HAL

archives-ouvertes.fr

Incremental learning for interactive sketch recognition

Achraf Ghorbel, Abdullah Almaksour, Aurélie Lemaitre, Eric Anquetil

▶ To cite this version:

Achraf Ghorbel, Abdullah Almaksour, Aurélie Lemaitre, Eric Anquetil. Incremental learning for interactive sketch recognition. Ninth IAPR International Workshop on Graphics RECognition, Sep 2011, Séoul, South Korea. hal-00646137

HAL Id: hal-00646137
https://hal.inria.fr/hal-00646137

Submitted on 29 Nov 2011

Incremental learning for interactive sketch recognition

Achraf Ghorbel*, Abdullah Almaksour*, Aurélie Lemaitre†, Eric Anquetil*

*INSA de Rennes, †Université de Rennes 2

UMR IRISA, Campus de Beaulieu, F-35042 Rennes

Université Européenne de Bretagne, France

{achraf.ghorbel, abdullah.almaksour, aurelie.lemaitre, eric.anquetil}@irisa.fr

Abstract—In this paper, we present the integration of a classifier, based on an incremental learning method, in an interactive sketch analyzer. The classifier recognizes the symbol with a degree of confidence. Sometimes the analyzer considers that the response is insufficient to make the right decision. The decision process then solicits the user to explicitly validate the right decision. The user associates the symbol to an existing class, to a newly created class or ignores this recognition. The classifier learns during the interpretation phase. We can thus have a method for auto-evolutionary interpretation of sketches. In fact, the user participation has a great impact to avoid error accumulation during the analysis. This paper demonstrates this integration in an interactive method based on a competitive breadth-first exploration of the analysis tree for interpreting the 2D architectural floor plans.

Keywords-sketch recognition; incremental learning; interactive recognition; 2D architectural floor plans;

I. Introduction

In this paper, we are working on mapping technical paper documents, like architectural floor plans, to numerical ones. We aim at offering a complete, interactive and auto-evolving solution to unify paper document recognition and pen-based sketch interpretation (for instance: with Tablet PC).

At present, structured documents can be very complex. Faced with this complexity, the various existing methods [1] [2] [3] [4] keep a margin of error. Therefore, very often, an a posteriori verification phase will be necessary to ensure there is no recognition error. In this phase the user browses the document to correct the errors due to the interpretation.

To avoid the verification phase on the one hand, and avoid error accumulation during the analysis step on the other hand, we proposed an interactive method of analysis of off-line structured document where the decision process solicits the user if necessary. In our previous work [5] the role of user was limited to validate the right hypothesis and then unlocks a situation where the decision process is not sure to make the right decision. In summary, the process can solicit the user to be sure to make the correct decision.

Now, we want to exploit the solicitation of the user during the analysis not only to unlock a situation but also to learn the process of analysis. In this context, we focus in this paper on improving the capacity of symbol recognition. In the sketch interpretation method, the classifier is responsible for symbol recognition.

The classification systems can be generally categorized into two types: static and evolving systems. Static systems are trained in batch mode using a predefined learning dataset, while incremental learning algorithms are used to train evolving classifiers, like for our symbol recognition system. In incremental learning algorithms, new instances from existing classes can be progressively introduced to the system to improve its performance. Moreover, new unseen classes can be added to the system at any time by the incoming data.

In this work, we present the advantage of soliciting the user to improve the recognition capacity of the classifier by incremental learning able to dynamically add new classes. The remaining of the paper is organized as follows. In the section II, we introduce our existing interactive analysis method. Section III describes principles of the incremental classifier. The coupling of this incremental classifier with our interactive analysis of sketches is described in section IV. Experimental results are reported in section V and finally, section VI concludes the paper.

II. Interactive breadth-first exploration

In this section, we summarize our interactive method of structured document interpretation [5] in which we propose to integrate our incremental classifier.

This analyzer is based on the following characteristics:

- a rule based analysis,
- a bidimensional descending breadth first analysis.
- the attribution of scores to each hypothesis,
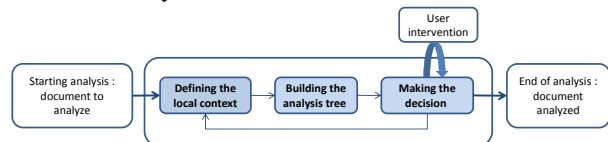- a spatial contextual focus of the exploration to limit the combinatory.

Figure 1. Analysis process

These characteristics were chosen in order to ensure the best interactivity with the analysis system. Figure 1 shows the complete process of analysis and the relationship between

the three parts of the analyzer. The analyzer begins by defining a spatial contextual focus. Once the context is well defined, the analyzer goes to the second stage. In this stage, the analyzer explores all possible hypotheses of interpretation in the spatial context using a set of bi-dimensional rules that describe the structure of the document. The next stage is the decision making. The role of the decision process is to validate the right hypothesis among a set of competing hypotheses generated with a descending breadth first analysis. Sometimes the decision process is not sure to make the right decision. In this case, it solicits the user. In practice, if the difference of scores between the top two branches is below a threshold of confidence and if these two branches are contradictory (at least one joint primitive is not consumed by the same rule production), the user intervention is required. When the correct hypothesis is validated, we return to the first step. The analysis is complete when no more production rule is applicable. This intervention of the user avoids a false decision and thus the propagation of errors during the analysis. This interactive recognition strategy allows lazy interpretation of complex structured documents.

In the current state, the information provided by the user is only used to unlock situations. In this paper, we want to exploit more widely this information by learning continuously during the analysis. In this context, we propose to integrate an incremental classifier which uses information supplied by the user to improve its capabilities during the analysis.

III. INCREMENTAL LEARNING OF A FUZZY INFERENCE SYSTEM

Our classification system is based on first-order Takagi-Sugeno (TS) fuzzy inference system[6]. It consists of a set of fuzzy rules of the following form:

Rule$_i$: **IF** $\vec{x}$ *is close to* P_i**THEN** $y_i^1 = l_i^1(\vec{x}), ..., y_i^k = l_i^k(\vec{x})$

(1)

where $l_i^m(\vec{x})$ is the linear consequent function of the rule i for the class m:

$$l_i^m(\vec{x}) = \vec{\pi}_i^m \vec{x} = a_{i0}^m + a_{i1}^m x_1 + a_{i2}^m x_2 + ... + a_{in}^m x_n \qquad (2)$$

where n is the size of the input vector. The Prototype P is defined by a center and a fuzzy zone of influence. To find the class of $\vec{x}$, its membership degree $\beta_i(\vec{x})$ to each fuzzy prototype is first computed. After normalizing these membership degrees, the sum-product inference is used to compute the system output for each class:

$$y^m(\vec{x}) = \sum_{i=1}^{r} \bar{\beta}_i(\vec{x}) \, l_i^m(\vec{x}) \qquad (3)$$

where r is the number of fuzzy rules in the system. The score of each class y^m is between 0 and 1. The higher is the score, the higher is the degree of confidence in that class to

be associated to the given input. The winner class is that with the maximum score. The membership degree is computed by the prototype center $\vec{\mu_i}$ and its variance-covariance matrix A_i using the multivariate Cauchy probability distribution.:

$$\beta_i(\vec{x}) = \frac{1}{2\pi\sqrt{|A_i|}} \left[1 + (\vec{x} - \vec{\mu_i})^t A_i^{-1} (\vec{x} - \vec{\mu_i}) \right]^{-\frac{n+1}{2}} \qquad (4)$$

The incremental learning algorithm of our model consists of three different tasks: the creation of new rules, the adaptation of the existing rule's premises, and the tuning of the linear consequent parameters. These three tasks must be done in an online incremental mode and all the needed calculation must be completely recursive.

The importance of a given new sample in an incremental clustering process can be evaluated by its *potential* value. The potential of a sample is defined as inverse of the sum of distances between a data sample and all the other data samples. A recursive method for the calculation of the potential of a new sample has been introduced in [7]. The recursive formula avoids memorizing the whole previous data but keeps - using few variables - the density distribution in the feature space based on the previous data (see [6] for more details). A Premise adaptation process allows to incrementally update the prototype centers coordinates according to each new available learning data, and to recursively compute the prototype covariance matrices in order to give them the rotated hyper-elliptical form. For each new sample, the center and the covariance matrix of the prototype that has the highest activation degree are updated recursive manner [6]. The tuning of the linear consequent parameters in a first-order TS model can be done by the weighted Recursive Least Square method (wRLS). More details are available in [6].

The incremental learning algorithm is supposed to be supervised. The recognition of each data sample must be followed by a validation or a correction action in order to learn it. If the system answer is validated, the data sample will reinforce the system knowledge associated to its class. If an external correction signal is sent, the confusion between the (wrong) winner class and the true class is solved by the incremental learning algorithm. A third scenario may take place when the input data sample is declared as the first sample from a new unseen class.

IV. USER INTERVENTION IN THE INTERACTIVE ANALYSIS PROCESS

In this section, we present the possibilities offered by the introduction of a classifier based on incremental learning in our interactive sketch recognizer. In particular we detail when and how the user can interact with the incremental classifier. During the analysis, each time the classifier is sought to identify a symbol, the decision process uses the confident degree given by the classifier to make its decision. If the decision process considers that the confidence degree is sufficiently high to make the right decision, it validates

the recognition. Otherwise, The decision process will solicit the user. The user is then in front of four possibilities (cf. Figure 2):
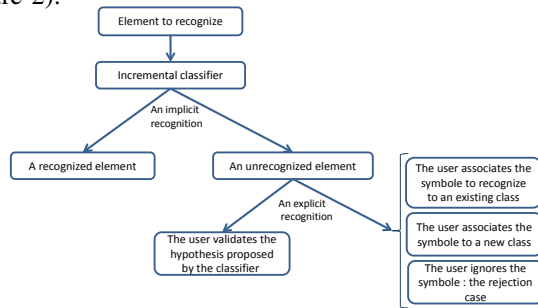

Figure 2.   Interaction scheme of symbol recognition

- The user validates the hypothesis proposed by the classifier in spite of the low degree of confidence given by the classifier. The classifier enhances the model of this class.
- The user associates the symbol to recognize to other existing class in the classifier. The classifier reduces the confusion between two classes.
- The user associates the symbol to a new class: the user considers that the symbol does not belong to an existing class. With this new information, the classifier starts to learn a new class of symbols.
- The user ignores the symbol to recognize: the rejection case. The user considers that the recognized symbol is an outlier (noise in the image). No action is done by the classifier.

With this interaction process, the classifier continuously learns to improve its interpretations. The more the analysis is going on, the more the classifier is accurate, the less the user is solicited. This incremental learning is able to deal with the recognition of new classes of symbols. It is a key point to absorb the great variability of symbols that can occur in a sketch.

Example

To illustrate the interaction process, we develop the analysis on 2D handwritten architectural floor plans. We focus the demonstration on cases where the decision process invokes user interaction. Figure 3 illustrates a handwritten architectural plan to interpret. The incremental classifier initially contains two classes: one kind of door and one kind of window.

Figure 4 shows a case where the solicitation of the user is judged necessary. In this intervention the user is in front of the four possible actions described in section IV. The system presents an interface that contains the hypothesis given by the classifier, the other available classes of the classifier and a field where the user can add a new class. Figure 4(a) shows a case in which the user indicates that the symbol to recognize is a classical window. Figure 4(b) shows a case

in which the user associates the symbol to a new class of windows (a sliding windows).
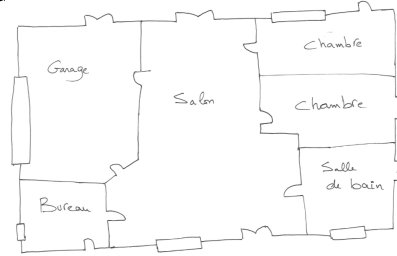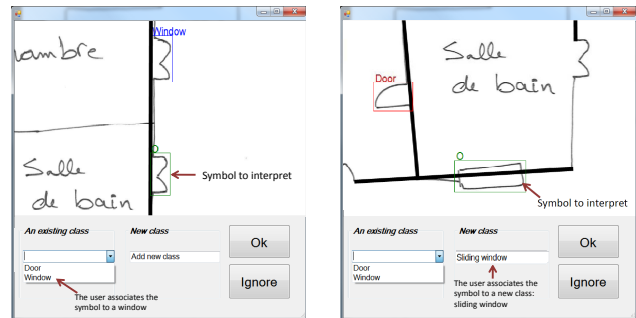

Figure 3.   Example of an architectural plan to interpret

(a) The user associates the symbol to a 'Window'

(b) The user associates the symbol to a new class (sliding window).

Figure 4.   User interventions. Four possibilities exist. The user associates the set of primitives located in the bounding box 'O' to a right class.

V. EXPERIMENTAL RESULTS

In this section we analyze the performance and the behavior of our model of incremental learning on the two types of experiements. The first experiment is conducted on an incremental learning problem related to our specific application context, the recognition of the openings (windows and doors) in architectural plans. The first experiment evaluates the performance of the model of incremental learning on a set of openings belonging in two classes (window and door).

For this evaluation, we choose:

- a classifier which contains two classes (door and window). This classifier is initially learned using 3 samples per class.
- an incremental learning database containing 10 doors and 10 windows.
- a test database containing 99 windows and 118 doors.

All symbols used in this experiment are extracted from 26 images of handwritten 2D architectural floor plans. The main idea of this experiment is to demonstrate the capability of incremental learning to improve the recognition rate of opening belonging in existing classes. For this, we use the test database after each incremental learning step. A learning step is a phase in which the classifier learns using one door and one window of the incremental learning database. Figure 6 illustrates the evolution of the error rate of opening (Door and window) during the incremental learning of a set

of samples. Thanks to the adaptation of the classifier, the error rate found on the test database decreases from 9.7% to 2.8%. The most of misrecognized symbols are badly drawn.
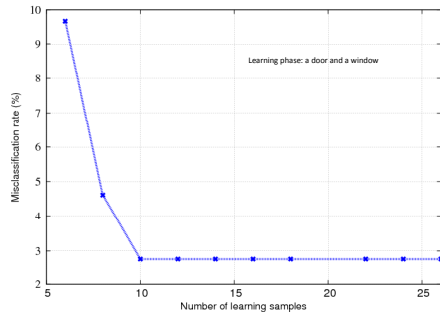

Figure 5. Evolution of performance during the incremental learning process

The second experiment describes the behavior of incremental classifier and its ability to add new classes during the learning phase. We start this experiment by:

- classifier which contains two classes (door and window). This classifier is the same used in the first experiment. It is initially similar to the classifier used in the first experiment (the same initial learning phase).
- incremental learning database containing 10 doors, 10 windows and 10 sliding windows.
- test database containing 99 windows, 118 doors and 14 sliding windows.

All symbols used in this experiment are extracted from 29 images of handwritten 2D architectural floor plans. This experiment has two phases of adaptation. The first phase is an existing class adaptation. This phase is similar to the first experiment. We use the test database after each learning step. Like the first experiment, an incremental learning step is composed of one door and one window. In the second phase, we introduce a new class (sliding window). Indeed, the classifier is incrementaly learned using at each step one door, one window and two sliding windows.

Figure 6 describes the evaluation of the classifier during this experimentation. The classifier begins to learn the symbol of existing classes. During this phase, the error rate decreases from 15.1% to 8.7%. During the second phase, the classifier needs to learn symbol of a new class (the pic in the curve of Figure 6). After a few samples of adaptation to this new class, the classifier finds its stability and absorbs this class. During this phase, the error rate decreases from 18.2% to 3.9%

We have shown in these first results the performance of a classifier based on incremental learning and its ability to integrate new classes during the analysis phase of structured documents. Our next work is to interpret more complex structured documents containing several classes of symbols.

We are rather confident in this new step because the incremental classifier have been already successfully tested
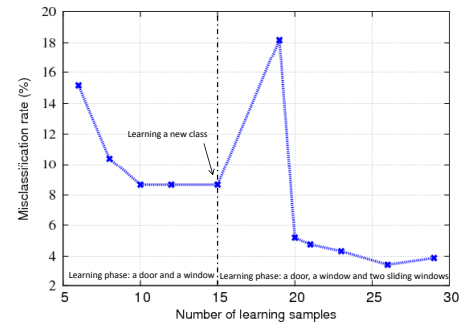

Figure 6. Evolution of performance during the incremental learning process (Integration of a new class)

on other database composed of several classes in others domains [6].

VI. Conclusion

In this paper, we have presented the integration of a classifier based on an incremental learning method, in an interactive method for interpreting the 2D architectural floor plans. The role of classifier is to recognize the symbol with a degree of confidence. If this degree of confidence is considered insufficient by the decision process to take the right decision, the analyzer solicits the user to validate the right hypothesis. The user is then in front of four possibilities. He can either confirm the recognition proposed by the classifier, or associates the symbol to an existing class, a new class, or ignores this recognition. The classifier is incrementally learned during the analysis phase. This strategy offers an auto-evolutionary method for sketch interpretation.

References

[1] K. Chan and D. Yeung, "An efficient syntactic approach to structural analysis of on-line handwritten mathematical expressions," *Pattern Recognition*, vol. 33, no. 3, pp. 375–384, 2000.

[2] J. Fitzgerald, F. Geiselbrechtinger, and T. Kechadi, "Mathpad: A fuzzy logic-based recognition system for handwritten mathematics," in *ICDAR 2007*, vol. 2, 2007, pp. 694 –698.

[3] S. Mao, A. Rosenfeld, and T. Kanungo, "Document structure analysis algorithms: a literature survey," in *Proc. SPIE Electronic Imaging*, vol. 5010, 2003, pp. 197–207.

[4] B. Coüasnon, "Dmos, a generic document recognition method: Application to table structure analysis in a general and in a specific way," *IJDAR 2006*, vol. 8, no. 2, pp. 111–122.

[5] A. Ghorbel, S. Macé, A. Lemaitre, and E. Anquetil, "Interactive competitive breadth-first exploration for sketch interpretation," *to appear in ICDAR 2011*.

[6] A. Almaksour and E. Anquetil, "Improving premise structure in evolving takagi-sugeno neuro-fuzzy classifiers," *Evolving Systems*, vol. 2, pp. 25–33, 2011.

[7] P. Angelov and D. Filev, "An approach to online identification of takagi-sugeno fuzzy models," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 34, pp. 484–498, 2004.