Learning spatial relationships in hand-drawn patterns using fuzzy mathematical morphology

Adrien Delaye, Eric Anquetil

Learning spatial relationships in hand-drawn patterns using fuzzy mathematical morphology

Adrien Delaye, Eric Anquetil
INSA de Rennes, Avenue des Buttes de Coësmes, F-35043 Rennes
UMR IRISA, Campus de Beaulieu, F-35042 Rennes
Université Européenne de Bretagne, France
{adrien.delaye,eric.anquetil}@irisa.fr

Abstract

We introduce in this work a new approach for learning spatial relationships between elements of hand-drawn patterns with the help of fuzzy mathematical morphology operators. Relying on mathematical morphology allows to take into account the actual shapes of hand-drawn patterns when modeling their spatial relationships, and thus to cope with the variability of handwriting signal. Extension of mathematical morphology to the fuzzy set framework further allows to handle imprecision of handwriting and to deal with the ambiguity of spatial relationships.
The novelty lies in the generative *aspect of the models we propose, in the sense that they can exhibit the region of space where the learnt relation is satisfied with respect to a reference object, and can thus be used for driving structural analysis of complex patterns. Experiments over on-line handwritten data show their performance, and prove their ability to deal with variability of handwriting and reasoning under imprecision.*

1. Introduction

Modeling relative positioning between objects plays a key role for many computer vision applications and is often necessary for processing tasks such as image segmentation, object detection and of course, scene understanding. More specifically, spatial relations provide useful information for solving the problem of general hand-drawn scheme interpretation, which involves the analysis of complex and highly structured two-dimensional handwritten patterns. For example, when dealing with handwritten mathematical equations, one has to analyze the relative positioning of isolated symbols in order to interpret the global meaning of the equation. In this specific case, some semantic operators such as *exponent* are only expressed through the *superscript* spatial relation between its two operands.

While human beings have the ability to intuitively think of relative positioning between objects with the help of concepts such as distance (*far from*, *close to*...) or directions (*on the left of*, *above*...), it is clear that these concepts are not precise by nature, and that they should be modeled under a flexible framework supporting reasoning under imprecision and dealing with ambiguity. Actually, applying soft computing for modeling spatial relations is not new, and many methods have been proposed following that idea (see for example the book [MS02]). A very good review of fuzzy-based methods for defining spatial relations in images was proposed by Bloch [Blo05].

Most of the methods used in computer vision achieve to provide a rich description of relative positioning by implicitly taking into account the shapes of the objects: indeed, they consider all the points from the two objects at hand for defining their spatial relations. This is the case for methods based on histogram of angles (compatibility, aggregation methods), or on linear sections (F-histograms)...

Other methods are notable for explicitly considering the shapes of the objects when modeling their relative positioning [Gad97], [Blo99]. The idea shared by Gader and Bloch is to describe spatial relations with the help of morphological operations processed in the image space, directly on the objects. Specifically, fuzzy structuring elements are proposed for modeling fuzzy spatial relations that were shown to better fit the intuition in comparison to other positioning descriptions [BR03]. We propose in this work to exploit this idea and introduce extended morphological models in the frame of hand-drawn pattern recognition application.

When dealing with hand-drawn patterns, we need to model the relative position of objects that are by nature noisy, imprecise, and subject to a strong variability according to the numerous input conditions: identity of the writer, nature of the input material, environment, space and time... If much research effort has been spent for improving the modeling of isolated shapes [PS00], most models for relative positioning rely on a rather simple description that consists in summarizing each hand-drawn object into a single point (centroid based method) or a virtual rectangle box (bounding box based methods), sometimes under the fuzzy set framework [ZBZ05]. This leads to poorly intuitive and poorly informative spatial positioning modeling that may penalize the overall recognition performances.

In this paper, we present a new method for learning

models for spatial relations based on fuzzy mathematical morphology operators. The original idea from Bloch is exploited further here : we propose to build models by automatically learning the fuzzy structuring elements supporting the morphological operations. The model (learned fuzzy structuring element) permits to deal well with the variability of hand-drawn patterns, as well as imprecision and ambiguity of their positioning. Besides, an interesting novelty is its generative property, in the sense that such a model can exhibit the region of the plane where the spatial relation is satisfied with respect to a given reference object. This feature make the model suitable for modeling spatial constraints in a context-driven parsing mechanism, or more generally make it usable in a predictive way for structural pattern recognition applications.

First section exposes the idea of using mathematical morphology for modeling relative positioning of objects such as introduced by Bloch. The formalization of new models and associated learning process is then presented in second section. Third section sums up some experimental results obtained on a database of handwritten gestures, showing the performance of the proposed models.

2. Relative positioning with fuzzy mathematical morphology

We first expose the general approach as introduced in [Blo99] for describing relative positions of objects. Although the method was designed to deal with fuzzy objects in images and can be extended to the 3D case, we only present here the concepts required for our target application: space is reduced to the plane, and we only consider crisp objects (classical sets of points). The limitation to crisp objects is not a heavy constraint when dealing with handwritten patterns, in contrast with image processing applications where uncertainty about the objects boundaries has to be modeled.

The fundamental principle is to see a spatial relation as a fuzzy set describing the adequacy of any point of the plane to the relation at hand, defined relatively to a reference object. Such type of relation can be for example *on the right of* R, where R is the reference object. In the sequel, this fuzzy function is referred to as the *fuzzy landscape* according to the vocabulary defined in [Blo99]. The evaluation of a positioning relation for a given object A with respect to R is processed in two steps : first step consists in defining the *fuzzy landscape* by operating a morphological operation over the reference object R, and second step consists in evaluating the adequacy of object A with the *fuzzy landscape*.

2.1. Definition of the fuzzy landscape

First step consists in derivating the absolute concept considered, for example *on the right of*, to the reference object

of interest, say R, in order to model the instantiated relation *on the right of* R. This is done by applying a morphological dilatation of the reference object with a structuring element (SE) modeling the absolute concept. For a given reference object R, located in the plane S, the dilatation of R with the SE ν is computed by:

$$\forall P \in S, \mu^R(P) = \max_{Q \in R} \nu(P - Q) \qquad (1)$$

(where P and Q are points of the plane, and Q belongs to the object R).

The resulting function μ^R is called *fuzzy landscape*, and it describes for each point of the plane its adequacy degree to the relation described with respect to R, illustrating the instantiated concept. Its definition explicitly takes into account the shape of the reference object R, by scanning all its points. It is the shape of the SE ν that defines the spatial relation modeled, and the fuzziness of the SE determines the fuzziness of the resulting landscape.

As an example, the following definition can be adopted for the fuzzy SE modeling the relational concept *in direction* α [Blo99]:

$$\forall P \in S, \nu_\alpha(P) = max\left(0, 1 - \frac{2}{\pi} \arccos \frac{\overrightarrow{OP}.\overrightarrow{u_\alpha}}{||\overrightarrow{OP}||}\right) \quad (2)$$

where O is the center of the SE and u_α is the unit vector of direction α. The SE fuzzy membership function ν assigns to each point P of the plane a degree from 0 to 1, varying linearly with the angle between $\overrightarrow{OP}$ and u_α.

Likewise, we can define an undirected fuzzy radial SE for modeling a distance relation *far from*, in the euclidean sense.

Figure 1(a) shows a SE for the relation *on the right of* defined according to equation 2, and figure 2 illustrates the fuzzy landscape obtained by dilatation of a reference object with this SE, according to equation 1.

Other definitions of structuring elements were proposed in [CA07], where the authors introduce several parameters controlling the flexibility of the radial and directional dimensions of the structuring element and thus provide a customized family of hybrid distance-direction fuzzy structuring elements.

2.2. Evaluation of the relationship

Once the fuzzy landscape for a spatial relation has been defined as described above (either based on directional or distance structuring elements), the second step consists in evaluating the adequacy degree of the whole object A with this fuzzy landscape, giving a global evaluation of the spatial relation between the two objects A and R. Several operators are proposed to aggregate the membership values of points of A to the fuzzy landscape for obtaining this global degree :

(a) *Structuring Element* ν_{right}

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

(c) *Structuring Element* ν_{dist}

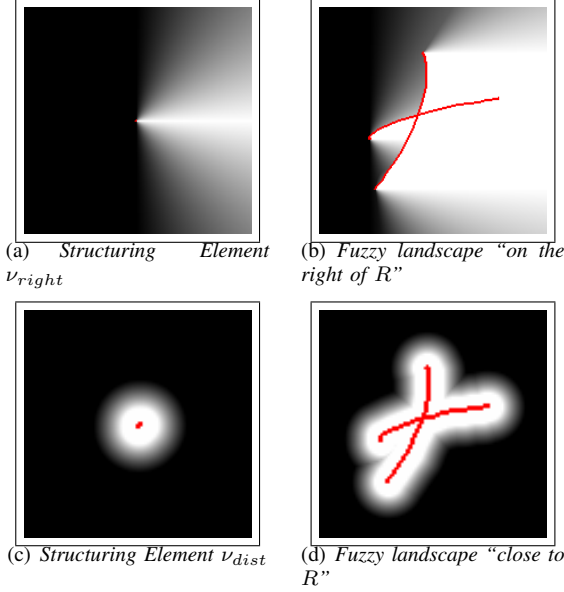(d) *Fuzzy landscape "close to R"*

Figure 1. Structuring Elements for spatial relation *on the right of* (a), and *close to* (c), and resulting fuzzy landscapes around a reference object R (b, d). Brightness reflects membership to the fuzzy functions, from 0 (black) to 1 (white).

mean measure, but also optimistic and pessimistic measures given by possibility and necessity respectively [Blo99]:

$$M^R(A) = \frac{1}{|A|} \sum_{P \in A} \mu^R(P) \qquad (3)$$

$$N^R(A) = \inf_{P \in A} \mu(R)(P) \qquad (4)$$

$$\Pi^R(A) = \sup_{P \in A} \mu(R)(P) \qquad (5)$$

As an illustration, the table 1 gives the three degrees measured for the object A embedded in the fuzzy landscape *on the right of R* from the figure 2. The mean score of 0.73 shows that the object A is mostly on the right of R. The two other measures traduce the fact that some points of A are *perfectly* on the right of R in the sense of the fuzzy landscape (hence possibility is equal to one), while some points of A are *not at all* on the right of R (hence necessity is null).
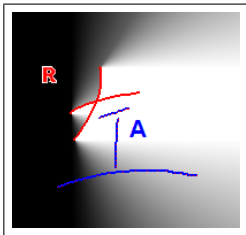

Figure 2. Relation *right of R*

Table 1. Evaluation of the relation

Measure	Value
$M^R(A)$	0,73
$N^R(A)$	0
$\Pi^R(A)$	1

3. Learning generative positioning models

As exposed in the previous section, the use of mathematical morphology permits to take into account the actual reference object in the definition of the fuzzy landscape, thus the description of the positioning naturally fits to the peculiarities of its shape. The fuzzy landscape modeling a relation such as *on the right of R* is suitable to a human-like description of the relation and fits the intuition very well, including when the reference object has singularities, concavities...

Moreover, thanks to its fuzzy definition, it deals well with the imprecise relative positioning of handwritten patterns and the spatial relationship evaluation can be included in a general soft computing framework for making recognition decision. Our idea is now to extend the idea presented above and to rely on fuzzy mathematical morphology operators for learning positioning models having a *generative* property, i.e. able to address two types of problems:

1) given a reference object and this positioning model, in what area of the plane (image space) is the argument object expected to be located ? (prediction problem)
2) given two objects, to what extent does their spatial relation fit with the model ? (evaluation problem)

Whereas all the methods encountered in the literature focus on the evaluation task, the prediction problem is very rarely addressed. However, predicting the location of objects prior to segmenting them can be very useful for example for driving constraint-based parsing or restricting the area of the plane where objects should be found. To our knowledge, only the works of Bloch [Blo99] consider the prediction task in the image space, but only for predefined relations (such as *on the right*), and propose no scheme for learning such models from samples.

In this section we present a method for automatically learning from samples generative models by relying on fuzzy morphology operators. For that, we propose to pick 5 axis (the 4 main directions and 1 distance), and to learn for each of them a fuzzy SE dedicated to the spatial relation being learned. We can then determine, with respect to each axis and for a given reference object, a fuzzy landscape describing the area admitted by the model. The global model is ultimately obtained by fuzzy intersection of these directional and distance fuzzy landscapes.

3.1. Structuring element learning

From several instances of objects $(R_i, A_i)_{i=1..N}$, denoting the training samples, that share the same positioning relation to be modeled, we want to model to what extent they satisfy the relationships carried by each of the 4 directional axis *up*, *down*, *left* and *right*, and the distance axis *far from*.

For each axis α, we consider the distribution of the degrees $(x_{i_k})_{i=1..N}^\alpha$ reached by the training points from N

objects A_i with respect to their associated reference objects R_i:

$$(x_{i_k})_{i=1..N}^\alpha = \left\{ x, \exists P \in A_i, \mu_\alpha^{R_i}(P) = x \right\}$$

We simply approximate this distribution by an histogram function H_α, normalized by the maximum frequency. This histogram describes a function from $[0,1]$ to $[0,1]$ modeling the degrees reached by training points with respect to the relation carried by axis α.

When combining the function H_α with the original SE corresponding to axis α, we can define a new fuzzy morphological operator that can be seen as a learnt SE, denoted $\hat{\nu}_\alpha$, and defined by :

$$\hat{\nu}_\alpha = (H_\alpha \circ \nu_\alpha)$$

Figure 3(a) shows this type of SE describing the positioning model for objects R and A of figure 2 according to the direction *right*. Associated fuzzy landscape $\hat{u}_\alpha(R)$ assigns to each point of the plane its adequacy degree with the learnt relation according to the direction. Thus, figure 3(b) represents the relation *to be on the right of R in the same extent than A according to the learnt model*. In other words, the brighter points are those for which the validity of relation *to be on the right of R* is the most conform to the degrees reached by the training points. The white area could be used to predict where it is expected to find an object positioned relatively to this given reference object, according to the model learnt and the considered axis (*right*).

(a) SE $\hat{\nu}_{right}$

(b) $\hat{\mu}_{right}^R$
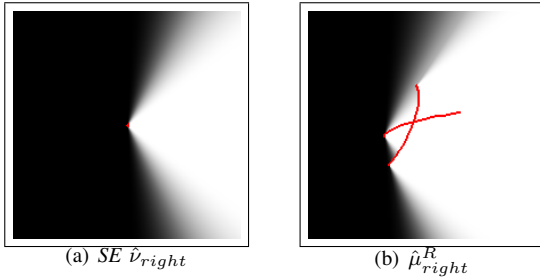
Figure 3. Learnt structuring element (a) and associated fuzzy landscape (b) for direction *right*

3.2. Fusion by fuzzy intersection

The process described in the previous part is repeated for each considered axis : 4 directions *up*, *down*, *left*, *right*, as well as on the additional distance axis. When given a new reference object, the five models are exploited separately, resulting in five fuzzy landscapes forming sub-parts of a global positioning model. This global model is then constructed by

(a) $\hat{\mu}_{up}^R$

(b) $\hat{\mu}_{down}^R$

(c) $\hat{\mu}_{left}^R$

(d) $\hat{\mu}_{right}^R$

(e) $\hat{\mu}_{dist}^R$
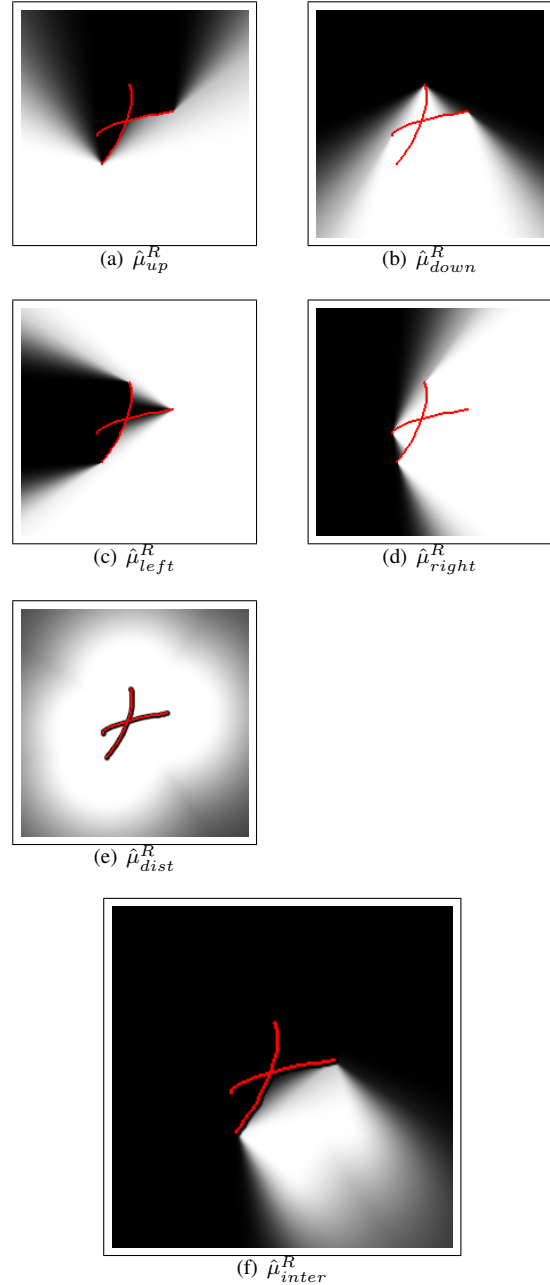
(f) $\hat{\mu}_{inter}^R$

Figure 4. Fuzzy landscapes from learnt SE for 4 directions (a,b,c,d) and distance (e), and global intersected landscape (f)

intersecting the fuzzy landscapes obtained considering each axis :

$$\hat{\mu}_{inter}^R = \top(\hat{\mu}_{up}^R, \hat{\mu}_{down}^R, \hat{\mu}_{left}^R, \hat{\mu}_{right}^R, \hat{\mu}_{dist}^R)$$

where $\top$ is a T-norm fuzzy operator. Intuitively, a point P is considered properly positioned with respect to the model if it fits with the landscape of each axis.

Figure 4 illustrates the fuzzy landscapes obtained for the

4 directions (a, b, c, d) and distance (e). The two fuzzy landscapes obtained for directions *up* (a) and *down* (b) are partly redundant. The first one models the fact that the object A is not above R, while the second one models that A is below R. However, this redundancy is necessary as the *up* fuzzy landscape is not completely included (in the fuzzy inclusion sense) in the *down* landscape, and it is then beneficial to intersect these two landscape for building a more accurate model. The same reasoning holds for *left* and *right* landscapes (c,d). Eventually, the global model computed by intersection is represented in figure 4(f) (we here the *product* as T-norm operator in our illustrations and experiments). This global landscape $\hat{\mu}_{inter}^{R}$ describes the area of the plane that satisfies the positioning relation with respect to R according to the model (prediction task), and can be exploited for evaluating the positioning of A (evaluation task) similarly to any of the fuzzy landscapes, by computing evaluation measures as exposed in section 2.2.

4. Experiments

We compared the performance of our models with respect to the evaluation task on a recognition problem of on-line gestures. The database consists of 18 classes of gestures, diacriticals, and punctuation symbols drawn relatively to a reference handwritten letter (see table 2).

Since several letters have a similar shape (e.g. comma and acute accent), modeling their positioning with respect to the reference letter is necessary for recognition. The experiment consists in classifying these gestures with a standard SVM classifier, by using different sets of positioning features combined with a fix set of shape features. Table 3 presents the features sets and the recognition rates obtained. Note that the improvement from *exp1* to *exp2* proves the interest of morphological description over bounding box based description, as expected. *exp3* further shows that learned models slightly improve the recognition, attesting their quality for the task of evaluating relative positioning of patterns.

Besides, the figure 5 illustrates the ability of the learned models to address the prediction task: i.e. exhibit the position of an argument object as fuzzy area of the plane that is adapted to the reference object R. Indeed, the results fit the intuition very well on the considered classes: *apostrophe*, *case switch* and *comma*.

In order to further prove the ability of the models to deal with different natures of handwriting data and to adapt properly to different shapes of reference objects, we present a model learnt for Chinese character structure analysis in figures 5 (d,e,f). In this case the same model is expanded over several reference objects, belonging to different classes of characters. The results show its ability to absorb strong variations in the shape of the reference while keeping a very good predictive performance.

Table 2. 16 selected classes from the online gestures database

acute	grave	circumflex	dieresis
accent deletion	cedilla	cedilla deletion	case switch
character deletion	last character deletion	caret return	insertion
space	apostrophe	comma	dot

Table 3. Comparison of recognition rates obtained with different sets of features

	features	%rec
exp0	(no positioning features)	55,35
exp1	4 bounding box based features	80,23
exp2	5 mean adequacy degrees $M_{left}^{R}, M_{right}^{R}, M_{up}^{R}, M_{down}^{R}, M_{dist}^{R}$	96,03
exp3	18 mean adequacy degrees to models $(\hat{\mu}_{inter}^{R})_{c=1..18}$	96,16

5. Conclusion

We have formalized in this paper a new method for automatically learning generative positioning models from samples that are designed to handle variability and imprecision of hand-drawn patterns. We proved in the experiments their quality in not only describing the positioning of handwritten patterns (evaluation task), but also predicting the position of an argument object with respect to a given reference object (prediction task). Future works will aim at quantitatively evaluate this prediction ability and exploiting it for driving the segmentation and the analysis of complex and highly structured hand-drawn patterns (such as Chinese characters and mathematical equations).

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.

[Blo05] I. Bloch. Fuzzy spatial relationships for image processing and interpretation: a review. *Image and Vision Computing*, 23(2):89–110, 2005.

(a) *comma*

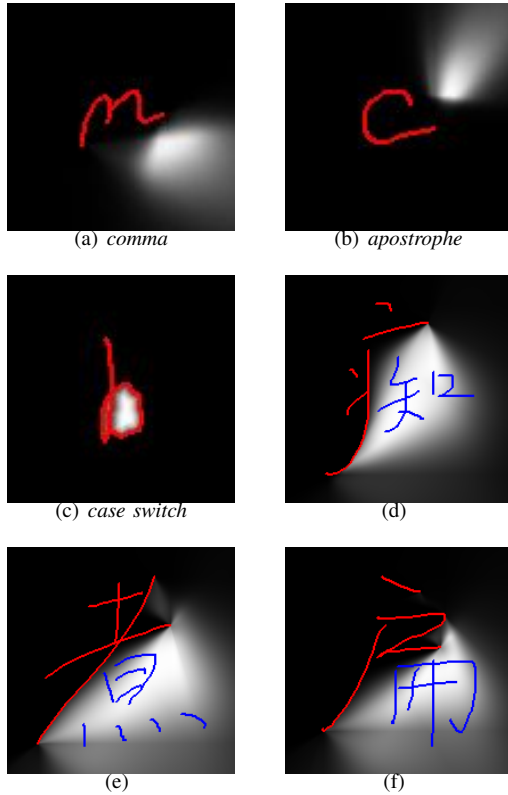(b) *apostrophe*

(c) *case switch*

(d)

(e)

(f)

Figure 5. Representation of learned models for classes "comma" (a), "apostrophe" (b), and "case switch"(c). Three instances (d,e,f) of the same positioning model for Chinese character structural template *bottom-right*

[BR03]   I. Bloch and A. Ralescu. Directional relative position between objects in image processing: a comparison between fuzzy approaches. *Pattern Recognition*, 36(7):1563–1582, 2003.

[CA07]   R.C. Cinbis and S. Aksoy. Relative position-based spatial relationships using mathematical morphology. In *Proceedings of the IEEE International Conference on Image Processing*, volume 2, pages 97–100. IEEE Computer Society, 2007.

[Gad97] P.D. Gader. Fuzzy spatial relations based on fuzzy morphology. In *Proc. Sixth IEEE International Conference on Fuzzy Systems*, volume 2, pages 1179–1183 vol.2, 1997.

[MS02]   P. Matsakis and L.M. Sztandera, editors. *Applying soft computing in defining spatial relations*. Springer, 2002.

[PS00]   R. Plamondon and SN Srihari. Online and off-line handwriting recognition: a comprehensive survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(1):63–84, 2000.

[ZBZ05] L. Zhang, D. Blostein, and R. Zanibbi. Using fuzzy logic to analyze superscript and subscript relations in