problems to face when designing and implementing physical agents like autonomous mobile robots. Using different kind of sensors a robot gets information from the environment that can be used either to control the robot or to create a model of the world. Indoor environments are usually structured so we can usually find generic objects like doors, walls or corridors and the ability of a robot to discover these places can be a key point for a robust navigation and model creation. In previous works, we have developed a fuzzy perceptual model based on ultrasound sensors able to detect those typical places and build a map of the environment that is used to aid our robot in its navigation (Aguirre y González 2002; Aguirre y González 2003). Unfortunately, this model is submitted to the typical weakness of ultrasound sensors.

Vision has been widely used to enlarge perception capabilities of robots (Adorni et al. 2001; Huber y Kortenkamp 1998; Paulino et al. 2001; Srinivasan et al. 1999; Wells et al. 1996). The use of 3D models has become very popular nowadays although it requires a great effort either to model the environment and to perform the matching process (Gasteratos et al. 2002; Stevens y Beveridge 2000; Zheng et al. 1993). Artificial landmarks have been used to perform a robust localization and navigation in the environment (Muñoz-Salinas et al. 2004; Desouza y Kak 2002; Katsuki et al. 2003; Li y Yang 2003; Scharstein y Briggs 2001; Tashiro et al. 1995). However it has the disadvantage of requiring to alter the environment to place the landmarks. It is usually preferable to use landmarks that can be found in the environment. Door is a common object that

can be found in indoor environments and it can be very useful for navigating, map-building and positioning tasks. In this paper we present a new approach to visual door-detection using a single camera and a visual fuzzy system.

Door-detection using artificial vision has been performed using different techniques in the literature. In (Cicirelli et al. 2003) a technique based on neural networks for detecting doors based on its components is presented. The system consist in two neural networks, one for detecting the upper corners of the door and another for detecting the lateral and upper part of the doorframe. Each net analyzes (for every pixel in the image) a subwindow of size 18x18 and decides if it belongs either to the corner of a door, to the lateral or upper part of the doorframe or to none of them. The input for the net is the hue and saturation components of the subwindows. Each net has a total of 648 inputs neurons, a hidden layer and an output one with one neuron in it. An analysis of the whole output is made after classifying each pixel of the image considering that there is a corner if the total number of pixels classified as corner by the net exceed a certain threshold. To detect the lateral bars of the door a similar process is made and all the information is combined properly to decide if the elements found form a door. The system is able to detect doors under partial occlusion conditions and from different perspectives, but it has three main drawbacks. First, it requires a high computational effort (3 seconds in analyzing an image). Second, it can not detect fully opened doors. And finally, it is dependent on the color of the door.

Using a functionality-based approach, a method for generic object recognition used for robot navigation is presented in (Kim y Nevatia 1998). A door is defined as an inverted U that can be crossed by people. A trinocular vision system is used in order to detect segments in the images of the environment. The segments are analyzed to check if they accomplish a set of size and height restrictions typical of its indoors environment doors. The trinocular vision system makes possible to know the real position of the segments in the space and thus check the imposed restrictions. The system has the disadvantage of the cost of the perceptual system.

In (Stoeter et al. 2000) a method for detection of doors limited to corridors is explained. First they capture an image of the corridor and enhance the edges. After a dilation followed by an erosion, vertical stripes are selected. Possible doors are detected taking into account the distance and direction of the walls in respect to the robot based on their expected dimensions. However, it is not very clear how the vertical stripes are classified in doors. Furthermore, the technique limits the detection to corridors and it does not consider deformations caused by changes of perspective.

In the approach developed by (Monasterio et al. 2002), a simple technique for detecting a door is used in order to aid an autonomous robot to cross it. The detection of the door is based either on ultrasound and vi-

sual information. The visual detection is based on the extraction of the lateral bars of the doorframe. Edges in the image are enhanced and then dilated, eventually, columns wider than 35 pixels are considered as doors. This method does not take into account perspective deformations and it is only applicable when the door is at a distance that makes its doorframe to be seen wider than 35 pixels. Furthermore, the method has only been tested placing the robot at 1 meter in front of the door and with relatives angles to the door not exceeding 30^o degrees.

In this work a new approach to visual door-detection is presented. Our approach leads on the detection of the segments that form the frame edges of a doorframe in a grey-level image. In order to detect if there is any doorframe in a image, edges are extracted using Canny (Canny 1986) detector. The result is used to calculate the Hough Transform (Hough 1962) and then segments are extracted (Foresti 2000). Features like size, direction, or distance between segments are used by a fuzzy system to analyze if the relationship between them reveals the existence of a door. The membership functions of the system have been designed based on expert knowledge. In order to adapt the fuzzy system to the environment conditions in which the robot works, it has been tuned. We have employed a tuning process using a genetic algorithm (GA) to maximize the performance of the system over a large database of doors of our environment.

The proposed system allows to detect doors under strong perspective deformations caused by the two degrees of freedom (DOF) allowed for our camera. It also allows the detection of completely opened and closed doors using a single camera independently of the color of the door. It has been proved that the system successfully detects the doors of our environment under a wide range of distances and angles. Our experiments also show that the system is fast enough to be used for real-time applications in our mobile robot. We have used for our experiments a Nomad 200 mobile robot that has been enhanced with Pentium IV laptop computer to perform the visual processing. Fuzzy logic (Zadeh 1975) brings several advantages when dealing with the problem. It allows us to define concepts in a flexible way. Properties like *vertical* or *horizontal* are defined as linguistics variables allowing to manage perspective deformations and vagueness in the segment extraction in a natural way. Another advantage of using fuzzy logic is the facility for combining the information provided by the visual system with other perceptual model previously developed based on ultrasound (Aguirre y González 2003).

The rest of the work is structured in the following sections. Section 1.3 explains the visual fuzzy system. Section 1.4 presents the tuning algorithm using a GA. Section 1.5 shows the experiments carried out. Finally, Section 1.6 exposes some conclusions and future work.

1.3 Visual Fuzzy Door-Detection

In this section it is explained how the doors are detected using artificial vision and a visual fuzzy system that examines the segments of the image. We must define what it is considered as *door*. Our approach consists in looking for doorframes. For that purpose, the segments of an image are examined analyzing if they belong to the frame edges of a doorframe.

According to our experience, after the segment extraction, it is possible to consider three cases of study. First case, only appears one of the frame edges of the doorframe. Second case, there appears the two frame edges of the doorframe, the internal and the external. Third case, there appears one of the frame edges of the doorframe and an incomplete part of the other. These three cases can be better understood seen Figure 1.1. Although in the figure the three cases are depicted with completely vertical and horizontal segments, our aim is to be able to detect them under perspective deformations that make the segments appear with different inclinations and sizes. In order to formally describe the method, we define three complex fuzzy concepts: *Frame Edge* (FE), *Complete Doorframe* (CDF) and *Frame Edge with Evidence* (FEE). The method analyzes the membership degree of the segments found in an image to those fuzzy concepts.

The fuzzy concept *Frame Edge* is used to represent a frame edge of a doorframe. It is defined as a pair of vertical segments (belonging to the lateral of a doorframe) joined in its upper part to a horizontal segment (belonging to the upper part of a doorframe). Schematically an FE can be represented as shown in Figure 1.1(a). The simplest approach would be to consider a door as a unique FE, but this could lead us to confusion with squared objects of the environment (like cupboards).

When we look at a door it is possible to see its two frame edges. One corresponding to the limit between the doorframe and the wall, and another caused by the grey-level difference between the doorframe and the scene behind it (if its is opened) or its leaf (if it is closed). We model this situation by the fuzzy concept *Complete Doorframe* (CDF) that is schematically represented in Figure 1.1(b).

Unfortunately, the two frame edges of a doorframe can not always be found. It happens when the door is seen open and the leaf hides a lateral bar of the doorframe. In this case it is usually possible to find one FE and *evidences* of the incomplete one around the former. This situation is represented by the fuzzy concept *Frame Edge with Evidence* (FEE) and is schematically depicted in Figure 1.1(c).

The method searches incrementally. First it looks for FE among the segments of an image. Then it analyzes if there are two frame edges that could belong to the same doorframe forming a CDF. Finally, it tries to find evidences around the remaining FE forming FEEs. In order to determinate the membership degree of the segments in the image to these complex fuzzy

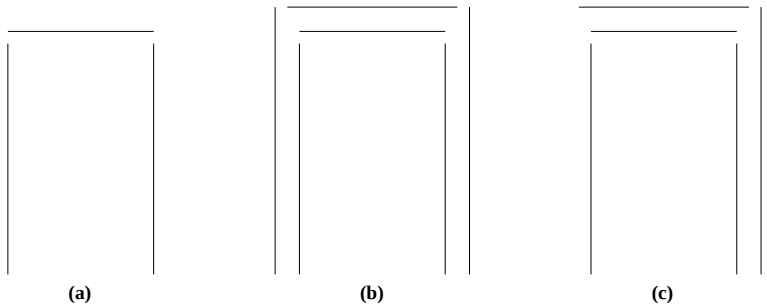


Figure 1.1: (a) FE (b) CDF (c) FEE

concepts, simpler fuzzy concepts like *vertical segment* or *parallel segments* are employed. Features of the segments like size, distance or direction are used to determinate if the segments accomplish certain restrictions typical of the frame edges of the doorframes. Fuzzy logic allows to analyze the segments taking into account the deformations caused by the perspective projection and to manage the inaccuracy in the segment extraction.

The detection process starts applying the Canny (Canny 1986) edge extractor to a grey level image. The result is a binary image $I(x, y)$ where pixels labeled as *true* belongs to edges in the original image. We assume, in order to ease notation, that $I(x, y)$ has size NxN. Using the edge pixels of $I(x, y)$ we apply the Hough Transform (Hough 1962) obtaining a set of segments (Foresti 2000) that we denote as $S = \{S^0, S^1, ..., S^n\}$. Each segment S^i is defined by two points in the image $S^i = \{p_0^i, p_1^i\}$ where $p_j^i = (x_j^i, y_j^i)$ (coordinates in the image plane). When the segments are extracted, the analysis to detect possible doorframes starts.

1.3.1 Candidate segment selection

The set of extracted segments is analyzed to select only those that could belong to frame edges. We must remember that we are looking for those segments that belong to the lateral and upper part of a doorframe. As we wish to make the detection under different distances and using a two DOF camera, the projection of the segments does not need to be necessarily parallel to the image plane. Therefore, doorframes do not necessarily are seen as rectangular objects in the image. This circumstance forces us to understand the concept vertical and horizontal segments belonging to a doorframe in a flexible way. Furthermore, when analyzing an image it is usual to extract segments that do not belong to any doorframe. To face both circumstances we define the two fuzzy concepts *Vertical Segment* (VS) and *Horizontal Segment* (HS).

If we analyze indoor environments, we could see that doors are relatively tall objects. Therefore their vertical segments (VS) are projected with a

large and relatively vertical aspect. On the other hand, the upper segments of the frame edges of a doorframe (HS) can be projected in a wide range of sizes and orientations around the horizontal direction. Nevertheless, we can assume that they are in upper positions of the image. Direction, size and height features of a segment S^i are used to establish its membership degree to the fuzzy concepts VS and HS.

The first feature is measured using a linguistic variable called $Direction(S^i)$ that has the three possible values (*horizontal, medium* and *vertical*) represented using the fuzzy sets shown in Figure 1.2(a). The direction is considered as *vertical* when the angle of the segment in respect to the x axis of the image is near to $\frac{\pi}{2}$ radians and *horizontal* when is near to zero. The input value to this variable, $directionS(S^i)$, is calculated as expressed in Equation 1.1. The function arctan returns the angle of the segment in the range $[0, \frac{\pi}{2}]$ but this value is escalated to the range $[0, 1]$:

$$directionS(S^i) = \frac{2}{\pi} \arctan\left(\frac{|y_1^a - y_0^a|}{|x_1^a - x_0^a|}\right) \tag{1.1}$$

The second feature is measured using the linguistic variable $SizeS(S^i)$ that has the three possible values represented using the fuzzy sets shown in Figure 1.2(b). The input value for this variable is the size of the segment $size(S^i)$ that is limited to $[0, 1]$ using Equation 1.2. Where $dist(p_0^i, p_1^i)$ is the Euclidean distance between the two points that define the segment S^i and $N\sqrt{2}$ is the maximum possible distance between two points in the image (the diagonal).

$$size(S^i) = \frac{dist(p_0^i, p_1^i)}{N\sqrt{2}} \tag{1.2}$$

Finally, the height of a segment in the image is measured using the linguistic variable $YPosition(S^i)$. It has the three possible values (*high, medium* and *low*), defined in Figure 1.2(c). Its input value, $YPos(S^i)$, is the position of the middle point of the segment in the vertical axis of the image limited to the interval $[0, 1]$. It is calculated using Equation 1.3. Low values indicates that the segment has its middle point in the upper part of the image and vice versa.

$$YPos(S^i) = \frac{y_0^i + y_1^i}{2N}. \tag{1.3}$$

Using these linguistic variables we can finally define the fuzzy concepts VS and HS. The fuzzy sets related to the concept VS are shown in Figure 1.2(d) and are identical to the fuzzy sets related to the concept HS. The two rule bases shown in Table 1.1 are used to calculate the membership degree of a segment S^i to the concepts VS and HS in the range $[0, 1]$. Both values are calculated by a fuzzy inference process and its corresponding defuzzification.

We shall denote the membership degree of a segment S^i to the fuzzy concepts VS and HS by $VS(S^i)$ and $HS(S^i)$ respectively.

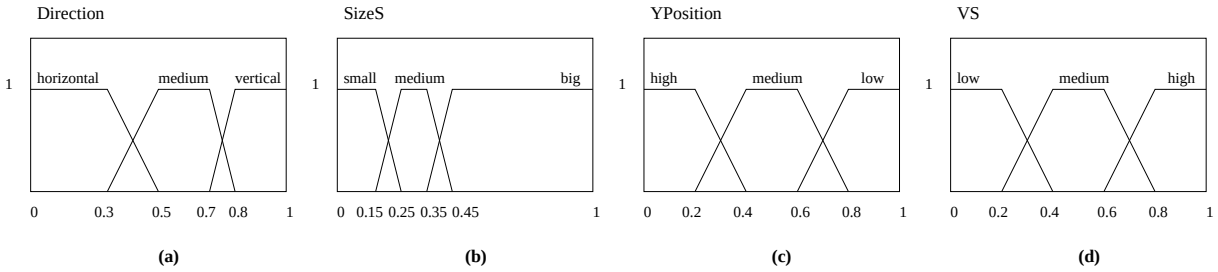


Figure 1.2: Linguistic variables for segment classification

IF			THEN	
Direction	SizeS	YPosition	VS(S^i)	HS(S^i)
horizontal	small	high	low	high
horizontal	medium	high	low	high
horizontal	big	high	low	high
medium	small	high	low	medium
medium	medium	high	low	high
medium	big	high	low	high
vertical	small		medium	low
vertical	medium		high	low
vertical	big		high	low

Table 1.1: Rule bases for classification of segments in vertical or horizontal

The selection of candidate segments for the following phase is made based on the membership degrees $VS(S^i)$ and $HS(S^i)$. Given a segment S^i, it is used in the following phase if: either $VS(S^i)$ or $HS(S^i)$ exceed a certain threshold α_1. In this case S^i is classified as vertical segment if $VS(S^i) > HS(S^i)$ and horizontal segment in the other case. Let us denote the set of vertical segments selected as $V = \{V^0, ..., V^n / VS(S^i) > \alpha_1 \wedge VS(S^i) > HS(S^i)\}$ and the horizontal one as $H = \{H^0, ..., H^n / HS(S^i) > \alpha_1 \wedge VS(S^i) \leq HS(S^i)\}$.

In the Figure 1.3(a) there is shown an image with a door in it. In Figure 1.3(b) there are all the segments extracted from the scene. The center column (Figures 1.3(c) and 1.3(d)) shows the vertical and horizontal segments selected for $\alpha_1 = 0.4$ and on the right column (Figures 1.3(e) and 1.3(f)) for $\alpha_1 = 0.7$. As it can be seen, higher values of α_1 reduces the number of segments for the next phase thus reducing the computing time required. Nevertheless, the segments that we could be looking for might not have the highest values. Therefore, the proper selection of α_1 is a non trivial

problem and this value is going to be automatically selected in the tuning process that will be explained in Section 1.4.

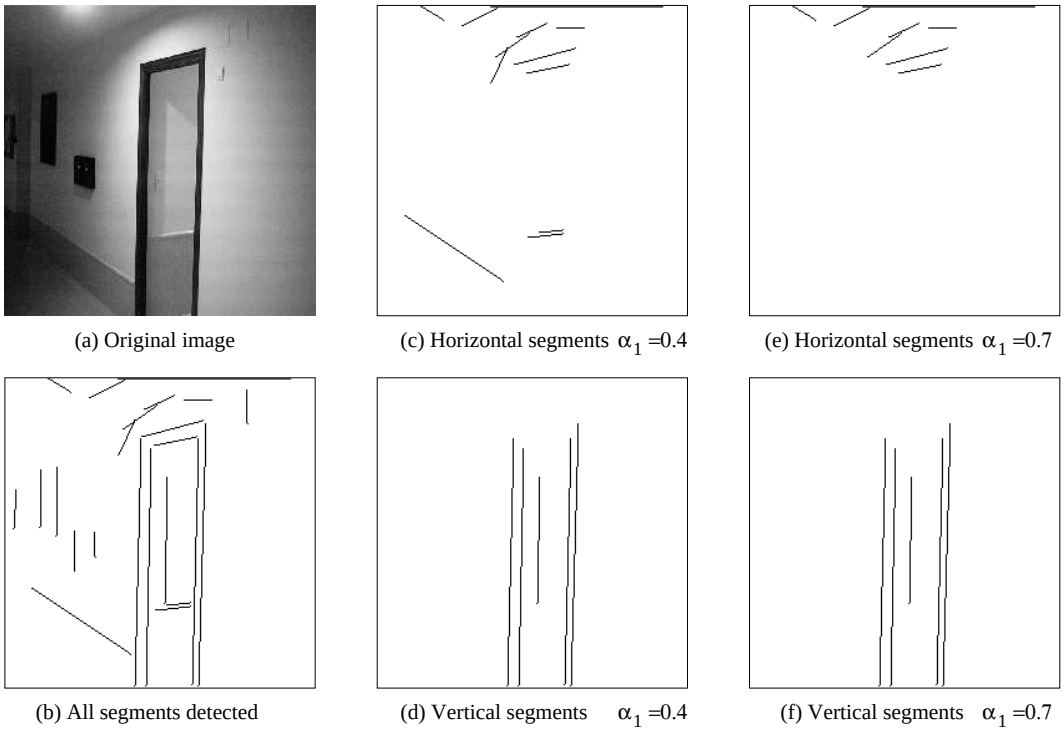


Figure 1.3: Classification of segments as vertical or horizontal for different α_1

1.3.2 Frame Edge

Next step is to analyze if there is any frame edge (FE) of a doorframe in the set of extracted segments. The concept FE expresses that there exists an horizontal segment joined in its upper part to two vertical segments (see Figure 1.1(a)). Therefore, we choose among the segments selected in the previous phase, one horizontal segment for each pair of vertical ones and the trio is analyzed. The horizontal and vertical concepts (HS and VS) have been previously calculated. In order to evaluate if the trio is properly joined, we define the fuzzy concept *Frame Edge Cohesion* (FEC) taking into account the distances between their extreme points as it is explained below.

The process starts selecting for each horizontal segment $H^i \in H$ a pair of vertical segments $\{V^j, V^k\} \in V$ and the trio is analyzed. Let us denote the trio by $F^i = \{L^i, Sup^i, R^i\}$, being $L^i \in V$ the leftmost vertical segment, $Sup^i \in H$ the horizontal segment and $R^i \in V$ the rightmost vertical segment of the trio.

If the trio were part of a frame edge, the upper points of the vertical segments should be very close to the extreme points of the horizontal one. Furthermore, the vertical segments should not be too closed. Both distances should be modeled taking into account that there can be errors in the segment extraction that makes appear the points of the vertical and horizontal segments not completely joined. There also must be considered that the distance between the vertical segments can vary depending on the perspective and the distance under which the door is seen. There have been defined two linguistic variables that measure these distances. $SegDistVH(F^i)$ measures the distance between the vertical segments and the horizontal one and $SegDistVV(F^i)$ measures the distance between the vertical segments. We denote Sup_l^i to the leftmost extreme point of the segment Sup^i and Sup_r^i to the rightmost one. Likewise we denote L_u^i to refer to the uppermost extreme point of the segment L^i and similarly R_u^i. In Figure 1.4(a) are depicted these points.

The linguistic variable $SegDistVH(F^i)$ can have the five possible values represented by the fuzzy sets shown in Figure 1.5(a) whose input value, $distVH(F^i)$, is calculated using the Equation 1.4. This equation calculates the maximum of the distances between the upper extreme points of each vertical segment and the horizontal one. The value is escalated to the range $[0, 1]$. The value 0 indicates that both elements are not separated and the value 1 that are completely separated. We have decided to use the maximum operator to force a small distance between both extreme points. These distances are shown in Figure 1.4(b).

$$distVH(F^i) = \frac{max\{dist(Sup_l^i, L_u^i), dist(Sup_r^i, R_u^i)\}}{N\sqrt{2}} \tag{1.4}$$

In order to evaluate the distance between the vertical segments, the linguistic variable $SegDistVV(F^i)$ is used. This variable has the five possible values represented in Figure 1.5(b). Its input value, $distVV(F^i)$, is calculated using Equation 1.5. It escalates the distance between the upper points of both segments to the range $[0, 1]$. This distance is also shown in Figure 1.4(b).

$$distVV(F^i) = \frac{dist(L_u^i, R_u^i)}{N\sqrt{2}} \tag{1.5}$$

The membership degree $FEC(F^i) \in [0, 1]$ of F^i to the fuzzy concept FEC is calculated using the rule base of Table 1.2 by a fuzzy inference process and its corresponding defuzzification. The fuzzy sets related to the possible values of the concept FEC are shown in Figure 1.5(c).

$FEC(F^i)$ indicates that the separation between the segments of F^i is appropriate to belong to a frame edge, but it does not take into account the corresponding membership degree of each individual segment to the concepts

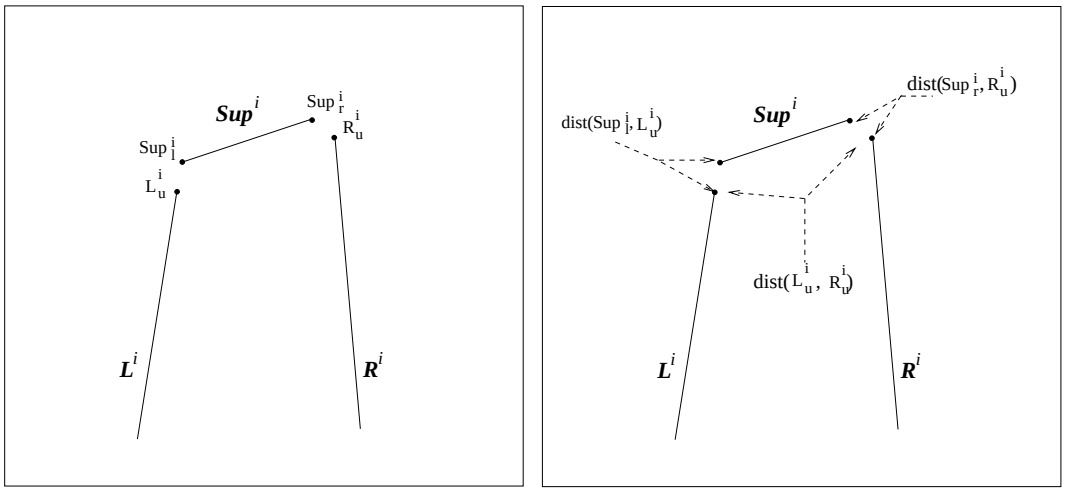


Figure 1.4: (a) Points used to calculate $SegDistVV(F^i)$ and $SegDistVH(F^i)$ (b) Distances used to calculate $distVV(F^i)$ and $distVH(F^i)$

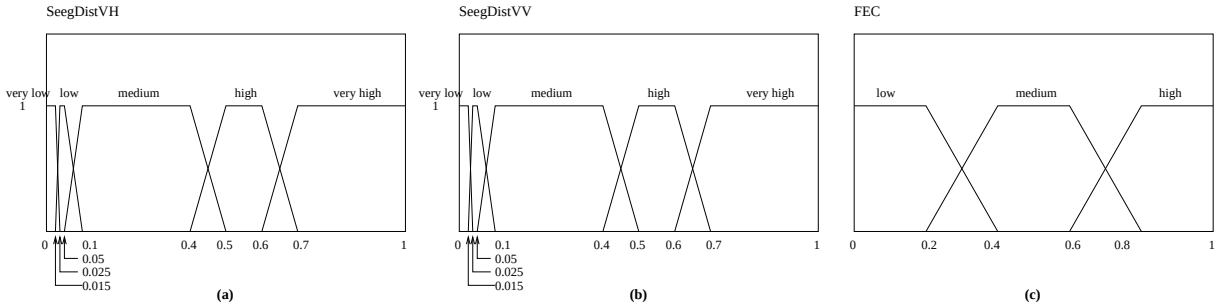


Figure 1.5: Linguistic variables for detecting frame edges

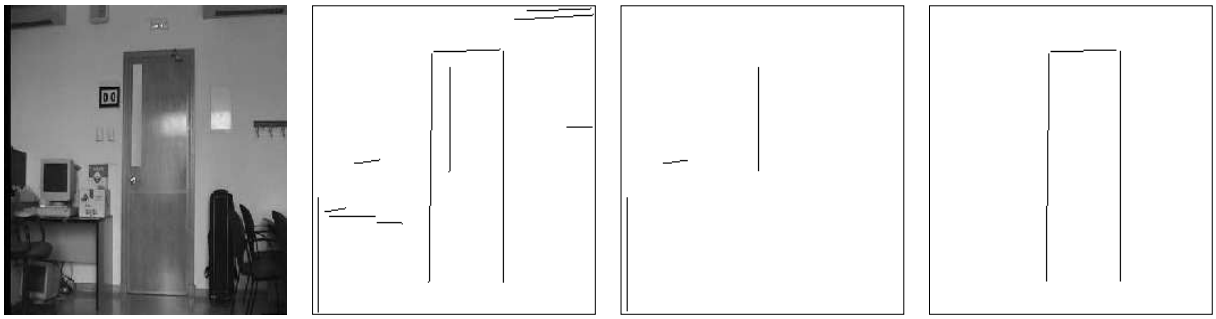
VS and HS. Finally, the membership degree $FE(F^i) \in [0,1]$ of F^i to the fuzzy concept FE is calculated as expressed in Equation 1.6. $FE(F^i)$ does not only expresses how well the trio forms a frame edge, but also how well its segments accomplish the corresponding fuzzy concepts HS and VS.

$$FE(F^i) = min\{FEC(F^i), VS(L^i), HS(Sup^i), VS(R^i)\} \qquad (1.6)$$

Only those F^i whose membership degree $FE(F^i)$ exceeds a threshold α_2 are used in the following phase. Let us denote this set as $F = \{F^0, ..., F^n / FE(F^i) > \alpha_2\}$. In the Figure 1.6 there is shown an image and the frame edges detected for different membership degrees to the fuzzy concept FE. As in the previous case, the selection of the parameter α_2 is going to be automatically selected in the tuning process.

$SegDistVV(F^i)$	$SegDistVH(F^i)$				
	VL	L	M	H	VH
VL	L	L	L	L	L
L	H	M	M	L	L
M	H	H	M	L	L
H	H	M	L	L	L
VH	M	L	L	L	L

Table 1.2: Rule base for linguistic variable $FEC(F^i)$



(a) Original image (b) Segments selected α_1 =0.6 (c) FE detected for FE(F^i) in [0,0.7] (d) FE detected for FE(F^i) in [0.7,1]

Figure 1.6: Frame Edges detected in a image

1.3.3 Complete Doorframe

In some cases, it is possible to see the two frame edges of a doorframe. For example when the door is seen open from the side that does not contains its leaf, or when the door is closed but the color of its leaf is different from the color of the doorframe. Therefore, the set of frame edges F selected in the previous phase is analyzed in order to see if two of them belong to the same doorframe. The fuzzy concept *Complete Doorframe* (CDF) evaluates if two frame edges F^i and F^j belong to the same doorframe (see Figure 1.1(b)). The previous phase allows the detection of frame edges F^i and F^j. The new fuzzy concept, *Frame Edges Similarity* (FES), evaluates the degree in which two frame edges are parallel and near as it is explained below.

If two frame edges belong to the same doorframe then their segments should be parallel and be relatively near. We define the linguistic variable $FEDist(F^i, F^j)$ to measure the distance between the two frame edges F^i and F^j. This variable has the five possible values represented in the Figure 1.7(a). The input value for this variable, $distF(F^i, F^j)$, is calculated using Equation 1.7. It is limited to the range $[0, 1]$ and it measures the distance of two frame edges as the maximum distance between its homonyms segments.

$$distF(F^i, F^j) = max\{maxDist(L^i, L^j), maxDist(Sup^i, Sup^j), maxDist(R^i, R^j)\} \tag{1.7}$$

The function $maxDist(S^i, S^j)$ (see Equation 1.8) calculates the distance between two segments as the maximum of all the possible combinations of distances between their extreme points. The value is escalated to the range $[0, 1]$ dividing by the maximum possible distance.

$$maxDist(S^i, S^j) = \frac{max\{dist(p_0^i, p_0^j), dist(p_0^i, p_1^j), dist(p_1^i, p_0^j), dist(p_1^i, p_1^j)\}}{N\sqrt{2}} \tag{1.8}$$

On the other hand, the linguistic variable $Paralellism(F^i, F^j)$ (whose possible values are shown in Figure 1.7(b)) measures the parallelism grade between two frames edges. In order to calculate the input value for this variable we need to express the inclination of a segment by the variable S_ϕ^i defined in Equation 1.9.

$$S_\phi^i = \arctan\left(\frac{y_1^i - y_0^i}{x_1^i - x_0^i}\right) \tag{1.9}$$

The input value for the variable $Paralellism(F^i, F^j)$ is calculated as expressed in Equation 1.10. The term $cos(S_\phi^a - S_\phi^b)$ measures the parallelism of two segments S^a and S^b in the range $[0, 1]$. If cosine is 0 it means that the segments forms an angle of $\frac{\pi}{2}$ radians and if it is 1 means that there is no difference in the angle between them. Therefore, Equation 1.10 expresses the minimum grade of the parallelism between the homonyms segments of two FEs.

$$parallelismF(F^i, F^j) = min\{|cos(L_\phi^i - L_\phi^j)|, |cos(Sup_\phi^i - Sup_\phi^j)|, |cos(R_\phi^i - R_\phi^j)|\} \tag{1.10}$$

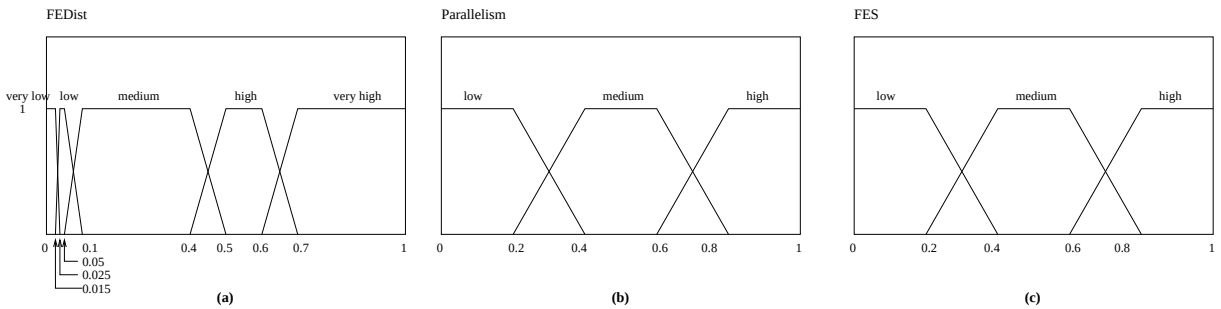


Figure 1.7: Linguistic variables for detecting complete doorframes

Using those fuzzy variables, we define $FES(F^i, F^j) \in [0, 1]$ (see Figure 1.7(c)) to express the membership degree of F^i and F^j to the fuzzy concept FES. The concept FES expresses the degree in which two frame edges are parallel and near taking into account its relative distance and parallelism. The rule base of the Table 1.3 is used for that purpose. As in previous cases, the value is calculated by a fuzzy inference process and its corresponding defuzzification.

Finally, the membership degree $CDF(F^i, F^j) \in [0, 1]$ of the two frame edges to the fuzzy concept CDF is calculated using Equation 1.11. CDF takes into account either the individual membership of each frame edge to the concept FE and the membership of the pair to the concept FES. If the membership degree $CDF(F^i, F^j)$ exceeds a certain threshold α_3 both frame edges are considered to be belong to the same doorframe. The appropriate value for α_3 will be automatically calculated in the tuning process. In Figure 1.8 there are shown several images of doors and below there are the complete doorframes detected in them.

$Paralellism(F^i, F^j)$	$FEDist(F^i, F^j)$				
	VL	L	M	H	VH
L	L	L	L	L	L
M	M	M	M	M	L
H	H	H	H	M	L

Table 1.3: Rule base for linguistic variable $FES(F^i, F^j)$

$$CDF(F^i, F^j) = min\{FES(F^i, F^j), FE(F^i), FE(F^j)\} \tag{1.11}$$

1.3.4 Frame Edge with Evidences

In some situations, the two frame edges of a doorframe can not be seen complete because there is a missing segment. It may happen, for example, when the door is seen open and a lateral bar of the door is hidden by its leaf. In that case, it is usual to find evidences of the incomplete frame edge around a previously detected one. This situation has been considered by the fuzzy concept *Frame Edges with Evidence* (FEE) (see Figure 1.1(c)). The evidence is a pair of connected vertical and horizontal segments (let us call it *junction*) that is near and relatively parallel to a detected frame edge. Therefore, the first step consist in detecting the set of junctions in the image. Then, the distance and parallelism between junctions and frame edges are analyzed to check if they belong to the same doorframe. Junctions are searched among the segments that have no been selected as belonging to any frame edge in the previous phases. The next two sections explains this

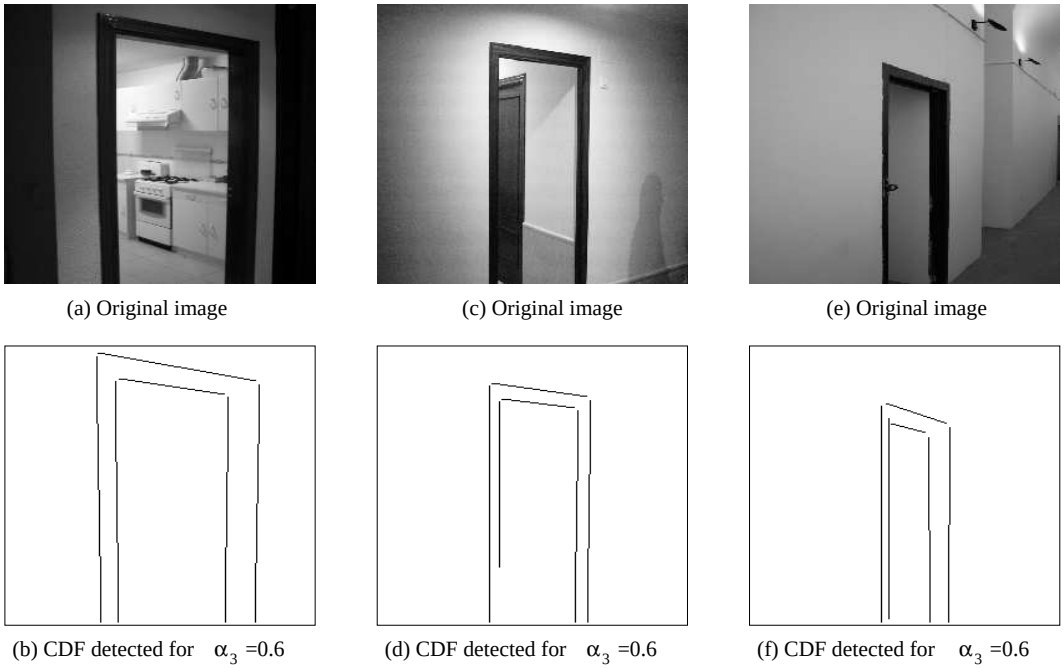


Figure 1.8: Complete Doorframes detected in images

process. Section 1.3.5 explains how the junctions are detected and Section 1.3.6 explains how the analysis of the distance and parallelism is performed.

1.3.5 Junction Detection

In order to detect the set of junctions in an image, the fuzzy concept *Junction* (JC) is defined. JC expressed that a pair of segments (one vertical and one horizontal) are joined in its upper part. The vertical and horizontal concepts have been previously computed (VS and HS). In order to evaluate that the distance between the extreme points is appropriate, the fuzzy concept *Junction Cohesion* (JCC) is defined as it is explained below.

We analyze if any segment $S_l \in V$ (that has not been selected as part of an frame edge) is joined to any remaining segment $S_h \in H$ measuring the distance between them. Let us denote a junction by $J^i = \{S_l, S_h / S_l \in V \wedge S_h \in H\}$. In order to measure the distance between the segments, the linguistic variable $DistJ(J^i)$ whose fuzzy sets are identical to those shown in Figure 1.7(a) is used. Its input value, $minDist(S^i, S^j)$, calculates the distance between two segments as the minimum of all the possible combinations of distances between its extreme points (see Equation 1.12) . If $minDist(S_l, S_h) = 0$ it means that there is no distance between two of the extreme points of the segments. If this value is 1 it means that the distance is maximum.

$$minDist(S^i, S^j) = \frac{min\{dist(p_0^i, p_0^j), dist(p_0^i, p_1^j), dist(p_1^i, p_0^j), dist(p_1^i, p_1^j)\}}{2\sqrt{N}} \tag{1.12}$$

The membership degree $JCC(J^i) \in [0,1]$ of J^i to the fuzzy concept JCC is calculated using the rule base of Table 1.4 by a fuzzy inference process and its corresponding defuzzification. The fuzzy sets related to the possible values of JCC are identical to those shown in Figure 1.7(c). Finally, the membership degree $JC(J^i) \in [0,1]$ of J^i to the fuzzy concept JC is calculated using Equation 1.13. The concept JC takes into account either the membership degree of each individual segment to the previously defined concepts HS and VS and to fuzzy concept JCC. If this value exceeds a threshold α_4 then J^i is considered for the next phase. Let us call the set of those junctions $J = \{J^0, ..., J^n / J^i = \{S_l^i, S_h^i\} \wedge JC(J^i) > \alpha_4\}$. The value α_4 will be automatically selected in the tuning process.

IF	THEN
$DistJ(J^i)$	$JCC(J^i)$
VL	H
L	M
M	L
H	L
VH	L

Table 1.4: Rule base for linguistic variable $JCC(J^i)$

$$JC(J^i) = min\{JCC(J^i), V(S_l), H(S_h)\} \tag{1.13}$$

In Figure 1.9(a) there is a door whose two frames edges can not be detected because the opened leaf hides one of them. In Figure 1.9(b) there are depicted all the segments selected for $\alpha_1 = 0.6$ and in Figure 1.9(c) there is the unique frame edge detected for $\alpha_2 = 0.6$. As it can be seen, there are not the segments that form the second frame edge but there are evidences of it around the detected one. In Figures 1.9(d), 1.9(e) and 1.9(f) there are shown the junctions detected for different values of $JC(J^i)$.

1.3.6 Frame Edge and Junction Analysis

Next step is to analyze if there is a J^i that is part of a doorframe. If so, it must be near a detected frame edge F^j and relatively parallel. The fuzzy concept *Frame Edge with Evidence* (FEE) expresses that a junction and a frame edge are part of the same doorframe. Both frame edge and junction has been previously detected (FE and JC). A new fuzzy concept, *Frame and*

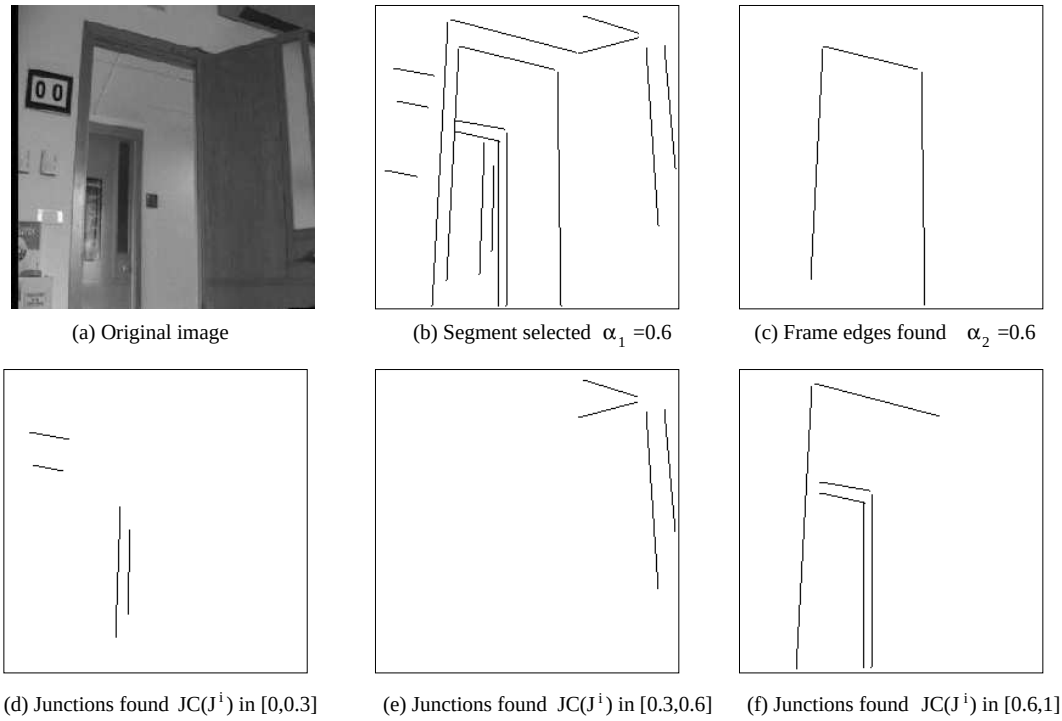


Figure 1.9: Examples of junctions detected with different membership degrees

Junction Cohesion (FJC), analyzes if they are near and parallel analyzing the distance and parallelism between them as it is explained below.

Analyzing a junction, it can be seen that there can be two possible situations. The vertical segment can be either at the left or at the right side of the horizontal segment. In the former case, the junction should belong to the left side of the frame, thus it should be placed at the left side of an existing frame edge and vice versa. If a junction J^i is part of the same doorframe than F^j then its segments must be parallels. In order to measure the parallelism between J^i and F^j the linguistic the variable $JParallelism(J^i, F^j)$ is used. Its fuzzy labels are identical to those shown in Figure 1.7(b). The input value for this variable is $parallelismJ(J^i, F^j)$ and it is calculated using Equation 1.14. It measures the parallelism between the junction and the frame edge in the range $[0, 1]$. If $parallelismJ(J^i, F^j) = 0$ indicates that there is an angle of $\frac{\pi}{2}$ radians between any two segments and if it is 1 indicates that there is no difference in the angle between the segments. $S^i_{l\phi}$ and $S^i_{v\phi}$ are the angles of the segments of J^i calculated as expressed in Equation 1.9.

$$parallelismJ(J^i, F^j) = \begin{cases} max\{|cos(S^i_{h\phi} - Sup^j_\phi)|, |cos(S^i_{l\phi} - L^j_\phi)|\} & if\ S^i_l\ is\ at\ left\ side\ of\ S^i_h \\ max\{|cos(S^i_{h\phi} - Sup^j_\phi)|, |cos(S^i_{l\phi} - R^j_\phi)|\} & otherwise \end{cases} \tag{1.14}$$

If J^i belongs to the same doorframe than F^j then the distance between them must be relatively small. Therefore the linguistic variable $DistJF(J^i, F^j)$ (whose fuzzy labels are identical to those shown in Figure 1.7(a)) is used in order to measure the distance between J^i and F^j. The input value for this variable is $distJ(J^i, F^j) \in [0, 1]$ and it is calculated using Equation 1.15. It measures the distance between the corresponding segments of the junction and the frame edge. The value $distJ(J^i, F^j) = 0$ indicates that there is no distance between J^i and F^j and the value $distJ(J^i, F^j) = 1$ indicates that the distance is the maximum possible.

$$distJ(J^i, F^j) = \begin{cases} \frac{max\{minDist(S^i_h, Sup^j), minDist(S^i_l, L^j)\}}{2\sqrt{N}} & if\ S^i_l\ is\ at\ left\ side\ of\ S^i_h \\ \frac{max\{minDist(S^i_h, Sup^j), minDist(S^i_l, R^j)\}}{2\sqrt{N}} & otherwise \end{cases} \tag{1.15}$$

With those variables we establish a membership degree, $FJC(J^i, F^j) \in [0, 1]$ on the fact that J^i is an evidence of the existence of a doorframe with frame edge F^j. The value is calculated by a fuzzy inference process using the rule base of Table 1.5 and its corresponding defuzzification. Finally, $FJC(J^i, F^j)$ is employed to calculate a membership degree $FEE(J^i, F^j) \in [0, 1]$, on the fact that the F^i and J^i belong to the same doorframe taking into account the membership degree of each element using Equation 1.16.

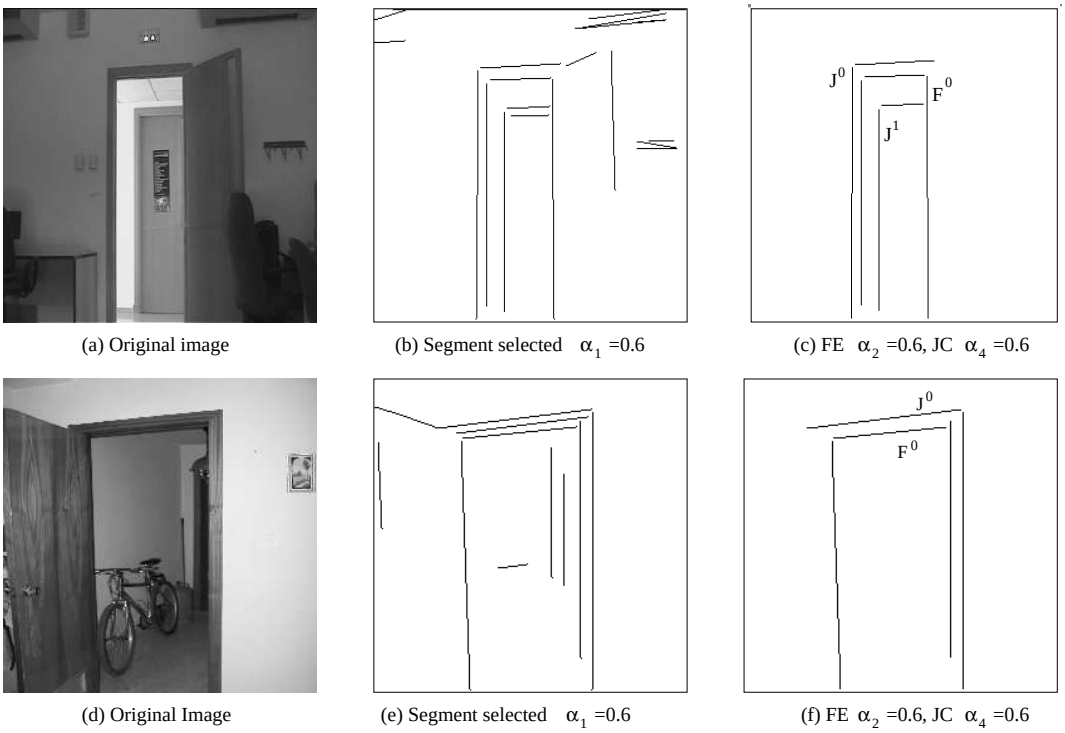
$JParallelism(J^i, F^j)$	$DistJF(J^i, F^j)$				
	VL	L	M	H	VH
L	L	L	L	L	L
M	M	M	M	M	L
H	H	H	H	M	L

Table 1.5: Rule base for linguistic variable $FJC(J^i, F^j)$

$$FEE(J^i, F^j) = min\{FJC(J^i, F^j), FE(F^i), JC(J^i)\} \tag{1.16}$$

Figure 1.10 shows two images in the left column. In the center column there are all the segments detected for $\alpha_1 = 0.6$. In the right column there are the junctions detected in each image labeled as J^i and the frame edges labeled as F^j. In Figure 1.10(c) there appears two junctions with $\alpha_4 = 0.6$ but only J^0 belongs to the door represented by the frame edge F^0. The value of $FEE(J^0, F^0) = 0.82$ while $FEE(J^1, F^0) = 0.3$. In Figure 1.10(f) there are only one junction J^0 and one frame edge F^0 and $FEE(J^0, F^0) = 0.74$.

As in previous cases, if $FEE(J^i, F^j)$ exceeds a threshold value α_5, the pair $\{J^i, F^j\}$ is considered as belonging to the same doorframe. The appropriate value for α_5 will be automatically selected in the tuning process.



(a) Original image (b) Segment selected α_1 =0.6 (c) FE α_2 =0.6, JC α_4 =0.6

(d) Original Image (e) Segment selected α_1 =0.6 (f) FE α_2 =0.6, JC α_4 =0.6

Figure 1.10: Images with frame edges and junction evidences

1.4 Tuning of the System

In this section it is explained how the tuning process has been carried out. The tuning process consists in adjusting the membership functions of the fuzzy sets and selecting appropriates values for $\{\alpha_1, ..., \alpha_5\}$. The aim of this process is to adjust the visual fuzzy system to the range of sizes, orientations and heights at which the robot has to detect the doors of its environment.

We have decided to use a GA for tuning the system instead of classical tuning approaches (Moreno-Velo et al. 2003) because it is possible to consider the whole visual fuzzy system as a unique system to optimize. Therefore, it is only necessary to define one error function to evaluate the performance of the whole system, instead of five error functions (one for each of the fuzzy systems). GAs are search algorithms based on natural evolution (Goldberg 1989; Holland 1975). They consist in the creation of an initial random population of *chromosomes* (each one representing a possible solution) that evolves over time in a controlled way looking for good solutions for

a specific problem (Goldberg 2002). GAs have been successfully applied to fuzzy systems (Cordon et al. 2004; Cordon et al. 2001) either for creating the knowledge database or tuning its fuzzy labels. The tuning process consists in adapting the original system by moving, stretching and narrowing its labels. The GA selected for the tuning process is *CHC* (Eshelman 1991), an evolutionary algorithm that introduces an appropriate trade-off between diversity and convergence. We have opted for a real coding of the chromosomes because it represents the problem in a more natural way than using binary coding.

A large database of images with and without doors, taken from different angles and distances at the fixed height of our camera, has been created. The performance of a visual fuzzy system is evaluated and used as fitness function for the GA. Below, the basis of the CHC algorithm and the proposed coding scheme are explained.

1.4.1 CHC

The CHC algorithm (Eshelman 1991) is an evolutionary approach that introduces an appropriate trade-off between diversity and convergence. For that purpose, it uses a high selective pressure based on an elitist scheme in combination with a highly disruptive crossover and a re-start when the population is stagnated. It is based on four distinguishing components:

- Elitist selection. The new population is composed of the best individuals of the parent and offspring populations.
- Uniform crossover operator. The original algorithm was designed to be used with binary encoding. For real coding we have used the $BLX\alpha$ operator (Eshelman y Schaffer 1993).
- Incest prevention. Two parents are not crossed over if they are too close. This ensures diversity.
- Re-start. When the population reaches an stagnated state, it is restarted keeping the best individual.

The pseudo-code of Figure 1.11 shows the algorithm. In a first step, the population is created using a perturbation operator over the initial chromosome. In our case, the initial chromosome is created from the visual fuzzy system previously explained (based on expert knowledge). Then, the algorithm measures the mean distance of the population (D_{mean}) in order to estimate when two individuals are too closed to be crossed and in this way avoid a possible incest. Incest prevention forces to cross separated elements causing an exploration of the area covered by the individuals of the population. D_{mean} is decremented when the incest prevention mechanism

CHC:

1. P_0=Create initial population.
2. D_{mean}=Calculate mean distance of the population.
3. D_{max}=Calculate maximum distance between individuals.
4. $Decrement = DecFactor * D_{max}$.
5. For $i = 0$ to number of desired iterations.

    5.1 Pair up randomly the individuals of the population to be used as parents of the new offspring population.
    5.2 Create a new offspring population $Child_i$ with M individuals ($M \leq N$) using BLX_α. If the distance between two parents is smaller than D_{mean} do not generate the offspring.
    5.3 Replace the individuals of P_i with the individuals of $Child_i$ that are better than them.
    5.4 If no new offspring are generated $D_{mean} = D_{mean} - Decrement$.
    5.5 If $D_{mean} < 0$ re-start keeping the best individual and recalculate D_{mean} and $Decrement$.

Figure 1.11: CHC evolutionary algorithm

makes impossible to generate new offspring. *DecFactor* allows us to select the degree of convergence of the algorithm, the smaller it is, the greater level of convergence is allowed and vice versa. When D_{mean} is below zero, the population is re-started keeping the best individual found. The new population is generated using the perturbation operator on the best individual with probability 35%. It means that only the 35% of the chromosome is altered.

1.4.2 Coding scheme and operators

In our approach, the whole visual fuzzy system is represented using a single chromosome composed of the aggregation of its membership functions and the values $\{\alpha_1, ..., \alpha_5\}$. Each fuzzy variable is encoded based on the crossing points of its membership functions and the separation between them.

We shall denote the set of variables of the visual fuzzy system by $V = \{V_0, ..., V_n\}$ and let us denote the set of membership functions of the variable V_i as $L_i = \{L_i^0 ... L_i^{n+1}\}$. As our system is entirely composed by trapezoidal functions, a membership function can be defined by four parameters as in Equation 1.17, where $[Left_i, Right_i]$ is the range of the input variable V_i.

$$L_i^j = [a_i^j \ b_i^j \ c_i^j \ d_i^j] \ / \ a_i^j, b_i^j, c_i^j, d_i^j \in [Left_i, Right_i] \tag{1.17}$$

In order to reduce the search space, we limit the set of possible configurations forcing the membership functions to cross at level 0.5. This approach has been previously employed in (Karr 1991) with triangular labels. In our case, the use of this constraint forces to be accomplished the following re-

striction:

$$a_i^{j+1} = c_i^j \ \wedge \ b_i^{j+1} = d_i^j$$

Consequently, the crossing point P_i^j of two consecutive labels L_i^j and L_i^{j+1} can be calculated as expressed in Equation 1.18.

$$P_i^j = \frac{c_i^j + d_i^j}{2} = \frac{a_i^{j+1} + b_i^{j+1}}{2} \tag{1.18}$$

Therefore, for each variable V_i there is a set of n crossing points $P_i = \{P_i^0, .., P_i^n\}$ that have to be selected by the GA to obtain the maximum performance. A range of allowed positions $[Pleft_i^j, Pright_i^j]$ is defined for each crossing point P_i^j. This restriction has two purposes. On one hand, it limits the search space of a crossing point to a range around the initial system. On the other hand, it avoids either the relative displacement and the overlap of the membership functions given by the initial system. This range is calculated using Equation 1.19. In Figure 1.12(a) there can be seen an example of a fuzzy variable V_i with three membership functions $\{L_i^0, L_i^1, L_i^2\}$. The variable has two crossing points $\{P_i^0, P_i^1\}$. As it can be seen, the search space for each P_i^j is independent thus avoiding either the overlap and the relative displacement between the membership functions. The range of each crossing point is calculated at the start of the process and remain the same during the evolution of the population.

$$Pleft_i^j = \begin{cases} Left_i & if\ j = 0 \\ \frac{P_i^j + P_i^{j-1}}{2} & otherwise \end{cases} \qquad Pright_i^j = \begin{cases} Right_i & if\ j = n \\ \frac{P_i^j + P_i^{j+1}}{2} & otherwise \end{cases} \tag{1.19}$$

In order to increase the capability for adjusting the membership functions, a parameter that indicates the separation between them is used. It represents the separation of the upper points (c_i^j and b_i^{j+1}) of two consecutive membership functions L_i^j and L_i^{j+1} to their crossing point P_i^j. We shall call this parameter $s_i^j \in [0, 1]$ and it is calculated as expressed in Equation 1.20. When $s_i^j = 0$, there is no separation between c_i^j and b_i^{j+1} ($c_i^j = b_i^{j+1}$). On the other hand, if $s_i^j = 1$ it means that the separation between c_i^j and b_i^{j+1} is the maximum allowed by the limits $[Pleft_i^j, Pright_i^j]$ of the crossing point P_i^j. To clarify the utility of this parameter, Figure 1.12(b) shows the example of a crossing point P_i^0 and the corresponding membership functions for two different values of s_i^0. While the membership functions represented by the solid lines correspond to $s_i^0 = 0.5$, the membership functions represented by dashed lines correspond to $s_i^0 = 0.25$.

$$s_i^j = \frac{P_i^j - c_i^j}{min\{P_i^j - Pleft_i^j, Pright_i^j - P_i^j\}} = \frac{d_i^j - P_i^j}{min\{P_i^j - Pleft_i^j, Pright_i^j - P_i^j\}} \tag{1.20}$$

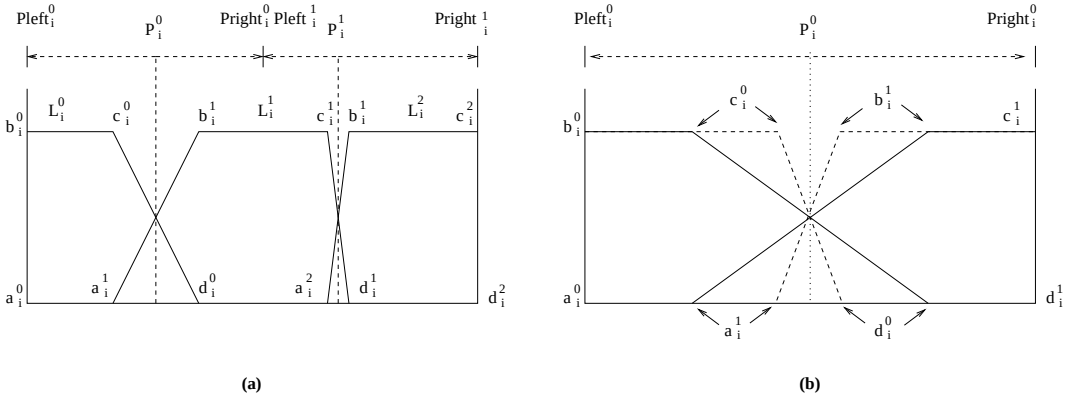


Figure 1.12: (a) A fuzzy partition with 3 labels (b) Two membership functions for different values of s_i^j

Tuning of trapezoidal membership functions using real encoding has also been performed in (Herrera et al. 1995), but the approach explained in this work reduces the number of parameters employed for each variable and thus the search space for the evolutionary algorithm is reduced. Using this coding scheme, a variable V_i can be represented as $\{P_i^0, s_i^0, ..., P_i^n, s_i^n\}$. The inverse step, given a variable V_i represented by its encoding obtaining its membership functions, is made using Equation 1.21.

$$L_i^j = \begin{cases} a_i^j = Left_i & b_i^j = Left_i & c_i^j = P_i^j - s_i^j MaxRange_i^j & d_i^j = P_i^j + s_i^j MaxRange_i^j & if\ j = 0 \\ a_i^j = c_i^{j-1} & b_i^j = d_i^{j-1} & c_i^j = Right_i & d_i^j = Right_i & if\ j = n \\ a_i^j = c_i^{j-1} & b_i^j = d_i^{j-1} & c_i^j = P_i^j - s_i^j MaxRange_i^j & d = P_i^j + s_i^j MaxRange_i^j & otherwise \end{cases} \tag{1.21}$$

A complete visual fuzzy system is represented by a chromosome C_i. It is encoded aggregating the set of variables and $\{\alpha_1, ..., \alpha_5\}$ as in Equation 1.22. In order to ease the notation, we will also denote the elements of each C_i by $(c_i^0, ..., c_i^n)$.

$$C_i = (\alpha_1, \alpha_2, \alpha_3, \alpha_4, \alpha_5, P_0^0, s_0^0, ..., P_0^n, s_0^n, ..., P_n^0, s_n^0, ..., P_n^n, s_n^n) \tag{1.22}$$

Either a crossover and a perturbation operator are required to use CHC. BLX_α (Eshelman y Schaffer 1993) has been selected as crossover operator. This operator works generating a random value in an extended range (given by a parameter α) of the parents. Let us suppose that we want to cross two chromosomes $C_a = (c_a^0, ..., c_a^n)$ and $C_b = (c_b^0, ..., c_b^n)$ to obtain an offspring $C_o = (c_o^0, ..., c_o^n)$. The BLX_α operator generates for each c_o^i a random value in a extended range $[BLX_{Inf}^i, BLX_{Sup}^i]$ given by the parent values c_a^i and c_b^i as shown in Figure 1.13. Nevertheless, we must remember that

the element c_o^i can not take any value, their values are limited due to the coding scheme employed. In order to keep the coherence in the values of the offspring, the range $[BLX_{Inf}^i, BLX_{Sup}^i]$ can not exceed the range imposed by the coding scheme for each element c_o^i. It means that if c_o^i is either an $\{\alpha_1, ..., \alpha_5\}$ or a s_k^j parameter then $[BLX_{Inf}^i, BLX_{Sup}^i]$ must not exceed the range $[0, 1]$. Similarly, if c_o^i is a crossing point P_j^k then $[BLX_{Inf}^i, BLX_{Sup}^i]$ must not exceed the range $[Pleft_j^k, Pright_j^k]$. Therefore, if some of the limits calculated by BLX_α exceeds the limits imposed by the coding scheme it is truncated to a valid value. On the other hand, the perturbation operator of the CHC algorithm must alter the 35% of a chromosome C_a. It has been implemented by selecting randomly the 35% of the elements of C_a. For each c_a^i selected, it is assigned a random value in the range of possible values according to the coding scheme.

Figure 1.13: $BLX\alpha$ operator.

An evaluation data set and a fitness function are required in order to run CHC. There has been created a database that contains 401 images with doors seen at different distances and angles captured with the camera placed on our Nomad 200 mobile robot. The database also contains 139 images of the environment where no doors are present. Let us denote by $I = \{I^0, ..., I^n\}$ the images in the database. For each image, all the segments have been extracted and those segments that belong to the doorframe have been manually labeled. We shall denote by IS^i all the segments extracted in the image I^i and by $DS^i \in IS^i$ the segments that have been manually labeled as belonging to the doorframe. When IS^i is passed to a visual fuzzy system, it returns only the set of segments that considers belonging to a door. Let us denote them by SS^i. If the system correctly classifies all the segments then $SS^i = DS^i$. Otherwise, there may be a subset of SS^i with segments that really belong to the doorframe, and there may be another subset of SS^i with segments that have been incorrectly considered as belonging to a doorframe. Let us denote the former subset by $CC^i = \{SS^i \bigcap DS^i\}$ and the latter by $IC^i = \{SS^i - DS^i\}$. There has been defined a fitness function in the range $[0, 1]$ (see Equation 1.23) to evaluate the individuals of the population in the evolution process. $Fitness(C^j)$ evaluates an individual C^j according to the number of segments that correctly and incorrectly returns.

$$Fitness(C^j) = \lambda \frac{1}{n} \sum_{i=0}^{n} \frac{|CC^i|}{|DS^i|} + (1 - \lambda) \frac{1}{n} \sum_{i=0}^{n} \left(1 - \frac{|IC^i|}{|IS^i - DS^i|}\right) \quad (1.23)$$

$Fitness(C^j)$ is composed by two sums. The first one has value 1 when C^j correctly returns all the segments of the database that really belong to doorframes. If the first sum is 0, it means that C^j does not returns any correct segment. The second sum is 1 when no incorrect segments are returned by C^j and it is 0 in the opposite case. Therefore, if $Fitness(C^j) = 1$, it means that C^j represents an optimal visual fuzzy system and vice versa. The parameter λ is used to weight independently the importance of a correct detection and the importance of avoiding false doors.

1.5 Experimental Results

In this section, both the performance of the proposed method and the results of the tuning process are shown. We shall note that the fitness function used in the evolution process (defined in Equation 1.23) does not gives an intuitive idea of the performance of a visual fuzzy system. It measures the number of successfully or incorrectly detected segments. Instead of that, the performance of a fuzzy system is going to be expressed in this section based on the number of successfully or incorrectly detected doors.

The initial visual fuzzy system (based on expert knowledge) has a *Total Success Fraction* (TSF) of 0.8363. TSF considers either the success in the images with doors and without doors. Therefore, it means that the systems successfully detects the presence or absence of doors in the 83.63% of the images of the database. Nevertheless, a separate analysis of both cases is important because it can help us to understand the performance of the system in the different situations. The *True Positive Fraction* of the initial system (TPF) is 0.7755, it means that the system detects the doors in the 77.55% of the images with doors. The *False Positive Fraction* (FPF) is 0.0287, it means that the system detects false doors in the 2.87% of the images that does not contain doors. These three values (TSF, TPF and FPF) are going to be employed to analyze each one of the genetically tuned visual fuzzy systems.

In order to run the CHC algorithm it is necessary to set five parameters: the number of individuals of the population, the number of iterations of the algorithm, the parameter α of $BLX\alpha$ operator, $DecFactor$ and λ. For our experiments, the number of individuals have been set to 50, the number of iterations to 300 and α to 0.3. On the other hand, in order to specify different values of convergence for the algorithm, two different values of $DecFactor$ have been used: 0.008 and 0.08. Finally, five different values have been employed for λ: $0.05, 0.25, 0.5, 0.75$ and 0.95. Consequently, the CHC algorithm has been executed ten times for these different combinations of $DecFactor$ and $lambda$. In Table 1.6 there can be seen the results of the experiments.

Among all the generated solutions, we consider as the best visual fuzzy

$DecFactor$	λ	TPF	FPF	TSF
0.008	0.05	0.8628	0.0719	0.8980
0.008	0.25	0.9351	0.1294	0.9176
0.008	0.5	0.8428	0.3597	0.8757
0.008	0.75	0.8329	0.2158	0.8723
0.008	0.95	0.9226	0.1510	0.9027
0.08	0.05	0.8329	0.0000	0.8781
0.08	0.25	0.9700	0.1438	0.9392
0.08	0.5	0.9775	0.2014	0.9290
0.08	0.75	0.9775	0.2589	0.9134
0.08	0.95	0.8877	0.5395	0.7720

Table 1.6: Results of the tuning process.

system the generated for $DecFactor = 0.08$ and $\lambda = 0.25$. Although it has a relatively high FPF, it also has a high TPF and we consider that the solution has a good trade-off between those parameters. In Figure 1.14 and Figure 1.15 there are shown some of the images with doors of the database. The output segments of the tuned visual fuzzy system are superimposed in white. As it can be seen, the system detects correctly doors with strong perspective deformations. In Figure 1.16 there are shown several of the false positives generated by the visual fuzzy system with images of the database. Some of the false positives are due to the presence of objects whose rectangular shape is similar to frame edges. These false positives could be avoided with the inclusion of more information to the system, i.e., color or ultrasound information.

Finally we want to mention the time consumed by the final visual fuzzy system. The visual fuzzy process is performed on the Pentium IV (2.4 Ghz) laptop computer added to our Nomad 200 mobile robot. A temporal analysis of the system shows that the average time consumed to analyze images of size $320x280$ pixels is 160 ms. Therefore it is possible to perform the visual processing at a frame rate of 6 fps. This frame rate has proved to be enough for the use of the method in real-time applications in our mobile robot.

1.6 Conclusions and Future Work

In this work a visual door-detection technique based on fuzzy logic that is genetically tuned is presented. The technique can be used for map-building, navigation or positioning tasks in autonomous mobile robots. It is able to detect typical doors in grey-level images and can be used for real-time applications. Several fuzzy systems are employed to analyze the segments extracted from the image looking for doorframes in the three different situations represented by the fuzzy concepts: *Frame Edge, Complete Doorframe*

Figure 1.14: Correct detections of the visual fuzzy system (I)

and *Frame Edge with Evidences*.

The system created is able to detect doors under the strong perspective deformations caused by the two DOF allowed for the camera of our robot. The proposed method is valid for different image sizes because all the parameters have been set with independence of the size of the image.

The use of fuzzy logic has allowed to manipulate several concepts like *Vertical Segment*, *Horizontal Segment*, *Frame Edge Cohesion* or *Parallelism* in a natural way. It also allows to manipulate the ambiguity and uncertainty in the segment extraction. Furthermore, the use of fuzzy logic will allow to combine the information provided by the visual fuzzy system with the information provided by our fuzzy perceptual model based on ultrasound (Aguirre y González 2003).

All the parameters of the visual fuzzy system has been tuned using a GA. The system has been adapted to the range of distances and orientations at which our robot has to detect the doors of its environment. A large database of images has been used and the final tuned system obtains a success percentage of 93.92%.

As future work two aspects are pointed out. First, the integration of this visual fuzzy model with the previously developed method for detection of typical places of indoor environments based on ultrasound. The combination of both methods could be used to obtain an unique belief degree on the existence of a door. As second aspect, we consider that the use of color could aid the detection process. Color information can be used to discriminate between objects that has been selected as doors and discard possible false positives like paintings or cupboards.

Figure 1.15: Correct detections of the visual fuzzy system (II)

Figure 1.16: False positive detections of the visual fuzzy system

Acknowledgment

This work has been partially supported by the Spanish MCYT under Project TIC2003-04900. We would like to thank Dr. Oscar Cordon for his suggestions on the development of this paper.

Bibliography

[Adorni et al. 2001] Adorni, G., Cagnoni, S., Enderle, S., Kraetzschmar, G. K., Mordonini, M., Plagge, M., Ritter, M., Sablatnög, S. and Zell, A. 2001. Vision-based localization for mobile robots. *Robotics and Autonomous Systems*, 36(2-3):103–119.

[Aguirre y González 2002] Aguirre, E. and González, A. 2002. Integrating fuzzy topological maps and fuzzy geometric maps for behavior-based robots. *International Journal of Intelligent Systems*, 17(3):333–368.

[Aguirre y González 2003] Aguirre, E. and González, A. 2003. A fuzzy perceptual model for ultrasound sensors applied to intelligent navigation of mobile robots. *Applied Intelligence*, 19(3):171–187.

[Canny 1986] Canny, J. 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analsysis and Machine Intelligence*, 8:679–698.

[Cicirelli et al. 2003] Cicirelli, G., D'orazio, T. and Distante, A. 2003. Target recognition by component for mobile robot navigation. *Journal of Experimental and Theoretical Artificial Intelligence*, 15(3):281–297.

[Cordon et al. 2004] Cordon, O., and. F. Herrera, F. G., Hoffman, F. and Magdalena, L. 2004. Ten years of genetic fuzzy systems: current framework and new trends. *Fuzzy sets and systems*, 141:5–31.

[Cordon et al. 2001] Cordon, O., Herrera, F., Hoffman, F. and Magdalena, L. 2001. *Genetic Fuzzy Systems: Evolutionary Tuning and Learning of Fuzzy Knwoledge Bases*. World Scientific Publishing.

[Desouza y Kak 2002] Desouza, G. and Kak, A. 2002. Vision for mobile robot navigation: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24:237– 267.

[Eshelman 1991] Eshelman, L. 1991. The chc adaptive search algorithm: How to have safe search when engaging in nontraditional genetic recombination. En *First Workshop on Foundations of Genetic Algorithms*, pág. 265–283. Morgan Kaufmann.

[Eshelman y Schaffer 1993] Eshelman, L. and Schaffer, D. 1993. Real-coded genetic algorithms and interval-schemata. En *Second Workshop on Foundations of Genetic Algorithms*, pág. 187–202. Morgan Kaufmann.

[Foresti 2000] Foresti, G. 2000. A real-time hough-based method for segment detection in complex multisensor images. *Real-Time Imaging*, 6:93–111.

[Gasteratos et al. 2002] Gasteratos, A., Beltran, C., Metta, G. and Sandini, G. 2002. Pronto: A system for mobile robot navigation via cad-model guidance. *Microprocessors and Microsystems*, 26:17–26.

[Goldberg 1989] Goldberg, D. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley.

[Goldberg 2002] Goldberg, D. 2002. *The design of Competent Algorihmts: Steps Towards a Computational Theory of Innovation.* Kluwer Academic Publishers.

[Herrera et al. 1995] Herrera, F., Lozano, M. and Verdegay, J. 1995. Tuning of fuzzy controllers by genetic algorithms. *International Journal of Approximate Reasoning*, 12:299–315.

[Holland 1975] Holland, J. 1975. *Adaptation in Natural and Artificial Systems.* University of Michigan Press.

[Hough 1962] Hough, P. 1962. Method and means for recognizing complex patterns. *U.S. Patent 3069654.*

[Huber y Kortenkamp 1998] Huber, E. and Kortenkamp, D. 1998. A behavior-based approach to active stereo vision for mobile robots. *Engineering Applications of Artificial Intelligence*, 11(2):229–243.

[Karr 1991] Karr, C. 1991. Genetic algorithms for fuzzy controllers. *AI Expert*, 6(2):26–33.

[Katsuki et al. 2003] Katsuki, R., Ota, J., Mizuta, T., Kito, T., Arai, T., Ueyama, T. and Nishiyama, T. 2003. Design of an artificial mark to determine 3d pose by monocular vision. En *IEEE International Conference on Robotics and Automation, 2003. Proceedings ICRA '03*, vol. 1(14-19), pág. 995–1000.

[Kim y Nevatia 1998] Kim, D. and Nevatia, R. 1998. Recognition and localization of generic objects for indoor navigation using functionality. *Image and Vision Computing*, 16(11):729–743.

[Li y Yang 2003] Li, H. and Yang, S. 2003. A behavior-based mobile robot with a visual landmark-recognition system. *IEEE/ASME Transactions on Mechatronics*, 8:390–400.

[Monasterio et al. 2002] Monasterio, I., Lazkano, E., Rañó, I. and Sierra, B. 2002. Learning to traverse door using visual information. *Mathematics and Computer in Simulation*, 60:347–356.

[Moreno-Velo et al. 2003] Moreno-Velo, F., Baturone, I., Senhadji, R. and Sanchez-Solano, S. 2003. Tuning complex fuzzy systems by supervised learning algorithms. En *The 12th IEEE International Conference on Fuzzy Systems*, vol. 1, pág. 226–231.

[Muñoz-Salinas et al. 2004] Muñoz-Salinas, R., Aguirre, E., García-Silvente, M. and Gómez, M. 2004. A multi-agent system architecture for mobile robot navigation based on fuzzy and visual behaviours. *To appear in Robotica.*

[Paulino et al. 2001] Paulino, A., Batista, J. and Araújo, H. 2001. Maintaining the relative positions and orientations of multiple robots using vision. *Pattern Recognition Letters*, 22(12):1331–1335.

[Scharstein y Briggs 2001] Scharstein, D. and Briggs, A. J. 2001. Real-time recognition of self-similar landmarks. *Image and Vision Computing*, 19:763–772.

[Srinivasan et al. 1999] Srinivasan, M. V., Chahl, J. S., Weber, K., Venkatesh, S., Nagle, M. G. and Zhang, S. W. 1999. Robot navigation inspired by principles of insect vision. *Robotics and Autonomous Systems*, 26(2-3):203–216.

[Stevens y Beveridge 2000] Stevens, M. and Beveridge, J. 2000. Localized scene interpretation from 3d models, range, and optical data. *Computer Vision and Image Understanding*, 80:111–129.

[Stoeter et al. 2000] Stoeter, S. A., Mauff, F. L. and Papanikolopoulos, N. P. 2000. Real-time door detection in cluttered enviroments. *Proceeding of the 15th IEEE International Symposium on Intelligent Control (ISIC 200)*, pág. 187–191.

[Tashiro et al. 1995] Tashiro, K., Ota, J., Lin, Y. and Arai, T. 1995. Design of the optimal arrangement of artificial landmarks. *IEEE International Conference on Robotics and Automation*, pág. 407–413.

[Wells et al. 1996] Wells, G., Venaille, C. and Torras, C. 1996. Vision-based robot positioning using neural networks. *Image and Vision Computing*, 14(10):715–732.

[Zadeh 1975] Zadeh, L. 1975. The concept of linguistic variable and its applications to approximate reasoning. *Parte I Information Sciences vol. 8, pag. 199-249, Parte II Information Sciences vol. 8, pag. 301-357, Parte III Information Sciences vol. 9, pag. 43-80.*

[Zheng et al. 1993] Zheng, N., Fu, X.-D. and Liu, H. 1993. 3 cad-based 3d robot vision. *Intelligent Robots and Systems'93, IROS '93. Proceedings of the 1993 IEEE/RSJ International Conference on*, 3:1905 – 1910.