Fuzzy relative positioning templates for symbol recognition

Adrien Delaye, Eric Anquetil
IRISA-INSA - Université Européenne de Bretagne
adrien.delaye, eric.anquetil@irisa.fr

Abstract—Relative positioning between components of a structured object plays a key role for its interpretation. Fuzzy relative positioning templates are a description framework for 2D handwritten patterns, that is based on positioning models specifically designed for dealing with variability and imprecision of handwriting. In this work, we present fuzzy positioning templates and investigate the idea of recognizing structured handwritten symbols by considering the relative positioning of the components, rather than the shapes of the components themselves or the global shape of the symbol. The templates are automatically trained from data without requiring any prior knowledge. Experiments on a database of on-line symbols prove that this original strategy is a promising approach for interpretation of structured patterns.

Keywords-on-line handwriting; symbol recognition; segmentation; fuzzy spatial relations;

I. Introduction

An interesting ability of human vision is to be able to recognize structure at the global level, by primarily considering the layout of structured objects, rather than focusing on the actual shapes of the inner components. For instance, one might have no difficulty in recognizing a complex structured symbol even if the shapes of primitives components have heavily distorted shapes. This illustrates the huge significance of spatial relationships and the predominance of their role in the recognition of structured objects.

Sketched symbols, complex east-Asian characters, and more generally bidimensional handwritten patterns often have a strong structural nature, where the role of spatial relationships is of primary importance. However, for their automatic recognition, most approaches rely primarily on shape information. The global shape of patterns is described, but the inner structure is only indirectly reflected in holistic approaches. Even in structural approaches, the shapes of primitives is often of more importance than the structural description itself, in the sense that shape is utilized for identifying segmentation hypothesis besides its role in the matching evaluation [LVSM02].

On the contrary, we propose in this work to exploit a new representation framework for structured handwritten symbols based on the primary importance of spatial relationships. We introduce structural fuzzy templates for describing 2D structures as a set of spatial relationships described in a flexible and adaptive fashion by the use of fuzzy mathematical morphology [Blo99], [DA10]. The idea is to investigate how the mere structure, described in terms of relative positioning and relative dimensions of primitives, conveys the essence of bidimensional structural patterns.

In previous work, we introduced a new type of spatial positioning models that make use of fuzzy mathematical morphology to deal with variability of the shapes of objects and with the imprecision of the spatial positioning itself [DA10]. The properties of adaptivity and flexibility qualify them for dealing with handwritten objects, a type of objects that can suffer large shape distortions and imprecise positioning.

In this paper, we first introduce the underlying theoretical framework of fuzzy spatial positioning, and predictive fuzzy spatial positioning models. We propose an extension of positioning models that permit to take into account not only the positioning, but also the *extent* of objects with respect to each other. The following section gives the details about structural fuzzy templates built upon these positioning models. The matching of templates with input handwritten pattern is detailed, and the learning process is presented. In the final section, we show by an experiment over a set of on-line Chinese characters, taken as an example of 2D symbols, that structured handwritten patterns can be recognized by application of fuzzy templates, while no modeling of the components shapes is included.

II. Predictive fuzzy models for spatial positioning description

In this section, we first introduce the idea of using fuzzy mathematical morphology for describing relative positioning between objects, and then the definition of predictive fuzzy positioning models. Eventually, we propose to improve these models by taking into account the *extension* of objects within the same descriptive framework.

A. Fuzzy spatial positioning with mathematical morphology

Following the works from Bloch about relative positioning of objects in images [Blo99], we make use of morphological fuzzy operators for describing spatial relationships between objects. The fundamental idea is to operate a morphological dilatation over a given reference object, by application of a well chosen 2D fuzzy function, called a fuzzy structuring element, to represent some directional positioning relation with respect to this reference. In our application case, the objects are handwritten strokes sampled from on-line signal, but the method is obviously applicable to images.

As an example, figure 1(a) shows a structuring element describing the relation *on the right* in the plane, while figure 1(b) represents the result of dilatation applied with this operator on a reference object. The dilatation is processed by moving the center point of the structuring element over all the points of the reference object, and keeping for each point of the plane the highest reached value. The result of the process is a 2D function that associates to each point of the plane a fuzzy degree between 0 and 1 describing its adequacy with the relation *to be on the right* of the reference object. In the sequel, we will call this function a *fuzzy landscape*, and denote it by $\mu_{right}(R)$, where R is the reference object.

(a) *Structuring element*

(b) *Fuzzy landscape "on the right of R"*

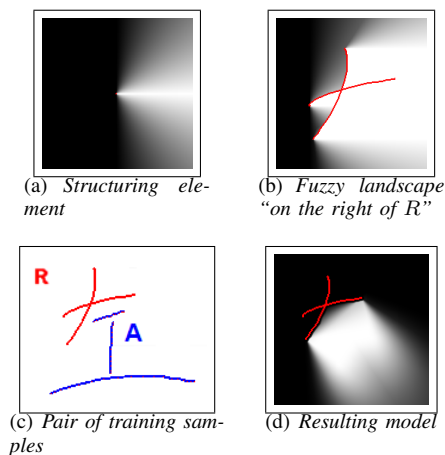(c) *Pair of training samples*

(d) *Resulting model*

Figure 1. Structuring Element for spatial relation *on the right of* (a), and resulting fuzzy landscape for reference object R (b). Example of a training pair reference-argument (c) and resulting model trained from samples and applied on a the same reference (d). Brightness reflects membership to the fuzzy functions, from 0 (black) to 1 (white).

The main interest of using such morphological operators is that the resulting description takes into account the exact shape of the reference object, since all the points of the sampled handwritten signal are considered in the morphological operation. Thus, the fuzzy landscape fits perfectly with the specificities of the shape of the reference object, which is a great advantage when dealing with handwritten strokes that are subject to large shape distortions.

B. Predictive spatial positioning models

For accurately modeling relative spatial positioning between objects, we proposed in previous work to rely on the description presented above for defining new types of positioning models and a procedure for automatically training these models from data [DA10]. Given a set of pairs of handwritten objects, each pair consisting of a reference object and an argument object, we can build fuzzy morphological operators (structuring elements) that model the admitted area of the space for training argument objects, with respect to their associate reference object. Several operators are learned, covering several points of view from which the argument is considered with respect to the reference (for example *on the right*, *on the left*, *above*...). By combining these points of view, a global positioning model can be learned, and the application of this model to a new given reference object results in a new fuzzy landscape that represents the expected location of argument object according to the learned spatial relation.

Figure 1(c) shows a training pair (made of a reference object R and an associated argument object A), and figure 1(d) shows the resulting learned model applied over the same reference object by morphological dilatation.

Detailed model formalization and learning strategy can be found in [DA10]. The interesting property of these models is that, being based on morphological operations, they benefit from the associate adaptivity and flexibility. It means that the same model will result in different fuzzy landscapes when applied to different reference objects, and that it will adapt to the singularities of the reference shape. Moreover, these models have a very useful prediction ability. Since the fuzzy landscape is defined in the plane, it can qualify the admissibility of any individual point with respect to the relationship at hand and the given reference object. We will show in the next section that this is an interesting property for solving the issue of stroke segmentation in the matching of structural models.

C. Integration of extension modeling

Fuzzy relative positioning models are able to describe accurately the position of elements with respect to a reference, and dimensions of elements are also somehow captured in the models. Indeed, if the training samples have very small dimensions and have a very stable positioning with a low variance, the resulting model will be very accurate and only allow a limited part of the plane. However, if the elements have a high variance in their relative positioning, the resulting trained model will be more lose and admit a larger part of the plane.

Consequently, a large, extended model (i.e. a model that considers a large part of the plane as acceptable area for positioning with respect to a reference) does not necessarily models the positioning of a large, extended object. This phenomenon is illustrated by figure 2. Two fuzzy positioning models are represented with associated learning reference and (imaginary) argument objects. While the dimensions of the argument object differs a lot between the two images, the models are similarly extended, probably because the argument object (a) has a larger variance in vertical position.

This can lead the model to be well activated even by objects that differ completely from the training samples, in terms of dimensions, or *extension*. To overcome this, an additional extension model is added, which evaluates the adequacy between extension of training samples and that of
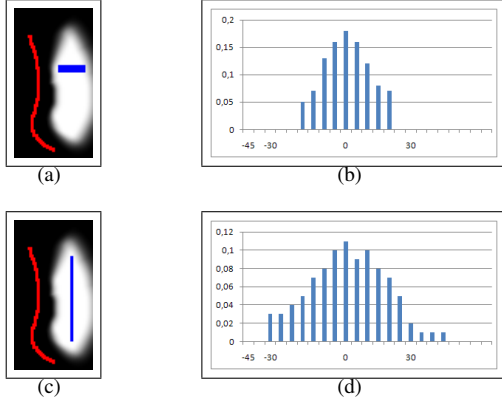
Figure 2. Objects of different dimensions can lead to similar positioning models (a) and (c). Extension models (density of points over angular directions) (b) and (d) allow to differentiate between the two cases by modeling the extension of training objects with respect to considered directions.
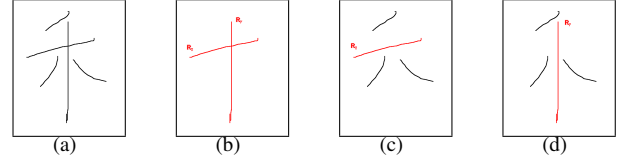

Figure 3. Example of a Chinese character (a). Two reference strokes (b). Illustration of the positioning relation of the character parts with respect to horizontal reference (c) and vertical reference (d).

the test samples. It simply consists in a histogram model, expressing the density of points from training samples in angular directions seen from the reference. It makes use of the same representation space as for positioning description, where angles are measured from the reference. For an argument object, all his points are projected in the angular representation, where angles are measured relatively to the angle from the reference to the centroid of the argument. A small object, even with a large position variation, has a small variation of angles within its points, and will thus have a narrow extension model. The two extension models corresponding to the two models are depicted in figure 2(b) and (d), for consideration of direction *down* with respect to the reference. The models can be differentiated according to this extension information.

D. Evaluation of models

For a given reference object, and a learned positioning model, the adequacy of an argument object is evaluated by combining the positioning and the extension aspects.

The global evaluation of positioning of an object A is computed as the mean of activation of all its points by the fuzzy landscape developed over the reference R. The extension score is evaluated by computing the distances between extension histograms of the model and normalized extension histograms measured on the object A (always with respect to R). A simple histogram distribution distance is used for distance computation. The sum of distances over each considered direction is then converted to an adequacy score, from 0 to 1, by a Cauchy-shaped similarity function.

The overall adequacy to a model is then computed as

$$a(R)(A) = \top(\mu(R)(A), \nu(R)(A)) \tag{1}$$

where ν is the extension adequacy and $\top$ is a fuzzy conjunctive operator (t-norm).

III. FUZZY TEMPLATES FOR SYMBOL DESCRIPTION

Fuzzy positioning models offer a tool for describing the spatial relations of objects with respect to a reference. In order to provide a global point of view of a structure, a strategy is needed for organizing positioning models and defining the references.

A fuzzy relative positioning template is defined as a set of several fuzzy positioning models, each one describing the positioning of one of the primitive parts of the pattern, with respect to a global reference. The global reference should be a part of the pattern itself, because the descriptive power of positioning models lies in their adaptivity to handwritten strokes as references. Obviously, the reference stroke should be a stable and easily identifiable part of the input pattern. In order to offer an optimal coverage of the pattern in terms of spatial relations, and thus to maximize the expressiveness of positioning models, we propose to select two distinct global references : an "horizontal" one and a "vertical" one. This way, a good support is provided for describing positions regarding the two dimensions of the plane.

Example of figure 3 illustrates the interest of choosing references that are extended in x and y directions. In this case, the independent parts of the pattern can be easily positioned with respect to the references, depicted as red strokes.

A. Strategy for reference extraction

We can assume with no lack of genericity that any type of 2D handwritten pattern contains identifiable strokes, sequence of strokes, or substrokes that are extended enough for each main direction of the plane. For finding hypothetical reference strokes, we first look for candidate segmentation points in the on-line handwritten signal, such as strokes beginning and ending points, as well as points extracted from polygon-based estimation of strokes, or points that have extreme x or y coordinates. From pairs of candidate segmentation points we make a list of candidate segments (defined by the part of strokes between the two points in on-line signal) and rank the segments according to a scoring function based on the *expected ability to express relationships*. The candidate segments are ordered by increasing values of a scoring function defined, for a "horizontal" candidate segment R of a pattern S, by
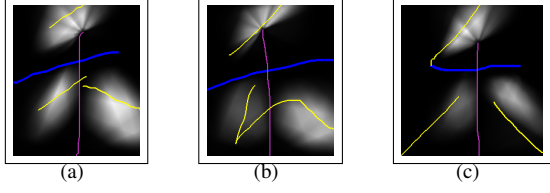
Figure 4. Fuzzy relative positioning template applied on three occurrences of the same symbol, with segmentation variants.

$$v_x(R) = \frac{1}{|S \setminus R|} \sum_{p \in S \setminus R} \max(\mu_{up}(R)(p), \mu_{down}(R)(p)) \tag{2}$$

where $\mu(R)$ is defined as explained in section II-A. A similar scoring function is defined for ordering "vertical" candidate segments. The score lies in $[0,1]$, and it reaches its maximum value when the primitive R perfectly spans every point of the pattern when considering either one of "up" or "down" points of views. A good reference is then a reference from where all of the pattern can be "seen" by considering the two orthogonal cardinal directions.

The design of scoring functions is consistent with the chosen positioning method definition, and it insures that chosen reference strokes maximize the quality of the positioning description. Candidate primitives for horizontal and vertical references are sorted according to the scoring functions above, and a limited number of best primitives are kept for further process.

B. Model formulation

A fuzzy template model t is a 3-tuple $< a^H, a^V, M = (a_h^i, a_v^i)_{i=1..n} >$, where
- a^H is a fuzzy positioning model describing the position of horizontal reference with respect to the vertical one;
- a^V is a fuzzy positioning model describing the position of vertical reference with respect to the horizontal one;
- $M = (a_h^i, a_v^i)_{i=1..n}$ is a set of n pairs of positioning models, each describing the position of one element of the template with respect to horizontal reference (a_h^i) and to the vertical reference (a_v^i).

The images from figure 3 implicitly show the different aspects of a template: positioning of the two references with respect to each other (b), and the positioning of parts of the pattern relatively to horizontal (c) and to vertical (d) references. Figure 4 gives a visual representation of the template corresponding to the same symbol.

C. Model inference

Formally, the adequacy of an input pattern s with respect to a template t is expressed by:

$$t(s) = \max_{(s_h, s_v) \in h_h \times h_v} f(t_0(s_h, s_v), t_1(s_h, s_v, s|_{hv})) \tag{3}$$

where h_h and h_v denote the list of candidate primitives for horizontal and vertical references, respectively, and $s|_{hv}$ is the remaining pattern, i.e. $s|_{hv} = s \setminus \{s_h \cup s_v\}$. f is a function from $[0,1] \times [0,1]$ to $[0,1]$, strictly increasing with respect to its arguments, and serves as a fusion function between two partial evaluations given by t_0 and t_1. $t_0(s_h, s_v)$ evaluates how the positioning and shapes of s_v and s_h with respect to each other matches the model. $t_1(s_h, s_h v, s|_{hv})$ gives a score for the matching of the elements of $s|_{hv}$, i.e. the remainder of the pattern, with respect to the template supported by the references s_h and s_v. The two evaluation functions are detailed below.

1) Evaluation of references positioning: The first part of the evaluation focuses on the good adequacy of references positioning with each other and on their shapes. This evaluation is processed by combining the score from positioning models a^H and a^V, and combined with a t-norm operator implementing a fuzzy conjunction operator.

$$t_0(s_h, s_v) = \top(a^H(s_h)(s_v), a^V(s_v)(s_h)) \tag{4}$$

where a^H and a^V are the positioning models for each reference with respect to the other, and $a(x)(y)$ denotes the global evaluation of the model over the reference x and argument y, as expressed in equation 1.

2) Model-driven segmentation of the pattern: The online input signal has a natural segmentation, in which segments are delimited by pen-up points. $r = s|_{hv}$ is then a sequence of strokes $r = r_0..r_m$. The segmentation of input pattern into strokes may be highly unstable and suffer large differences among writing styles. A "canonical" segmentation of the pattern has to be found before permitting the structural matching with the template.

In our case, we can make use of the predictive ability of the positioning models for finding hypothetical segmentation points. Indeed, the positioning models can be used to detect where the boundaries of each element should lie. In other words, the positioning models can "select" the part of the pattern of interest to them.

We first compute the adequacy degrees of each point of the input pattern with respect to every model available in the template $M = (a_h^i, a_v^i)_{i=1..n}$. Then we select as candidate segmentation point points where occur significant changes of *dominant model* over the strokes, where significance is defined in terms of minimum number of points or minimal length of candidate segments.

3) Branch and bound search for optimal assignment: Evaluation of the remainder of pattern r with respect to its references s_h and s_v and to the template t consists in finding the best segmentation and the best assignment of candidate segments to sub-models of the template. Let us denote $r = r_1..r_N$ the list of candidate segments such that each candidate segment r_k is delimited by two successive candidate segmentation points found in the model-driven

segmentation. Now we are looking for an assignment of $r_k, k = 1..N$ to the sub-models of the template: $M^i = (a_h^i, a_v^i), i = 1..n$. This function $\sigma : \{1..N\} \rightarrow \{1..n\}$ does not necessarily have to cover all the n models of the template, since a template model should be evaluable even when one of its models is not provided any segments.

A branch and bound algorithm is employed in order to find the best assignment

$$\sigma^* = \arg\max_\sigma \sum_{i=1..n} \top(a_h^i(s_h)(\{r_{\sigma^{-1}(i)}\}), a_v^i(s_v)(\{r_{\sigma^{-1}(i)}\}))$$
(5)

with $\{r_{\sigma^{-1}(i)}\}$ the set of segments assigned to model i by σ. Starting from an empty assignment function σ_0, the algorithm generates candidates assignment functions by adding associations for non-matched segments of r. The search for σ^* benefits from an efficient heuristic function that is computed by assigning each point p_j of non-matched segments to their best model. This ensures to give an optimistic evaluation of what remains to be matched and thus to be an admissible heuristic, i.e. the optimal solution for σ is always found. The score optimized by the finding of σ^*, denoted $t_1(s_h, s_v, s_{|hv})$, allows to complete the evaluation of the template matching.

D. Learning procedure

A process for automatically training templates from symbols samples is set up. The main characteristic of this learning procedure is that is does not need structural labeling of the training patterns, nor any prior information about symbol domain. Information about the number of parts, boundaries of segments, and the determination of references strokes are all discovered automatically. For this, the template is initiated with the references that maximize the scoring function defined in equation (2), and by considering the natural segmentation of strokes in the samples. Then a number of iterations permit to optimize the definition of the template, by successively matching training samples with the template, discovering their segmentation through inference procedure, and learning new template model based on these segmented patterns. At each step, the template is is evaluated over a set of validation samples. The process stops after a predefined number of iterations, insuring a good quality of models with respect to their intrinsic evaluation over the validation set.

IV. Experiments

The performance of fuzzy relative positioning templates for description and recognition of handwritten symbols was experimented over a set of Chinese characters, which are highly structured by nature and can be considered as "symbols" in the generic sense. For the 50 classes modeled, fuzzy templates were automatically learned from 40 samples (25 training samples + 15 validation samples), and recognition
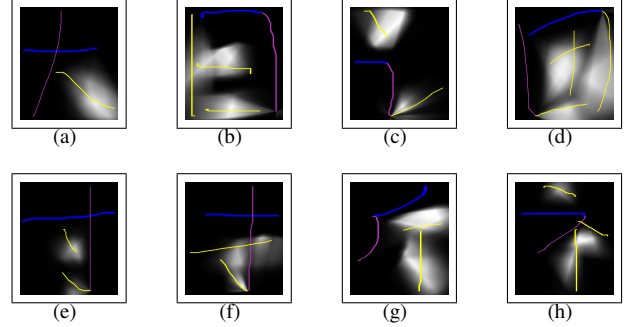

Figure 5. Example of templates applied on Chinese characters.

rates were measured over 20 more samples. The overall recognition rate over this set of symbols is about 86 percent. It is to note that the main source of error is confusion between symbols whose structure is very similar, or even indistinguishable. In some cases the method is expected to fail, because the confused symbols only differ by the shape of one of their primitives. For a qualitative assessment of the fuzzy templates, figure 5 presents several examples of symbols with application of the associated fuzzy templates.

V. Conclusion

The approach presented in this work follows a very original idea, based on the observed ability of human vision: how possible is it to recognize structured symbols by only taking into account the structure itself, in terms of relative positions and dimensions, while no information about the shapes of handwriting stroke is modeled. For this, we introduced advanced modeling of relative positions, based on fuzzy relative models with mathematical morphology, and an associated procedure for automatic learning of fuzzy structural templates. The results prove on a modest database of samples that the approach can open new perspectives for recognition of structured symbols.

References

[Blo99]   I. Bloch. Fuzzy relative position between objects in image processing: a morphological approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(7):657–664, 1999.

[DA10]   A. Delaye and E Anquetil. Learning spatial relationships in hand-drawn patterns using fuzzy mathematical morphology. In *Proceedings of the 2nd International Conference of Soft Computing for Pattern Recognition*, 2010.

[LVSM02] Josep Llados, Ernest Valveny, Gemma Sanchez, and Enric Marti. Symbol recognition: Current advances and perspectives. In Dorothea Blostein and Young-Bin Kwon, editors, *Graphics Recognition Algorithms and Applications*, volume 2390 of *Lecture Notes in Computer Science*, pages 104–128. Springer Berlin / Heidelberg, 2002.