Identifying Ability of a Recognition Method Based on the Field of Induction

Michihiro NAGAISHI

ATR Auditory and Visual Perception Research Laboratories
2-2 Hikaridai, Seika-cho, Soraku-gun, Kyoto 619-02, Japan

Abstract

A new recognition method based on the field of induction is proposed. This method is based on observations of the field as well as on psychological and physiological knowledge. To adapt this method for use with handwritten characters, it is necessary to determine its ability to identify these characters. This paper clarifies that this method has a good ability to identify handwritten characters.

1: Introduction

Much work has been focused on improving the present recognition methods for handwritten characters. However, it is difficult to recognize unconstrained handwritten characters by using the present recognition methods, because unconstrained handwritten characters vary greatly in style and shape [5]. Therefore, a new recognition method was considered based on observations of the field of induction. Actually, the field around figures has been reported as a physiological and psychological phenomenon (Figure 1). There has been past physiological research on the field of induction [3], as well as past psychological research. Such research includes a theory on the field of induction by Yokose [8]. His theory shows that field strength is calculated in terms of electromagnetism. He also suggested that the field was strongly related to human character recognition. Therefore, the recognition method that was proposed focused on Yokose's theoretical property. To develop this method for recognition of handwritten characters, it is necessary to clarify its identifying ability. This paper first illustrates a recognition method based on the theory of the field of induction. It then presents analysis results of the method's ability to identify *kana* (the syllabic Japanese) and handwritten numbers. Finally, advantages and the problems are discussed.

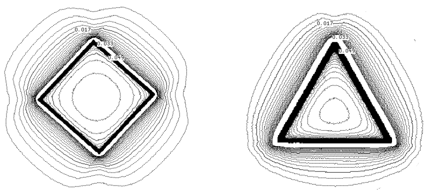


Figure 1 Examples of the filed of induction.

2: Recognition Based on the Induction Field

2.1: Calculation Method of the Induction Field

The definition of the field of induction is formulated as follows [6]. We assume that +1 electric charges are distributed uniformly on curves $f_1(s)$, $f_2(s)$, .., $f_n(s)$ in Figure 2. If $R(f_i(s))$ is the distance from point P to any point on curve i, strength M_p of the field of induction created by the curves on point P is

$$M_p = \sum_i^n \int \frac{ds}{R(f_i(s))}. \tag{1}$$

The field created on a digital image is extended as follows [6]. It is assumed that the field on any point P is created by a curve $f(s)$ consisting of n dots in Figure 3. We can assume that each dot on curve $f(s)$ has +1 electric charges. If r_i is the distance from point P to dot i on curve $f(s)$, strength M_p of the field on point P is

$$M_p = \frac{1}{n} \sum_i^n \frac{1}{r_i}. \tag{2}$$

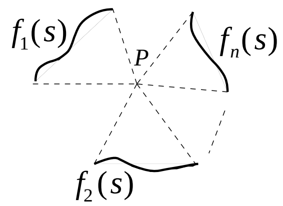


Figure 2 A Induction Field created by curves.

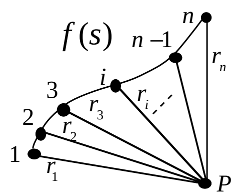


Figure 3 The Induction Field of an array of dots.

2.2: Relationship between the Field of Induction and Character Similarity

Most previous recognition methods using the 'field' focused on the effect of the local field [1][4]. This paper focuses on the similarity of the field distribution. From the definition of the induction field, it is obvious that the induction field is neither a local field nor a gradation with a Gaussian filter, but has a long range interaction. Furthermore, psychological experimental results suggest that the similarity of field distributions is very important for character recognition among different categories of characters [7][8]: If two characters belong to the same category, their field distributions are similar. It is possible to deform and match one field to the other field and still keep the topology of the character.

2.3: Quantifying Differences in Field

From the similarity of field distributions for character recognition, it is considered that the induction field has dynamic action like an elastic body. A thin membrane, like rubber extended on a plane, deforms and its shape is changed when it is pushed up to a certain height. If we consider the height of the membrane to be an electrical potential, the shape of a deformed membrane corresponds to an electrical field distribution. Based upon its formulation, the induction field is considered to be an electrical field. The distribution of the induction field can therefore, be considered to be the shape of a thin membrane. Since a deformed field is restored to its original shape after it is deformed to match another field, strain is generated on the induction field. The greater the field deforms, the greater the strain. By evaluating the degree of strain, we can determine the quantitative difference among fields. When a field with an unknown pattern is deformed to match the field of a standard pattern (such as A and B in Figure 4), the unknown pattern is considered to correspond to the standard pattern which has the smallest strain between it and the unknown pattern.

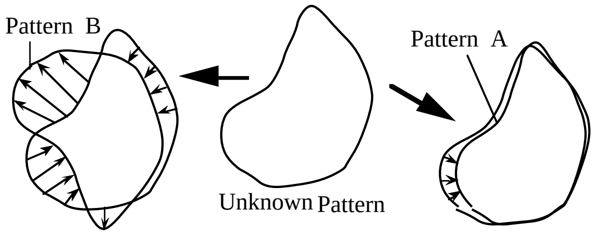


Figure 4 Recognition based on the induction Fields.

2.4: Elasticity Energy

To quantify the difference in fields, we measure the elasticity energy when one contour is deformed to match another contour. The elasticity energy U generated when a thin membrane is pushed up from the equilibrium state to a certain height, equals the summation of energy U_1 (generated by the strain of the membrane) and U_2 (generated by the strain of the contour). We do not assume any external force. When tension τ is uniform around the membrane and u_x, u_y are minute displacements of the membrane, the energy U_1 is

$$U_1 = \alpha \iint (u_x{}^2 + u_y{}^2)\, dxdy \quad (\alpha = \frac{\tau}{2}). \tag{3}$$

When u is the displacement of contour s and ρ is the density of the membrane, the energy U_2 is

$$U_2 = \frac{\beta}{2}\int u_{ss}{}^2\, ds \quad (\beta = \frac{\tau^2}{\rho^2}C). \tag{4}$$

2.5: Recognition using Elasticity Energy

Since the induction field is a scalar field on a two-dimensional plane, its distribution can be considered to be a three-dimensional figure. This figure is considered to be a gathering of many potential planes, so it is possible to sum up all contour strains on each equational potential plane as elasticity energy using Equations (3) and (4).

Fields for various character categories are entered into the dictionary. The recognition method presented in this paper is based on the total strain elasticity energy in the field. When the field of an unknown pattern is deformed to match the field of dictionary pattern i on the equational potential plane with potential value p , $e_i(p)$ is the elasticity energy and total elasticity energy E_i is

$$E_i = \int e_i(p)dp. \tag{5}$$

When the field of an unknown pattern is more similar to the field of dictionary pattern i, E_i decreases and recognition becomes equal to finding the category i which satisfies the following equation for all dictionary :

$$\min E_i \ . \tag{6}$$

3: Identification Ability Analysis

3.1: Displacement Calculation Method

All characters are normalized by centering each character in a circumscribing rectangular frame (64 x 64 dots); after removing noise, each character is enlarged or reduced to fill the frame. The induction field is calculated using Equation (2) . The fields were smoothed by using a median filter and an adaptive gradient filter. The borders extracted from the smoothed field are then smoothed by a Gauss function convolution. As an example is shown in Figure 5, we calculated the displace-ments that all points on border 'ìy' move to normal points on border 'ém' on equational potential plane. This is determined from the tangent line determined by using Hough tarnsformation. The elasticity energy between unknown patterns and dictionary patterns is calculated on each equational potential plane (from 0.03 to 0.39 in 0.01 steps). Characters are recognized using Equation (6).

To clarify the identifying property of the induction field, similarity on LDCD (Local direction contributivity density) feature [2] is used as a reference of evaluation measure for characters. It is reported that this feature has good ability to recognize handwritten characters. The more a pattern is similar another pattern, the closer the similarity approaches 1.

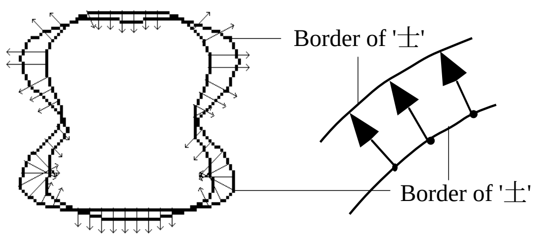


Figure 5 Example of the displacement.

Experiment (1): This experiment attempted to clarify the identifying property for artificial patterns shown in Figure 6. The induction field for these patterns were themselves used as unknown patterns and as dictionary patterns to calculate elasticity energy and similarity on LDCD feature.

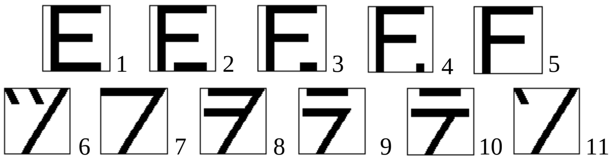


Figure 6 Artificial sample patterns.

Experiment (2): This experiment attempted to clarify the identifying property for forty-six types of *kana* characters (the syllabic Japanese). These character images were taken from a dictionary of standard writing style for Japanese characters (Mokuji-sha, Japan). They were converted by an image scanner (Ricoh, IS-50, 400 DPI) into black and white images. The fields of these patterns were used as both unknown patterns and as dictionary patterns to calculate elasticity energy and the similarity on LDCD feature.

Experiment (3): This experiment attempted to clarify the identifying property for handwritten characters. As unknown patterns, one hundred good samples of numbers '1' and '3' were selected from ETL1 (the database of the Electrotechnical Laboratory in Tukuba, Japan). They were converted to black and white images. As dictionary patterns, ten numerical patterns (0-9), were also selected from ETL1. The elasticity energy and similarity on LDCD feature were calculated for the one hundred samples of number '1' and '3'.

4: Experimental Results

Experiment (1): Figure 7 shows five of the calculated fields in Figure 6 .

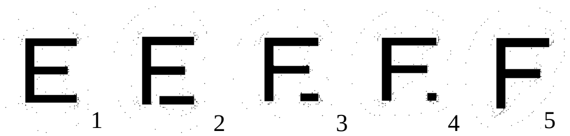


Figure 7 The field of induction for pattern 1 - 5.

Figure 8 shows a comparison of LDCD similarity and elasticity energy for the patterns shown in Figure 6. The dictionary pattern in Figure 8 is pattern 1 and 6. Pattern 2 is more similar to pattern 1 than to pattern 3. Pattern 4 is similar to pattern 5, but not to pattern 1. Therefore, most people read pattern 2 as 'E' and pattern 4 as 'F', but they read pattern 3 as neither 'E' nor 'F' [7]. Even though 'E' and 'F' are different characters, Figure 8 (a) indicates that the range of similarity from pattern 1 to 5 is limited compared to the range of elasticity energy. In particular, the energy in pattern 5 is about 100 times larger than that in pattern 1. Since patterns 6 to 11 are similar, many Japanese may err in reading them. However, in Figure 8 (b) the range of similarity from 6 to 11 is very broad. The elasticity energy has a similar range for patterns 1 to 5, but it is sufficient to distinguish these patterns.

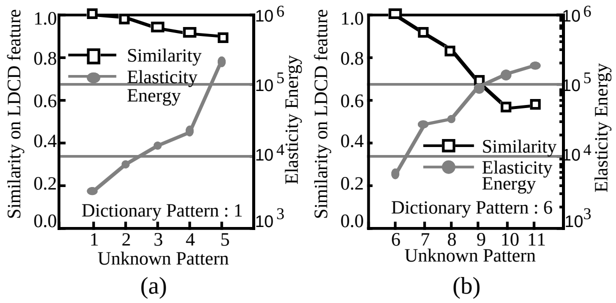


Figure 8 Comparison of similarity on LDCD feature and elasticity energy: Energy is the summation of strain on each equational potential plane from 0.03 to 0.39 in 0.01 steps. (a) 1 versus 1 (itself), 2, 3, 4 and 5 (b) 6 versus 6 (itself), 7, 8, 9, 10 and 11

Experiment (2): Figure 9 shows examples of *kana* characters and a comparison of LDCD similarity and elasticity energy for these samples. These dictionary characters are similar to the unknown characters 'あ' in the order 'お,' 'ぬ', 'め', 'の', 'わ', 'ろ'. There is little similarity between the unknown characters 'あ' and the other characters. Corresponding to similarity in terms of the physical aspects of characters, the elasticity energy for the patterns increases from 'あ' to 'ろ'. By using the elasticity energy, 26 *kana* characters can be distinguished from each other. The energy of the second candidate character was about 100 times larger than the first. When the second candidate was similar to the first candidate, such as 'あ' and 'お,' the energy of the second candidate was about 10 or less times larger than the first.

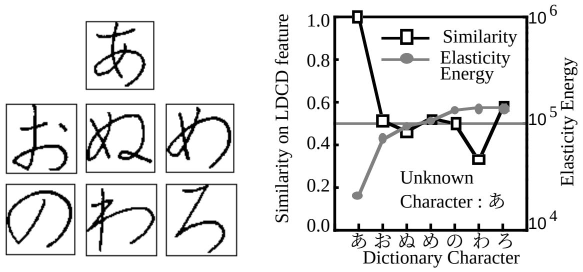
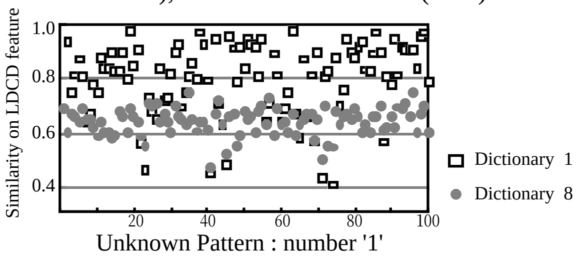


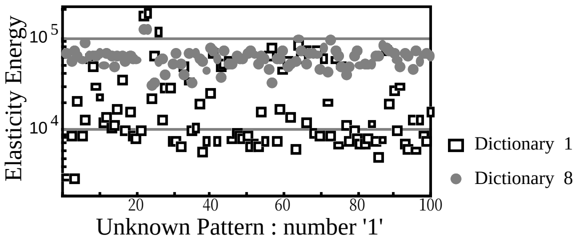
Figure 9 Examples of Kana characters and Comparison of Similarity on LDCD feature and elasticity energy.

Experiment (3): From an evaluation using elasticity energy, it was found that the probability of finding '8' as the second candidate is high for all numbers, so similarity on LDCD feature and elasticity energy were compared for dictionary patterns '1' and '8.' Figure 10 (a) shows a comparison of similarity on LDCD feature for 100 '1' patterns from ETL1 with dictionary patterns '1' and '8.' Figure 10 (b) shows a comparison of elasticity energy. Figure 10 (a) indicates that the border which distinguishes '1' and '8' based on the similarity is unclear, because '1' is more common than '8' (the mean of similarity for '1' is 0.804 and for '8' 0.641).

On the other hand, Figure 10 (b) indicates that the elasticity energy distribution of '1' and '8' has clear separation and that the difference in elasticity energy is large (the mean of energy for dictionary '1' is 2.62×10^4 and for '8' 7.26×10^4). The same distribution tendency for '1' and '8' was also found for 100 patterns of '3' (mean energy for dictionary '3' is 2.92×10^4 and that for '8' is 3.28×10^4). The recognition rate for '1' is 80% (94% of which includes the second candidate), and that for '3' is 87% (93%).



(a)



(b)

Figure 10 Comparison of similarity on LDCD feature and elasticity energy for '1' : (a) Similarity of characters between '1' (ETL1 100 patterns) and a dictionary '1' and '8' . (b) Elasticity energy of the field values are calculated between '1' (ETL1) and the fields '1' and '8'.

5: Discussion

Experiments (1) and (2) clearly show that a recognition method based on the theory of the induction field, that uses elasticity energy as a measure to evaluate characters, has better ability to identify similar characters than the method involving LDCD feature, because the energy difference is larger than the similarity difference, especially for similar patterns. When characters are similar, the energy difference is small; it is still large enough however, to identify characters. When characters are quite different, on the other hand, the energy difference is large. We can easily understand this property from the field distribution (such as the examples shown in Figure 7). When we compare the difference in fields between two patterns, we can learn where and how the difference or character variation occurs visually from the displacement (such as the example shown in Figure 5). Consequently, an evaluation using elasticity energy may possibly be closer to our sense of recognizing the similarity of characters in terms of physical aspects.

Experiment (3) suggests that a recognition method based on the field of induction has a good ability for also identifying handwritten characters. This is because by using elasticity energy, the distribution of '1' is separated from the distribution of '8', which is the most similar category to '1' (Figure 10).

The displacement and elasticity energy for the recognition of error patterns indicated that the calculated displacement between dictionaries was not sufficient, leading to a large strain. This paper defined the elasticity energy by considering the displacement as minute. The displacement was calculated only in the normal direction on a border. These conditions, however, may not hold for recognition error patterns. The method for calculating displacements must be able to find suitable displacements. Nonetheless, the method presented in this paper sufficiently identifies printed-style characters.

6: Conclusion

The identifying ability of a recognition method based on the theory of field of induction was analyzed. The analysis clarified that this recognition method has sufficient ability to identify printed-style characters compared with methods involving LDCD feature. This method was also found to have a good ability for identifying handwritten characters. From elasticity energy difference and visualization of character variation, an evaluation using elasticity energy may possibly be closer to our sense of recognizing the similarity of characters. However, the method of calculating displacements is not suitable for handwritten characters at present. It remains to consider improving the method calculating suitable displacement.

Acknowledgments

I am grateful to Eiji Yodogawa, President, and Keiichi Ueno, Head of the Visual Perception Department, ATR Auditory and Visual Perception Research Laboratories, for providing the opportunity to do this research. I would also like to thank Hiromi Kakitani, Kaoru Mishima, and Sigeru Mukaida for their cooperation in preparing and analyzing the character data.

References

[1] H., Kazmierzak,"The potential field as an aid to character recognition", Proc., *International Conference on Information Processing*, UNESCO, Paris, pp. 244 - 247, 1959.

[2] Hagita, N., Naito, S., and Masuda, I., "Recognition of Hand printed Chinese Characters by Global and Local Direction Contributivity Density-Feature", *Trans., IEICE*, J66-D, 6, pp. 722 - 729, 1983.

[3] Motokawa, K., "Field of retinal induction and optical illusion", *Neurophysiology*, **13**, pp. 413 - 426, 1950.

[4] Mori, T., Mori, S., and Yamamoto, K., "Field Effect Method for Feature Extraction from Patterns", *Trans., IEICE*, 57 - D, 5, pp. 812 - 819, 1974.

[5] Nagaishi, M., "Analysis of Written Styles and Shape Variation in Unconstrained Handwritten Characters", *IEICE FALL CONFERENCE*, D - 234, pp. 236, 1991.

[6] Nagaishi, M., "A Proposal of Character Recognition using Theory of Field of Induction on the Retina", *TECHNICAL REPORT OF IEICE*, PRU 92 - 46, pp. 7 - 14, 1992.

[7] Yokose, Z., "Pattern recognition and reading of characters by machine", *Fap. F. Ergonomics*, 2, 3, pp. 10 - 16, 1966.

[8] Yokose, Z., "A Study on Character-Patterns Based Upon the Theory of Psychological Potential Field", *Japanese Psychological Research*, 12, 1, pp. 18 - 25, 1970.