VICAL: Visual Cognitive Architecture for Concepts Learning to Understanding Semantic Image Content

Yamina Mohamed Ben Ali

Research Group on Artificial Intelligence, Department of computer Science,
Badji Mokhtar University BP 12, Annaba, Algeria
e-mail: benaliyam2@yahoo.fr

Summary. In this paper, we are interested by the different sides of the visual learning and the visual machine learning, as well as the development of the "visual cognitive" evolution cycle. For this purpose, we present an expected cognitive architecture framework to highlight all the visual learning functionalities. Despite the fact that our investigations were based on the conception of a cognitive processor as a high interpreter of object recognition tasks, we strongly emphasize on a novel evolutionary pyramidal learning. Indeed, this elaborated learning approach based on association rules enables to learn highest concepts induced from concepts of lower level in order to progressively understand the highest semantic content of an input image.

1 Introduction

Visual learning is an attractive trend in object recognition investigation because it seems to be the only way to build vision systems with the ability to understand a broad class of images. Indeed, visual learning is a complex task that usually requires problem decomposition, big amount of data processing, and eventually an expensive time consuming. In most approaches to visual learning reported in the literature, learning is limited to parameter optimization that usually concerns a particular processing step, such an image segmentation, feature extraction, etc. Only, a few contribution attempt to close the feedback of the learning process at the highest (i.e. recognition) level [4, 12, 13, 14, 15, 16, 17, 19, 21].

Acknowledging the need for providing image analysis at semantic level, research efforts set focus on the automatic extraction of image descriptions matching human perceptions. The ultimate goal characterizing such efforts is to bridge the so called semantic gap between low-level visual features that can be automatically extracted from the visual content and the high-level concepts capturing the conveyed meaning. In the followed cognitive architecture we put emphasize on the fact that several levels of description are based on

R.S. Choraś (Ed.): Image Processing & Communications Challenges 2, AISC 84, pp. 15–30.
springerlink.com

the widely held belief that computational vision cannot proceed in one single step from signal-domain information to spatial and semantic understanding.

In this paper, we are interested by the different sides of the visual learning and the visual machine learning, as well as the development of the "visual cognitive" evolution cycle. For this purpose, we have focused our investigations on (i) a novel cognitive architecture able to outperform all visual tasks in the domain of object recognition, and (ii) the approach to interpret from object recognition the higher image understanding.

2 Visual Cognitive Framework VICAL

Though proven the layered VICAL conceptualization from input image to its understanding, we propose a visual machine learning that implements: a *cognitive* behavior for *visual modality*, and a *visual learning* for object recognition. In below, we point out all fundamental features, structural and behavioral, needed in the image preprocessing (features detection) and post-processing (image understanding content-based).

2.1 Eye Processor

With respect to human sensorial organ, eye processor designs the input sensory channel which filters visual problems with the ability of perceiving what must be effectively done relating to the given problem. Moreover, in the elaborated visual architecture, eye processor represents the interface which connects the environment surrounding the *problem space* to the target cognitive behavior.

Its primary functions are limited to analyzing, processing, and interpreting visual information using some elementary operators that refer to low-level image processing. Thus, provided that visual information is somewhere in the image, eye processor encapsulates more than one elementary processor (blind detector processors) to in carry out the segmentation process in a parallel way. This strategy enables an effective exploration without any priori knowledge and reduces the segmentation time consuming.

(i) At coarse processing granularity, the planning schedule of eye processor consists in the following activities:
   - Selection of a visual problem from problems space.
   - Decomposition of the "visual problem space" in more then one visual constraint (object) in order to extract from each visual sub-space the features vector.

(ii) At fine processing granularity, a blind detector processor elaborates a contour detection, extracts relevant features, and then constitutes the geometric model relative to the corresponding visual constraint. The blind detector is therefore capable to:

- Performs an arithmetic mental process (reasoning process) using a buffer memory as a working memory to store, at each detecting step, relevant information needed to encode the geometric model at the next step.
- Returns the corresponding geometric model G_i as a symbolic vector of the smallest object O_i having significance with human comprehension.

Fig. 1 clearly showns that there is no interaction between blind detectors what explains the absence of a communication scheme. This is substantially correct if we consider that detector processors perform independently blind explorations with regard to the whole visual context. Relating to the fixed

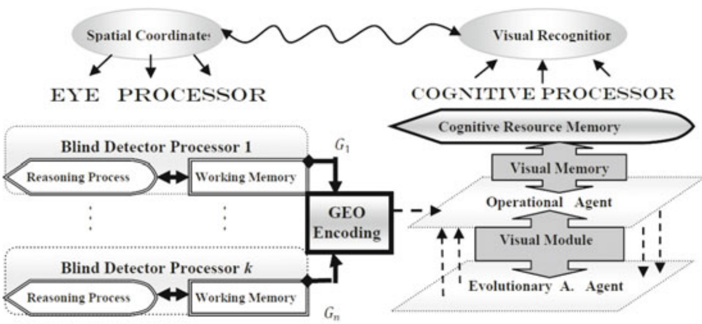


Fig. 1. Cognitive visual memory architecture

aim, the description of the geometric is subjective, because our investigations are concentrated at the cognitive level in which the bridge between the low-level processing and the high-level processing fills the semantic gap. A geometric model can regarded as a vector of interest points of an object in image processing.

2.2 Cognitive Processor

Cognitive processor is a vulgar imitation of the brain in solving problem since logical, complex, and high functionalities such; deduction recognition and understanding are borrowed from the brain functionalities. The cognitive processor is the core of VICAL that lead to desired goal as a result of the interaction occurring at different levels of organization and at different time scales and involving not only the elementary elements (i.e. cognitive resource memory, visual module) but also the behavioral properties emerging from the interactions. Among these properties, we can cite:

- Adaptive behavior, where each input image requires its own resolution scheme improved by the cognitive processor

- Decentralized organization guided by the role of each component.
- Distributed processing with both communicative behavior and cooperative behavior between concerned components (agents).

Cognitive resource memory

Memory is one of the primary domains examined by cognitive psychologists, since encoding, storage, and retrieval of information constitute a significant portion of our cognitive activity. In the context of computer vision, and based on the fact that the memory is context-independent, we define the dominant source of knowledge as a cognitive resource memory.

Definition 1 *(Cognitive resource memory): The cognitive resource memory is a resource repository of multiple connected memories areas whereas each area is roughly expected for a cognitive modality (e.g. visual modality).*

This definition explains that the cognitive resource memory is not presented like a compact whole but in the form of resources network composed by several semantic nodes, each node models one modality of the cognitive memory. Indeed, the cognitive processor as a supervisor disposes of many sensorial agents (captors) to proceed with the environment resources like images, and speech. The acquisition of knowledge is done through cognitive channels under a selective activation. The distribution of knowledge implies that multiple recognition process could be undertaken in parallel and in a collaborative way. Sine the role of each area memory is to enclose a priori knowledge and posteriori conclusions the cognitive processor disposes of specialized cognitive modules to reason about distinct cognitive activities.

In VICAL, we only refer to the visual modality as the effective behavior without worrying about other interesting emergent behaviors. This means a real interest only to basic components required by the visual application.

Visual memory

It is the more concerned area belonging in the cognitive resource memory. We define the visual memory as the visual information repository activated by its corresponding selective channel that points out one of the multi cognitive modalities assigned to the cognitive processor. The basic elements stored and retrieved from visual memory are concepts.

Definition 2 *(Concept): A concept is a symbolic formalization of concrete information in the image possessing semantic content. A concept can be either elementary like: eye, nose, arm, square, etc. or composed like: face, human, cat, house, etc. So, all knowledge acquired through eye processor enriches both visual memory, because of the direct connection established between the two reliable components, and others memory areas in case where the needed information can be useful in the search process. Thus, two meaningful searching strategies are viewed:*

- *Intra search memory. Retrieval process of concepts consists in local interactions between some knowledge domains representing features (concepts) according to appropriate representation formalism.*
- *Extra search memory. Retrieval process of concepts consists in consulting one or more sensorial memories to locate the appropriate database representing the claimed concepts. This operation needs to create semantic interactions as communication between different cognitive resources.*

Visual Module

Before considering objectives and components of the visual module, let us briefly see again the claimed goal: *from an input image, recognize all the salient objects in order to establish the concepts mapping involved to accurately provide the high meaning to the content-image.* At first sight this seems easy to accomplish but actually it is much deeper than that appear. For this reason, we propose to capture the meaning of the image toward a novel learning approach named pyramidal learning. The understanding philosophy behind pyramidal learning is to point out the top of the pyramid in order to achieve not only the highest concept assigned to the image but also the semantic content able to furthermore indexing this image.

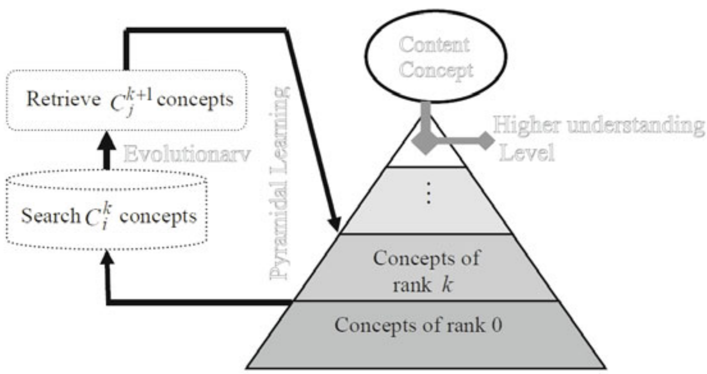


Fig. 2. Evolutionary pyramidal learning cycle

The following algorithm presents broadly the main activities elaborated by the visual module.

1. Select initially formal concepts of rank $k (k = 0)$
2. Locate the appropriate database (used for learning concepts) according to the similarity measure between the given concepts and the keywords used for indexing this database.

3. In the case where such matching is possible, thus apply the evolutionary procedure to learn high concepts and go to step 4. Otherwise, the final formal concepts are proposed as the semantic content of the proceeded image.
4. Replace the formal concepts of rank k by the new learned concepts of rank $k+1$.
5. Compute distances between new concepts.
6. Compare these distances with the empirical threshold to ensure the fusion of some concepts and constitute the new sets of concepts.
7. Go to step 2.

Fig. 2 explains in more detail the pyramid of learning. The basis of the pyramid defines the first learning level with concepts of rank '0' These concepts (provided by the eye processor) are obtained after decoding geometric models of objects such: $\{G_1, G_2, \ldots, G_n\} \rightarrow \{C_1^0, C_2^0 \ldots, C_n^0\}$. After a learning step, the concepts of superior rank ($k = 1$) are retrieved and construct the next bricks of the pyramid. This process is reiterated until the top of the pyramid is achieved by given the highest concept attributed to the image.

The above description concerns only the behavioral aspect of the visual module, remains now to describe its organizational aspect. Thus, Fig. 3 outlines the eventual multi-agent description where agents shared the same visual memory. The main activities clustered under two distinct sub-goals give more details on the specialization of the agents and their thoughtful occurrence in the cognitive processor to tackle visual recognition problems.

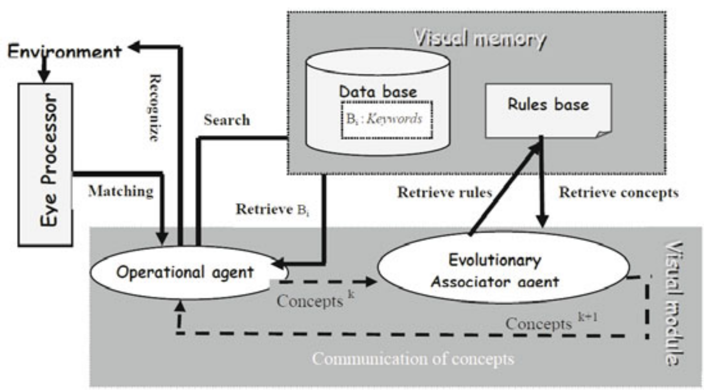


Fig. 3. Diagram of interactions between agents for solving visual problem

3 Structural Abstraction of VICAL

This section discusses the adequate organizational structure of VICAL. Effectively, based upon the fact that the recognition process is strongly influenced by two cognitive properties: evolutionary concept learning and restructuring, the suitable organization is therefore a *role-guided* structure. A comprehensible solution that makes this possible is to settle on a *multi-agent* organization. This choice is not trivial since raise a distributed processing guided by two distinguish roles. For more comprehension, we propose the followings definitions.

Definition 3 *(Agent): An agent is a cognitive entity active by its deterministic role modeled by the cognitive property.*

Definition 4 *(Role): A role is a cognitive property that encloses some related activities.*

In agreement with all considerations, an organizational structure with two agents was being approved. The cooperation between the *operational agent* and the *evolutionary agent* through the communication of concepts enables the good progression of the iterative pyramidal learning. Thus, at each macro[1] evolution, the operational agent intend to select the target database with the predicting concepts in order to facilitate the work of the evolutionary agent, which needs the training examples to extract the relevant set of rules required to predict the target higher concept.

3.1 Operational Agent

Although decoding salient objects represents initial performed task, the operational agent pays more attention at restructuring, at each time, available data (all C_i^{k+1}) provided by the evolutionary agent. Thus, operational agent accentuates its interest on the geometric aspect of a concept rather than its semantic aspect. Thus, it manipulates objects regarded as vectors of integer coordinates, and carries out its restructuring according to the position of objects in the image.

Broadly, the cognitive activity of restructuring encloses two complementary tasks: reduction and clustering. The *reduction task* is the setting in correspondence of the nearest objects which enables the reduction of the visual space. Thus, instead of considering detected objects as spatially independent, *clustering task* aims at regrouping objects in clusters according to a vicinity relation. Such relation is considered as the inclusion (only at beginning) and becomes a relation of nearness throughout the recognition process.

Reduction task

The reduction principle is to put away the nearest objects those which are either dependent from the global context or those which belong to a local context. The reduction task enables to reduce the visual space as well as the understanding limited to a local context. Thus, we define a local context

[1] A macro generation comes after learning a concept, and then takes several micro generations needed for one evolutionary learning.

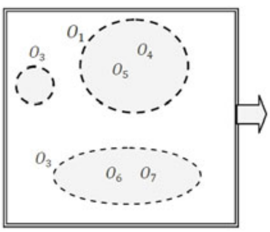
as a reduced area which provides a sense to a cluster of objects. Formally, the reduction task consists at covering nearest objects by the construction of a binary matrix $M(n \times n)$ were n represents the number of all available objects at this learning phase. The rows and columns of the matrix are then the objects $O_i(i = 1 \ldots n)$ surrounded by their shapes boundaries such that $O_i = \{(Row_{min}, Row_{max}), (Col_{min}, Col_{max})\}$, and the elements are either '0', in the case where no correspondence exists between two objects, or '1' otherwise. The aimed correspondence consists at finding a logical relation between two objects of the form inclusion or neighborhood relation.

- **Inclusion**

Let us assume that the number of salient objects detected in the image (before we lunch the learning algorithm) is m and the set of objects is $O = \{o_1, o_2, \ldots, o_m\}$. The operational agent tries then to fill in the matrix Maccording to the following properties:

- R is Reflexive $\begin{cases} O_i \subseteq O_j & \text{Then } \{(i = j) \wedge M(i,j) = 1\} \\ O_i \sqsubset O_j & \text{Then } (i \neq j) \end{cases}$
- R is Asymmetric $\begin{cases} O_i \subset O_j \\ O_i \in O_j \end{cases}$
- R is Transitive $\begin{cases} (O_i \sqsubset O_j) \wedge (O_j \subset O_k) \\ (O_i \subset O_k) \end{cases}$

As is depicted in Fig. 4, the correspondence between the image and its relative matrix provides the creation of three clusters.



$\subset$	O_1	O_2	O_3	O_4	O_5	O_6	O_7
O_1	1						
O_2		1					
O_3			1				
O_4	1						
O_5	1						
O_6			1				
O_7			1				

$S_1 = \{O_1, O_4, O_5\}$
$S_2 = \{O_2\}$
$S_3 = \{O_3, O_6, O_7\}$

Fig. 4. Cognitive visual memory architecture

- **Neighborhood**

Let is the number of reduced objects detected in the image (after at least one learning phase) is p and the set of objects is $O = \{o_1^k, o_2^k, \ldots, o_p^k\}$. The index k indicates that the objects are extracted after k^{th} learning phase.

The operational agent tries then to fill in the matrix Maccording to the measure of the distance between pairs of objects, where in general this distance has the following properties:

$$d(O_i^k, O_j^k) = 0$$

$$d(O_i^k, O_j^k) = d(O_i^k, O_j^k)$$

Since the objects represent edges or contours we then assume that the appropriate distance is the Hausdorff distance which computes the maximum distance of a set to the nearest point in the other set [10]. More formally, Hausdorff distance between two edge points sets O_i and O_j is the maximum function (Equation 1), defined as:

$$H\left(O_i^k, O_j^k\right) = \max_{a \in O_i^k} \left\{ \min_{b \in O_j^k} \{d(a,b)\} \right\} \tag{1}$$

Thus, the operational agent assigns to the matrix elements different values according to Equation 2 as follows:

$$M(i,j) = \begin{cases} 1 \text{ If } H(O_i^k, O_j^k) \leq \omega \\ 0 \text{ If } H(O_i^k, O_j^k) > \omega \end{cases} \tag{2}$$

The parameter ω is a fixed threshold. Once Mconstructed by filling in its elements, we can then pass to the next step of clustering.

Clustering task

The clustering principle consists at regrouping in a cluster all the nearest objects. This step strongly depends on the reduction step and effectively enables the extraction of all clusters available at this pyramidal learning phase. The clustering task is then lunched after the construction of M. According to a column k, the existence at least of an element with value 1 at a row l yields the construction of a new cluster j. For this same column k, if there exits more than elements with values 1, then all other objects representing the rows are included in the this same cluster j (Equation 3). In the case, where all the elements of a column are set to zero this implies that the object of this column is already contained in a previous cluster.

$$\text{If } M(k,l) = 1 \text{ Then } \{O_k \in S_j \wedge O_l \in S_j\} \tag{3}$$

When all matrix columns were visited, therefore the optimal number of clusters is determined. The following algorithm enables the formation of clusters:

The result provided from this task is the number of clusters constructed as well as the covered objects contained in each cluster. It is important to claim that each cluster S_j possesses a *reference object* S_j^{ref} which represents the identifier of both the cluster and this local context. The reference object

Algorithm 1. Clustering Algorithm.

// In [M:matrix ($n \times n$)]
begin
$j \leftarrow 0; k \leftarrow$; // k is the column index
while $k \leq n$ **do**
  if (($k > 1$) and $O_k \in \{S_e/r = 1, \ldots, j-1\}$) **then**
    $k \leftarrow k + 1$
  else
    $j \leftarrow j + 1; S_j = \emptyset$;
    for each row l **to** n **do**
      if $M(k, l) = 1$ **then**
        $S_j = S_j \cup O_l$;
        if ($k = l$) **then**
          $S_j^{ref} = \{O_l\}$;
        end if
      end if
    end for
  end if
end while
// **Out [S: set]**
end.

is useful to compute the different distances between different clusters (local contexts). In the case of inclusion relation, S_j^{ref} is only one object that which matches to the encompassing object. In the case of a neighborhood relation, the reference object will be the union of all objects present in the current cluster.

3.2 Evolutionary Associator Agent

The main characteristic of learning-based approaches is their ability to adjust their internal structure according to input and respective desired output data pairs in order to approximate the relations (rules) implicit in the provided training data, thus elegantly simulating a reasoning process. Consequently, the use of some approaches like classification rules or even association rules provides a powerful method for discovering complex and hidden relationships for a variety of applications domains.

Association rules

Association rules are used at a high interpretation and representation granularity to emerge strongly associated objects in order to get a more compact representation of the data. We define association rules [2] according to the context of perceptual objects. Let $O = \{c1, c2, \ldots, cn\}$ be a set of objects concepts, and D be a database (a set of formal concepts hierarchies), where each formal concept hierarchy h is a set of objects such that $h \subseteq O$. We say that a formal object hierarchy h contains X, a set of objects in O, if $X \subset h$. An association rule is an implication of the form $X \Rightarrow Y, X \subset O, Y \subset O$, and $X \cap Y = \emptyset$.

Many important algorithms were be used to discover association rules [3, 18]. This formulation yields a thought of two basic aspects for data representation:

- How to discover the appropriate association rules from a set of concepts X detecting some relationship between the predicting concepts and the goal concept$Y$. Actually, with regarded to all consideration cited above, the problem is tackled by considering it as a concept learning problem. In this case, each association rule is subjected to learn one high formal concept.
- If we consider the vector space whose dimensions are the concepts objects, concept hierarchy can be represented as object vectors with binary coordinates denoting the occurrence of objects in the hierarchy.

Visual evolutionary-based learning

Representing concepts as sets of rules has long been popular in machine learning, because, among other properties, rules are easy to represent and humans can interpret them easily. In evolutionary algorithms there are two main ways to represent rule sets. In the "Michigan" approach exemplified by Holland. classifier [9], and the "Pittsburgh" approach exemplified by Smith LS-1 system [5, 20]. Systems using the Michigan approach maintain a population of *individual rules* that compete with each other for space and priority in the population. This approach is simpler and syntactically shorter. This tends to reduce the time taken to compute the fitness function and to simplify the design of the genetic operators. A number of systems based on concept learning with Michigan approach have been proposed COGIN [6] and REGAL [7].

In contrast, systems using Pittsburgh approach maintain a population of variablelength rule sets that compete with each other with respect to performance on the domain task. This approach takes into account rule interaction when computing the fitness function of an individual. Consequently, some systems based on concept learning with Pittsburgh approach have been proposed like GABIL [5] GIL [11] and DOGMA [8].

Individual (rule) representation
In our studies, we adopt the Michigan approach to encode a rule. Thus, an individual in its general form is a set of items of $attribute_i \in [l_i, u_i]$, where $attribute_i$ is the i^{th} numeric attribute in the rule template from the left to the right. Fig. 5 outlines a general example:

This rule specification could induce to some problems in recombination process. Indeed, if two individuals have different representations the crossover applied to them will generate invalid offspring. Hence, to avoid to have invalid offspring due to the absence of some attributes in a rule, and in order to

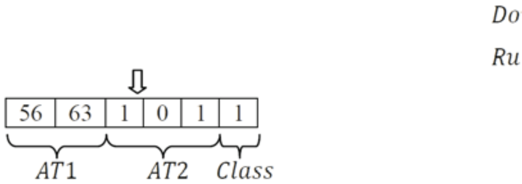


Fig. 5. Hybrid encoding of individual

use the usual crossover without specific modifications, each attribute in the individual has the same position (rank) that it occur in the dataset (in the database).

Evolutionary learning process
With respect to the used dataset where all attributes are nominal, the absence of an attribute in the dataset is refereed to a gene with '0' value. The evolutionary learning process is performed by the function EvoAlg as shown in Fig. 6. Thus, EvoAlg has a set of examples as its input parameter. It returns a rule that is the best individual of the last generation. The initial population P is built randomly by the function *InitializePopulation*. Some examples are randomly selected and individuals that cover such examples are generated. After initializing the population, the for-loop repeats the evolutionary process max-generations. At each iteration, the individuals of the population are evaluated according to defined fitness function, thus each individual acquires goodness. The best individual of each generation is replicated (*Replicate*) to be included in the next generation. Later, a set of individuals are selected through the roulette wheel method and replicated. Finally, another set of individuals are recombined and the offspring are included in the next generation.

```
Function EvoAlg
// In [T: file of encoded-examples ]
Begin
 InitializePopulation(P)
 For i=1 to max-generations do
  Evaluate(P);
  next.P:=SelectTheBestOf(P);
  next.P:=next.P+Replicate(P);
  next.P:=next.P+Recombine(P);
  P:=next.P
 End;
 Evaluate(P);
return SelectTheBestOf(P) // Out [ r:Rules]
End EvoAlg
```

Fig. 6. Pseudocode of evolutionary learning rules

Equation 3 gives the fitness function $f(\tau)$ used during the evaluation process. The greater the value, the better the individual is.

$$f(\tau) = N - CE(\tau) + G(\tau) + \mathrm{cov}erage(\tau) \tag{4}$$

Where τ is an individual, N is the number of examples being proceeded; CE is the concept error, i.e. the number of examples belonging to the region defined by the rule τ, which they do not have the same concept (class); $G(\tau)$ is the number of examples correctly "classified" by τ, and $\mathrm{cov}erage(\tau)$ gives the size proportion correctly "classified". For more detail about the influence of $\mathrm{cov}erage(\tau)$ on fitness see [1].

4 Experimental Results

To test our image theory we created synthetic images as a starting point in showing the feasibility of concept learning as a technique to understanding content image. The images database contains combinations of plane geometric shapes, where each shapes belongs to the set $S = \{$ triangle, circle, rectangle, hexagon, ellipse $\}$. The number of images in the database generated for the test phase corresponds to 30 images, where each image is identified by a label. As depicted in Figure 7, the majority of images have only one understanding level. This constraint is necessary because of the prediction of concepts. But to be able to validate at least two learning levels, we have drawn four images (20, 24, 27, 30) which represent semantic objects, constructed from S (primitives concepts), and correspond respectively to face, bicycle, chair, and cat. The association rules shown to the environment have the form:

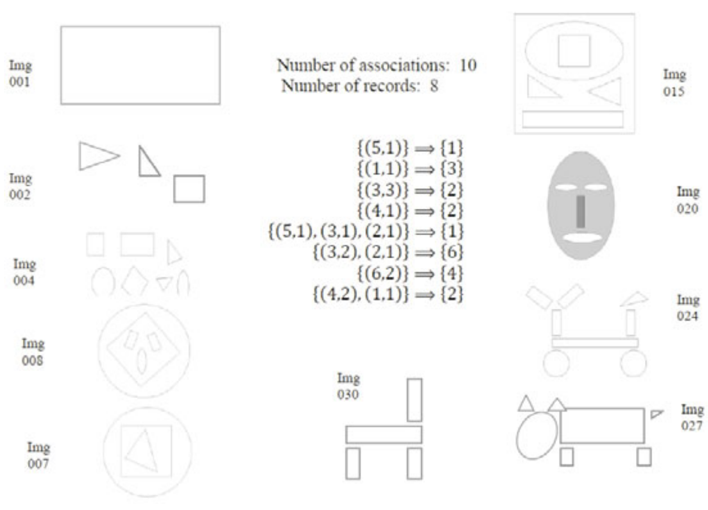


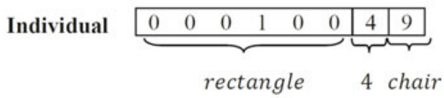
Fig. 7. Geometric shapes and association rules as selected training examples

$$\{(AT1_1, AT2), (AT1_2, AT2), \ldots, (AT1_k, AT2), \} \Rightarrow \{AT1_{k+1}\}$$

In concordance with the chosen images, we define the new application context:

$$Domain \begin{cases} AT1_i \in \{\text{triangle, circle, rectangle, hexagon, ellipse}\} \\ AT2 \in [1, 10] \end{cases}$$

Example of coding rule. Rule : If $AT1_1 \in \{\text{rectangle}\}$ and $AT2 \in [1, 5]$ Then chair



Each image in the database stores both the low-level features (first descriptor) and the high-level concept attributes (second descriptor). In fact, at this postprocessing level, we are only interested by the second descriptor. Thus, the first rule $\{(5, 1)\} \Rightarrow \{1\}$ tells us that if there is a hexagon there is triangle. This result is validated according to a high fitness attributed to this rule. Other rules indicate the same reasoning such that the rule $\{(1, 1)\} \Rightarrow \{3\}$ where if there is a triangle there is a square. Also, the rule $\{(4, 1)\} \Rightarrow \{2\}$ illustrates that if there is a rectangle there is necessary a circle. Furthermore, to show the ability of the system to learn more complex objects not composed only by primitive concepts, we have introduced four specialized rules. Each rule yields to each corresponding concept (face:7, bicycle:8, chair:9, and cat:10).

$$\begin{array}{ll} R1 : \{(6, 4), (4, 1)\} \Rightarrow \{7\}; & R2 : \{(4, 5), (1, 1), (2, 2)\} \Rightarrow \{8\} \\ R3 : \{(4, 4)\} \Rightarrow \{9\}; & R4 : \{(4, 3), (1, 3), (6, 1)\} \Rightarrow \{10\} \end{array}$$

In the current studies, the number of occurrence of each object must appear in order to avoid the redundancy of the same objects. identifiers as conditions in the left side of a rule. We have deliberately omitted to reserve a field of distance in the object representation since the coverage of objects in a cluster is a task performed before learning rules. Then, we suppose that all nearest objects could as a whole specify a new concept. Obviously, we want to stress that our synthetic images are not as complex as images from the real world. But as a first attempt to mine association rules with evolutionary learning in image processing seems a good trend to improve the emergent behaviour from the cognitive architecture.

5 Conclusion

We present in this paper a cognitive architecture inspired from human characteristics to deal with the visual cognitive modality. The system includes many reasoning paradigms such as image processing, inductive learning, evolutionary computing and image mining. Our efforts were based on the conceptualization of the cognitive framework for detecting and understanding objects in several contexts. Obviously, the system relies on knowledge discovering using association rules from databases. Results obtained so far look promising but we need to improve several aspects in our research efforts. We can currently working on the some evident tasks like image mining, and image retrieval to outperform the semantic content-based image. These visual tasks will help in the automation of such behaviours. Also, we should give more attention to other kind of attributes (discrete and continuous) and how to determine the optimal representation and even the appropriate genetic operators. In the future, we look to improve the cognitive aspect of the architecture by introducing more than one cognitive modality, and to find the optimal way to connect the aimed specialized components to this general cognitive framework. Regarded to the fixed objective, it results therefore a good opportunity to develop both intelligent and autonomous system in the domain of pattern recognition.

References

1. Aguilar-Ruiz, J.S., Riquelme, J.C., Toro, M.: Evolutionary learning of hierarchical decision rules. IEEE Transactions on Man and Cybernetics, Part B 33(2), 324–331 (2003)
2. Agrawal, R., Imielinski, T., Swami, A.: Mining association rules between set of items in large databases. In: Proc. ACM SIGMOD International Conference on Management of Data, Washington, DC, pp. 207–216 (1993)
3. Agrawal, R., Srikant, R.: Fast algorithms for mining association rules in large database. In: Proceedings of 20th International Conference on Very Large Data Bases, Santiago, Chile (1994)
4. Bhanu, B., Lin, Y., Krawiek, K.: Evolutionary synthesis of pattern recognition systems. Springer, New York (2005)
5. De Jong, K.A., Spearsm, W.M., Gordon, D.F.: Using genetic algorithms for concept learning. Machine Learning, 161–188 (1993)
6. Giordana, A., Neri, F.: Search intensive concept induction. Evolutionary Computation 3(4), 376–416 (1996)
7. Greene, D.P., Smith, S.F.: Competition-based induction of decision models from examples. Machine Learning 13, 229–257 (1993)
8. Hekanaho, J.: GA-based rule enhancement concept learning. In: Proc. Third International Conference on Knowledge Discovery and Data Mining, New Port Bearch, CA, USA, pp. 183–186 (1997)
9. Holland, J.: Escaping brittleness: The possibilities of general-purpose learning algorithms alied to parallel rule-based systems. In: Michalski, R., Carbonell, J., Mitchell, T. (eds.) Machine Learning: An Artificial Intelligence Aroach (1986)
10. Huttenlocher, D.P., Rucklidge, W.J.: A multi-resolution technique for comparing images using Hausdorff distance, Dpt. of CS, Cornell University (1994)
11. Janikow, C.Z.: A knowledge intensive genetic algorithm for supervised learning. Machine Learning 13, 189–228 (1993)

12. Krawiec, K.: Learning high-level visual concepts using attributed primitives and genetics programming. In: Rothlauf, F., Branke, J., Cagnoni, S., Costa, E., Cotta, C., Drechsler, R., Lutton, E., Machado, P., Moore, J.H., Romero, J., Smith, G.D., Squillero, G., Takagi, H. (eds.) EvoWorkshops 2006. LNCS, vol. 3907, pp. 515–519. Springer, Heidelberg (2006)
13. Krawiec, K.: Evolutionary learning of primitive-based visual concepts. In: Proc. IEEE Congress on Evolutionary Computation, Vancouver, BC, Canada, pp. 1308–1315 (2006)
14. Krawiec, K., Bhanu, B.: Visual learning by coevolutionary features synthesis. IEEE Transactions on Systems Man and Cybernetics Part B 35, 409–425 (2005)
15. Maloof, M., Langley, P., Binford, P., Nevatia, R., Sage, S.: Improved rooftop detection in aerial images with machine learning. Machine Learning 53, 157–1991 (2003)
16. Ogiela, M., Tadeusiewicz, R.: Nonlinear processing and semantic content analysis in medical imaging - a cognitive aroach. IEEE Transactions on Instrumentation and Measurements 54, 2149–2155 (2005)
17. Rizki, M., Zmuda, M., Tamburino, L.: Evolving pattern recognition systems. IEEE transactions on Evolutionary Computation 6, 594–609 (2002)
18. Savasere, A., Omiecinski, E., Navathe, S.: An efficient algorithm for mining association rules. In: Proceedings of the VLDB Conference, Zurich, Switzerland, pp. 432–444 (1995)
19. Segen, J.: GEST: a learning computer vision system that recognizes hand gestures. In: Michalski, R., Tecuci, G. (eds.) Machine Learning, A Multistrategy Aroach, vol. IV, pp. 621–634. Morgan Kaufman, San Francisco (1994)
20. Smith, S.: Flexible learning of problem solving heuristics through adaptive search. In: Proc. Eight International Joint Conference on Artificial Intelligence, pp. 422–425 (1983)
21. Teller, A., Veloso, M.: PADO: a new learning architecture for object recognition. In: Ikeuchi, K., Veloso, M. (eds.) Symbolic Visual Learning, pp. 77–112. Oxford Press, New York (1997)