

Explicit fuzzy modeling of shapes and positioning for handwritten Chinese character recognition

Adrien Delaye - Eric Anquetil - Sébastien Macé
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,sebastien.mace}@irisa.fr

Abstract

In this paper, we present a new method for on-line Chinese character recognition that relies on an explicit description of characters structure. Contrary to most of known structural approaches, this model can describe characters written in a fluent style, thanks to a flexible fuzzy modeling of shapes and positioning of their structural components (primitives and radicals). We designed a process for incremental training of the models cooperated with automatic structural labeling for minimizing the required manual task in model design. First experiments show that the method is able to recognize non-regularly written characters and has a convincing generalization ability.

1 Introduction

Chinese characters are known to be naturally structured patterns, and it is commonly admitted that considering structural information would be beneficial in the perspective of building interpretable recognition systems, in opposition to the many “black box” holistic approaches that have been proposed in the recent years for their automatic recognition [4]. Nevertheless, statistical methods are currently offering the highest performances since they are more suited to handle imprecision and variation of writing styles, including stroke connections, stroke order variability and shape distortions. Besides, statistical models are easier to learn from samples in comparison with structural models that often require a tedious manual design task.

In this paper, we introduce a new hybrid statistical-structural method for on-line Chinese character recognition, based on an explicit description of their structure. While our previous work was limited to the modeling of sub-characters from structural primitives [3], we extend here this approach to the modeling of characters by adding a sec-

ond structural level.

We model characters as a combination of morphological information (shapes of sub-components constituting the character), and spatial configuration information (relative positioning and sizes of sub-components within the character). In order to handle the imprecision of handwritten data that affects both shapes and positioning, we formalize these two types of information under the formalism of fuzzy sets [6]. Characters are described by two levels of fuzzy inference rules explicitly reflecting their natural structural properties, and combining two types of fuzzy conditions about elementary shapes and spatial configuration. A Fuzzy Inference System (FIS) aggregates the rules and provides a global decision-making process in a completely homogeneous way, by taking into account the two types of information.

For model training, we focused on minimizing the manual task to be completed by human expert, by automatizing the learning of spatial relationships. Besides, we designed a process for automatic incremental labeling and learning that considerably reduces the manual task needed for training.

Experiments show that the method is able to recognize on-line handwritten Chinese characters written in a fluent style, and confirm the interest of structural description: shape and positioning models that are shared between trained characters do generalize very well to the description of non-trained characters.

This paper is organized as follows. In the first part we introduce our explicit structural FIS model, and its exploitation. Automatic learning process is described in part 3, before we present experimental results and conclude.

2 Explicit modeling of characters with Fuzzy Inference System

Chinese characters are much more complex than Latin alphabet letters : their writing often involve a large num-

ber of strokes (1-30) organized in a 2-dimensional structure. However, these complex characters can be decomposed with intermediate components called *radicals*, that consists of several strokes each. While there is a huge number of character classes (6,763 in the GB2312 specification), radicals are only in a limited number of 214 (as they are defined in the reference KangXi dictionary).

2.1 Radical description

Theoretically, Chinese radicals can be described as a canonical combination of well-separated *fundamental strokes* drawn in a specific order. However, when it comes to handwritten characters, strong variations can occur, including strokes connections, distortions, and no specific stroke order is respected. Strong shape distortions make segmentation of handwritten strokes into corresponding theoretical fundamental strokes a difficult and error-prone task. To overcome this problem, we propose to categorize directly handwritten strokes as the primitives of our modeling, as we presented in [3]. Indeed, these primitives are very stable among different writers, allowing us to build an exhaustive dictionary \mathcal{P} presented in table 1. From this definition principle we can proceed recognition of handwritten character without the need of a segmentation step.

We define 31 elementary primitives : primitives numbered p_1 to p_{20} are Chinese fundamental strokes, whereas primitives p_{21} to p_{31} result from frequent merging or warping of fundamental strokes that occur in handwriting. These primitives are pure morphological elements, defined out of any context or scale consideration.

p_1	p_2	p_3	p_4
p_5	p_6	p_7	p_8
p_9	p_{10}	p_{11}	p_{12}
p_{13}	p_{14}	p_{15}	p_{16}
p_{17}	p_{18}	p_{19}	p_{20}
p_{21}	p_{22}	p_{23}	p_{24}
p_{25}	p_{26}	p_{27}	p_{28}
p_{29}	p_{30}	p_{31}	

Table 1. Primitive dictionary \mathcal{P}

Based on this dictionary, we are able to describe different admitted variations for writing the radicals. However, the

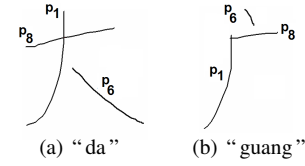


Figure 1. Primitive decomposition of radicals “da” and “guang”

list of primitives constituting a radical is not a sufficient information: a *spatial configuration* is also needed, that specifies how these primitives are organized within the radical. Our definition of spatial configuration comprises relative positioning and relative dimensions of primitive strokes.

As an illustration, figure 1 shows two different radicals (“da” and “guang”) that are made of the same three primitives (p_1 , p_6 and p_8) but having different spatial configuration. Unlike the primitives shape models that are common to each radical, spatial configurations are specific to radicals, and we do not categorize them in a model dictionary.

It follows that a radical r_A ($r_A \in \mathcal{R}$, the set of all radicals) can be expressed as:

$$r_A : \underbrace{p_{A_1} p_{A_2} \dots p_{A_n}}_{shapes} + configuration_A \quad (1)$$

with $p_{A_i} \in \mathcal{P}, \forall A_i, i = 1..n$.

2.2 Character description

The second level of Chinese character structure expresses a complex character from its radicals. This description is natural and immediate, as it is traditionally used to index characters in Chinese dictionaries.

Likewise, the list of radicals constituting a character alone is not sufficient, and an inter-radicals spatial configuration is needed for a complete character description. In this case, however, we can draw a list of generic positioning models of radicals within characters. This positioning models are the 9 well-known character patterns presented by figure 2 [5]. The structure of any composed character can be described by recursively applying these positioning models, resulting in a binary tree as the one of character “qi” presented by figure 3.

A character c_A ($c_A \in \mathcal{C}$, the set of all characters) can thus be described by a unique expression such as (2).

$$c_A : \underbrace{r_{A_1} r_{A_2} \dots r_{A_n}}_{shapes} + configuration_A \quad (2)$$

with $r_{A_i} \in \mathcal{R}, \forall A_i, i = 1..n$.

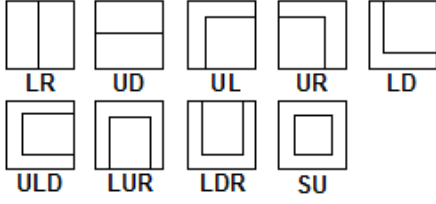


Figure 2. The 9 composed character patterns

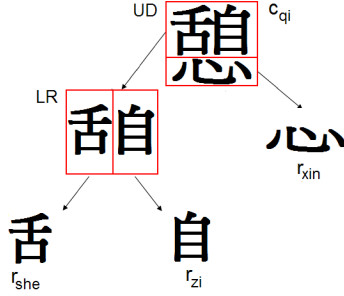


Figure 3. Radical decomposition of “qi”

2.3 Modeling with Fuzzy Inference System

Based on the two structural levels presented before, we model characters with two levels of fuzzy inference rules, gathered in a FIS. Being based on fuzzy sets theory, a FIS offers an efficient way for modeling reasoning with imprecise data, such as handwritten data. It also constitutes an explicit model that provides interpretable results, and the decision-making by inference mechanism is free of early-stage decisions, allowing a global fusion of information.

In our model, each handwritten variation of a radical r_A is formalized by a level-1 fuzzy inference rule of the form:

$$\underbrace{(F_{A_1}^p \wedge \dots \wedge F_{A_n}^p)}_{shapes} \wedge \underbrace{(Q_{A_1}^p \wedge \dots \wedge Q_{A_{n-1}}^p)}_{configuration} \Rightarrow r_A \quad (3)$$

where $F_{A_i}^p$ and $Q_{A_i}^p$ are fuzzy conditions about the shapes of primitives and their expected spatial configuration within the radical, respectively. A radical composed of n primitives is described by n shape conditions and $n - 1$ spatial configuration conditions, because the configurations are described relatively to one reference primitive. The two types of fuzzy conditions are modeled by fuzzy sets and their degree of truth for a given input is calculated as an membership degree to the corresponding set (see section 3).

Similarly, a level-2 fuzzy rule describes a character c_A :

$$\underbrace{(F_{A_1}^r \wedge \dots \wedge F_{A_n}^r)}_{shapes} \wedge \underbrace{(Q_{A_1}^r \wedge \dots \wedge Q_{A_{n-1}}^r)}_{configuration} \Rightarrow c_A \quad (4)$$

where $F_{A_i}^r$ and $Q_{A_i}^r$ are fuzzy conditions about shapes of expected radicals and inter-radicals spatial configuration within the character c_A , respectively. For a character containing n radicals, the premise has n shape fuzzy conditions and $n - 1$ spatial configuration conditions, each of them modeling the relative positioning of two sub-parts of the character, just like one node of the binary tree presented in figure 3. We chose to use the same modeling for inter-radicals spatial configuration as we use for inner-radical spatial configuration, for ensuring homogeneity in decision making process (see next paragraph).

2.4 Decision by homogeneous inference

The decision making relies on a inference process that we prove to be completely homogeneous : given an input sequence of strokes to be recognized, all the models considered for recognition are perfectly comparable. This is ensured thanks to a rule selection based on our primitives definition principle, and the choice of homogeneous modeling of inner-radical and inter-radicals spatial configurations.

During the recognition step, the two levels of fuzzy rules are dynamically merged into a single level of *extended rules* that completely models characters from the basic primitives. These rules are generated from level-2 inference rules where each “ $F_{A_i}^r$ ” fuzzy condition is replaced by the premises of a level-1 inference rule describing the radical r_{A_i} . By combining all the level-1 rules describing each radical r_{A_i} , we obtain a set of all the admitted writing variations for the full character c_A .

Let us consider a handwritten sequence of n strokes $T = t_1 \dots t_n$, to be recognized as a character. Recognition is made by inference process over extended rules, according to the three following steps.

1. Rule selection. Based on primitives definition principle (see section 2.1), we assume that the character to be recognized is made of n primitives (there is no segmentation of input strokes). Thus we filter out inconsistent extended rules, keeping only the ones that model characters with n primitives shapes models F^p . We can also show that the number of spatial configuration models (Q^p and Q^r) in these rules is always equal to $n - 1$, no matter the number of radicals in the considered character. Consequently, all the activated rules consists of n shape and $n - 1$ spatial configuration conditions.
2. Rule inference. For each activated rule, we compute adequacies of input data with shape and configuration

models. By using a *product* T-norm operator for both conjunction and inference, truth degree of conclusion for character c_K is computed as:

$$\mu_{K_i}^c(T) = \underbrace{\left(\prod_{i=1}^n \mu_{K_i}^p(t_{\sigma_o(i)}^p) \right)}_{\text{shapes}} \times \underbrace{\left(\prod_{i=1}^{n-1} \nu_{K_i}^q(t_{\sigma_o(i)}^q) \right)}_{\text{configuration}}^\alpha \quad (5)$$

where $\mu_{K_i}^p$ and $\nu_{K_i}^q$ are membership functions of shape fuzzy sets ($F_{K_i}^p$) and spatial configuration fuzzy sets ($Q_{K_i}^p$ and $Q_{K_i}^r$) respectively, and α is a coefficient allowing to balance the weight of spatial configuration scores with respect to shape scores in decision making. σ_o denotes the best permutation of input strokes for this rule, which is found by Branch-and-Bound algorithm under the temporal-adjacency constraint for the strokes belonging to the same radical.

3. Aggregation. The final output is decided by *maximum* aggregation operator: recognized character is given by the highest conclusion score.

3 Learning

Given the large number of character classes, it is crucial to make the training of models as automated as possible. We achieved to limit the required manual task to very a reasonable level : an expert only has to describe the radicals as lists of primitives (in a simple composition rule such as (1)) without any positioning information, and to describe characters from the list of their radicals and the structural patterns involved (such as in (2)).

From this manual basis, fuzzy sets that model both shapes and spatial configurations can be automatically trained. However, training models require to have a sufficient amount of labeled samples for each component of character: labeled samples of primitives are needed for shape models training, and labeled samples of radicals are needed for inner-radical and inter-radicals spatial configuration models training. The task of manually labeling these components would itself be very tedious and time consuming. For that reason, we designed an automatic process for labeling characters structural components.

3.1 Primitive shape models learning

Primitives presented in table 1 are modeled as fuzzy prototypes defined in a 20-dimensional space, using qualitative geometrical features. We cluster training data into fuzzy prototypes by unsupervised fuzzy C-Means run independently on each class of primitives, leading to intrinsic modeling of the class. Prototypes are defined by estimate mean

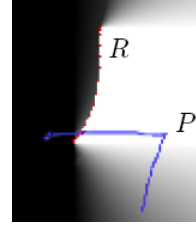
vector μ and covariance matrix σ , and their membership characteristic function P :

$$P(x) = \frac{1}{1 + d_\sigma(x, \mu)} \quad (6)$$

where d_σ denotes the Mahalanobis distance. A primitive model consists of several prototypes and a shape adequacy degree $F_j^p(x)$ is computed as the maximum activation score of the prototypes modeling primitive j for input x .

3.2 Spatial configuration models learning

We use a similar method for evaluating the relative positioning and dimensions of primitives within a radical or radicals within a character. This method, that describes fuzzy relative positioning by extending the mathematical morphology concepts to fuzzy framework, was first introduced by Bloch [1], and then adapted to the context of on-line handwritten objects in our previous work [2]. It consists in evaluating the fuzzy adequacy degree of an object P with respect to a reference R considering a given direction (see an example in figure 4).



The background grey levels (from black to white) express the fuzzy degree (from 0 to 1) of “to be on the right of R ” for points of the plan. Global degree for primitive P is computed by averaging degrees of its points.

Figure 4. Relative positioning of P w.r.t. R

Spatial configuration of elements are represented in a 6-dimensional space using 4 of these measures of directional positioning (for the 4 main directions) and 2 measures of relative dimensions. Fuzzy sets are defined by their mean vector and covariance matrix in this 6-features space, and have the same membership function than the shape fuzzy prototypes (6). Their parameters are estimated from training samples. Since inner-radical configuration models Q^p are specific to radicals, they need to be trained from samples of each radical, whereas inter-radicals models Q^r can be trained with samples of any characters sub-parts that respect a given pattern.

3.3 Automatic labeling-learning process

The principle of automatic labeling-learning process is to make models and data cooperate in order to incrementally build accurate models while automatically labeling data. A rejection threshold allows to control labeling confidence, that can be gradually lowered as the model confidence is

increasing with successive loops. This process can be described as a three-steps mechanism, as follows.

1. Initialization. Given simple composition rules written by an expert for the two structural levels, and primitive shape models learnt beforehand from a set of manually labeled training samples, the initialization consists in identifying and labeling radicals and primitives in the training dataset. This is done by applying the recognition process on ground truth labeled character samples, by only taking into account the shape information in decision-making (since the spatial configuration models have not been learnt at this step).
2. Learning. Shape models, inner-radical and inter-radicals configuration models are learnt from the available labeled primitives and radicals samples.
3. Labeling. Thanks to updated models, more primitives and radicals samples can be labeled by applying the usual recognition process.

Steps 2 and 3 can be run in loop in order to progressively improve the quality of models as long as more and more labeled samples are available.

4 Experimentations

We evaluated the performance of our approach on the CASIA¹ database, consisting of 6,763 classes of characters written each by 60 writers, in a regular/fluent style.

We modeled 36 radicals with 57 level-1 fuzzy rules, and 125 character classes built from these radicals with one level-2 fuzzy rule each. Training of shape and configuration models was processed with samples collected from 40 first writers, using different amounts of training samples (depending on the number of labeled samples for each model). Training data was labeled with the automatic labeling-learning process described in section 3.3, which made the task very fast and efficient. We were able to label about 88.4% of the radicals from a dataset of 6,200 characters, with a high confidence level (a visual control confirmed that labeling error was very low).

The weighting α (from equation (5)) was determined experimentally, its value being set to $\alpha = 3.5$, showing that it is beneficial to trust more positioning scores than shapes scores for a better recognition. Recognition test was conducted with data from 20 different writers, and overall recognition rate obtained was 92.6%.

In order to evaluate the generalization capacity of our model, we described 30 more classes of characters (made

from same radicals as the 125 previous ones), by just writing the corresponding decomposing rules. A similar correct recognition rate was obtained (93.1%) when recognizing these characters that had not participated in the training process. This confirms the relevance of our approach that shares common positioning and radicals models between many characters, allowing to model and recognize correctly some characters that are not in the training dataset.

5 Conclusion

We presented a new method for recognition of handwritten Chinese characters, based on the combination of shapes modeling and spatial configuration modeling. Natural structure of characters is explicitly modeled by a fuzzy inference system that processes global decision for recognition through homogeneous inference. The automatic process we designed efficiently reduces the human implication in structural model learning. With an original definition of primitives and the flexibility of fuzzy models, this method is able to recognize non-regularly written characters, as that was confirmed by experimental results. Finally, the structural models are shown to generalize well on non-trained characters, which is encouraging when considering our future goal of modeling the whole set of Chinese characters.

6 Acknowledgments

The authors would like to thank Professor G. Lorette for his precious advice as well as Professor C.-L. Liu for providing us the CASIA database.

References

- [1] I. Bloch. Fuzzy relative position between objects in image processing: a morphological approach. *IEEE Tr. on PAMI*, 21(7):657–664, 1999.
- [2] F. Bouteruche, S. Mace, and E. Anquetil. Fuzzy Relative Positioning for On-Line Handwritten Stroke Analysis. In *Proceedings of the 10th International Workshop on Frontiers in Handwriting Recognition (IWFHR'06)*, 2006.
- [3] A. Delaye, S. Macé, and E. Anquetil. Hybrid statistical-structural on-line chinese character recognition with fuzzy inference system. In *Proceedings of the 19th International Conference on Pattern Recognition (ICPR'08)*, 2008.
- [4] C.-L. Liu, S. Jaeger, and M. Nakagawa. Online recognition of chinese characters: the state-of-the-art. *IEEE Tr. on PAMI*, 26(2):198–213, 2004.
- [5] A.-B. Wang and K.-C. Fan. Optical recognition of handwritten Chinese characters by hierarchical radical matching method. *Pattern Recognition*, 34(1):15–35, 2001.
- [6] L. Zadeh. Fuzzy sets as a basis for a theory of possibility. *International Journal on Fuzzy Sets and Systems*, 1(1):3–28, 1978.

¹Chinese Academy of Science, Institute of Automation