A Novel Approach for Graphics Recognition based on Galois lattice and Bag of Words representation

Amani Boumaiza
LORIA UMR 7503 - University of Nancy 2
Nancy, France
boumaiza@loria.fr

Salvatore Tabbone
LORIA UMR 7503 - University of Nancy 2
Nancy, France
tabbone@loria.fr

***Abstract*—This paper presents a new approach for graphical symbols recognition by combining a concept lattice with a bag of words representation. Visual words define the properties of a graphical symbol that will be modeled in the Galois Lattice. The algorithm of classification is based on the Galois lattice where intentions of its concepts are visual words. The words as visual primitives allow to evaluate the classifier with a symbolic approach that no longer need a signature discretization step to build the Galois Lattice. Our approach is compared to classical approaches on different graphical symbols and we show the relevance and the robustness of our proposal for the classification task.**

***Keywords*-Graphics Recognition, Symbol Classification, Shape descriptors, Local Descriptors, Visual words, Bag of Features, Concept Lattices.**

I. INTRODUCTION

The aim of any computer vision application is the recognition of information which is contained into images. This information allows to describe, index and search for images. In this perspective, symbol recognition is at the center of many recognition systems. The drawings, maps and diagrams use graphical notations which are dependent on their domain of application. The automatic interpretation of graphical document requires a process that recognizes the corresponding alphabet for each symbol. In general, a symbol can be defined as a significant graphical entity specifically for a domain of application. For example, an architectural plan consists of different types of symbols depending on whether it represents doors, windows or tables. Many representation methods have been proposed [4]. Choosing a method of representation is generally related to the type of application. This choice has a direct influence on graphics recognition results. More precisely, local methods which are based on the extraction of interest points showed a robustness to scale transformations and occlusions. There are many descriptors using this representation such as: SIFT [8], SURF [2] and GLOH [9]. Supervised classification is a task of mining data that consists on building a classifier from samples which are labeled by their class (learning phase), and then predict the class of new examples with the classifier (classification phase). Classical approaches for image recognition use classical classifiers such as KNN, decision trees or Bayesian networks [1] that have already been successfully applied to many learning and pattern recognition problems. On the other hand, the growing interest in Formal Concepts Analysis (FCA) since 2000, either in the field of data mining or in knowledge representation has risen the use of Galois lattice structure. The lattice of concepts is a graph, its intents are related to objects via a binary relation $\{AttributesxObjects\}$. The nodes of the graph are concepts, a concept is a grouping of items with attributes. The Galois lattice which is composed by a set of concepts related by inclusion, provides a very intuitive representation of data. A previous study [5] provides a comparison between different methods of supervised classification which are based on a Galois lattice, and where experiments clearly show that the concept lattice provides an interesting framework for classification, despite the exponential complexity in some cases. Navigala method [6] was designed to recognize symbols from technical documents with discrete signatures extracted from graphical symbols.

The originality of our approach is an adaptive method of symbol classification based on the Galois lattice and using a bag of words representation. For the vectorization phase and data representation, we exploit the approach of bag of words where symbols are represented by vectors of frequencies as visual words. Local features are extracted from images and then clustered using a k-means algorithm where each cluster represents a visual word. Then, a vector of visual words frequencies is defined and for each image we associate the cluster to its corresponding visual word. This method is used for image classification and it can replace the classical approach that is based on numerical signatures. Furthermore, the lattice is built using classes rather symbols. That is, each node of the Galois lattice corresponds to a class represented by its extension (instances of class) and intention (the common properties of the class presented with all visual words extracted previously). Visual words are defined as symbolic concepts, and do not need as in [6] to be discretized to built the concept lattice which is used as a classifier in our work.

1520-5363/11 $26.00 © 2011 IEEE
DOI 10.1109/ICDAR.2011.170

II. Proposed Approach

The overall method is described in Figure 1. The process contains three major steps: the first one is to extract salient image regions (interest points), and represent it with feature vectors. The next step is to quantify cluster features into visual words. Then, the bag of features is used to build the visual Galois Lattice in order to classify symbols.

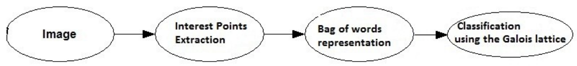

Figure 1: Overall algorithm.

A. Interest Point Detection

Object detection is a preprocessing step based on the extraction of interest points which are usually used to define correspondences between images, recognize textures or classify objects. We evaluated several corner detectors used in the literature and choose the Harris detector. Then, for the representation we use the SURF algorithm [2] which describes a distribution of responses of the Haar wavelet in the neighborhood of the interest point. SURF is made of 64 gradients and the image of the object to be characterized is divided into several cells counting the occurrences of gradient orientations in a histogram. So, each point of interest is described by a local descriptor which uses the oriented gradients histogram based on a calculation of the gradient in a simple and effective way. We choose to use this algorithm instead of the SIFT method since it is not only used to detect but also to characterize, in order to recognize these areas by matching points of interest in other images of the same scene. This algorithm has a very important success in the community of vision, but also outside the community, and numerous versions have been proposed.

B. Bag of Words representation

To represent an object, we define a signature that summarizes relevant information and characteristics of the symbol. That is, after the detection of interest points, we construct the signature for each symbol as a vector of visual words.

Visual dictionary. To construct the visual dictionary [7], each symbol in the training or test set is represented by means of visual words in the dictionary. In practice, building the visual vocabulary is up to quantify the local descriptors space of objects through a clustering method. The number of clusters chosen is actually the size of the dictionary. The center of each cluster represents the visual words in the dictionary (in our experiment set to 300 visual words). A symbol is no longer a visual descriptor but represents a cluster and each object in the database is represented by an histogram of frequencies of visual words. After building the visual dictionary of the whole database, we can proceed to extract vectors of visual words representing each symbol. The bag of words consists on finding occurrences of visual words in the dictionary for each symbol. For each image we associate a vector of visual words and weights which are the frequencies of these words in each symbol.

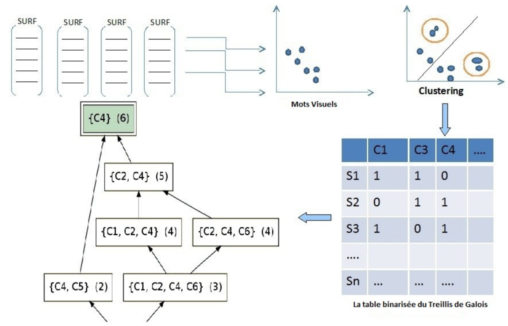

Figure 2: Architecture of the proposed approach.

C. Symbol classification using the Galois lattice

In all previous approaches [3], [6], graphical symbols are represented by a numerical descriptor. Each value of this feature is discretized into a number of intervals following a cutting criterion for example the minimum or maximum entropy. Given a cutting value dividing a feature into intervals, each symbol has a relation with one of the intervals.

Our approach consists on eliminating the step of attributes discretization by the use of the symbolic representation of symbols. An image will be represented by a vector of visual words and the concept lattice is built from the binary relation between the visual dictionary and the graphical symbols and not from the binary relation between intervals and symbols. Our approach is used to represent the relation *Object-Attributes* (see Table 1):

If the object O contains the visual word X then R(O,X)=1 else R(O,X)=0.

When a symbol has to be classified, a bag of features is extracted and the search in the Galois lattice results in the class of the queried symbol. This approach is robust to noise, occlusions or even pieces of symbols. The weights defined in Table 1 assess the importance of a visual word into an image. The importance increases with the number of occurrences of this visual word. In order to reduce the size of the Galois lattice, we opt for grouping symbols sharing same properties.

Grouping symbols into classes. A pattern is a set of properties or attributes. Symbols are grouped into classes corresponding to these basic types. The first idea is to build the concept lattice, using classes rather than symbols.

The new approach is to move from a table containing symbols and attributes to a table of symbols grouped into classes. The clusters obtained in the previous step, allow us to determine classes of symbols which share the same visual features. It should be noted that the step of discretization is not necessary in our approach (based on the bag of features)

Table I: Classification using visual words and its occurrences as attributes.

	$C1^2$	$C2^3$	$C3^2$
Class1			
Symbol1	1	1	**0**
Symbol2	1	**1**	0
Symbol3	0	**1**	1
Class2			
Symbol4	0	1	**1**
Symbol5	0	**1**	0
Symbol6	1	**1**	0
Class3			
Symbol7	0	1	**1**
Symbol8	0	**1**	0

that allows to decrease the time processing and the size of memory. Indeed, for symbolic data, objects can be directly distinguished from each other following its visual words.
However for numerical data in this case [6], discretization is required and involves the creation of disjoint intervals. Objects will be distinguishable following its signatures belonging to such or such intervals.

How to build the Galois Lattice. There is no criteria or parameter to consider in the construction of this graph given that it represents all possible combinations of objects and visual words which are related by a binary relation R.

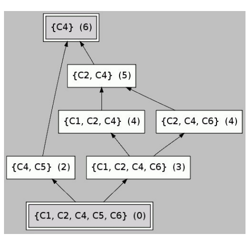

Figure 3: The Galois Lattice.

For example, let's see Table 1: $C1^2$, $C2^3$, $C3^2$ represent 3 visual words which are extracted for each symbol S. We define a number k (the exponent in Ci) which represents the occurrence of a visual word in a symbol. A binary relation is created between the set of symbols and the set of visual words.
If Symbol1 contains the visual word C1, a relation R=1 is created else R=0. We iterate this process for all symbols in the database. The size of the concept lattice constructed with the method using numerical descriptor as attributes is very high while the concept lattice built with the technique of bag of features is smaller (see the number of concepts in Table 3) because it has a reduced number of concepts which limit the problem of complexity related to the large size of a Galois lattice.
After the two phases described above which are respectively data preprocessing and building the concept lattice, we proceed to the step of classification that consists in assigning a class to the requested symbol. To summarize, the main steps of the method of indexing, searching and objects classification that have been adopted in our work are shown in Figure 2. The method involves three steps:

1) Representation of symbols using visual words instead of discrete signatures.
2) Construction of the Hass diagram based on the binary relation between visual words (visual features) and the set of graphical symbols.
3) Classification of the requested symbol using the Galois lattice.

III. Experimental Results

We perform our experiments on the GREC2003 database of segmented symbols (see Figure 4). The basic symbols of GREC2003 contains 10 symbols of each class and 9 levels of degradation. First, we compare our approach with one based on discrete signature where the Hass Diagram of the Galois lattice is built using numerical descriptor discretization.

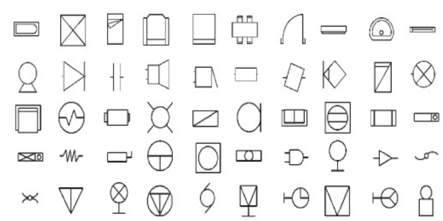

Figure 4: Samples from GREC2003: http://www.cvc.uab.es/grec2003/

As mentioned previously, the Galois lattice is generally based on a phase of attributes discretization, this step is important if we use the numerical signature, but it is eliminated when we work with the visual dictionary. We compared some discretization algorithms in order to evaluate its performance (see Table 2). We remark that the concept lattice gives the best results, when associated with a symbolic approach (97.08%) and does not require a discretization algorithm as the Bin-log: 91.02% or the Entropy distance: 90%. The Bayesian network is more relevant in some cases when combined with numerical signatures than the Galois lattice but our approach is based on this classifier (the Galois lattice) since it can generate useful associations rules which will serve in future work for documents mining.

The bag of words representation, which assigns a vector of frequencies (visual words) for each symbol, permits to group symbols into classes by maximizing the intra-classes similarity and minimizing the inter-class similarity. These classes of symbols reduce the size of the Galois lattice since it is represented by a table containing classes of symbols sharing the same properties instead of isolated symbols.

Table II: Comparison between different algorithms of discretization.

Recognition rate		Discretization Algorithm	
Classifier	**continuous**	**Bin-logl**	**Entropy**
Naive Bayes	96.14	84.29	98.57
BFTree	95.83	90	88.35
J48	92.86	84.29	95.71
NBTrees	92.8	84.2	89.57
Galois Lattice	**97.08**	**91.02**	**90**

The association between the bag of words representation to extract visual features from graphical symbols with the Galois lattice as a classifier provides a high recognition rate, in addition it eliminates the step of the signature discretization used in classical approaches (see the number of discretization steps in Table 3). The concepts of the Galois lattice are no longer an association between symbols and intervals representing discretized signatures but they are made up of symbols and the visual words describing each symbol. Our results are better because a good representation of visual features of graphical symbols offers a reasonable size (see the number of concepts in Table 3), since symbols of the same class are grouped in the same concepts. Thus a Galois lattice with a small size (see Figure 5) means that there are few isolated symbols and the bag of words representation gives better results with a reduced time complexity (see Figure 6).

Table III: Comparison between the approach based on the Galois lattice (GL) with/without the Bag Of Features (BoF).

	Concepts Number	Discretization steps	Time (s)
GL with BoF	3263	0	142.158
GL without BoF	3701	63	229.242

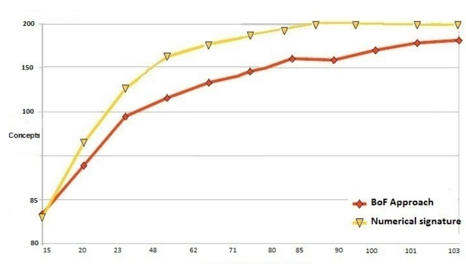

Figure 5: Evolution of the size of the Galois lattice of the approach with/without the bag of features as a function of the number of attributes.

The evaluation of the precision/recall curves show a good performance for our approach (see Figure 7) compared to the classical approach which is based on numerical signatures defined as concepts of the Galois lattice. To summarize, the combination between the Galois lattice as a classifier and the bag of words as a features representation technique shows good results in comparison with the use of numerical

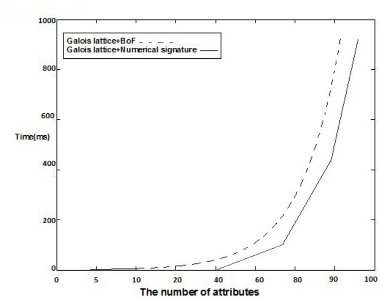

Figure 6: Evolution of the calculation time of the approach with/without the bag of features as a function of the number of attributes.

signatures.

Now, we compare our approach with the one defined in [6]

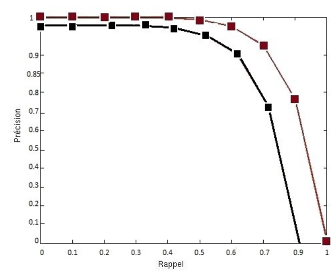

Figure 7: Precision/recall curves: Galois lattice+BoF (in Red) and the Galois lattice+ discretized signatures(in Black).

where the authors used three signatures (Zernike moments, Fourier descriptors and the R-signature). The association between the visual words representation and the Galois lattice (see Table 4) gives a good result (94.4%) in comparison to the Fourier Descriptors (87%), zernike moments (59.3%) or the R-signature (88.1%). The Galois lattice is robust in the classification task in comparison with the decision tree also.

Table IV: Recognition rate by classifier using different signatures.

Recognition Rate(%)	Galois Lattice	kmeans	Decision Tree
SIFT Algorithm	79.7	85.8	70.2
BoF+SURF Algorithm	**94.4**	82.9	87.1
Radon signature	88.1	87.6	75.9
Fourier Descriptors	87	75.8	62
Zernike Descriptors	59.3	60.4	46

A. *Discussion*

Increasing the size of the learning set. We evaluated the recognition rate versus the size of the vocabulary. This experiment aims to evaluate the performance of our recognition system based on the size of the training set. We do a cross validation selecting randomly n symbols of each class for the learning set, the remaining symbols represent the test set. We observe that the recognition rate improves with the increase of the number of learned symbols.

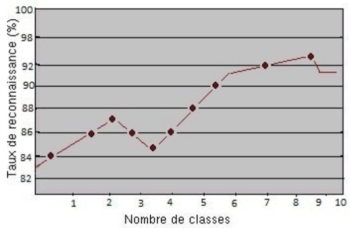

Figure 8: Evolution of the recognition rate (%) versus the number of symbols by class in the learning phase.

Table V: Evolution of the recognition rate (%) versus the size of the dictionary.

Number of class	Dictionary size	Recognition rate (%)
2	100	87.2
5	200	88.1
8	280	93
9	300	94.4
10	500	91.3

However, above a certain number of learned symbols, the recognition rate decreases. The size of the Galois lattice increases with increasing the size of the training set. In conclusion, for a fixed number of classes, the size of the training set improve the recognition rate until a certain level. When the number of classes of symbols reaches a specified threshold, we note that the recognition rate drops and leads us to assume that this is a phenomenon of over fitting (see Figure 8 and Table 5). This problem will be eliminated in our future work by applying a pruning for the Galois Lattice and focus predominantly on frequent items issued from the learning phase.

Features extraction phase. One of the major problems of symbol recognition is to combine segmentation and recognition. This problem is known as the segmentation/recognition paradigm in the literature: a system should segment before recognizing and conversely. The extraction of interest points is a robust processing step required by many visual tasks. Interest points are used to compute local visual descriptors. A local descriptor is calculated from a neighborhood of a sample point. Then, for the classification step, the system has to recognize symbols, giving their labels and their localizations and in our case, we want to check if interest points are efficient to describe symbols. For this reason, we compare a random selection (RS) of points (this method consists on selecting a set of points chosen randomly) and the Harris detector. We can remark that the recognition rate (94.4%) with the Harris detector is good in comparison with the random selection (89.9%) and that argues our choice in this work (see Table 6).

Table VI: Comparison between different methods for Interest Points(IPs) Extraction.

	Harris Detector	RS
Number of PIs	175	175
Calculation Time (ms)	1025	360
Recognition rate(%)	94.4	89.9

IV. CONCLUSION AND FUTURE WORK

In this paper, we propose a novel approach for graphics recognition based on the Galois lattice and bag of words representation. A comparison between the use of discretized signature for the construction of the Galois lattice and the symbolic approach based on visual words has been proposed. The experimental results show the relevance and robustness of our approach. A perspective of this work is to take into account the spatial relations between visual words that represent each symbol and improve the step of features extraction.

V. ACKNOWLEDGMENT

This work is supported by the European project Eureka SCANPLAN.

REFERENCES

[1] S. Barrat and S. Tabbone. A bayesian network for combining descriptors: application to symbol recognition. *International Journal on Document Analysis and Recognition*, 13(1):65–75, 2010.

[2] H. Bay, A. Ess, T. Tuytelaars, and L. J. Van Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346–359, 2008.

[3] M. Coustaty, S. Guillas, M.Visani, K. Bertet, and J. Ogier. On the joint use of a structural signature and a galois lattice classifier for symbol recognition. In *GREC*, pages 61–70, 2007.

[4] D. Frejlichowski and P. Forczmanski. General shape analysis applied to stamps retrieval from scanned documents. In *AIMSA*, pages 251–260, 2010.

[5] H. Fu, H. Fu, P. Njiwoua, and E. M. Nguifo. A comparative study of fca-based supervised classification algorithms. In *ICFCA*, pages 313–320, 2004.

[6] S. Guillas, K. Bertet, and J. M. Ogier. A generic description of the concept lattices classifier: Application to symbol recognition. In *GREC*, pages 47–60, 2005.

[7] F. Jurie and B. Triggs. Creating efficient codebooks for visual recognition. In *ICCV*, pages 604–610, 2005.

[8] D. Lowe. Object recognition from local scale-invariant features. In *ICCV*, pages 1150–1157, 1999.

[9] K. Mikolajczyk and C. Schmid. A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell.*, 27(10):1615–1630, 2005.