CHARACTERIZING IMAGE SETS USING FORMAL CONCEPT ANALYSIS

Emmanuel Zenou[†‡] and Manuel Samuelides[†]

zenou@supaero.fr

[†]National School of Aeronautics and Space (SUPAÉRO)
10 avenue Edouard Belin, BP 4032, 31055 Toulouse Cedex - FRANCE
[‡] LAAS -CNRS
7, avenue du Colonel Roche, 31077 Toulouse Cedex 4 - FRANCE

ABSTRACT

This article presents a new method for supervised image classification. Given a finite number of image sets, each set corresponding to a place of an environment, we propose a localization strategy, which relies upon supervised classification. For each place the corresponding landmark is actually a combination of features that have to be detected in the image set. Moreover, these features are extracted using a symbolic knowledge extraction theory, "formal concept analysis". This paper details the full landmark extraction process and its hierarchical organization. A real localization problem in a structured environment is processed as an illustration. This approach is compared with an optimized neural network based classification, and validated with experimental results. Further research to build up hybrid classifier is outlined on discussion.

Keywords. Supervised classification, Visual landmarks, Galois lattices, Concept lattices, Computer vision, Localization.

1. INTRODUCTION

Characterizing and recognizing a place in a -structured or not- environment, using only a set of views attached to each place to characterize, is a difficult challenge to take up for a machine (computer or robot) today. To do this, the machine needs to find "something" that 1. characterizes a considered place, and 2. distinguishes it from the others. This "something", under specific conditions, is called a (visual) *landmark*. What is a landmark ? How to find it ? And how to select it ?

This paper presents a new method to answer these questions. All the images issued from one place are regrouped into a set. Thus, the machine has to recognize one original place from *some* images of the associated set. At first, during a learning stage, the relationships between sets of images and features are structured and organized into a hierarchy, through a formalism called *Galois Lattices*, or *Concept lattices*. The use of such mathematical structures

allows the machine to determine its own landmarks attached to each place. Subsequently, once this initial characterization has been performed, the machine is able in a second stage to recognize the corresponding place thanks to the landmarks it has learned.

The choice of the application we have done makes the connection between one set of images and one room of a structured environment. Thus we expect that there will be more or less common properties between images of one set. But the theory we have developed here considers only sets of images without any restriction.

This paper is organized as follows: sections 2 introduces landmarks, primitives and features; section 3 gives an outline of Formal Concept Analysis; sections 4 shows how we use it to define and to build landmarks; section 5 exposes the results of this approach on an experimental setup, before conclusion and perspectives (section 6).

2. LANDMARKS, PRIMITIVES AND FEATURES

2.1. The classical notion of landmark in autonomous mobile robotics

As defined in the Cambridge Dictionary, a landmark is *a building or place that is easily recognized, especially one which you can use to judge where you are.* This original definition, applied to the mobile robotics field, has several versions such as "distinctive templates from one image which can be readily recognized in a second image acquired from a different viewpoint" [1], or more simply "identifiable visual objects in the environment" [2]. Usually landmarks are not introduced according to a formal definition but through some specific properties such as "easily distinguishable" [3] or "locally unique" [3]. In concrete terms, a landmark could be an object [4], a color [5], interest points [6], *etc.* In our case, landmarks are not restricted to one kind of elements, but could be a combi-

nation of elements. For instance, a landmark of a place A could be a *"big blue object"*, even if there is a *"big object"* in the place B and some *"blue"* in the place C. Nevertheless, it is essential that a landmark checks the two following characteristics: first it should discriminate between locations, and second it should be stable to allow robust identification against variations of the observer position and time [1, 7]. Several classifications of landmarks, as static/dynamic [8], already exist, still we propose here another classification based upon the learning ability and the autonomy of the recognition system. We do separate landmarks into three categories:

(i) fully pre-defined landmarks: the machine is given a database of objects [4, 1, 9] which are "just" to be recognized;

(ii) partially pre-defined landmarks: such potential landmarks are specified by a common structure. For instance, in [10], the authors use planar quadrangular forms (typically, posters) characterized with interest points [11] and Hausdorff distance. Observations which could fit into the specified framework are then dynamically chosen as landmarks;

(iii) non pre-defined landmarks: no hypothesis is assumed about potential landmarks. The main approaches with such landmarks are biologically inspired [12, 13, 14].

Our approach deals with the last category: we want the machine to choose the most relevant landmarks in an autonomous and dynamic way. Notice the connection between landmark localization and supervised classification. When the landmark is predefined the classifier is designed by hand using expert knowledge about robustness of object shape and so on. In case (iii), the landmark is defined through a learning process which is similar to learning a supervised classifier. An important difference still remains between landmark-based localization and supervised classification: in our case, if a landmark is not found in the current image, the robot visual system is requested to provide additional information to the localization system through a new picture. This "no answer" event decreases classification error. We will develop the landmark selection process further in this paper according to a learning approach, still we will keep the "no answer" event.

2.2. Primitives and Features

Different pictures are extracted from each room of the environment; thus, a set of images is attached to each room. From these different pictures, *primitives* are extracted to build *features* of images, to help the robot to find *properties* of each place. We do distinguish feature from

properties by the fact that features are attached to images, whereas properties are attached to the place. Three kinds of primitives are extracted from the different pictures:

- structural primitives: segments with their size and orientation (they are issued from polynomial contour extraction), interest points [15], ... ;

- colorimetric primitives: extraction of red, green, blue, cyan, magenta or yellow pixels with joint histograms, objects, contrast, ... ;

- photogrammetric primitives, issued from pixels intensity: contours, texture, *etc.*

From all these primitives, features are extracted in all set of images. Notice that our definition of feature is extensive and includes any potential feature, whatever it is present in an image or not. For instance, with colorimetric primitives, potential features could be *"there is some yellow here"* or *"there is such texture"*. Notice that we include features that are invariant against rotation, translation and scaling. For instance, using segments (primitives) extracted from contours, one feature could be: *"there is a large number of identical (orientation and size) segments"* (Typically, this feature may be issued from a bookcase that is present in the considered place).
Let us also note that our system is "open", that means that any other (visual or not) feature could be included to increase efficiency of our learning process.

2.3. Raw display of visual information

Once all primitives are extracted from images and features are detected, information is organized into a look-up table, that displays the presence or not of a feature in an image (see figure 1).
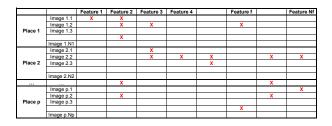
		Feature 1	Feature 2	Feature 3	Feature 4		Feature f		Feature Nf
Place 1	Image 1.1	X	X						
	Image 1.2		X	X			X		
	Image 1.3								
			X						
	Image 1.N1								
Place 2	Image 2.1		X						
	Image 2.2		X	X	X		X		X
	Image 2.3				X				
	Image 2.N2								
...	...		X				X		X
Place p	Image p.1								X
	Image p.2		X				X		
	Image p.3								
							X		
	Image p.Np								

Figure 1: Visual Information Look-Up Table.

3. FORMAL CONCEPT ANALYSIS

Galois -or concept- lattices have been widely used in Artificial Intelligence in the past 20 years. This theory has been developed as *FCA: Formal Concept Analysis*, and

several lattice building algorithms appeared since then, more and more efficient [16]. Still few concrete applications recently appeared mainly in data mining topics such as machine learning [17, 18] or in the aeronautic field [19]. We outline here an application to localization in the autonomous mobile robotics field.

3.1. Mathematical Formalism [20, 21]

Definition 1 *A **lattice** is defined as an ordered set in which any couple of elements has a least upper bound (lub) and a greatest lower bound (glb). A complete lattice is a lattice where any set has a lub and a glb.*

For instance, the set $\mathcal{P}(\mathcal{O})$ of all subsets of a set $\mathcal{O}$ ordered by the inclusion $\subset$ is a complete lattice.

Definition 2 *A **context** $\mathcal{K}$ is a triple $(\mathcal{O}, \mathcal{F}, \zeta)$ where $\mathcal{O}$ is a set of **objects**, $\mathcal{F}$ is a set of **attributes** and ζ is a **mapping** from $\mathcal{O} \times \mathcal{F}$ into $\{0, 1\}$.*

In our application, objects are images taken by the robot, attributes are features and the mapping ζ is defined by $\zeta(o, f) = 1$ if and only if feature f is present in image o. The graph of this mapping is the look-up table of figure figure 1.

Definition 3 *Given a context $\mathcal{K} = (\mathcal{O}, \mathcal{F}, \zeta)$ let us define two mappings from $\mathcal{P}(\mathcal{O})$ into $\mathcal{P}(\mathcal{F})$ and from $\mathcal{P}(\mathcal{F})$ into $\mathcal{P}(\mathcal{O})$ using the same notation $'$ by the formula*

$$\forall \mathcal{A} \subset \mathcal{O}, \mathcal{A}' = \{f \in \mathcal{F} \mid \forall o \in \mathcal{A}, \zeta(o, f) = 1\} \quad (1)$$
$$\forall \mathcal{B} \subset \mathcal{F}, \mathcal{B}' = \{o \in \mathcal{O} \mid \forall f \in \mathcal{B}, \zeta(o, f) = 1\} \quad (2)$$

*These mappings are called the **Galois connections** of the context; $\mathcal{A}'$ is called the **dual** of $\mathcal{A}$, similarly $\mathcal{B}'$ is called the **dual** of $\mathcal{B}$.*

Clearly, $\mathcal{A}'$ is the set of common attributes to all objects of $\mathcal{A}$, and $\mathcal{B}'$ is the set of objects which share all attributes belonging to $\mathcal{B}$.
The properties of the Galois connections can be found in [22]. Let us recall the basic following properties:

Property 1 $\mathcal{A}_1 \subset \mathcal{A}_2 \Rightarrow \mathcal{A}'_2 \subset \mathcal{A}'_1$

Property 2 $\mathcal{A} \subset \mathcal{A}"$

Property 3 $\mathcal{A}' = \mathcal{A}"'$

We are now able to state the definition of a concept:

Definition 4 *Given a context $\mathcal{K} = (\mathcal{O}, \mathcal{F}, \zeta)$, the pair $\mathcal{C} = (\mathcal{A}, \mathcal{B})$ is called a **concept** of $\mathcal{K}$ if and only if $\mathcal{A}' = \mathcal{B}$ and $\mathcal{B}' = \mathcal{A}$.*

Definition 5 *$\mathcal{A}$ is called the **extent** of the concept $\mathcal{C}$ and $\mathcal{B}$ is called its **intent**. One notes $\mathcal{A} = \mathtt{extent}(\mathcal{C})$ and $\mathcal{B} = \mathtt{intent}(\mathcal{C})$.*

The set of all concepts of a context $\mathcal{K}$ is denoted $\mathcal{L}(\mathcal{K})$ or simply $\mathcal{L}$ if the context is clear. One proves [21] the following theorem:

Theorem 1 *Let $\mathcal{C}_1 = (\mathcal{A}_1, \mathcal{B}_1)$ and $\mathcal{C}_2 = (\mathcal{A}_2, \mathcal{B}_2)$ be a couple of concepts, then $\mathcal{C}_1 \vee \mathcal{C}_2 = ((\mathcal{A}_1 \cup \mathcal{A}_2)", \mathcal{B}_1 \cap \mathcal{B}_2)$ and $\mathcal{C}_1 \wedge \mathcal{C}_2 = (\mathcal{A}_1 \cap \mathcal{A}_2, (\mathcal{B}_1 \cup \mathcal{B}_2)")$ are concepts.*

This result may be extended to any set $\mathcal{I}$ of concepts. We shall note $\mathcal{C}_\mathcal{I} = (\mathcal{A}_\mathcal{I}, \mathcal{B}_\mathcal{I}) = \bigvee_{i \in \mathcal{I}} \mathcal{C}_i$ and similarly $\mathcal{C}^\mathcal{I} = (\mathcal{A}^\mathcal{I}, \mathcal{B}^\mathcal{I}) = \bigwedge_{i \in \mathcal{I}} \mathcal{C}_i$
Thus, the set of concepts $\mathcal{L}$ when it is endowed with the order relation $\subset$ of its extents is a complete lattice and we can set:

Definition 6 *The complete lattice $\mathcal{L}(\mathcal{K})$ of concepts of the context $\mathcal{K}$ is called the **Galois lattice** or the **concept lattice**.*

3.2. Lattice Building Algorithm

Concept lattice building algorithms are divided into two families: incremental algorithms and non-incremental algorithms. See [23] for a complete description. The most appropriate algorithm for our application is the *Norris* algorithm [24] (the complexity is $O(|\mathcal{O}|^2.|\mathcal{F}|.|\mathcal{L}|)$ with $|\mathcal{L}|$ the number of concepts [23]). It is practically efficient to process middle size problems with time constraints for this application in spite of the worst-case exponential complexity as shown in paragraph 5.4.

3.3. Finding Landmarks with Concept Lattices

From now on, we shall use the term "Concept Lattice". The extent of a concept is an object subset that is completely defined by a set of attributes that are simultaneously checked by the elements. The intent of a concept is a set of attributes that are a maximal characterization of an object set.
The context in our application being defined with a set of images (objects), a set of features (attributes), and a mapping, here the presence or not of a feature f in an image i, the general lattice is built and landmarks are extracted thanks to the following definition:

Definition 7 *Given a context $\mathcal{K} = (\mathcal{O}, \mathcal{F}, \zeta)$ and a subset of objects $\mathcal{A} \subset \mathcal{O}$. A subset $\mathcal{B} \subset \mathcal{F}$ is said to be a **landmark** of $\mathcal{A}$ if and only if*

- $\mathcal{B}' \subset \mathcal{A}$.

- $\mathcal{B}" = \mathcal{B}$,

By this way, a landmark is a combination of features of a concept (intent) that respects the above conditions. The complete process is detailed in next section.

Let us note that the first property ($\mathcal{B}' \subset \mathcal{A}$) could be enough to define a landmark. However, $\mathcal{B}'$ would not always correspond to a specific subset of objects, so the combination is not optimized. Thus, to avoid an explosion of possibilities, and to restrict the number of landmarks to a minimal number, it is necessary to fit with concept intents. The choice of concept as the basis element to build classification rules is hoped to provide robustness to classification and to improve generalization properties.

4. BUILDING A LANDMARK-BASED CLASSIFIER

In this section, we do expose the complete reasoning first to extract landmarks from a set of images, and second to label an image to a set.

Let us detail our basic application. We have at our disposal a set of images from a structured environment. Each image is labelled by the room from which it was shot. Our objective is to provide a mobile robot equipped with a camera a decision rule to allow it to find its localization in a topological map[1]. It is basically a supervised classification problem. The decision rule is provided by a maximal partial landmark. Note that we are in a typical learning situation. The decision rule is extracted from a set of labelled examples, the learning base of images. This rule is formalized for each set by concepts that shall be defined as maximal landmarks. Some images of the learning set may escape from the decision rule. Thus, due to the image preprocessing (primitive extraction) and the complexity of the environment, learning failing may occur.

There are actually two phases: the first phase deals with landmark extraction (*learning phase*), and the second phase deals with the use of these landmarks to find the place a new image comes from (*generalization phase*). Let us give first some definitions useful for our particular application.

4.1. Formal Definitions in a Partitioned Context

Given a context $\mathcal{K} = (\mathcal{O}, \mathcal{F}, \zeta)$, a partition $(\mathcal{O}_\theta)_{\theta \in \Theta}$ of the object set is available. So We have:

$$\mathcal{O} = \bigoplus_{(\theta \in \Theta)} \mathcal{O}_\theta$$

Definition 8 *θ is called a* **site** *and Θ the set of sites.*

---

[1] A topological map of a structured environment is a graph for which, most of the time, a node is a room and an edge is a connection between two rooms [25]

More generally, a semantic label can be considered instead of a site in a general classification context.

4.1.1. Landmarks

Definition 9 *Let $\mathcal{B}_\theta$ be a subset of $\mathcal{F}$. $\mathcal{B}_\theta$ is said to be a* **landmark** *of a site θ if and only if*

- $\mathcal{B}'_\theta \subset \theta$.
- $\mathcal{B}''_\theta = \mathcal{B}_\theta$,

A landmark is thus a set of attributes for which the simultaneous presence is effective in some image of the site to characterize.

4.1.2. Full Landmarks

In particular, if the landmark $\mathcal{B}_\theta$ is a set of attributes present simultaneous in *all* images of the site, $\mathcal{B}_\theta$ is called a full landmark.

Definition 10 *$\mathcal{B}_\theta \subset \mathcal{F}$ is said to be a* **full landmark** *of a site θ if and only if*

- $\mathcal{B}'_\theta = \mathcal{O}_\theta$
- $\mathcal{B}''_\theta = \mathcal{B}_\theta$,

4.1.3. Maximal Landmarks

If a full landmark $\mathcal{B}_\theta = \mathcal{O}'_\theta$ exists, it is sufficient to define a decision rule for localization with respect to site θ. Of course, that issue does not occur very often in practical applications. If there is no full landmark, it is interesting to limit the number of landmarks by introducing maximal landmarks.

Definition 11 *A* **maximal landmark** *$\hat{\mathcal{B}}$ is a landmark of minimal intent in a set of landmarks of a given site .*

4.1.4. Coverage

The coverage of a site by a landmark or a set of landmarks specifies if every image of the site contains some of landmarks or not.

Definition 12 *let $\{\mathcal{B}_{\theta,i}\}_{i=1...N_\theta}$ be the N_θ landmarks of a site θ. This site is said to be* **covered***, or the landmarks* **cover** *the site, if and only if*

$$\bigcup_{i=1...N_\theta} \{\mathcal{B}'_{\theta,i}\} = \theta$$

If there is a full landmark in a site, the coverage is obvious. If not, the set of mages from a site may not be covered by landmarks. Note that if such a full coverage exists, it is provided by maximal landmarks.

4.2. Learning Phase: Extracting the Landmarks

The first step is to extract primitives from each image. The algorithms used to do this are quite classical. For instance, to obtain segments, the contours are extracted with a Canny-Deriche algorithm, then they are approximated with polynomial figures. Eventually segments are extracted by a fusion process. Other primitives are found through image color or texture segmentation.

The second step is to find features with these primitives, and to fill up the look-up table. The third step is the building of the associate lattice. The last step is to "read" the lattice, *i.e.* to select landmarks attached to each class (each place). Let us detail this last process.

Following the strict definition of a landmark, the general lattice is built and concepts are put into a hierarchy. Considering all concepts $\{C_\theta\}$ relative to a site θ, *i.e.* all concepts whose extent are made with images from the site θ (and only from this site), landmarks are intent of these concepts. We precise the definitions from previous section:

Definition 13 *A* **landmark-concept** *relative to a class θ is a concept whose extent are made with objects belonging to $\mathcal{O}_\theta$.*

Definition 14 *A* **landmark** *of a class θ is the intent of a landmark-concept relative to a class θ.*

Definition 15 *Considering the set of all landmark-concepts relative to a class θ, a* **maximal landmark-concepts** *is a landmark-concept whose extent has no no parent in the considered object set $\mathcal{O}_\theta$.*

Definition 16 *A* **maximal landmark** *of a class θ is the intent of a maximal landmark-concept relative to a class θ.*

The general algorithm of the landmark selection method is presented figure 2.

---

1. Extract primitives from each image
2. Determine the presence or not of features and Fill up the cross table
3. Build the corresponding lattice
4. For each place θ,
   4.1 select landmark-concepts,
   4.2 read concepts from landmark-concepts,
   4.3 select maximal landmarks.

Figure 2: General Algorithm of the Landmark Selection Method

4.3. Generalization Phase: Image (or Object) Classification

Once the landmarks selected, we consider now a new image issued from any place. Primitives and attributes are extracted from this image. Two cases should be considered:

- if the image contain at least one landmark of a class θ and no landmark of any other class $\theta' \neq \theta$, then the image is classified in the class θ;

- if no landmark is included in the image or if several landmarks, from several class, are included, the classifier gives no-response. In this case, the lattice has to be updated.

5. EXPERIMENTATIONS AND RESULTS

Different experimentations have been managed to confirm our approach. The general frame of these experimentation is the navigation of a robot in a structured (human) environment. The goal, for the robot, is to extract visual landmarks with the aim to locate itself in view.

66 potential features could be detected in our images: number of pixels of the primary and secondary colors greater than 1000, black, white and colored small, medium and big objects detected thanks to morphological operators, bio-inspired color contrasts such as black-white, red-green and yellow-blue contrasts, small, medium and large oriented (12 directions) segments issued from image derivation.

The first experimentation consists on a *classical* classification process: some images from four different classes have been analyzed to build the classifier. Next this classifier has been tested with other images from the same places. This approach is validated through a comparison with an optimized neural network. Next, a real robotics experimentation have been processed to fit closely with our general research context. Finally, an experimentation has been carried out with a much bigger context.

5.1. Image Classification

First, we state results in terms of image classification with landmarks. 177 images have been taken for the learning stage, in four different places of the laboratory environment. The feature extraction process gives a 177×66 look-up table. The corresponding 5265 concept lattice is computed in 25 seconds on a Spark 100 machine. For the four classes, 883 concept-landmarks are extracted, there is no full landmarks and 42 maximal (partial) landmarks are kept : 9 for the first place, 8 for the second one, 17 for the third one, and 8 for the fourth one (see table 1).

Place	Ldks	Full Ldks	Maximal Ldks
Place #1	194	0	9
Place #2	316	0	8
Place #3	291	0	17
Place #4	82	0	8
Total :	883	0	42

Table 1: Learning Phase : Landmarks Extraction.

During the generalization phase, 32 images are issued from the place #1. These images are *different* from those of the learning phase. Landmarks are searched on all images: 1 image contains 2 ambiguous landmarks (one of the place #1, one of the place #3) and 14 no landmark ; 16 images contain only landmarks of the place #1, and 1 image contains a place #4 landmark. There is thus a response rate of 53.1%, an absolute well situated image rate of 50% on all images, more important a relative well situated image rate of 94.1% on (well or not) located images, an absolute error rate of 3.1% and a relative error rate of 5.88%. The results of the full analysis for all places are displayed on table 2.

Place	NI	NR	NGR	NFR
Place #1	32	17 (53.1%)	16 (94.1%)	1 (5.88%)
Place #2	50	13 (26%)	12 (92.3%)	1 (15%)
Place #3	31	10 (32.3%)	10 (100%)	0 (0%)
Place #4	38	20 (52.6%)	19 (95%)	1 (5%)
Total :	151	60 (39.8%)	57	**3 (5%)**

Table 2: Generalization Phase using Landmarks: Results (NI : Number of Images, NR : Number of Responses, NGR : Number of Good Responses, NBR : Number of False Responses).

Let us note that the classification rule has been tested with the learning set of images to assess the equivalent of learning error. Of course, by definition, for each place, there is no landmark from another place, however the response rate is not 100% (88%, 43.1%, 85.7% and 54.8% for respective places #1, #2,#3 and #4) : there are some images with *a posteriori* no useful information, *i.e.* images whose features are shared with some pictures of other sets.

5.2. Comparison with a Optimized Neural Network

Comparison with a classical neural network classification under MATLAB has been processed to appraise our ap-

proach on the same data basis. To improve neural network results, several experimentations with different architectures have been computed to obtain the best network as possible.

The optimized network is composed of 66 neurons in the first layer (corresponding to our 66 features), 66 neurons in the middle layer and 4 neurons (corresponding to the 4 places) in the last layer. The training function is a Backpropagation gradient training with an adaptive learning (`taingda`), with an hyperbolic tangent sigmoid transfer function for each layer of the network. Other comparisons have been done with different number of layers, different number of neurons in the middle layer, different training process and/or different transfer functions, but with worse results. The Levenberg-Marquardt and Bayesian regularization algorithms fail due to the high number of entries.

With the number of 700 training epochs, the smallest learning rate is 4.10^{-2} and more significantly the smallest error rate (false response compared to all response) we obtained is 5% on the learning set of images, and 30% on the testing set...

More over, the variability of responses of a network is very different from one learning process to one another, with the same learning database. Best results cited above are reached once on five or six tries.

To fit with our technique and to have comparable results (see table 3), a program has been developed to allow the neural network to give some "no-responses". In a practical way, the classification answer is validated if and only if the difference between the greatest probability to be in one place and the second greatest probability to be in another place is above a threshold, that is adjusted to have the same rate of no-responses.

Place	NI	NR	NGR	NFR
Place #1	32	17 (53.1%)	17 (100%)	0 (0%)
Place #2	50	17 (26%)	17 (100%)	0 (0%)
Place #3	31	14 (45.2%)	10 (71.4%)	4 (28.6%)
Place #4	38	12 (31.6%)	10 (83.3%)	2 (16.6%)
Total :	151	60	54	**6 (10%)**

Table 3: Generalization Phase using a Optimized Neural Network.

5.3. Mobile Robot Localization

This experimentation has been done with a real mobile robot in our laboratory. There is also four places in this process but they are different from the previous experimentation. Yet, features are identical. The strategy here

is different : during the learning phase and the generalization phase, the robots moves alone with its own speed, and images are issued from a continuous flow of images ("continuous" means here that the robot do not jump from one place to another, there is some ambiguous transition zones, difficult to classify).

The robot moves thus over the structured environment; 295 analyzed images give a 295×66 look-up table, the lattice of which is built with 8020 concepts. A total of 649 landmark-concepts are extracted and 48 of them are isolated to be maximal partial landmarks (17 from the place #1, 16 from the place #2, 9 from the place #3 and 6 from the place #4).

	NI	NR	NGR	NFR
Results :	161	54	50	4 (8%)

Table 4: Result with a real mobile robot experimentation.

During the generalization phase, the robot move also over the same environment. 161 images are analyzed, 50 are well located in their respective place, and 4 no. The global error is thus 8%, and the response rate is 33.5%. The reason of such a *low* response rate is that the robots moves through a white corridor, that has very few features and landmarks, and a lot of white images *pollute* the analysis rate. However, the number of (well or not) located images has no impact on our application : either the robot may give an answer (the place it is) with an heuristic based on all image responses of the considered set (passive vision process), or the robot may look by itself for landmarks by moving around (active vision process). This is one of our next working orientations.

5.4. Experimentation with a bigger context

Another experiment has been carried out with a higher number of features. The use of the HSV color space allow us to divide the whole spectrum into as many bands as wanted, and by this way we have increased the number of feature up to 153. With a Pentium4 (2.4GHz) PC, under 850 images, the lattice update time process is inferior to the image analyze time process (half a second about). After 850 images, time processes are quite similar (contrary to the image analyze process, the lattice update time process depends on the image itself and the extracted features), and after 1100 images, the update process is longer if new images appear. In a place already visited, new combinations of features become scarce, so the update time process decreases. However, in a bigger environment, other techniques have to be implemented. A possible way to reduce the processing time is to split the environment representation into local lattices. For instance

a lattice may cover a place and its topological neighbors. We are currently investigating this approach.

6. CONCLUSION AND PERSPECTIVES

In this paper, a new original supervised classification method has been developed to classify images *w.r.t.* to the place they have been taken. This method is strongly based on visual landmarks, anybody or *anything* needs to locate oneself.

Our algorithms have been validated first through real images issued from four different places of a structural environment, second through a comparison with an optimized neural network that gives lower quality results with a lot of instability, and finally through a real experimentation with a autonomous mobile robot.

In this last case, a lot of heuristics could be developed to improve results, especially in introducing local constrains such as connected -or not- places, probabilities of transition, *etc.* However, our objective here was to validate our algorithms in the worst case, *i.e.* in a pure classification problem without any *a priori* knowledge.

Our system is open, *i.e.* other attributes from any captor could be used, or high-level attributes depending on the final purpose (*e.g.* "rectangles" for building in outside urban scenes). Thus we may incorporate "partially predefined landmark" in the sense of paragraph 2.1. Such an approach will be probably needed to process more complex tasks such as outdoor localization in partially unknown environment. However, in our applicative context, it was not necessary and this is worth to be noted.

Four main directions will lead our further research program. First, we have to improve our primitives and features, in order to obtain a more stable and wider range of landmarks for the different classes. Second, we have to find a way to associate a symbolic classifier such as the Concept Lattice classifier we developed herein and a numerical classifier such as Neural Network to improve results. Indeed, results from these two techniques seem to be complementary, and probably Galois lattices could preprocess a neural network classifier through pre-selecting features. Afterwards, it would be valuable to introduce recent classification techniques such as "support vector machines". Classification failures occur often on topological boundaries of the sites. Support vector techniques are supposed to help getting a more robust classification. Notice that the concept of margin is closer in its spirit from our "no decide" symbolic classifier. It is also important to investigate unsupervised classification methodologies to induce the creation of new classes *i.e.* nodes of the topological map.

Support vector techniques are supposed to help getting a more robust classification and/or to induce the creation of

new classes *i.e.* nodes of the topological map. Finally, in a more applicative way, our goal is to allow a robot to build a topological map of structured -or not- environment, in a fully autonomous process.

7. REFERENCES

[1] M. Knapek, R.-O. Oropeza, and D. Kriegman, "Selecting promising landmarks," *IEEE Int. Conf. on Robotics and Automation*, 2000.

[2] G. Bianco and A. Zelinsky, "Biologically-inspired visual landmarks learning and navigation for mobile robots," *IEEE/RSJ, IROS*, 1999.

[3] Z. Dodds and G. Hager, "A color interest operator for landmark-based navigation," in *AAAI/IAAI*, 1997, pp. 655–660.

[4] G. Dudek and D. Jugessur, "Robust place recognition using local appearance based methods," *IEEE Int. Conf. on Robotics and Automation*, 2000.

[5] I. Ulrich and I. Nourbakhsh, "Appearance-based place recognition for topological localisation," *IEEE Int. Conf. on Robotics and Automation*, 2000.

[6] Robert Sim and Gregory Dudek, "Learning visual landmarks for pose estimation," in *Int. Conf. on Computer Vision (2)*, 1999, pp. 1217–1222.

[7] M. Mata, J. M. Armingol, A. de la Escalera, and M.A. Salichs, "A visual landmark recognition system for topological navigation of mobile robots," *IEEE International Conference on Robotics and Automation (ICRA01)*, 2001.

[8] S. Thompson, T. Matsui, and A. Zelinsky, "Localisation using automatically selected landmarks from panoramic images," *ACRA*, 2000.

[9] D. Lowe, "Object recognition from local scale-invariant features," *Int. Conf. on Computer Vision*, 1999.

[10] J.B. Hayet, F. Lerasle, and M. Devy, "A visual landmark framework for indoor mobile robot navigation," *Int. Conf. on Robotics and Automation*, 2002.

[11] C. Harris and M. Stephens, "A combined corner and edge detector," *Alvey Vision Conference*, december 1988.

[12] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 11, pp. 1254–1259, 1998.

[13] A. Arleo, F. Smeraldi, S. Hug, and W. Gerstner, "Place cells and spatial navigation based on 2d visual feature extraction, path integration, and reinforcement learning," *Advances in Neural Information Processing Systems 13, MIT-Press*, 2001.

[14] P. Gaussier, C. Joulin, J.P. Banquet, S. Leprêtre, and A. Revel, "The visual homing problem : An exemple of robotics/biology cross fertilization," *Robotics and Autonomous Systems*, 2000.

[15] C. Schmid, R. Mohr, and C. Bauckhage, "Comparing and evaluating interest points," *ICVV*, 1998.

[16] S. Kuznetsov and S. Obiedkov, "Algorithms for the construction of concept lattices and their diagram graphs," *Principles of Data Mining and Knowledge Discovery (PKDD 2001), Freiburg, Germany*, 2001.

[17] R. Godin, G. Mineau, and R. Missaoui, "Incremental concept formation algorithms based on galois (concept) lattices," *Computational Intelligence, vol 11*, 1995.

[18] M. Liquière and J. Sallantin, "Structural machine learning with galois lattice and graphs," *International Conference on Machine Learning*, 1998.

[19] L. Chaudron, N. Maille, and M. Boyer, "The (cube) lattice model and its applications," *Applied Artificial Intelligence, vol 19, no 3*, 2003.

[20] M. Barbut and B. Monjardet, *Ordre et Classification*, Hachette Université, 1970.

[21] B. Ganter and R. Wille, *Formal Concept Analysis*, Springer-Verlag, 1999.

[22] M. Boyer, *Induction de régularités dans une base de connaissance, application au phénomène bruit/Gène et ses extensions*, Ph.D. thesis, École Nationale Supérieure de l'Aéronautique et de l'Espace, 2001.

[23] S. Kuznetsov and S. Obiedkov, "Comparing performance of algorithms for generating concept lattices," *Journal of Experimental and Theorical Artificial Intelligence*, 2002.

[24] E.M. Norris, "An algorithm for computing the maximal rectangles in a binary relation," *Revue Roumaine de Mathématiques Pures et Appliquées 23 (2)*, 1978.

[25] P. Ranganathan, J.B. Hayet, M. Devy, S. Hutchinson, and F. Larasle, "Topological navigation and qualitative localization for indoor environment using multisensory perception," *Int. Symp. on Intelligent Robotic Systems, Toulouse*, 2001.