2004 IEEE Workshop on Machine Learning for Signal Processing

LEARNING-BASED VISUAL LOCALIZATION USING FORMAL CONCEPT LATTICES

Manuel Samuelides† and Emmanuel Zenou†‡

†Department of Applied Mathematics, Information and Control Lab,
SUPAERO, 10 av. Edouard Belin, BP4032, 31055 Toulouse Cedex, France
‡LAAS-CNRS, 7 avenue du Colonel Roche, 31077 Toulouse Cedex 4, France
E-mail: {samuelid,zenou}@supaero.fr

Abstract. We present here a new methodology to perform active visual localization in the context of autnomous mobile robotics. The robot is endowed with a topological map of its environment. During the learning phase, the robot takes a lot of pictures from the environment; each picture is labelled by its origin place in the topological map. After the learning phase, the robot is supposed to locate itself in the learnt environment using the visual sensor. Since the discriminating information is sparse, the usual supervised classification techniques as neural networks are not sufficient to perform efficiently this task. Therefore, we propose to use a symbolic learning approach, the "formal concept analysis". The relevant information is gathered into one concept lattice. A formal classification rule is proposed to achieve localization on the topological map. In order to improve the response rate of the decision process, the original formal landmark set is extended to plausible landmarks for a given confidence level. Experimental results in a structured environment support this approach. Perspectives for implementing active strategy to look for visual information and to improve on-line learning and localization process are presented in the final discussion.

INTRODUCTION

Finding its place in an environment is a difficult challenge today for an autonomous mobile robot. This robot needs to know how to characterize and to recognize a place by itself, to be considered as fully autonomous in term of orientation and navigation. We study here the case where the robot uses mainly visual sensors to locate itself. The most classical approach in that case is to select objects in the environment and to provide to the localization system the object characteristics and their localization in order for the robot

0-7803-8608-6/04/$20.00 ©2004 IEEE

to recognize them. These objects are called "landmarks". In [11] the authors discussed the concept of landmark to answer the following questions : what is a landmark ? How to select it ? A taxonomy of landmark is considered (fully pre-defined landmarks, partially pre-defined landmarks and non pre-defined landmarks).

Once landmarks are recognized, the robot uses them to determine its position more or less precisely. Generally, the landmarks are associated to a localization on a map. The map is a global model of the environment. For a synthetic survey about robot localization through mapping, the reader is referred to [10]. There are several kinds of maps. A metric map seems more precise. The robot infers its localization straightforward from the snapshot of the environment that is currently shot. There is an alternative approach which is called the "topological mapping" approach [6]. In that approach, the robot recognizes the place where it is and not exactly its metric position. This approach is thus more qualitative than quantitative. Indeed, metric localization is efficient in a local context, but as soon as the environment is growing up, or as soon as objects could move (typically in a human environment), metric localization may become very costly and less stable than the topological approach. Moreover modern processing tend to be hybrid and modular combining the two approaches [9].

Recently, the landmark approach was criticized for its drawbacks with respect to flexibility and autonomy. Learning approaches were introduced to enable the robot to build by itself, in an incremental way or in a batch way, its own environment representation which will provide the basis for localization. Some approaches are more biologically oriented and are based upon template matching and movement fields as [2],[4]. Others are based on Bayesian localization and optimize feature extraction with respect to average posterior localization error as [8]. These approaches can be applied in more general applicative contexts than classical landmark-based approaches. They are presented in the litterature in a metric based localization framework.

We propose here a learning-based localization approach in a topological based localization framework. We believe that topological mapping is an useful mediation for difficult localization problems in a structured or semi-structured environment. Our learning process is based on a data base which consists in a set of snapshots that are labelled by their places of origin. The use of snapshots is similar than in [8] with this significant difference: in our case, the snapshots are labelled with their topological localization (qualitative), whereas in [8], they were labelled with their metric position (quantitative). Instead of designing optimal fiter of features for the average localization error, we propose here a symbolic discrimination approach: the formal concept analysis. Experimental results show it is more efficient than a neural classifier. Afterwards, we show how it is possible to combine this approach with statistical discrimination to improve the performance of the system. We precise future research directions in the final discussion. In the next section, we sketch the basis of formal concept analysis and how we have developped it for classification purpose.

FORMAL CONCEPT ANALYSIS

Concept lattices [1][3]

Definition 1 *A* **lattice** *is defined as an ordered set in which two any elements have 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 inclusion $\subset$ is an order relation on the set $\mathcal{P}(\mathcal{O})$ of all subsets of a set $\mathcal{O}$. Each family of subsets has a lub, which is the intersection and a glb which is the union. Thus $\mathcal{P}(\mathcal{O})$ is a complete lattice. Notice that in that classical case, the lub and the glb are distributive with respect to each other but this is not the general case. The lattices that will be considered hereafter are generally not distributive.

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 $\{0, 1\}$.*

The interpretation of ζ is that $\zeta(o, f) = 1$ if and only if object o is endowed with the attribute ζ.

Definition 3 *Given a context $(\mathcal{O}, \mathcal{F}, \zeta)$ 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*

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

and

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

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 any attribute belonging to $\mathcal{B}$.

The properties of the Galois connections can be found in [3]. Let us recall the algebraic following properties that are commonly used:

$\mathcal{A}_1 \subset \mathcal{A}_2 \Rightarrow \mathcal{A}_2' \subset \mathcal{A}_1'$; $\mathcal{A} \subset \mathcal{A}''$; $\mathcal{A}' = \mathcal{A}'''$;

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 $\mathcal{A}' = \mathcal{B}$ et $\mathcal{B}' = \mathcal{A}$.*
$\mathcal{A}$ 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 $\mathcal{B} = \mathtt{intent}(\mathcal{C})$.*
A feature set (resp. an object set) is called a **concept-based feature set** *(resp. a concept-base object-set) if it is the intent (resp. the extent) of a concept.*

A concept is thus a particular association between a set of images $\mathcal{A} \in \mathcal{O}$ and a set of features $\mathcal{B} \in \mathcal{F}$. $\mathcal{B}$ is included inside all images of $\mathcal{A}$, but not in any other image. In the same time, $\mathcal{B}$ is the larger set of features present in all images of $\mathcal{A}$. One proves [3] the following:

Theoreme 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.*

These lattice operators may be extended to any concept set. 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. The set of concepts $\mathcal{L}$ is endowed with the order relation $\subset$ of its extents, it is a complete lattice. We can set:

Definition 5 *The complete lattice $\mathcal{L}(\mathcal{K})$ of concepts of the context $\mathcal{K}$ is called the* **concept lattice** *of the context*[1].

The extent of the smaller element of $\mathcal{L}$ is the set of objects that share all features of $\mathcal{F}$. So it is not generally $\emptyset$. Actually, the formal definition of a context that is given above is too general and includes unnecessary redundancies. It is possible to cancel trivial features which bring no information.

Definition 6 *Let $\mathcal{K} = (\mathcal{O}, \mathcal{F}, \zeta)$ be a context. If $(\emptyset, \mathcal{F})$ and $(\mathcal{O}, \emptyset)$ are concepts, then the context $\mathcal{K}$ is said* **proper**.

In case the context is not proper, at least one feature is present in all images, and/or at least one image has all features. If a feature f is present in all images on the learning set, its utility is nil. In this case, either the set of images in the learning set has to be enlarged, or the feature f has to be cancelled. If one image of any class θ has all features, no combination of features could discriminate θ from another class. In this case, classification cannot be done properly, and the set of features has to be enlarged.

In case the context is proper, $(\emptyset, \mathcal{F})$ is the smaller concept (the more restrictive), and $(\mathcal{O}, \emptyset)$ is the greater concept (the more general). There is no canonical complementary in Galois lattice: if $\mathcal{C} = (\mathcal{A}, \mathcal{B})$ is a concept, generally neither $(\mathcal{B}^{c\prime}, \mathcal{B}^c)$ nor $(\mathcal{A}^c, \mathcal{A}^{c\prime})$ are concepts.

Formal classification in a partitioned Context

Definition 7 *Let $\mathcal{K} = (\mathcal{O}, \mathcal{F}, \zeta)$ be a context and suppose that the object set $\mathcal{O}$ is divided according a partition $(\mathcal{O}_\theta)_{\theta \in \Theta}$. An element θ of Θ is called a* **class** *and Θ is the set of classes.*
Given a class θ, a concept-based feature set $\mathcal{B}_\theta$ is said to be a **criterion** *for θ if and only if*

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

Thus a formal criterion is the intent of a concept, which extent is included in the class of interest. Any image, from the learning set, checking this

[1] Originally it was called in [1] the **Galois lattice** from the idea of the duality between objects sets and attribute sets, e.g. a set of numbers and a set of polynomials that are annulated by these numbers.

criterion belongs to the class θ, we shall use such a criterion to decide whether a new image from the testing set belongs to θ or not.

Once all the criteria are extracted from the partitioned context, a global classification rule is proposed.

Definition 8 *The **formal classification rule** is a partial application from $\mathcal{O}$ to Θ , which is defined as follows :*

- *given $o \in \mathcal{O}$, if there exists a class θ and a criterion $\mathcal{B}_\theta$ for which $o \in \mathcal{B}'_\theta$, then the formal classification rule associates θ to o,*
- *if there is no criterion $\mathcal{B}$ such that $o \in \mathcal{B}$ then the formal classification rule gives no response to classify o.*

Notice, the formal classification rule is well defined in the sense that no object can belong to the intersection of extents of criteria for distinct classes since such an intersection is void.

In particular, if the criterion $\mathcal{B}_\theta$ is a set of attributes that are simultaneously present in *all* the objects of the class, it is called an *exhaustive criterion.* The existence of an exhaustive criterion is of course welcome to provide a mild characterization of the objects of the relevant class. If there is no exhaustive criterion, it is interesting to limit the number of criteria to be considered and to restrict oneself to *maximal criteria.*

Definition 9 *A* **maximal criterion** $\hat{\mathcal{B}}$ *for the class θ is a criterion of minimal intent among all criteria for θ.*

Learning and generalization for the formal classification

Up to now, the classification rule is defined only on the learning set from which the concepts have been extracted. Notice that for that classification rule, **by construction**, there cannot be any learning error.

Of course, it is possible to retrieve the learnt classification rule in a generalization phase. The formalization is obvious. One has to consider a partitioned context $(\mathcal{O}, \mathcal{F}, \zeta, (\mathcal{O}_\theta))$. Let $\mathcal{O}'$ be a subset of $\mathcal{O}$ that is called the **learning set**. Then we defined the restricted partitioned **learning context** $(\mathcal{O}', \mathcal{F}, \zeta', (\mathcal{O}'_\theta))$ where ζ' is the restriction of ζ' on $\mathcal{O}' \times \mathcal{F}$ and where $(\mathcal{O}'_\theta = \mathcal{O}' \cap \mathcal{O}_\theta)$ is the trace of the partition on $\mathcal{O}'$. We extracted the formal concept lattice from this learning context and we define accordingly the formal classification rule which will be implemented on the whole object set $\mathcal{O}$. Of course, there may be generalization errors for the formal classification rule. However, due to the existence of "no response" ability and the restriction on the classification rule that was imposed by the formal concept analysis, one can expect that the generalization error will be smaller than with a usual statistical-based classifcation rule. The experimentations that are presented below show this is the case for image classification focused onto visual localization.

Visual landmark learning using formal concept analysis

During the learning phase, all pictures are shot from the robot camera in the different places of the environment. Thus, a set of images is attached to each place of the environment. From these pictures, primitives (segments from polynomial contour extraction, color from the HSV space, objects quantified with morphological operators. . .) are extracted to find features, which are supposed to be more robust than primitives.

The algorithms used to extract primitives from images are quite classical. For instance, to obtain segments, the contours are extracted with a Canny-Deriche algorithm and approximated with polynomial figures, from which segments are extracted. Features are issued from primitives. Other primitives are found through image color or texture segmentation. For instance, "there is a yellow object" or "there is a large number of identical (orientation and size) segments" (shelf) are typical features, more robust against rotation, translation and scaling than primitives. To quantify such an information, morphological operators are used in the six binary images. Indeed, given for instance one image and the associated real "red" image in the normalized HSV space, we clean it with opening and closure operators and we finally get a minimum rectangle red object in the original image if and only if the result of an erosion is not an empty image.

At the end of this preprocessing phase, we are able to provide a proper partitioned context for the formal concept analysis. The objects are the images, the classes are the the places from which these images have been shot, and the feature extraction has just been previously described. Thus the learning basis consists in the occurence table of features into labelled images. In this visual localization application, we shall call **landmark** a criterion issued from the formal concept analysis of the partitioned context. Thus, this type of landmark refers to an abstract combination of sensor features and not necessarily to specific objects.

To implement formal concept analysis, one has to implement efficient lattice building algorithms. The complexity of the algorithms is exponential *w.r.t.* the size of the context, but for some algorithms is linear in the number of concepts modulo some polynomial of the input size. See [5] for a complete overview on algorithms. We have used in our experimentation an incremental algorithm called the *NORRIS* algorithm [7]. We report hereafter the full occurence table processing results. We shall present in the following section an improvement using Bayes discrimination of the concepts.

Experimental Results

To support this approach, we have used the environment of SUPAERO Information and Control laboratory. We used the robot *Pekee* that was equipped with a CCD color camera. A collection of 177 snapshots have been obtained, originating from four rooms of the lab. It constituted the learning base.

66 Potential features have been searched in the 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 and contour extraction.

The corresponding 5265 concept lattice was computed in 8 seconds on a AMD Athlon 2400+ machine. For the four classes, 883 landmarks were extracted, and 42 maximal landmarks were kept : 9 for the first place, 8 for the second one, 17 for the third one, and 8 for the fourth one.

Then an additional collection of 151 labelled images was used as the test set to measure the performances of the various classifiers. The results of the formal classification rule are reported in table 1.

For instance, among 32 images issued from place 1, 1 image contains 2 contradictory landmarks (one fore place 1, one for place 3) and 14 no landmark at all; 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%, and a false response rate of 3.1%.

Place	**#Images**	**#Responses**	**#Good**	**#False**
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 (95%)	**3 (5%)**

TABLE 1: TEST RESULTS OF THE FORMAL CLASSIFICATION WITH THE NUMBER OF IMAGES, THE NUMBER OF RESPONSES, THE NUMBER OF GOOD RESPONSES AND THE NUMBER OF FALSE RESPONSES

FORMAL ANALYSIS VERSUS BAYES STATISTICS

To assess the formal classification, we compared it with a statistical-based classification using neural networks. We provided the same learning base to a feed-forward neural network. The Matlab Neural Network toolbox was used for that issue. Several experimentations have been achieved 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 adaptative learning rate, with an hyperbolic tangent sigmoid transfert function for each layer of the network. Other simulations have been tempted 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.

Second-order algorithms failed due to the high number of entries. 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 were reached once on five or six tries.

The smallest error rate we obtained is 5% on the learning set of images, and 30% on the testing set. The obvious cause of such a distortion of the results is due to the allowance of non response for the formal classification rule. This property is acceptable in the context of mobile robot localization and active vision. If there is no response, the robot may be programmed to look for additional information, changing the attitude of the camera and shooting other pictures. Indeed, it is easy to alleviate the error rate of the neural based classifier allowing an adjustable non response rate. If we tune this rate to be identical to the previously quoted formal classification one, we obtain results that are reported in table 2

Place	**#Images**	**#Responses**	**#Good**	**#False**
Place 1	32	17 (53.1%)	17 (100%)	0 (0%)
Place 2	50	17 (26%)	17 (100%)	0 (0%)
Place 3	31	14 (32.3%)	10 (71.4%)	4 (28.6%)
Place 4	38	12 (31.6%)	10 (83.3%)	2 (16.6%)
Total :	151	60 (39.8%)	60 (90%)	**6 (10%)**

TABLE 2: TEST RESULTS OF THE OPTIMIZED NEURAL NETWORK ALLOWING THE SAME NON RESPONSE RATE THAN THE FORMAL CLASSIFIER.

There is still an important difference of the error rates between the two methodologies and it seems that the formal classification rule is more able that the MSE optimizer to select a good classification rule including a non response issue. However, the formal classification rule that was described above does not allow adjusting the non-response rate. The trade-off between non-response and false response is likely to be non-optimal. In order to obtain a smaller non-response rate, we introduce an extension of formal landmark with the notion of plausible landmark:

Definition 10 *Let us denote* $\pi(. \mid \mathcal{B})$ *the* ***empirical distribution*** *of classes conditionned by a concept* $\mathcal{B}$, *i.e.* $\pi(\theta \mid \mathcal{B}) = \frac{\#(\theta \cup \mathcal{B}')}{\#\mathcal{B}'}$

A concept $\mathcal{B}$ *is a* ***plausible landmark for*** θ ***with confidence level*** α *if and only if*

- $\theta = \arg\max_{\vartheta \in \Theta} \pi(\vartheta \mid \mathcal{B})$
- $\pi(\theta \mid \mathcal{B}) > \alpha$

The set of all plausible landmarks with confidence level α *is denoted* $\mathcal{L}_\alpha$.

Notice it is possible to embed the empirical distribution computation into the original Norris incremental algorithm [7] to build the formal concept lattice. Thus, the additional computational charge to get the empirical distribution for each concept is small. So, we get a new landmark set which is larger than the previous formal landmark set. Actually the formal landmark set is exactly $\mathcal{L}_1$ the plausible landmark set with confidence level 1.

Then it is possible to apply the same classification rule while using $\mathcal{L}_\alpha$, which was previously implemented with $\mathcal{L}_1$. However, we adopt here a different approach. There are two reasons for no response: either there is no landmark present in the current image or on the contrary, landmarks for different locations may be present in the same image. By enlarging the landmark set from $\mathcal{L}_1$ to $\mathcal{L}_\alpha$, we restrict the first cause but we enlarge the second one. So we shall consider among all the present landmarks in an an image the more representative one. For doing so, we introduce an empirical choice rule by associating to image $\mathcal{I}$ the plausible concept $\mathcal{B}(\mathcal{I}) = \arg\max_{C \in \mathcal{L}_\alpha} \frac{\#(C \cap \mathcal{I}')}{\#(C \cup \mathcal{I}')}$ We recall that $\mathcal{I}'$ is the set of features that are present in image $\mathcal{I}$. Then the class of $\mathcal{I}$ is defined straightfroward from $\mathcal{B}(\mathcal{I})$.

With confidence level of 0.75, using the same data set than has been presented above we got the following experimental results:

Place	**#Images**	**#Responses**	**#Good**	**#False**
Place 1	32	25 (53.1%)	24 (96%)	1 (4%)
Place 2	50	14 (28%)	11 (71,4%)	3 (28,6%)
Place 3	31	12 (38.7%)	12 (100%)	0 (0%)
Place 4	38	22 (57.9%)	22 (100%)	0 (0%)
Total :	151	73 (48.3%)	69 (94.5%)	**4(5.5%)**

TABLE 3: TEST RESULTS OF THE PLAUSIBLE LANDMARK BASED CLASSIFICATION RULE.

Even if the final result is quite identical in term of good response rate (94.5% against 95%), the number of well localized images is much better (73 against 60, i.e. 21.7% much better), that was the goal of introducing plausible landmarks.

DISCUSSION

Our original approach using Galois lattices gives very good results compared to classical classification techniques such as neural network for learning "visual landmarks" or "feature filters". Moreover, the extension of our original landmark concept to plausible landmarks with a given confidence level brings improvements of the experimental results. The final reported experiment is promising.

However neural networks may perform better in places where there is less discriminating landmarks (place #1 and #2 in our application). Indeed, neural networks structurally determine the optimal weighted combination of *all* features, whereas concept-based landmarks concentrate on particular discriminating feature subsets. This is thus a disparity between an optimal average approach and an optimal specific approach. Recent statistical machine learning theory seems to favor the second approach. Moreover, neural network performances decrease while the number of neurons in the first layer (features) increases.

Up to now, the formal concept analysis algorithm was not limiting for hundred of images and about a hundred features. Still, we used incremental algorithms that support on-line learning and we had not used the additional trick of cancelling non informative feature. We may face computational complexity problems when we shall implement on-line learning.

We consider the present study as a necessary step to turn to active vision and autonomous learning. We are currently experimenting an active vision strategy where the robot will shoot new pictures to get a response.

As far as learning is concerned, the next step of this research will consider the situation where only a few topological nodes of the maps are given and where the localization system has to decide to create new nodes to fill the blanks of this map.

REFERENCES

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

[2] G. Bianco, A. Zelinsky and M. Lehrer, "Visual Landmark Learning," in **Proceedings of IEEE/RSJ Intelligent Robots and Systems**, 2000.

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

[4] P. Gaussier, C. Joulin, J. Banquet, S. Leprêtre and A. Revel, "The Visual Homing Problem : An Exemple of Robotics/Biology Cross Fertilization," **Robotics and Autonomous Systems**, 2000.

[5] S. Kuznetsov and S. Obiedkov, "Comparing Performance of Algorithms for Generating Concept Lattices," **Journal of Experimental and Theorical Artificial Intelligence**, 2002.

[6] M.Mataric, "Integration of representation into goal-driven behavior-based robots," **IEEE Transactions on Robotics and Automation**, vol. 8, no. 3, pp. 304–312, 1992.

[7] E. Norris, "An Algorithm for Computing the Maximal Rectangles in a Binary Relation," **Revue Roumaine de Mathématiques Pures et Appliquées 23 (2)**, 1978.

[8] S. Thrun, "Bayesian Landmark Learning for Mobile Robot Localization," **Machine Learning**, vol. 33, no. 1, pp. 41–76, 1998.

[9] S. Thrun, "Learning Metric-Topological Maps for Indoor Mobile Robot Navigation," **Artificial Intelligence**, vol. 99, no. 1, pp. 21–71, 1998.

[10] S. Thrun, "Robotic Mapping: A Survey," in G. Lakemeyer and B. Nebel (eds.), **Exploring Artificial Intelligence in the New Millenium**, Morgan Kaufmann, 2002.

[11] E. Zenou and M. Samuelides, "Characterization of Image Sets : The Galois Lattice Approach," in **International Conference on Advanced Concepts on Intelligent Vision Systems, Ghent, Belgium**, 2003.