# Character Image Reconstruction from a Feature Space Using Shape Morphing and Genetic Algorithms

Chihiro Iga and Toru Wakahara
*Faculty of Computer and Information Sciences*
*Hosei University*
*3-7-2 Kajino-cho, Koganei-shi, Tokyo, 184-8584 Japan*
*wakahara@k.hosei.ac.jp*

## Abstract

*This paper proposes a powerful method that realizes image reconstruction from a feature space in optical character recognition. Due to the invisibility of a high-dimensional feature space, it is difficult to fully understand advantages and disadvantages of the given feature space and search for more robust features. The proposed method consists of two parts. The first part is 2D shape morphing based on a mesh model via bilinear transformation. The second part is use of genetic algorithms for determining optimal morphing parameters. Given an arbitrary feature vector in a feature space the proposed method deforms each category's template to yield the maximal fitness value against the given feature vector and the deformed template thus obtained is considered as a reconstructed image from a feature space. In experiments we use the public handwritten numeral database IPTP CDROM1B and a gradient feature space. We first demonstrate a high matching ability of the proposed mesh model. Then, we show promising experimental results of image reconstruction from a feature space and discuss how to use this technique to improve recognition performance.*

## 1. Introduction

Many handwriting recognition algorithms have been proposed [1], [2], and some of them have achieved high recognition accuracy, in particular, in handwritten numeral recognition [3]. Given a feature space, we can make full use of statistical or probabilistic pattern recognition techniques, including neural networks, hidden Markov models, and support vector machines [4]. However, we can say that there is no definite guideline for generating a "good" feature space itself. Therefore, we have to search for more robust features by examining and analyzing misclassified samples in a rather heuristic manner. This problem arises partly from the invisibility of a high-dimensional feature space. From this viewpoint, image reconstruction from a feature vector in a feature space would be helpful in understanding advantages and disadvantages of the given feature space.

Sakano et al. [5] proposed an interesting technique that aimed at reconstructing a character image from a feature vector by using genetic algorithms (GA) [6] as applied to 2D crossover of randomly selected character images. However, 2D crossover of character images generated jaggy, discontinuous patterns with background noise. Also, mutation in GA was not used because a flip-flop of each binary pixel only increased noise in a reconstructed image. In this sense, this technique was not completely successful in visualizing a high-dimensional feature vector although it was highly challenging and promising.

In this paper, we propose to use a combination of 2D shape morphing and genetic algorithms. 2D shape morphing is to generate a variety of continuous character patterns with sufficient freedom of deformation as applied to a template. Here, we adopt a mesh model via bilinear transformation. Also, in GA application we implement full operations of selection, crossover, and mutation to search for optimal morphing parameters. We take a high-dimensional gradient feature space [3] widely used in handwritten numeral recognition. The proposed method deforms each category's template by 2D shape morphing linked with GA to achieve a maximal fitness value against an arbitrary given feature vector in a feature space and generates a reconstructed image in an image space corresponding to the given feature vector.

Section 2 explains the handwritten numeral database used in our experiments. Section 3 addresses the intrinsic problems in image reconstruction from a feature space from the viewpoint of an inverse problem with a one-to-many relation. Section 4 introduces 2D shape morphing based on a mesh model and bilinear transformation. Section 5 describes a gradient feature space we used. In Section 6, we explain genetic algorithms that search for optimal morphing parameters to generate a reconstructed image from a given feature vector. Section 7 demonstrates successful experimental results and discusses how to use this technique to improve recognition performance.

Proceedings of the 9th Int'l Workshop on Frontiers in Handwriting Recognition (IWFHR-9 2004)
0-7695-2187-8/04 $20.00 © 2004 IEEE

## 2. IPTP CDROM1B character database

The handwritten numeral database IPTP CDROM1B provided by Institute for Posts and Telecommunications Policy of Japan [7] is used in our experiments. The IPTP CDROM1B contains binary images of handwritten digits. These binary digit images were manually segmented through binarization from 8-bit gray-scale images of three digit ZIP codes optically scanned from real Japanese New Year greeting cards. The size of each binary image is 120 dots $\times$ 80 dots in height and width. This database consists of two groups of 17,985 samples used for training and 17,916 samples used for test. We use training samples to generate templates for each of 10 digits. These templates are deformed to generate a reconstructed image from an arbitrary given feature vector using 2D shape morphing and GA. Also, we show a high matching ability of the proposed 2D shape morphing technique by reproducing real test samples from their corresponding feature vectors.

Figure 1 shows sample digit images of the IPTP CDROM1B database.

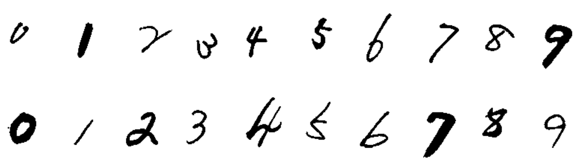

**Fig. 1: Sample digit images of IPTP CDROM1B database.**

## 3. Image reconstruction as an inverse problem

Here, we discuss an intrinsic difficulty in image reconstruction from a feature space viewed as an inverse problem.

By means of feature extraction a real character pattern is projected onto a point in the feature space. This is a one-to-one relation. On the other hand, inverse projection of a point in the feature space onto a character pattern in an image space has a one-to-many relation. This is totally because feature extraction is a kind of dimension-reduction process.

From the above consideration, we can point out the following two characteristics of this inverse problem.

(1) Image reconstruction from a feature space can obtain many solutions that have almost the same feature vectors and we cannot say one special solution is superior to other solutions. In other words, these reconstructed images which might be quite different in appearance cannot be discriminated in the given feature space.

(2) We can use a similarity measure only in a feature space not in an image space. Hence, reasonable image reconstruction heavily depends on a good selection of similarity measures in a feature space. However, considering the above-mentioned characteristics of a one-to-many relation, it is not straightforward to evaluate the goodness of those similarity measures.

## 4. 2D Shape morphing

Shape morphing as deformable models is divided into two categories. One uses learned models consisting of curved segments or splines whose shape is governed by a small number of control points [8]. The other is a deformable template consisting of a number of square blocks and deforms the template by displacing edges of each block [9].

We use 2D shape morphing belonging to the second category or a mesh model. The proposed mesh model consists of only $3 \times 2$ blocks and deforms a template by displacing a total of twelve corner points and using bilinear transformation of quadrilateral regions.

### 4.1. Template generation

A single template for each digit is generated using training samples provided by IPTP CDROM1B. The number of training samples for each digit is as follows: 0: 2535, 1: 1691, 2: 1658, 3: 1736, 4: 1188, 5: 1996, 6: 1977, 7: 1410, 8: 2366, and 9: 1428.

Each digit image has 120 dots $\times$ 80 dots in height and width. First, as preprocessing, position and size normalization by moments is applied to the image. Namely, the center of gravity of black pixels is shifted to the center of the image, and the image is expanded or shrunken so that the second moment around the center of gravity is equal to the predetermined value. The normalized image has the same size of 120 dots $\times$ 80 dots.

Figure 2 shows one example of position and size normalization.

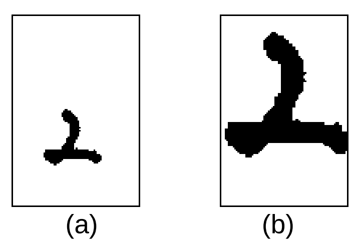


**Fig. 2: Position and size normalization. (a) Original image. (b) Normalized image.**

Then, we generate overlapped images of normalized patterns for each of 10 digits, respectively.

Figure 3(a) shows those overlapped images for 10 digits. From Fig. 3(a), it is found that the simply overlapped images are considerably blurred due to a wide handwriting variation.

Hence, we use the GAT correlation technique [10] to adaptively normalize each pattern and reduce the entropy of overlapped images. Finally, we generate templates through binarization of these enhanced overlapped images.

Figure 3(b) shows templates for 10 digits thus obtained.

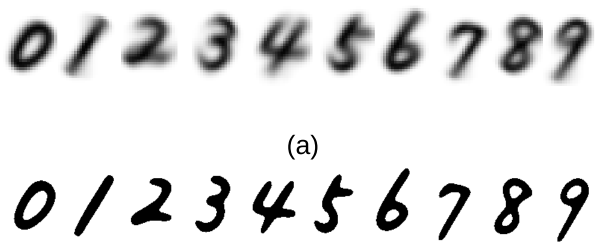
(a)
(b)

**Fig. 3: Templates for 10 digits. (a) Simply overlapped images. (b) Templates obtained via enhanced normalization.**

### 4.2. Mesh model

A character image is divided into 3 × 2 square blocks. Then, we move each corner point of respective blocks in X and Y directions independently to generate a deformed image.

Figure 4 shows the proposed "mesh model." The total number of corner points is twelve. Therefore, our mesh model has 24 morphing parameters.

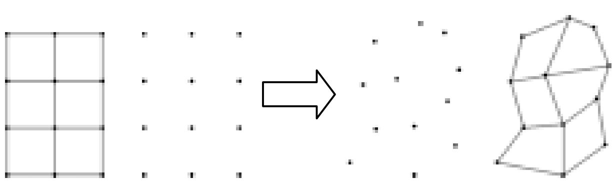

**Fig. 4: Mesh model.**

### 4.3. Bilinear transformation of quadrilaterals

We determine both X and Y coordinates of every point within a quadrilateral region using a well-known technique of bilinear transformation [11]. As shown in Figure 5, four corner points of a square block move from $(X_1, Y_1)$, $(X_2, Y_2)$, $(X_3, Y_3)$, and $(X_4, Y_4)$ to $(U_1, V_1)$, $(U_2, V_2)$, $(U_3, V_3)$, and $(U_4, V_4)$, respectively. Here, we assume that the transformed block is convex.

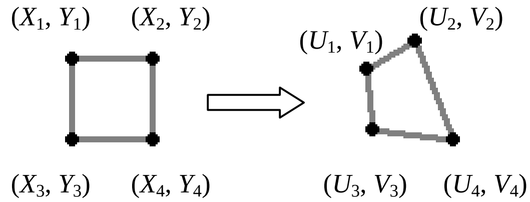


**Fig. 5: Bilinear transformation of a quadrilateral.**

The bilinear transformation is given by

$$\begin{aligned} U &= a_1 X Y + a_2 X + a_3 Y + a_4 \\ V &= b_1 X Y + b_2 X + b_3 Y + b_4 . \end{aligned} \tag{1}$$

By substituting the correspondence relation of four corner points for *X*, *Y*, *U*, and *V* in Eq. (1), we obtain eight simultaneous linear equations for eight unknown coefficients, $a_i$ and $b_i$ ($1 \le i \le 4$). These simultaneous linear equations are easily solved.

Then, we can determine X and Y coordinates of every internal/boundary point of a quadrilateral region by means of Eq. (1) using the obtained coefficients, $a_i$ and $b_i$.

Figure 6 shows examples of transformed images by bilinear transformation as applied to an original character image of digit "2."

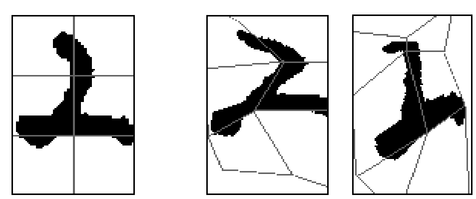
(a) (b)

**Fig. 6: Examples of shape morphing by bilinear transformation. (a) Original image. (b) Transformed images.**

Here, we should notice that a random choice of morphing or displacing parameters is likely to generate invalidly transformed images with some gaps as shown in Figure 7. This is totally because interpolation by bilinear transformation does not work properly in the case of a concave quadrilateral.

Therefore, we need to select a valid set of morphing parameters that generate only convex quadrilateral regions from randomly generated ones. By using the convex hull calculation algorithm called "Package wrapping algorithm" [12] we determine whether four corner points of each transformed block form a convex quadrilateral or not. When applying GA to an optimization problem of morphing parameters we require those parameters to pass this convex hull test.

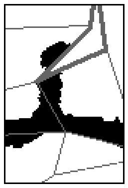

**Fig. 7: Transformed image with a gap.**

## 5. Gradient feature space

In this section, we explain a gradient feature space used in our experiments. Gradient-based features have been proven to be very effective in handwritten numeral recognition [3].

The procedure of gradient feature extraction we used is described below.

First, a neighbor averaging filter of size $3 \times 3$ is repeatedly applied ten times to the binary character image $b(i, j)$ to generate a gray-scale image $g(i, j)$, ($0 \le i < 80$, $0 \le j < 120$).

Next, Roberts cross-gradient operator [11] given by Eq. (2) is applied to each pixel $g(i, j)$ to calculate the gradient direction $\theta(i, j)$ at the pixel as follows.

$$\begin{aligned} \Delta u &= g(i+1, j+1) - g(i, j), \\ \Delta v &= g(i+1, j) - g(i, j+1), \\ \theta(i, j) &= \tan^{-1}\left(\frac{\Delta v}{\Delta u}\right) - \frac{\pi}{4}. \end{aligned} \qquad (2)$$

The gradient direction $\theta(i, j)$ is quantized with $\pi/4$ interval to eight-directional codes from 0 to 7, as shown in Fig. 8.

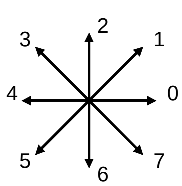


**Fig. 8: Eight-directional codes.**

We divide the entire image into $5 \times 5$ blocks. Then, in every block we count up the number of pixels that have each of eight-directional codes, respectively. As a result, we obtain a gradient feature vector consisting of $5 \times 5 \times 8$ elements each of which stores the corresponding count. This feature vector represents a 2D distribution of gray-scale gradients of the character image.

## 6. Morphing using GA

We use genetic algorithms (GA) [6] to search for optimal morphing parameters that deform a template so as to yield the maximum fitness value between an arbitrary given feature vector and a feature vector of the deformed template in a high-dimensional feature space. Here, the fitness value serves as a similarity measure.

### 6.1. Gene encoding and specifications of GA

First, we explain gene encoding that represents a set of 24 morphing parameters described in **4.2**. Because the entire image with 120 dots $\times$ 80 dots is divided into $3 \times 2$ square blocks, each square block has 40 dots $\times$ 40 dots. Then, we set the range of each corner point's displacement in X and Y directions from –30 dots to +30 dots around its initial loci. Hence, 6 bits are adequate for representing one morphing parameter, and the total length of each chromosome amounts to 144 bits.

Next, specifications of GA are as follows. The initial population of a size 200 is randomly generated. We adopt the roulette selection rule based on the fitness values in each generation. Regarding crossover, we use the one-point crossover method with the rate of 80%. Mutation is a bit-wise reversal within a chromosome with a rate of 2%. We provide each generation with 200 chromosomes that pass the convex hull test described in **4.3**.

Finally, we stop the GA process when the maximal fitness value against the given feature vector attained by an elite chromosome exceeds the predetermined threshold value. The corresponding set of optimal morphing parameters generates a deformed template. This completes a process of image reconstruction from the given feature vector.

However, it is to be noted that the obtained image is only one possible solution and use of other templates, in particular topologically different ones, might generate totally different solutions against the same feature vector.

### 6.2 Fitness value in a feature space

We calculate a fitness value between a given feature vector, $\boldsymbol{t} = (t_i)$, and a feature vector, $\boldsymbol{f} = (f_i)$, extracted from a deformed template for each of chromosomes in the same generation.

First, we calculate the city block distance, $d(\boldsymbol{f}, \boldsymbol{t})$, between these two feature vectors given by

$$d(\boldsymbol{f},\boldsymbol{t}) = \sum_{i=1}^{5\times5\times8} | f_i - t_i |. \tag{3}$$

Then, the fitness value, $F(\boldsymbol{f}, \boldsymbol{t})$, with the range of [0, 1] is calculated as follows.

$$F(\boldsymbol{f},\boldsymbol{t}) = \exp(-d(\boldsymbol{f},\boldsymbol{t})/2.0). \tag{4}$$

A chromosome that has the maximum value of $F(\boldsymbol{f}, \boldsymbol{t})$ is selected as an elite one in the given generation.

This definition of a fitness value is a simple choice. As mentioned in **3**, an appropriate selection of fitness values is a key to reasonable image reconstruction.

# 7. Experimental results

In this section, we demonstrate two kinds of experimental results.

One is to show a high matching ability of 2D shape morphing based on a mesh model and bilinear transformation. Here, we select a gradient feature vector extracted from a real test sample as a given feature vector. Then, we deform the correct category's template to yield the maximal fitness value against this given feature vector. Finally, we check if the deformed template and the original test sample are very alike or not.

The other is to show limited but encouraging results of image reconstruction form a feature space. First, we reconstruct an image from each of mean feature vectors of 10 digits. Next, we generate the midpoint of two mean feature vectors and reconstruct two images from this midpoint using two corresponding templates.

## 7.1. Shape matching ability

2D shape morphing based on a $3 \times 2$ mesh model and bilinear transformation is considered a too simple model. However, we obtained successful experimental results to show a high matching ability of the proposed model.

Figure 9 shows examples of deformed templates against feature vectors extracted from real test samples.

From Fig. 9, it is clearly found that not only global/local shape deformation but also line width variation were faithfully reproduced to a considerable extent by the proposed model of 2D shape morphing.

These results strongly suggest that the proposed morphing model can serve as a stable and powerful tool of shape matching in a feature space between templates and real images belonging to the same digit category. Of course, we need multiple templates in order to deal with topologically different images within the same category. In that case, we can use unsupervised clustering techniques [4], for example, $k$-means or mixture models, in a feature space to generate multiple templates for each digit category.

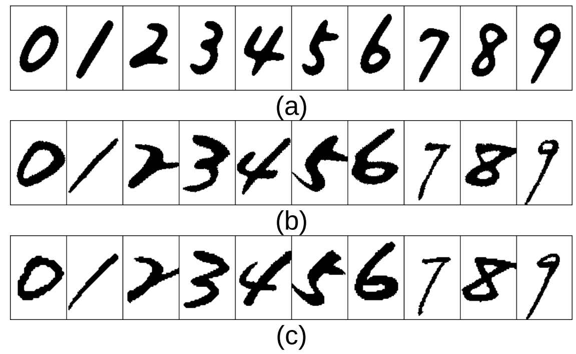


**Fig. 9: Examples of shape morphing via feature vectors. (a) Templates. (b) Results of morphing. (c) Real test samples.**

## 7.2. Image reconstruction from a feature space

In this subsection, we deal with a real inverse problem from a feature space to an image space. Namely, we reconstruct an image from an arbitrary given feature vector that has no original counterpart in an image space.

First, we select a mean feature vector of each digit category and deform its corresponding template to yield the maximal fitness value in a feature space against the mean feature vector.

Figure 10 shows image reconstruction results from mean feature vectors of 10 digits.

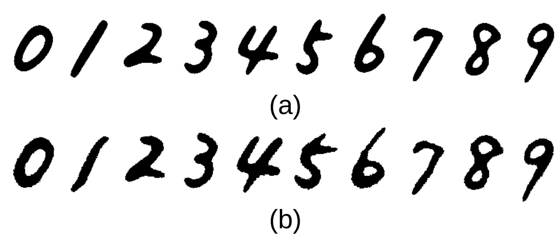


**Fig. 10: Image reconstruction results from mean feature vectors. (a) Templates. (b) Reconstructed images.**

From Fig. 10, it is found that reconstructed images are similar to their corresponding templates, but are slightly different from them. These differences are important and might provide a vital clue to analyzing advantages and disadvantages of the gradient feature space and searching for more robust features.

Second, we select the midpoint of two mean feature vectors in a feature space. In this case, it is reasonable to


0-7695-2187-8/04 $20.00 © 2004 IEEE

take two corresponding templates to be deformed. As a result, we obtain two different reconstructed images from the midpoint.

This is a typical example of a one-to-many relation as discussed in **3**.

Figure 11 shows an example of image reconstruction from the midpoint of "4" and "6" in a feature space.

(a) (b)

**Fig. 11: Image reconstruction from the midpoint of two mean feature vectors.**
**(a) Reconstructed image using "4."**
**(b) Reconstructed image using "6."**

It is very interesting that two reconstructed images from the midpoint fairly differ in appearance while they have almost the same values in a feature space. Based on the difference of their appearances, we can say which of two categories the midpoint being considered should belong to. Moreover, it might be possible to classify any intermediate point on the line connecting two mean feature vectors into either of two categories.

From these results we can say that the proposed technique of image reconstruction from a feature space will be useful for providing a classifier with artificial learning samples near the category boundaries in supervised learning. Hence, it is challenging and promising to reinforce such classifiers like multi-layer neural networks and support vector machines with this technique.

Furthermore, as a tool for analyzing misclassified samples, it is interesting to reconstruct images from their feature vectors using rival categories' templates. It is clear that we cannot discriminate these reconstructed images from the corresponding misclassified samples in the given feature space. This kind of analysis of the given feature space can help us toward searching for more robust features.

## 8. Conclusion

This paper presented a powerful technique for reconstructing a 2D character image from a high-dimensional feature space by means of 2D shape morphing and genetic algorithms.

Experimental results first demonstrated a high shape morphing ability of the proposed technique. Then, we successfully generated reconstructed images from both mean feature vectors and their midpoints in a gradient feature space. We can say that this technique is promising as a powerful tool for analyzing a given feature space and searching for more robust features. Moreover, visualizing any intermediate feature vector between two or more of mean feature vectors will be useful for examining category boundaries.

Future work is to apply this technique not only to analysis of a given feature space in pursuit of more robust features but also to artificial generation of learning samples near category boundaries in supervised learning.


## Acknowledgement

The authors would like to thank Dr. Yozo Tamura for many fruitful discussions regarding this work.


## References


[1] R. Plamondon and S. Srihari, "On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-22, pp. 63-84, 2000.

[2] H. Bunke, "Recognition of Cursive Roman Handwriting – Past, Present and Future," *Proc. of 7th Int. Conf. on Document Analysis and Recognition*, pp. 448-459, 2003.

[3] M. Shi, Y. Fujisawa, T. Wakabayashi, and F. Kimura, "Handwritten Numeral Recognition Using Gradient and Curvature of Gray Scale Image", *Pattern Recognition*, vol. 35, pp. 2051-2059, 2002.

[4] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Second Edition, John Wiley & Sons, 2001.

[5] H. Sakano, H. Kida, and N. Mukawa, "Seeing the Character Images that an OCR System Sees – Analysis by Genetic Algorithm -", *Proc. 13th Int. Conf. Pattern Recognition*, pp. 411-416, 1996.

[6] D.E. Goldberg, *Genetic Algorithms – in Search, Optimization & Machine Learning*, Addison-Wesley, 1989.

[7] K. Osuga, T. Tsutsumida, S. Yamaguchi, and K. Nagata, "IPTP Survey on Handwritten Numeral Recognition," *IPTP Research and Survey Report*, R-96-V-02, 1996.

[8] M. Revow, C.K.I. Williams, and G.E. Hinton, "Using Generative Models for Handwritten Digit Recognition," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-18, pp. 592-606, 1996.

[9] A.K. Jain and D. Zongker, "Representation and Recognition of Handwritten Digits Using Deformable Templates," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. PAMI-19, pp. 1386-1390, 1997.

[10] T. Wakahara, "Shape Matching Using GAT Correlation against Nonlinear Distortion and its Application to Handwritten Numeral Recognition" *Proc. of 7th Int. Conf. on Document Analysis and Recognition*, pp. 54-58, 2003.

[11] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, Second Edition, Chaps. 5 & 7, Prentice Hall, 2002.

[12] R. Sedgewick, *Algorithms in C*, Addison-Wesley, 1990.




0-7695-2187-8/04 $20.00 © 2004 **IEEE**