

ON-LINE HANDWRITING RECOGNITION BASED ON K-NEIGHBOR LOCAL FEATURE

K 次近傍局所特徴量によるオンライン文字認識

Wenjie CAI* Seichi UCHIDA** Hiroaki SAKOE**
{*Graduate School, **Faculty} of Info. Science and Elec. Eng., Kyushu Univ.

1. Introduction

DP matching, a dynamic programming based time alignment algorithm, is widely used in online character recognition systems as a reliable tool for comparing the strokes. Strokes are represented as sequences of local features which convey geometric information of primitive points on the strokes. As a measure of the difference between two local features of two strokes, a local distance is employed between them. Conventionally, the local distance is evaluated just by using of the features of individual point [1]. The problem was that the local shape information around the points hadn't been used effectively. In this paper, we propose a new definition of local feature which incorporates the local shape information. Experimental results showed the effectiveness of the proposed local feature.

2. The proposed local feature and distance

We define a new local feature of the i th point of an input stroke as

$$a_i = (xy_i \mid \delta xy_i, \dots, \delta xy_{i-k+1}, \dots, \delta xy_{i-K+1}) \quad (1)$$

where $xy_i = (x_i, y_i)$ is the coordinate of the i th point, and δxy_{i-k+1} is the moving direction of the k th point back from xy_i (Fig.1). In a similar manner, we define the local feature of the j th point of a reference stroke as

$$b_j = (uv_j \mid \delta uv_j, \dots, \delta uv_{j-k+1}, \dots, \delta uv_{j-K+1}) \quad (2)$$

As a measure of the difference between the input feature a_i and the reference feature b_j , we define the local distance d as

$$d(i, j) = \|a_i - b_j\| \quad (3)$$

$$= w \|xy_i - uv_j\| + (1-w) \left(\frac{1}{K} \sum_{k=1}^K \|\delta xy_{i-k+1} - \delta uv_{j-k+1}\| \right)$$

where w is a weighting coefficient. When $K=1$, Equations (1), (2) and (3) are reduced to the conventional definitions of local features and local distance, respectively.

Unlike the conventional definition, the newly defined local feature explicitly includes the local shape information around the individual penpoint. We call the proposed local feature "K-Neighbor Local Feature (KNLF)".

3. Experiment and Results

In order to test the effectiveness of KNLF, we conducted recognition experiment. The proposed KNLF was implemented in a cube search-based online kanji recognition system [1]. From an online Kyouiku kanji (882 classes) handwritten character

database written by 30 persons, abandoning the connected stroke characters, we selected total 18041 sample characters as the test patterns.

Fig.2 shows the experimental results. When $K=4$, KNLF had achieved the best recognition rate of 99.49% (17949/18041). Compared with the result of $K=1$ (conventional definition) of 99.28% (17911/18041), the recognition rate is improved by 0.21%.

4. Conclusion

We proposed a novel definition of local feature for online character recognition and demonstrated its effectiveness by the experiment. Instead of using direction of individual point, we used feature of local shape information around the individual point to evaluate the distance. In conclusion, the proposed local feature significantly improved the DP matching based recognition system.

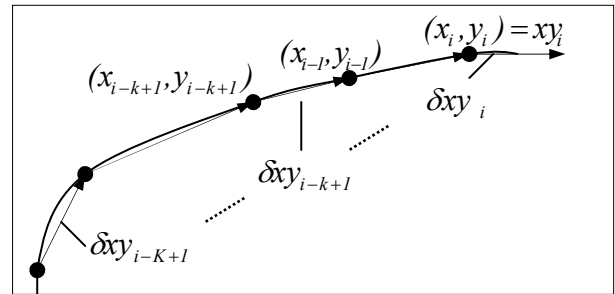


Fig. 1. Definition of K-neighbor local feature

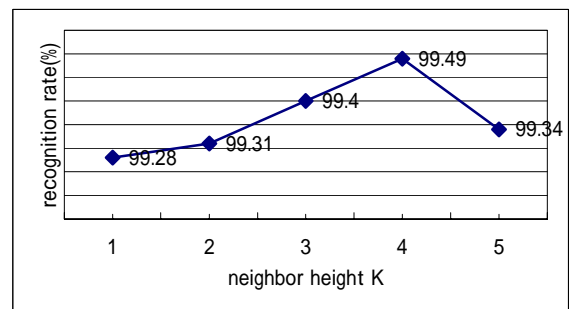


Fig. 2. Experimental result

References

[1]Hiroaki SAKOE, Jung-Pil SHIN, "A Stroke Order Search Algorithm for Online Character Recognition," Research Reports on Information Science and Electrical Engineering of Kyushu University, Vol.2, No.1, March 1997.