HAL

archives-ouvertes.fr

Use of semantic and physical constraints in Bayesian Networks for Form Recognition

Emilie Philippot, Yolande Belaïd, Abdel Belaïd

▶ To cite this version:

HAL Id: inria-00629334

https://hal.inria.fr/inria-00629334

Submitted on 5 Oct 2011

Use of semantic and physical constraints in Bayesian Networks for Form Recognition

Philippot Emilie
Loria - Nancy University - France
emilie.philippot@loria.fr

Belaïd Yolande
Loria - Nancy University - France
yolande.belaid@loria.fr

Belaïd Abdel
Loria - Nancy University - France
abdel.belaid@loria.fr

Abstract—In our previous research, we worked on on-line form recognition by exploiting semantic constraints between fields using Bayesian networks. The semantic constraints allowed us to check the co-existence of fields filled up by hand by users. In this paper, we propose to test the use of architectural constraints for a design problem related to the modelling of shower areas. The proposed method exploits the physical dependencies between different parts of a space shower. The tests are performed on a database composed of 500 forms representing 5 models. The first results reach a recognition rate of 96.7%.

Keywords-Forms, Electronic ink, Handwritten strokes, Bayesian networks

I. INTRODUCTION

The work presented in this paper is related to form recognition by exploiting the architectural constraints in a Bayesian network. The application concerns the design of a shower area. The idea is to propose a method to an artisan, who arrives on a future building site, to quickly model the layout of a shower space using simple features, easily identifiable in the building context. A single form should define all the possibilities present in several constructor catalogues.

The Figure 1 shows the general form used to enter the customer's choice and the architectural constraints. Once this form filled using an electronic pen, the objective is to propose a shower space model best suited to the context and needs of the customer.

A. Technology used

The writing tool used in this application is a clip coupled with an electronic pen. All the traces made by the user are transmitted through the clip device. A trace is represented by a sampling of the composed point coordinates. The traces are collected in XML format. This data collection method has the advantage to focus on the main data but does not transmit the frame of the form which can be considered as a handicap for model identification.

B. Problem overview

The challenge is to provide a method able to identify the model of a shower space from some traces, sketched by the user on a form, using an electronic pen. It was therefore necessary to conceive a general form with all the shower

Figure 1. General form of a shower space

space configurations and then to define a model form for each shower space configuration.

In [6], we proposed a method to classify handwritten and on-line forms, based on Bayesian networks. As for our new application, only the electronic ink file is sent to the system. We lose the form woof and therefore the type of the completed form. Thus, we analyse the electronic ink file to determine the class of the completed form. To do this, we separated the form into three parts (header, body and foot) to achieve local recognition, then per part, and global.

We relied on these concepts to address the problem of shower area design. Unlike the first approach where all the forms are different, only a single form is considered, the general one, and as many models as shower areas.

Table I summarizes the differences between the two classification applications.

	Form Classification	Shower Model Identification
Constraints	semantic	architectural
Form filled	one by class	one general
Model form	one by class	one by type of shower area

Table I
DIFFERENCES BETWEEN FORMS CLASSIFICATION AND SHOWER IDENTIFICATION

C. State of the art

The literature is not provided for the recognition of on-line forms. We can however mention some researches related more or less to this topic. In [3], handwritten on-line medical forms are treated based on form categorization coupled with a search system. The strength of the system is focused on the information retrieval from the input requests.

In [4], a search of handwritten on-line documents is proposed. The system uses a list of keywords related to documents and a tailored questionnaire. In our case, we use a single form similarly to the approach quoted in [3] to find a model. However, we discarded the methods [4] and [3], because insofar all our models have the same lexical characteristics, only the way to fill up the form allows to distinguish one model from another.

Regarding the classification using Bayesian networks in [5], the authors propose the use of sub-networks for the classification of gene expression data to reduce the computational complexity of learning. In our approach, we also use sub-networks to reduce the variables number by pooling the components common to several shower areas. But unlike [5] we cannot content ourselves with these sub-networks as we are not looking just for components but the shower area containing these components. The constraints are no longer semantic but physical.

The rest of the paper is organized as follows: section II describes the proposed approach. In section III, we remind some basic notions about Bayesian networks. In section IV, the identification approach of model space shower is detailed. Finally in section V, the first results are presented before we conclude and give some perspectives.

II. THE PROPOSED APPROACH

This is to analyse the electronic ink to determine the form class based on the information provided. Our approach is partly based on the study of the two following dependencies observations:

- dependencies between the form fields and dependencies between components of different parts of the shower area.
- dependencies between different parts of the form (i.e. the different parts of the shower area).

For example, the shape of the shower area enclosure will depend on the wall arrangement. Indeed, if the shower area has to be installed in the corner of a bathroom, the model will be composed of up to two shower enclosures. Similarly, if the shape of the shower area has an arc, so the shower tray will necessarily have the same arc.

To make best use of these dependencies, we decided to separate the shower area into three distinct parts: one corresponding to the shape of the shower area and consisting of walls and shower enclosures, a second part concerning the shower tray and finally a part for the door definition. For each part, a local Bayesian network is trained and then all the Bayesian networks are gathered in a global network in order to determine the best model suited to the context. This division provides less complex Bayesian networks and thus it is easier to train. Moreover, this solution also sets the same local Bayesian network for several shower area models.

Figure 2 shows an example of the dependencies between the walls and the shower enclosures of the shower area. The links and probabilities allow us to visualize and quantify the dependencies and thus define the model. If the existing walls W1 and W2 are selected, we will reach the probability of 0.96 that the shower enclosure E2 will be retained. Conversely, if a wall is not selected then we will have a probability of 0.96 that a shower enclosure E1 will be.

		W1 =	
		Y	N
E1 =	Y	0,06	0,96
	N	0,94	0,04

W1   W2

E1   E2

		W1 , W2=			
		Y,Y	Y,N	N, Y	N, N
E2 =	Y	0,96	0,9	0,97	0,96
	N	0,04	0,1	0,03	0,04

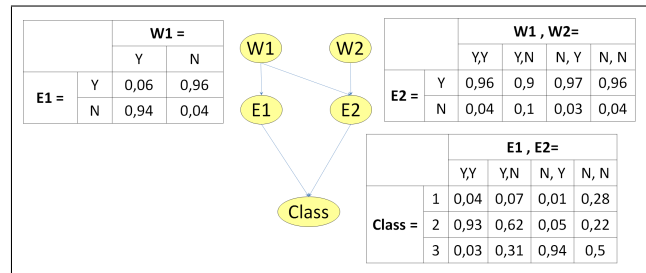|		E1 , E2=			
		Y,Y	Y,N	N, Y	N, N
	1	0,04	0,07	0,01	0,28
Class =	2	0,93	0,62	0,05	0,22
	3	0,03	0,31	0,94	0,5

Class

Figure 2. Example of Bayesian network for part of the shower area where each class represents a model of possible shape for the shower area

III. BAYESIAN NETWORKS

A Bayesian network is defined by a directed acyclic graph and a finite probabilistic space. Nodes represent a set of random variables and arcs the conditional dependencies between nodes.

A. Structure learning: MWST algorithm

Among the many learning algorithms of Bayesian network structure we have chosen MWST (Maximum Weight Spanning Tree) ([2]). This algorithm is part of the family of those based on a score. The goal is to find the tree that covers all the nodes of the network by maximizing a score defined for all possible arcs. The starting point of the algorithm is a set of n trees composed of a single node (as many trees as variables). Trees are then merged based on edge weights. The advantage of this algorithm is that all variables are connected and therefore come into account during the recognition step.

947

B. Feature learning

In our application, all variables are observed and have a value indicating that the corresponding field is filled. We have chosen to use the maximum likelihood ([2]) for feature learning. It is question to compute the frequency of an event in order to determine its probability, which gives the following equation:

$$\hat{p}(X_i = x_k | pa(X_i) = x_j) = \frac{N_{i,j,k}}{\sum_k N_{i,j,k}}$$

where $N_{i,j,k}$ is the number of events in the database for which the random variable X_i is in state x_k and his parents are in the configuration x_j.

C. Recognition

The recognition is based on inference whose objective is to spread the information known to the rest of the Bayesian network in order to modify the probabilities of random variables that have not been observed. In our case, it is to determine the best shower space model for the ground truth based on the filling of form fields. Initially, we transform the Bayesian network structure into a tree using the junction tree algorithm. Then we use the "message passing" to spread the information within the tree.

IV. EXPLOITING PHYSICAL CONSTRAINTS BY BAYESIAN NETWORKS

Figure 3 provides an overview of the proposed approach. The idea is to define for each shower space model, a form model. This model will serve as a basis for the entire system as well for learning as for recognition and a general form containing all model fields which will be filled by the user. The models are used to extract the form fields in order to use them for the Bayesian network creation during the learning phase and for selecting nodes in these networks for recognition. All the details concerning this procedure will be given later.

A. Model form

Each shower space model is represented by a model form. Created manually, a model form contains information about the division into parts and fields that represent the components of each part. These parts are defined by the shower area elements. In Figure 6, (A) and (B) are examples of model forms.

A form model is described in XML whose tags can help to highlight fields relevant to the shower space definition. For that purpose, we follow the traditional reading, from left to right and from top to bottom. Each field is defined by the coordinates of its bounding box, type (checkbox, string, number, drawing, etc.) and the part to which it belongs. A field can be divided into several bounding boxes to refine the
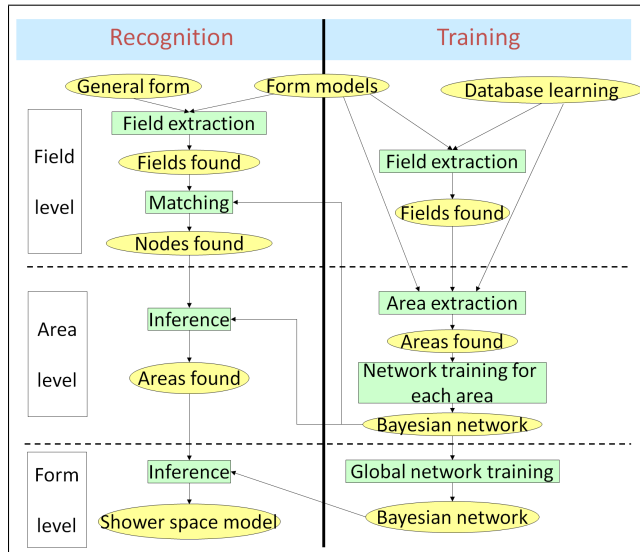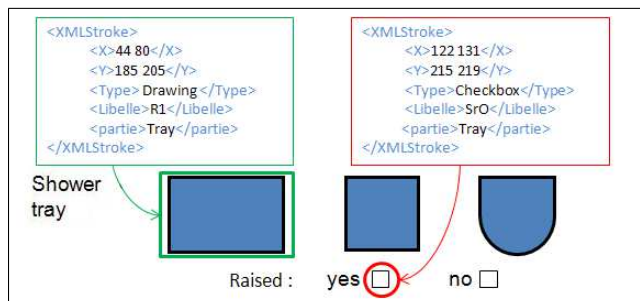

Figure 3. Global schema of the approach

Figure 4. Extracts of a model form contents and link with the corresponding frame in the general form

accuracy of the model. Figure 4 shows two excerpts form models and the corresponding fields in the frame.

Figure 5 shows an example of fields with two options for bounding boxes. We can observe that the cutting (2) is more accurate because every shower enclosure can be completed so that the three bounding boxes will be validated. Cutting (1) does not make the difference between the selection of only a shower enclosure and three shower enclosures 1, 2 and 3.
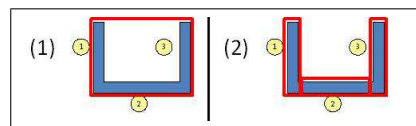

Figure 5. Example of possible bounding boxes for the same field

B. General form

The general form (Figure 1) is the form that the user has to complete. It contains all the fields included in the forms and templates necessary to model the shower area.

The difficulty was to establish an intuitive and simple form with all possible variations to reduce the filling ambiguities and mistakes.

C. Field extraction

Field extraction is the first step of both learning and recognition. It provides a list of the found fields. It is done by superimposing the model forms to electronic ink files. It is question to match the strokes with the form model fields. Each stroke is divided into as many vectors as coordinates it has. It then associates each stroke vector to the field it belongs by comparing the coordinates of this vector with those of the bounding boxes of the model fields. If a field has more than 85% of a stroke, we consider the stroke belonging to this field and therefore is present in the completed form. This percentage was defined experimentally to allow a stroke to slightly exceed a field. Once all the strokes performed, we dispose of the list of strokes per field and the list of strokes that could not be matched, either because they do not correspond to the model form field, or because their belonging rate to a field is too low. If the rate of unmatched strokes is high, then we may consider the model, with which we tried to match the electronic ink file, cannot satisfy the need of the user.

Figure 6 shows a matching example between an electronic ink file and two different model forms. The filled fields are those containing a stroke. We can see that the model (A) cannot match the form of the proposed model because the majority of strokes do not correspond to the fields of this model, unlike the model (B) where the strokes match with the model fields.
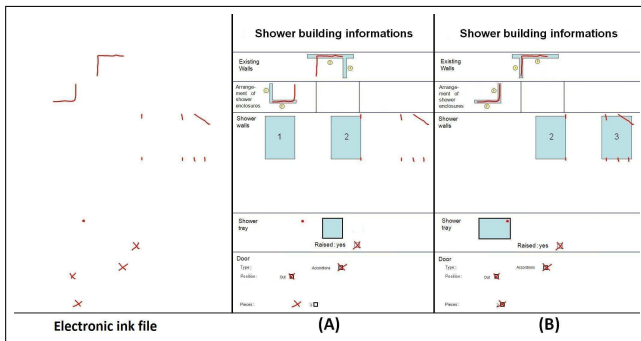

Figure 6. Matching example of an electronic ink file and two models of different shower spaces

D. Model learning

We work on a basis of electronic ink files. Learning takes place in two stages, one is about the learning of Bayesian networks of the identified parts in the forms. The other is about learning the overall Bayesian network to identify the best suited shower space.

1) Part learning:

Initially, we manually define the various parts to learn from the model forms: arrangement, door and shower tray. Then we associate each part to a list of extracted fields belonging to it. Each field corresponds to a node of the network part of its Bayesian form. From the list of a part nodes, we construct a fully connected graph as a basis to learn the Bayesian network structure. This learning is done using the algorithm MWST (see III-A). The probability learning uses the maximum likelihood algorithm (see III-B). Figure 7 shows an example of the Bayesian network structure and the corresponding part of the form. It highlights the dependencies that exist between the fields in this part. Such a correlation between fields Arrang1 (arrangement) and Shower_enclosure1, Arrang2 and Shower_enclosure2 and Arrang3 and Shower_enclosure3 can be observed.

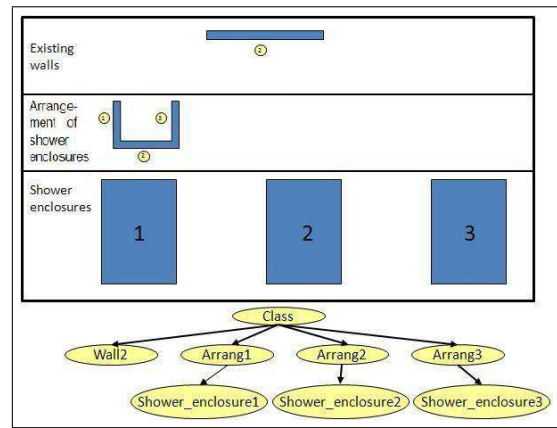
Figure 7. Example of the structure part of the model 5 and its corresponding Bayesian network structure

2) Global network training:

The global network regroups all parts of all shower space models and allows to link different parts of the forms and classes for all models. Initially, we use Bayesian networks to identify the model previously obtained in order to define for each part a probability distribution. Each distribution serves as a node to the global network. We then apply MWST algorithm to determine the structure of the global network. Figure 8 shows an example of the global network obtained for the classification of five types of shower space. We observe that the same part refers to several shower space models. Thus we have only two types of shower tray for five shower space models.
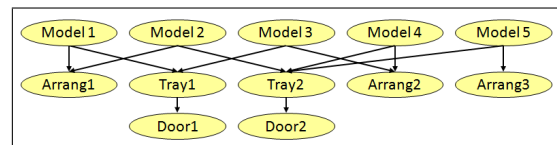

Figure 8. Example of global form structure classification for 5 models

E. Recognition

The recognition takes place in several stages. First as during learning, we extract the fields by matching strokes with model form fields. Then, for each part, we compute a belonging probability, i.e. the probability that the extracted fields correspond to the type of the given part using the networks defined for the parts. Then, we use the different probabilities obtained to define the type of space shower model using the global Bayesian network.

V. EXPERIMENTS

We tested our approach on a database of 500 forms. This database includes five shower models. In these models, we find three types of arrangements, two types of receivers and two types of doors.

We conducted tests using a cross-validation method. The test database was divided into four parts, each containing 25 forms of each class, which corresponds to 125 samples in total. From these subdivisions, we created four databases for learning and different recognition databases on which we applied our classification method. To validate our approach, we did a test by a simple matching. The results show a recognition rate of 81% against 96.7% when using Bayesian networks. Regarding the classification of forms using semantic constraints we obtained a recognition rate of 97.89%. These results show that our approach adapts to both physical and semantic constraints.

The tests were performed using the Matlab toolbox BNT ([1]). The results presented in Tables II, III and in figure 9 are encouraging.

	Arrangement			Shower tray		Door	
Class	1	2	3	1	2	1	2
Precision in %	94,1	96,2	96,3	94,7	95	95,9	96,8
Recall in %	97	96,8	97,7	97,1	96,5	98,3	96

Table II
RECALL AND PRECISION FOR THE THREE PARTS OF THE SHOWER SPACE

Class	1	2	3	4	5
Precision in %	96,8	92,7	95,4	95,7	96,9
Recall in %	97,1	95,6	97,9	97,5	95,3

Table III
RECALL AND PRECISION FOR THE TOTALITY OF THE SHOWER SPACE

The precision (number of documents correctly assigned to class / number of documents assigned to the class) and recall (number of documents correctly assigned to class / the class document number) are computed on three sides and on models in their entirety. We observed that the results are consistent regardless of the learning base. We can be satisfied with the search algorithm for Bayesian network structure. The accuracy rate for Model 2 is only 92.7%. This can be explained by its strong resemblance to model 1. Indeed, only two components differentiate the two models. Regarding the use of constraints, we can observe that the
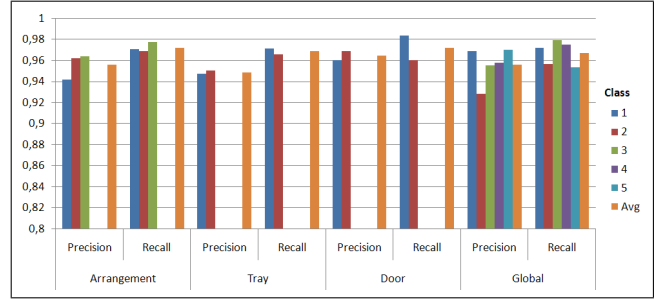

Figure 9. Result presentation

exploitation of architectural and syntactic constraints are successful. However, concerning the results of form parts show that the architectural constraints are more effective than semantic constraints. This can be explained by the complexity of the forms.

VI. CONCLUSION

We extended a method used primary for form classification using Bayesian networks with semantic constraints, to an architectural form identification method using physical constraints. The application chosen concerns the arrangement of a shower area. Early results are encouraging and pave the way for many opportunities. Initially, the application could be generalized to the complete arrangement of kitchen or bathroom spaces. Another way would be to expand the use of our method to problems involving other types of constraints, for example, logic or syntactic.

ACKNOWLEDGMENT

REFERENCES

[1] K. Murphy, *The Bayes Net Toolbox for Matlab*, Computing Science and Statistics, Proceedings of Interface, vol.33, 2001.

[2] R.E. Napolitain, *Learning Bayesian Netwokrs*, Prentice Hall, illustrated Edition, 2003.

[3] R.J.Milewski, V. Govindaraju and A. Bhaedwaj, *Automatic recognition of handwritten medical forms for search engines*, IJDAR, February 2009.

[4] G. Russell, M.P. Perrone, Y. Chee and A. Ziq, *Handwritten Document Retrieval*, IWFHR, Los Alamitos USA, 2002.

[5] P. Helman P, R. Veroff, SR. Atlas and C. Willman, *A Bayesian network classification methodology for gene expression data*, Journal Computer Biologie, 2004, vol 11(4), pp 581-615.

[6] E. Philippot, Y. Belaïd and A. Belaïd, *Bayesian Networks Learning Algorithms for Online Form Classification*, ICPR, Istanbul Turkey, August 23-26, 2010, pp 1981-1984.