# A CASE-BASED APPROACH TO INDOOR NAVIGATION USING SONAR MAPS

**Alessandro Micarelli*** **Alessandro Neri****
**Stefano Panzieri*** **Giuseppe Sansonetti***

** Dipartimento di Informatica e Automazione*
*Università "Roma Tre"*
*Via della Vasca Navale 79, 00146 Roma, Italy.*
*** Dipartimento di Ingegneria Elettronica*
*Università "Roma Tre"*
*Via della Vasca Navale 84, 00146 Roma, Italy.*

Abstract: The aim of this paper is to propose an alternative to the traditional approaches that are applied to address the complex problems generated by indoor navigation using sonar maps. Essentially, we present an architecture based on a reasoning method that is known as Case-Based Reasoning in the Artificial Intelligence domain. The system we have developed is capable of analyzing the maps obtained from a robot's ultrasonic sensors, of recognizing the represented object and consequently of making this information available for subsequent experiences. In this way, the robot acquires knowledge on a progressive basis and is therefore able to navigate autonomously in an environment of which it initially has no prior information.

Keywords: Robot Navigation, Image Recognition, Case-Based Reasoning.

## 1. INTRODUCTION

Indoor navigation for a mobile robot using only sonar sensors can be a difficult task when its mission is described in linguistic terms containing topological elements such as "go straight along the corridor, turn right at first corner, and follow the next corridor as far as the second door on the left". In this case an effective environment description is mandatory to a correct navigation and must describe all the essential features necessary for robot's self-localization. Unfortunately, in a dynamic environment, those features can vary and some unknown configurations could be found leaving to the robot the choice on several strategies: one could consist in finding the nearest matching topological element in a static library; an other one could include a supervised learning stage in which the new pattern is used to increase the base library itself. This second approach is often referred as Case-Based Reasoning (CBR) (Aamodt and Plaza, 1994; Kolodner, 1993) and tries to catch all the learning opportunities offered both by the environment and, in an initial phase, by an external supervisor, to improve robot's skill in analyzing its exteroceptive sensorial view.

In literature (Borenstein *et al.*, 1996) the way the world is represented is found to be grouped into two main classes: *metric maps*, giving absolute geometric information about objects, and *topological maps* that contain only relations between objects with no metric at all. In general, topological maps can be more flexible due to their abstract world representation and can be successfully employed when there are no metric information or their quality is extremely poor. Moreover, a planar graph can be used to describe a topological map, and metric information, when present, can be introduced as weights on arcs or nodes (Thrun, 1998; Fabrizi and Saffiotti, 2000). As a matter of fact, the navigation above de-

scribed implies a recognition phase for each step taken by the robot to estimate its position, or better, to understand the particular shape of the environment (the topological feature) inside its actual range of view. This can be done matching the actual sonar view coming from sensors with a reference view that is associated with the particular feature. As stated before, the association is done in most cases comparing the actual view with a static list of models obtained with *a priori* considerations on the environment itself (Fabrizi *et al.*, 2000). Instead, following a CBR philosophy, a learning approach can be devised in which several real-world cases are obtained from a supervised navigation and used to build a dynamic library. In this paper we want to show how such method may be successfully applied to this problem to help the robot during navigation in dynamic environments containing features that only partially correspond to the already-known cases.

In particular, the problem we intend to address concerns the recognition of a sonar-based digital image and its classification under one of the categories of a set of predetermined categories (Corridor, Corner, Crossing, End Corridor, Open Space).

Digital images that we want to analyze are *Fuzzy Local Maps* (FLM), i.e., *Fuzzy Maps* (Oriolo *et al.*, 1998) representing the surroundings of the robot. Those geometric maps, extremely useful in sensor fusion problems, are divided into cells and, for each cell, two values specifying the degree of membership to the set of empty cells and to the set of occupied cells are given. A FLM, usually derived at each step merging the last $n$ sets of collected data, is thereafter represented by two fuzzy sets: the empty cells set $\mathcal{E}$, and the occupied cells set $\mathcal{O}$. For example, in Fig. 1 is reported the $\mathcal{E}$ set of a FLM obtained in a Corridor.

Now, assume that we have a new image, termed "new digital image" in Fig. 2, to be classified. The initial operation performed relates to the extraction of "features" relevant for the purposes of the image recognition. This representation constitutes the "new case" of the proposed CBR system. The retrieval module shown in the figure will effect a search in the case library containing the old cases, with a <*problem representation*, *solution*> structure, which in this specific case will be <*image representation*, *topological feature represented by the image*>. The solution given in the old case can therefore be seen as a pointer to a new library, which we can name "Library of Objects", containing the categories (i.e., "topological features") that could be present in the images to be analyzed. The "Recognized Feature" is at this point taken into consideration by the robot navigation system to plan its motion. This object, which constitutes

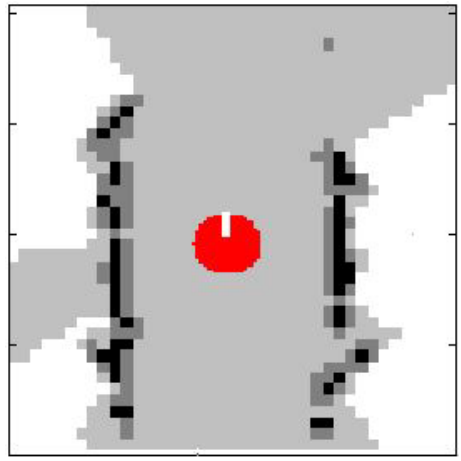

Fig. 1. Fuzzy Metric Map in a Corridor.

the old solution of the case retrieved from the Case Library, will also be considered as the solution of the new problem (basically, there is no need for an adaptation of the old solution to suit the new case) and if the human robot supervisor accepts it, the pair <*New Image*, *Recognized Feature*> can be inserted as a new case in the Case Library.

## 2. IMAGE RECOGNITION

Let's now analyze further in detail the architecture of the module shown in the Fig. 3. For sake of clarity, we provide a simplified version as to illustrate the problems addressed and the respective solutions proposed by us. The algorithm was implemented in C since the entire control procedure in respect to the robot's navigation was originally developed using this language. For a better understanding of the procedure, the algorithm reported in Fig. 4 is based on the use of simplified data structures as opposed to the ones actually implemented. We assume the use of a record consisting of two fields for the case representation, both as concerns the new case as well as any other in the Case Library. The first field is reserved to the representation of the image. In this specific case, to guarantee the applicability of this approach to our real-time control problem, a further simplification has been considered replacing the full 2-D fuzzy description with a polar representation (world mark), a new fuzzy set calculated on the map by associating to each direction a degree of the space around the robot (see Fig. 5). Therefore, the "new digital image" practically consists of an array of 360 real numbers between 0 and 1. Instead, the second field is reserved to the recognized object, where "object" is intended as an integer corresponding to one of the above possible categories. Accordingly, the set of instructions given in pseudo-code is as described in Fig. 4.

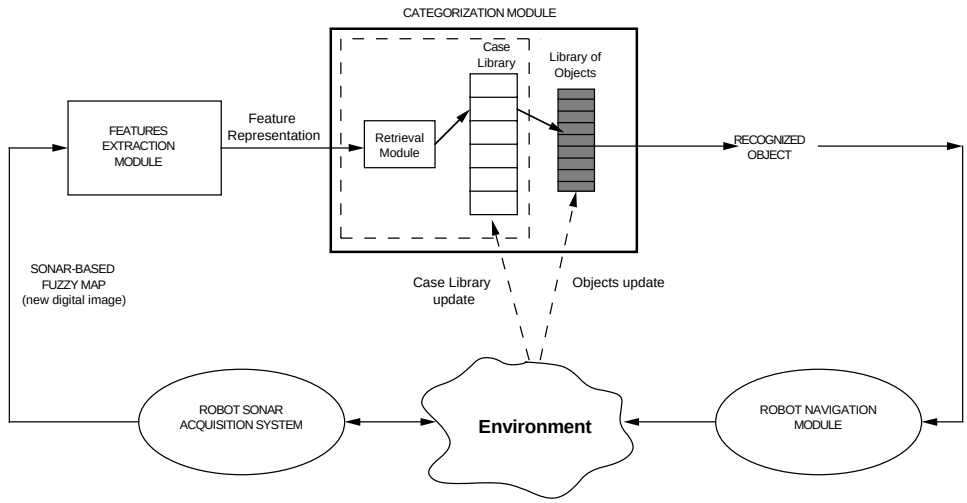


Fig. 2. Case-Based Approach to Image Recognition for Indoor Navigation.


Fig. 3. Image Recognition Module.

```
Function REC(NewImage) returns RecObject
inputs : NewImage; the input image
variables : CaseLib; the case library
        C_j; the generic old case
        S_a; the reliability threshold
        S_b; the identity threshold
local variables : D.image; the image representation
        D.object; the recognized object
        s_j; the metric value
        tempvalue; the temporary metric value
        tempind; the temporary case index

D.image ← DENOISE(NewImage)
D.object ← 0
tempvalue ← 0
tempind ← 0
for each old case C_j in CaseLib do
   begin
      s_j ← COMPARE_CASE(D.image, C_j.image )
      if (tempvalue < s_j) then
         begin
            tempvalue ← s_j
            tempind ← j
         end
   end
   if (tempvalue < S_a) then
      begin
         D.object ← HumanExpertSolution
         C_{n+1}.image ← D.image
         C_{n+1}.object ← D.object
      end
   else
   begin
      if (C_{tempind}.object = HumanExpertSolution) then
            D.object ← C_{tempind}.object
         else
            D.object ← HumanExpertSolution
      if (tempvalue < S_b) then
         begin
            C_{n+1}.image ← D.image
            C_{n+1}.object ← D.object
         end
   end
RecObject ← D.object
returns RecObject
```

Fig. 4. The Image Recognition Algorithm.

This structure envisages the possibility of intervention by a human expert, who is charged with the initial training phase of the robot as well as the checking phase of the retrieved solution. As shown, the use of two different threshold values $S_a$ and $S_b$ are postulated and the former controls whether the case solution retrieved from the

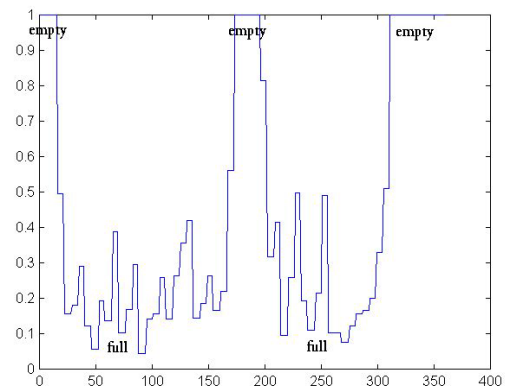


Fig. 5. World Mark.

library is, or is not, extended to the new case. Basically, $S_a$ is the "reliability threshold" that rates the possibility for the case extracted from the library to be considered as representative of the one under observation. Instead, the possibility of the insertion of the new case in the Case Library is dependent on $S_b$, so-called the "identity threshold". The reason for the existence of this second threshold is obvious. Indeed, a richer case library increases the possibility of retrieving the best matching case to the input one and, consequently, improves the accuracy of matching the found solution to the new situation.

On the other hand, an extensive library involves two drawbacks:

- more time is required to retrieve cases from memory;
- more memory capacity is needed.

In fact, if we imagine the Case Library as an array of records, each time a case is input to our module, the procedure is run through from the beginning to the end, therefore the computational complexity of this operation is linear to the number of available cases. As a matter of fact, these are actually underlying problems common to all CBR applications, as also demonstrated

by the considerable size of material generated by the Artificial Intelligence community on these concerns. Various solutions (Schank, 1982) have been proposed, which can be grouped into three categories:

- flat memory;
- discriminant nets;
- E-MOP (Episodic Memory Organization Packets).

Besides these technical aspects, we believe it useful to focus on the more problematic aspects of our architecture, which are basically the following:

- signal representation;
- denoising phase;
- similarity metric;

which are all closely interrelated.

### 2.1 *The Signal Representation*

For a most effective treatment of the above matters, we decided to take advantage of the potential offered by the Wavelet Transform theory, a relatively recent concept, since the first references in the literature date back about 10 years. It is a pure mathematics issue, which took only a few years to find extensive application to the most diverse scientific areas. Wavelets represent an alternative approach to the traditional signal processing techniques, such as Fourier's analysis, for the decomposition of the generic signal into its constituent parts. Their success is linked to their specific property whereby they may be located both in time (space), as well as in scale (frequency), thus providing a "time-scale map" of the signal, wherefrom it is possible to extract the features variable in time. Due to this special property wavelets are an ideal tool for the analysis of physical situations characterized by signals of a discontinuous nature and featuring sharp spikes, as is our case. The analysis procedure, by means of wavelets, is based on the use of a prototype function, so-called "mother wavelet", whose translated and extended (or compressed) versions constitute the basis functions for the series expansion that allows for the representation of the original signal through coefficients. Data operations may therefore be effected on the corresponding wavelet coefficients. In particular, if the mother wavelet is appropriately chosen, or if the coefficients under a certain threshold are eliminated, it is possible to "sparsely" represent the original data. This means that wavelets are a most useful tool in the context of data compression. However, due to the way it is introduced, the wavelet transform seems to be a rather unwieldy operation from a computational point of view. Indeed, the information it supplies, redundant in respect to the signal reconstruction,


Fig. 6. Data Analysis by Wavelets.

has definite negative effects on the length of the computation and on the quantity of resources available. These negative effects would render the wavelet transform unmanageable in practical applications, had Stephane Mallat not introduced the fast DWT, also known as Pyramid Algorithm, a paradigm that our architecture uses extensively (Mallat, 1998).

As may be easily verified, this algorithm is $O(n)$, where $n$ is the length of the array of data to be transformed; this is extremely advantageous compared to other algorithms (for example the FFT is $O(n\log n)$). Moreover, the DWT can be easily extended to multidimensional data, such as images, a factor that could turn out useful in view of a possible extension of our architecture to more complex data than those currently under review. It is therefore on the basis of the above considerations that we decided to make an extensive use of the wavelet transform for our experimentation purposes.

### 2.2 *The Denoising Phase*

The above-mentioned considerations explain why wavelets have been greeted by the international scientific community with such enthusiasm, in particular by those scientists who must often deal with the very signals that we are addressing herein, where the original signal is covered by noise. In fact, it should not be forgotten that "our" signal is built on the basis of readings effected by ultrasonic sensors, i.e., instruments having a measurement process riddled with uncertainty. It is to resolve these very noise problems that different techniques have been proposed, among which the one by David Donoho and Iain Johnstone found particular favor among the community, known as "wavelet shrinkage and thresholding", based on the following idea: the decomposition of a data set by means of wavelets involves both filters that act as "averaging filters" and others that generate the so-called "details"; if the latter appear of minor significance, they can be eliminated without causing a substantial alteration of the main features of our given data set. This method is exemplified in the diagram shown in Fig. 6.

Note that the signal is initially transformed, its coefficients are matched against a predefined threshold, and subsequently reutilized to reconstruct the original data set through the inverse DWT. Donoho e Johnstone propose a number of different denoising strategies, using predefined thresholds as well as adaptive ones. In actual fact, as we

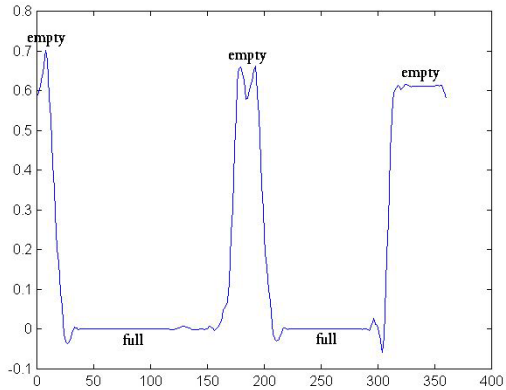


Fig. 7. Reconstructed Function after Soft Thresholding.

can see, there is no one strategy that always gives better results compared to the others, but only strategies whose validity depends on the type of application. Once the threshold is determined, it may then be matched against the wavelet coefficients using two different possible methods:

- *Hard Thresholding*, which implies the "keep or kill" principle, i.e., all the wavelet coefficients are checked at an absolute value against a predetermined threshold $t_h$: if the magnitude of the coefficient is less than $t_h$, the coefficient is replaced by zero, otherwise it is left unchanged:

$$c_{jk} = \begin{cases} 0, & c_{jk} < t_h \\ c_{jk}, & c_{jk} \geq t_h \end{cases}$$

- *Soft Thresholding*, in this case, the wavelet coefficients are shrinked towards the origin in accordance to the following expression:

$$c_{jk} = sign(c_{jk})(|c_{jk}| - t_s)_+$$

The figures 7 and 8, report the result of the application of the two different methods to the signal shown in Fig. 4, using the same threshold value ($t_h = t_s = 1.1$) in both cases (the 3-scales forward and inverse DWT are performed by using Symmlet with 8 vanishing moments as the mother wavelet).

On the basis of the obtained results we can affirm that the computational burden ensuing from the denoising operation appears compensated by the significant reduction of cases to be inserted in the library. This is due to the fact that by eliminating the overlying noise the similarity of two situations corresponding to the same topological condition may often be revealed, while their respective representations covered by noise misleadingly appear considerably different. Moreover, we have already seen how the computation length and the quantity of utilized resources are directly proportional to the number of cases kept in memory.

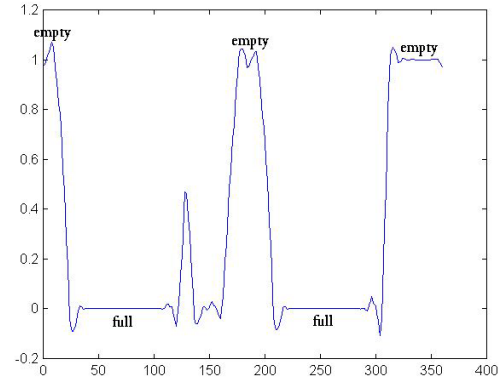


Fig. 8. Reconstructed Function after Hard Thresholding.

### 2.3 *The Similarity Metric*

We must now address the third problem, that is to say, finding a metric for the estimate of the "similarity" between input case $x$ and generic case $y$ belonging to the Case Library. Indeed, it is obvious that the quality of the retrieval phase, the crucial point of any CBR application, essentially depends on the correctness of this choice.

In the context of our experimentation, we adopted different approaches, each of which yielded its specific merits and shortfalls. Among these, the best overall results were given by the Cross-Correlation Factor metric, expressed as follows:

$$\frac{\max(\mathrm{R_{xy}}(\tau))}{(\mathrm{R_{xx}}(0)\mathrm{R_{yy}}(0))^{\frac{1}{2}}}$$

This quantity was calculated both in the time domain and in the Fourier's domain, obtaining important results, with reasonable computational lengths using calculation resources available on the market. Furthermore, this method allowed us to compute the other information available in the mono-dimensional map, that is, the polar information.

## 3. FUTURE DEVELOPMENTS

While producing results that justify our confidence in respect to subsequent experiments, clearly, our proposed architecture does not allow for an easy assessment of its room for improvement. Currently, we are working in different directions, in particular, as observed, the denoising phase envisages the application of the forward and inverse DWT, therefore one of our objectives is to optimize the identification of the mother wavelet and the number of scales to which the wavelet transform is applied. For now, this choice is predetermined, but our aim is to develop a model where this happens dynamically, on the basis of feedback on the obtained results. The same is true

for the two threshold values $S_a$ and $S_b$ introduced in the pseudo-code description of our procedure, indeed, these two parameters have a direct impact on the efficiency of the entire module. The last, and maybe the most important point, concerns the metric. Although the produced results can be considered significant, clearly, the computation of the cross-correlation between input signal and the models extracted from the library, albeit conceptually immediate, is not completely efficient.

In fact, the ability to resolve cases deriving from those in the library by applying simple geometric transformations, such as translations, rotations or scale changes, requires anyhow the calculation of the above-mentioned metric for an infinite number of possible variants. Therefore, in view of a possible application of our system to signals originating from more sophisticated sensors than ultrasonic ones, an alternative approach that is currently being explored, envisages the use of some kind of "invariant" signatures, investigated in the past decades in the context of classical pattern recognition. The major class of rotation invariant image classification techniques, originally introduced in the field of optical processing, is based on the extraction of the dominant circular harmonic components obtained expanding a given image in its angular Fourier's series. This circular harmonic decomposition could be further combined with scale invariant representations like the Fourier-Mellin Transform (FMT) to devise rotation and scale invariant pattern recognition algorithms. To preserve the ability of the system of inferring the current situation based on the whole set of collected cases, at the expenses of a reasonable computational complexity, in our architecture we could resort to a general purpose rotation and scale invariant indexing tool based on the cited mathematical representations, proposed in (Jacovitti and Neri, 1998). In essence, the index associated to each pattern stored in the case dictionary could be constituted by the dominant components of the Laguerre-Gauss Circular Harmonic Wavelet (CHW) decomposition, computed by means of a bank of linear operators.

## 4. CONCLUSIONS

Generally, the normal pattern recognition techniques require models of the objects that must be recognized and classified. The collection of models available to the classifier clearly reflects the original knowledge of the situation to be analyzed. However, in most cases, as for the robot's autonomous navigation, there exists practically no prior information whatsoever. Our proposed architecture includes traditional feature extraction algorithms incorporated into a CBR system, which allows a constant increase in the knowledge of the surrounding environment. Furthermore, it should not be forgotten that there is also the possibility of updating the Object Library as well as the Case Library. It is nevertheless obvious that, while in principle there is no limit to the number and complexity of information that may be collected with a system such as ours, the constraint of a real-time operation involves restrictions linked to the technology that is available today on the market.

Nevertheless, in view of the vast potential of our architecture and the results achieved so far, we are confident that the constantly evolving technology will allow for subsequent ulterior developments following the prospects provided by our work.

## 5. REFERENCES


Aamodt, A. and Plaza, E., "Case-Based Reasoning: Foundational Issues, Methodological Variations, and System Approaches". *AI Communications*, **7**(1), 1994, pp. 39-59.

Borenstein, J., Everett, H.R. and Feng, L. *Navigating mobile robot: sensors and techniques*, A. K. Peters, Ltd., Wellesley, MA., 1996.

Fabrizi, E., Panzieri, S. and Ulivi, G., "Extracting topological features of indoor environment from sonar-based fuzzy maps". To appear on *Proc. of The 6Th International Conference on Intelligent Autonomous Systems,* July 25-27, 2000, Venice, Italy.

Fabrizi, E. and Saffiotti, A., "Extracting topology-based maps from gridmaps", *Proc. of Int. Conf. on Robotics and Automation,* San Francisco, 2000.

Jacovitti, G. and Neri, A., "Multiscale Circular Harmonic Wavelets: a Tool for Optimum Scale-Orientation Independent Pattern Recognition". In: *Wavelets Applications V, Proc. of SPIE 3391,* 1998.

Kolodner, J., *Case-Based Reasoning.* San Mateo, Calif., Morgan Kaufmann, 1993.

Mallat, S.G., "A Theory for Multiresolution Signal Decomposition: the Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7), 1989, pp. 674-693.

Oriolo, G., Vendittelli, M. and Ulivi, G., "Real-Time Map Building and navigation for Autonomous Robots in Unknown Environments," *IEEE Transactions on Systems, Men and Cybernetics - Part B: Cybernetics*, **28**(3), 1998, pp. 316-333.

Schank, R., *Dynamic Memory: A Theory of Learning in Computers and People.* New York, Cambridge University Press, 1982.

Thrun, S., "Learning Metric-Topological Maps for Indoor Mobile Robot Navigation," *Artificial Intelligence*, **99**(1), 1998, pp. 21–71.