k-NN Classification of Handwritten Characters via Accelerated GAT Correlation

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

Yukihiko Yamashita
Graduate School of Engineering and Science
Tokyo Institute of Technology
2-12-1 O-okayama, Meguro-ku, Tokyo, 152-8550 Japan
E-mail: yamasita@ide.titech.ac.jp

***Abstract*—This paper addresses the problem of reinforcing the ability of k-NN classification of handwritten characters via distortion-tolerant template matching techniques with a limited quantity of data. We make a comparison of three kinds of matching techniques: the conventional simple correlation, the tangent distance, and the GAT (Global Affine Transformation) correlation. The k-NN classification method is straightforward and powerful, however, is very time-consuming. Hence, to reduce the computational cost of matching in k-NN classification we propose to accelerate the GAT correlation technique by reformulating its computational model and adopting efficient lookup tables. Recognition experiments made on the handwritten numerical database IPTP CDROM1B show that matching techniques of the simple correlation, the tangent distance, and the accelerated GAT correlation achieve recognition rates of 97.07%, 97.50%, and 98.70%, respectively. Also, the computation time ratios of the tangent distance and the accelerated GAT correlation to the simple correlation are 26.3 and 36.5 to 1.0, respectively.**

***Keywords*-affine-invariant template matching; k-NN classification; normalized cross-correlation; character recognition;**

I. Introduction

Benchmarking of state-of-the-art techniques in handwritten digit recognition has demonstrated that statistical pattern recognition techniques using a large amount of training data are most successful [1]. However, we have constant need of powerful matching algorithms in problems where advanced statistical modeling for standard classification methods is not possible due to a limited quantity of data.

To resolve this problem, we have two major approaches: distortion-tolerant template matching and k-NN (k-Nearest-Neighbor) classification. Of course, these two approaches can be combined effectively.

Regarding the first approach several promising techniques based on deformable models have been proposed. Revow et al. [2] and Jain et al. [3] reinforced their deformable models via probabilistic viewpoints. Also, Ronee et al. [4] enriched DP-based 2D warping. Especially, Bunke et al. [5] showed that most of handwriting distortion can be expressed by affine transformation and proposed the perturbation method based on affine transformation. The tangent distance by Simard et al. [6] and GAT correlation by Wakahara et al. [7] aimed to absorb affine transformation in a straightforward manner. Moreover, Wakahara et al. [8] extended GAT correlation to PAT (Partial Affine Transformation) correlation to deal with nonlinear distortion. By the way, in online handwriting recognition, stroke-based affine transformation [9], affine moments invariants [10], and affine integral invariants [11] have been proposed.

Regarding the second approach k-NN classifiers are memory-based, and require no model to be fit [12]. In particular, by incorporating invariances under certain natural transformations into the metric used to measure the distances between objects, the k-NN classifier can serve as a high-accuracy distortion-tolerant template matching technique [6]. However, one major drawback of k-NN classifiers in general is the computational load. Distortion-tolerant template matching techniques are also very time-consuming.

In this paper, we propose a new, powerful combination of accelerated distortion-tolerant template matching and k-NN classification. The key contribution is drastic reduction of the computational cost of the GAT correlation technique. By reformulating the computational model of the GAT correlation based on separation of variables, we generate 8-directional GAT correlation templates for calculating optimal affine parameters efficiently. Also, adopting lookup tables helps to further reduce the computational load.

Experimental results made on the handwritten numeral database IPTP CDROM1B show that k-NN classification via the accelerated GAT correlation achieves a much higher recognition accuracy than those obtained by k-NN classifiers using the conventional simple correlation and the tangent distance. Furthermore, the computational cost of the GAT correlation has been reduced to a roughly comparable level with that of the tangent distance even though the GAT correlation technique involves an iterative optimization process.

II. Simple correlation and the tangent distance

We adopt normalized cross-correlation as a matching measure or metric in k-NN classification.

First, we denote an input image and a template by $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$, $\boldsymbol{x} \in \mathcal{D}$, respectively. $\mathcal{D}$ specifies the common

978-0-7695-4774-9/12 $26.00 © 2012 IEEE
DOI 10.1109/ICFHR.2012.225

domain of the input image and the template.

Then, according to the procedure of definite canonicalization [13] with robustness against image blurring and additive random noise, both the input image and the template are linearly transformed as follows.

$$(f,1)=(g,1)=0, \quad \| f \| = \| g \| = 1, \tag{1}$$

where

$$(f,g) \equiv \iint_{\mathcal{D}} f(\boldsymbol{x})g(\boldsymbol{x})d\boldsymbol{x}, \quad \| f \| \equiv \sqrt{(f,f)}.$$

As a result, a normalized cross-correlation value calculated between the input image and the template is simply represented by an inner product (f,g). We call this value a simple correlation value, $C^{simple}(f,g)$, given by

$$C^{simple}(f,g)=(f,g). \tag{2}$$

On the other hand, as is well known, the tangent distance [6] gives an invariant metric with respect to a set of local transformations which do not affect the identity of the image. Concretely, seven such image transformations are identified: horizontal and vertical translations, rotation, scaling, shearing, squeezing, and line thickening or thinning. Then, we generate two sets of seven tangent vectors, $\left\{v_i^f(\boldsymbol{x})\right\}$ and $\{v_i^g(\boldsymbol{x})\}\,(i=1,\ldots,7)$, for the input image and the template, respectively. Finally, we define the tangent distance between the input image and the template by

$$\begin{aligned} dist^{TD}(f,g) &= \min_{\boldsymbol{\alpha}^f,\boldsymbol{\alpha}^g} \left\|\tilde{f}-\tilde{g}\right\|^2, \\ \tilde{f}(\boldsymbol{x}) &= f(\boldsymbol{x})+\sum_i \alpha_i^f v_i^f(\boldsymbol{x}), \\ \tilde{g}(\boldsymbol{x}) &= g(\boldsymbol{x})+\sum_i \alpha_i^g v_i^g(\boldsymbol{x}), \end{aligned} \tag{3}$$

where optimal coefficients, $\boldsymbol{\alpha}^f$ and $\boldsymbol{\alpha}^g$, are easily determined according to a linear least squares problem.

Now, we newly define a tangent distance correlation value, $C^{TD}(f,g)$, by

$$C^{TD}(f,g)=\iint_{\mathcal{D}} \tilde{f}(\boldsymbol{x})\tilde{g}(\boldsymbol{x})d\boldsymbol{x}=(\tilde{f},\tilde{g}), \tag{4}$$

where both $\tilde{f}(\boldsymbol{x})$ and $\tilde{g}(\boldsymbol{x})$ are definitely canonicalized according to (1) using optimal coefficients, $\boldsymbol{\alpha}^f$ and $\boldsymbol{\alpha}^g$.

We use these two kinds of correlation values, C^{simple} of (2) and C^{TD} of (4), in k-NN classification of handwritten characters in Section V.

III. Original computational model of GAT correlation

First of all, affine transformation in the 2D image plane is defined by

$$\boldsymbol{x}' = A\boldsymbol{x}+\boldsymbol{b}, \tag{5}$$

or

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a_{00} & a_{01} \\ a_{10} & a_{11} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_0 \\ b_1 \end{pmatrix}. \tag{6}$$

Accordingly, we have the affine-transformed input image, $f^*(\boldsymbol{x})$, given by

$$f^*(\boldsymbol{x})=\frac{1}{|A|}f(A^{-1}(\boldsymbol{x}-\boldsymbol{b})), \tag{7}$$

where $|A|$ is the determinant of the matrix A.

Then, we define a GAT correlation value, $C^{GAT}(f,g)$, by

$$\begin{aligned} C^{GAT}(f,g) &= \max_{A,\boldsymbol{b}} J_{GAT}(A,\boldsymbol{b}), \\ J_{GAT}(A,\boldsymbol{b}) &= \iint_{\mathcal{D}} f^*(\boldsymbol{x})g(\boldsymbol{x})d\boldsymbol{x} \\ &= \iint_{\mathcal{D}} \frac{1}{|A|}f(A^{-1}(\boldsymbol{x}-\boldsymbol{b}))g(\boldsymbol{x})d\boldsymbol{x} \\ &= \iint_{\mathcal{D}} f(\boldsymbol{z})g(A\boldsymbol{z}+\boldsymbol{b})d\boldsymbol{z}. \end{aligned} \tag{8}$$

The key idea of the original GAT computational model [7] is to introduce a Gaussian kernel, $G(\boldsymbol{x})$, so that a new objective function, $\tilde{J}_{GAT}(A,\boldsymbol{b})$, is differentiable with respect to A and $\boldsymbol{b}$ as follows.

$$\begin{aligned} \tilde{J}_{GAT}(A,\boldsymbol{b}) = \iint_{\mathcal{D}}\iint_{\mathcal{D}} G(A\boldsymbol{x}_1+\boldsymbol{b}-\boldsymbol{x}_2) \times \\ \delta\left(\nabla f(\boldsymbol{x}_1),\nabla g(\boldsymbol{x}_2)\right) f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2, \\ G(\boldsymbol{x})=\exp\left(-\frac{\|\boldsymbol{x}\|^2}{2D^2}\right), \end{aligned} \tag{9}$$

where gradients $\nabla f(\boldsymbol{x})$ and $\nabla g(\boldsymbol{x})$ are quantized into eight directions with the $\pi/4$ interval, and take integers ranging from zero to eight; the value of zero corresponds to no gradient. Also, the $\delta(i,j)$ is a kind of the Kronecker delta given by

$$\delta(i,j)=\begin{cases} 1, & \text{for } i=j\neq 0 \\ 0. & \text{for } i\neq j \text{ or } i=0 \text{ or } j=0 \end{cases}$$

The value of D of (9) that controls the spread of the Gaussian kernel is adaptively determined as a function of the disparity of the input image and the template with constraints on their gradients as follows.

$$\begin{aligned} D = \frac{1}{2}\mathop{Av}_{\boldsymbol{x}_1}\left\{\min_{\{\boldsymbol{x}_2|\nabla f(\boldsymbol{x}_1)=\nabla g(\boldsymbol{x}_2)\neq 0\}}\|\boldsymbol{x}_1-\boldsymbol{x}_2\|\right\} \\ +\frac{1}{2}\mathop{Av}_{\boldsymbol{x}_2}\left\{\min_{\{\boldsymbol{x}_1|\nabla f(\boldsymbol{x}_1)=\nabla g(\boldsymbol{x}_2)\neq 0\}}\|\boldsymbol{x}_1-\boldsymbol{x}_2\|\right\}, \end{aligned} \tag{10}$$

where D is the average minimum distance between two points, one in f and the other in g, with the same gradient direction.

By using the following relation:

$$\begin{aligned} & \|A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2\|^2 \\ = & \|A\boldsymbol{x}_1\|^2 + \|\boldsymbol{b}\|^2 + \|\boldsymbol{x}_2\|^2 \\ & + 2\langle A\boldsymbol{x}_1, \boldsymbol{b}\rangle - 2\langle A\boldsymbol{x}_1, \boldsymbol{x}_2\rangle - 2\langle \boldsymbol{b}, \boldsymbol{x}_2\rangle \\ = & \operatorname{tr}\left[A\boldsymbol{x}_1\boldsymbol{x}_1^T A^T + 2\boldsymbol{b}\boldsymbol{x}_1^T A^T - 2\boldsymbol{x}_2\boldsymbol{x}_1^T A^T\right] \\ & + \|\boldsymbol{b}\|^2 + \|\boldsymbol{x}_2\|^2 - 2\langle \boldsymbol{b}, \boldsymbol{x}_2\rangle, \end{aligned}$$

we obtain the derivatives with respect to A and $\boldsymbol{b}$ as

$$\begin{aligned} \frac{1}{2}\frac{\partial \|A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2\|^2}{\partial A^T} &= A\boldsymbol{x}_1\boldsymbol{x}_1^T + \boldsymbol{b}\boldsymbol{x}_1^T - \boldsymbol{x}_2\boldsymbol{x}_1^T, \\ \frac{1}{2}\frac{\partial \|A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2\|^2}{\partial \boldsymbol{b}} &= \boldsymbol{b} + A\boldsymbol{x}_1 - \boldsymbol{x}_2. \end{aligned} \tag{11}$$

Now, by setting the derivatives of $\tilde{J}_{GAT}(A, \boldsymbol{b})$ of (9) with respect to A and $\boldsymbol{b}$ equal to zero, we obtain a set of nonlinear equations given by

$$\begin{aligned} 0 = -\frac{1}{D^2}\iint_{\mathcal{D}}\iint_{\mathcal{D}} & \frac{\partial \|A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2\|^2}{\partial A^T} \\ & G(A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2)\delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) \\ & f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2, \\ 0 = -\frac{1}{D^2}\iint_{\mathcal{D}}\iint_{\mathcal{D}} & \frac{\partial \|A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2\|^2}{\partial \boldsymbol{b}} \\ & G(A\boldsymbol{x}_1 + \boldsymbol{b} - \boldsymbol{x}_2)\delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) \\ & f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2. \end{aligned} \tag{12}$$

Here, in order to remove the nonlinearity from these equations we adopt the 0th order approximation that sets $A = I$ and $\boldsymbol{b} = \boldsymbol{0}$ in the Gaussian kernel of (12).

As a result, we have a set of simultaneous linear equations given by

$$\begin{aligned} 0 &= A\overline{\boldsymbol{x}_1\boldsymbol{x}_1^T} + \boldsymbol{b}\overline{\boldsymbol{x}_1}^T - \overline{\boldsymbol{x}_2\boldsymbol{x}_1^T}, \\ 0 &= \boldsymbol{b}\overline{1} + A\overline{\boldsymbol{x}_1} - \overline{\boldsymbol{x}_2}, \end{aligned} \tag{13}$$

where

$$\begin{aligned} \overline{1} = \iint_{\mathcal{D}}\iint_{\mathcal{D}} & G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \\ & \delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2, \\ \overline{\boldsymbol{x}_1} = \iint_{\mathcal{D}}\iint_{\mathcal{D}} & \boldsymbol{x}_1 G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \\ & \delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2, \\ \overline{\boldsymbol{x}_2} = \iint_{\mathcal{D}}\iint_{\mathcal{D}} & \boldsymbol{x}_2 G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \\ & \delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2, \\ \overline{\boldsymbol{x}_1\boldsymbol{x}_1^T} = \iint_{\mathcal{D}}\iint_{\mathcal{D}} & \boldsymbol{x}_1\boldsymbol{x}_1^T G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \\ & \delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2, \\ \overline{\boldsymbol{x}_2\boldsymbol{x}_1^T} = \iint_{\mathcal{D}}\iint_{\mathcal{D}} & \boldsymbol{x}_2\boldsymbol{x}_1^T G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \\ & \delta\left(\nabla f(\boldsymbol{x}_1), \nabla g(\boldsymbol{x}_2)\right) f(\boldsymbol{x}_1)g(\boldsymbol{x}_2)d\boldsymbol{x}_1 d\boldsymbol{x}_2. \end{aligned} \tag{14}$$

Finally, we solve these simultaneous linear equations of (13) and obtain a sub-optimal solution of A and $\boldsymbol{b}$ as follows.

$$\begin{aligned} A &= \left(\overline{\boldsymbol{x}_2\boldsymbol{x}_1^T} - \frac{\overline{\boldsymbol{x}_2}\ \overline{\boldsymbol{x}_1}^T}{\overline{1}}\right)\left(\overline{\boldsymbol{x}_1\boldsymbol{x}_1^T} - \frac{\overline{\boldsymbol{x}_1}\ \overline{\boldsymbol{x}_1}^T}{\overline{1}}\right)^{-1}, \\ \boldsymbol{b} &= -\frac{A\overline{\boldsymbol{x}_1}}{\overline{1}} + \frac{\overline{\boldsymbol{x}_2}}{\overline{1}}. \end{aligned} \tag{15}$$

In order to obtain the true optimal solution of (8), we use the successive iteration method [14] by iterative application of sub-optimal affine parameters of (15) to the input image until the value of $J_{GAT}(A, \boldsymbol{b})$ of (8) arrives at a maximum.

IV. Accelerated computational model of GAT correlation

The key ideas to accelerate the original computational model of GAT correlation are twofold.

The first one is to generate 8-directional GAT correlation templates in advance by means of separation of variables in the original computational model. This drastically reduces the computation time to calculate a set of components of (14) necessary for determining sub-optimal affine parameters.

The second one is to generate lookup tables of distances from each point to its nearest point with the same gradient direction. These lookup tables are used to replace a runtime computation of D of (9) with a simpler array indexing operation.

A. Generation of GAT correlation templates

We reformulate the original computational model of GAT correlation described in Section III via separation of variables to generate 8-directional GAT correlation templates.

First, by using the notation:

$$\boldsymbol{x}_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} x_2 \\ y_2 \end{pmatrix},$$

$$G_1(x_1 - x_2) = \exp\left(-\frac{|x_1 - x_2|^2}{2D^2}\right),$$

we can easily verify that the Gaussian kernel of (9) has the following property given by

$$G(\boldsymbol{x}_1 - \boldsymbol{x}_2) = G_1(x_1 - x_2)G_1(y_1 - y_2), \tag{16}$$

which provides a basis for separation of variables in the accelerated GAT computational model.

Then, by making full use of the above property of the Gaussian kernel, we define a set of functions, $\{h_0(y_1, x_2, s)\}$ and $\{H_0(\boldsymbol{x}_1, s)\}$ $(s = 0, \ldots, 8)$, by

$$h_0(y_1, x_2, s) = \int G_1(y_1 - y_2) \times \delta(s, \nabla g(\boldsymbol{x}_2))\, g(x_2, y_2) dy_2, \tag{17}$$

and

$$\begin{aligned} H_0(\boldsymbol{x}_1, s) &= \iint_{\mathcal{D}} G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \delta(s, \nabla g(\boldsymbol{x}_2))\, g(\boldsymbol{x}_2) d\boldsymbol{x}_2 \\ &= \int G_1(x_1 - x_2) h_0(y_1, x_2, s) dx_2, \end{aligned} \tag{18}$$

respectively, where s specifies the gradient direction and takes an integer value ranging from zero to eight.

Similarly, we define another set of functions, $\{h_1(y_1, x_2, s)\}$ and $\{H_1(\boldsymbol{x}_1, s)\}$ $(s = 0, \ldots, 8)$, by

$$h_1(y_1, x_2, s) = \int y_2 G_1(y_1 - y_2) \times \delta(s, \nabla g(\boldsymbol{x}_2))\, g(x_2, y_2) dy_2, \tag{19}$$

and

$$\begin{aligned} H_1(\boldsymbol{x}_1, s) &= \iint_{\mathcal{D}} \boldsymbol{x}_2 G(\boldsymbol{x}_1 - \boldsymbol{x}_2) \times \delta(s, \nabla g(\boldsymbol{x}_2))\, g(\boldsymbol{x}_2) d\boldsymbol{x}_2 \\ &= \begin{pmatrix} \int x_1 G_1(x_1 - x_2) h_0(y_1, x_2, s) dx_2 \\ \int G_1(x_1 - x_2) h_1(y_1, x_2, s) dx_2 \end{pmatrix}. \end{aligned} \tag{20}$$

We call these functions of (17), (18), (19), and (20) 8-directional GAT correlation templates. It is clear that we can generate and store these templates in advance for each of templates to be excluded from a runtime computation.

Finally, by using these templates, we can calculate a set of components of (14) very efficiently by

$$\begin{aligned} \overline{1} &= \iint_{\mathcal{D}} H_0(\boldsymbol{x}_1, \nabla f(\boldsymbol{x}_1)) f(\boldsymbol{x}_1) d\boldsymbol{x}_1, \\ \overline{\boldsymbol{x}_1} &= \iint_{\mathcal{D}} \boldsymbol{x}_1 H_0(\boldsymbol{x}_1, \nabla f(\boldsymbol{x}_1)) f(\boldsymbol{x}_1) d\boldsymbol{x}_1, \\ \overline{\boldsymbol{x}_2} &= \iint_{\mathcal{D}} H_1(\boldsymbol{x}_1, \nabla f(\boldsymbol{x}_1)) f(\boldsymbol{x}_1) d\boldsymbol{x}_1, \\ \overline{\boldsymbol{x}_1 \boldsymbol{x}_1^T} &= \iint_{\mathcal{D}} \boldsymbol{x}_1 \boldsymbol{x}_1^T H_0(\boldsymbol{x}_1, \nabla f(\boldsymbol{x}_1)) f(\boldsymbol{x}_1) d\boldsymbol{x}_1, \\ \overline{\boldsymbol{x}_2 \boldsymbol{x}_1^T} &= \iint_{\mathcal{D}} H_1(\boldsymbol{x}_1, \nabla f(\boldsymbol{x}_1)) \boldsymbol{x}_1^T f(\boldsymbol{x}_1) d\boldsymbol{x}_1. \end{aligned} \tag{21}$$

B. Generation of lookup tables

It is very time-consuming to calculate the value of D of (10) based on a naive, raster-scan-based search for nearest-neighbor points, one in f and the other in g, with the same gradient direction.

To reduce the above-mentioned computational cost, we generate two kinds of lookup tables.

The first one is a lookup table of nearest-neighbor interpoint distances with respect to each gradient direction at every point, $\boldsymbol{x}_2$, in the template denoted by $\{dist(\boldsymbol{x}_2, s)\}$ $(s = 1, \ldots, 8)$.

When we know the gradient direction, $\nabla f(\boldsymbol{x}_1)$, of the input image, we can immediately pick up a value of $dist(\boldsymbol{x}_1, \nabla f(\boldsymbol{x}_1))$ as the nearest-neighbor interpoint distance to be determined.

However, we don't provide the input image with this type of lookup table because generation of the lookup table requires more time on the contrary.

Hence, we generate the second type of lookup table that stores nearby points' coordinates and corresponding distances sorted in the increasing order of interpoint distances for every point in the image plane. This lookup table realizes an efficient, fast search for the nearest-neighbor point in the input image from each point in the template with the same gradient direction.

V. Experimental results

We use the handwritten numeral database IPTP CDROM1B [15]. This database contains binary images of handwritten digits divided into two groups of 17,985 samples for training and 17,916 samples for test. Actually, the highest recognition rate ever reported for this database is 99.49% obtained via sophisticated discriminant functions in high-dimensional feature space [16].

First, position and size normalization by moments [17] is applied to each binary image so that the center of gravity of black pixels is located at the center of the image and the average distance of black pixels from the center of the image is set at the predetermined value of ρ $(= 6.0)$. Then, we transform all of binary images into grayscale images by Gaussian filtering and set the image size at 24×16 pixels.

In recognition experiments, we make a comparative study of k-NN classification of a total of 17,916 test samples using three kinds of matching measures: the conventional simple correlation of (2), the tangent distance correlation of (4), and the GAT correlation of (8). Here, templates are selected from a total of 17,985 training samples.

We have two major concerns: (1) relations between recognition rates of k-NN classification using each of three matching measures and the number of templates, and (2) the comparison of the computational cost among those three matching measures.

A. Recognition accuracy of k-NN classification

First, we investigated the recognition accuracy of k-NN classification using three kinds of competing matching measures as a function of the number of templates being sampled at random.

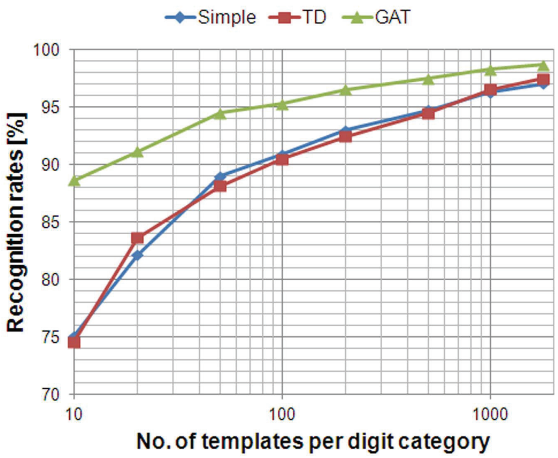

Figure 1. Relations between recognition rates via simple, tangent distance, and GAT correlations and the number of templates per digit category.

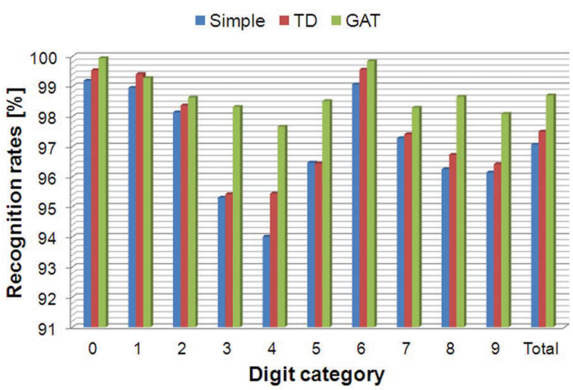

Figure 2. Recognition rates of each digit category using all training samples as templates in k-NN classification.

Figure 1 shows relations between recognition rates of k-NN classification via simple, tangent distance, and GAT correlations and the number of templates per digit category.

From Fig. 1, it is clear that the GAT correlation method is far superior in recognition accuracy to both the simple correlation and the tangent distance methods even when there are only a small number of templates.

From these results, we can say that k-NN classification of handwritten characters via the GAT correlation method is very powerful when we have only a limited quantity of training data.

Figure 2 shows recognition rates of each digit category using a total of 17,985 training samples as templates.

From Fig. 2, it is found that the accuracy of k-NN classification of handwritten characters via the GAT correlation method with a large number of templates can be comparable to that of the sophisticated statistical technique [16].

B. Computational cost

Figure 3 shows relations between processing times of k-NN classification via simple, tangent distance, and GAT correlations and the number of templates per digit category.

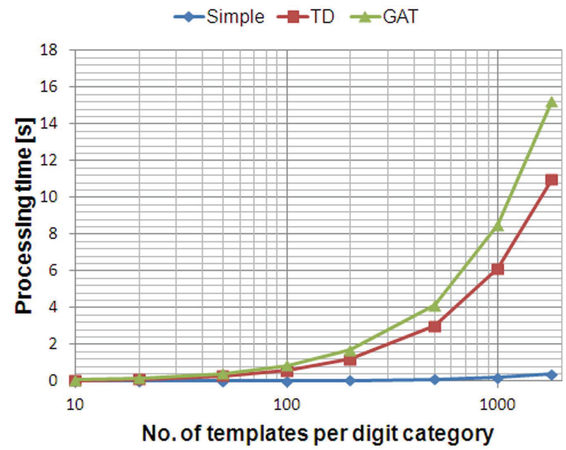

Figure 3. Relations between processing times via simple, tangent distance, and GAT correlations and the number of templates per digit category.

From Fig. 3, it is found that the average k-NN classification time per character via the accelerated GAT correlation using a total of 17,985 training samples as templates was 15.2s on a 2.67GHz Intel Xeon X5650 processor.

Table I shows the computation time ratios of the simple correlation, the tangent distance, and the GAT correlation in k-NN classification.

Table I
COMPUTATION TIME RATIOS OF THE SIMPLE CORRELATION, THE TANGENT DISTANCE, AND THE GAT CORRELATION.

Matching measure	Computation time ratio
Original GAT correlation	582.2
Accelerated GAT correlation	36.5
Tangent distance	26.3
Simple correlation	1.00

From Table I, it is first found that the computation time of the accelerated GAT correlation was greatly reduced to around six percent of that of the original GAT correlation. It is also found that the computational load of the accelerated GAT correlation is roughly comparable to that of the tangent distance while the former far surpasses the latter in recognition accuracy.

However, it is clear that the present classification time per character is still too long. Recently, engineers and scientists have increasingly studied the use of GPUs for non-graphical calculations because most of these computations involve matrix and vector operations. Here, we can point out that the accelerated GAT correlation basically consists of computations of inner products between vectors with weights. From this viewpoint, exploiting current state of the art processing schemes for implementation of the proposed method would make the relevance and viability of our approach stronger.

VI. Conclusion

This paper proposed a promising technique of k-NN classification of handwritten characters using the accelerated GAT correlation as a distortion-tolerant template matching technique.

Recognition experiments made on the handwritten numeral database IPTP CDROM1B showed that the proposed method achieved a much higher recognition rate of 98.70% than those of 97.07% and 97.50% obtained by the conventional simple correlation and the tangent distance, respectively. The superiority of the GAT correlation in recognition accuracy to the other two matching measures was held when there are only a limited number of templates in k-NN classification. Also, the computational cost of the accelerated GAT correlation was around six percent of that of the original GAT correlation.

Future work is to improve the recognition accuracy of k-NN classification of handwritten characters by combining the accelerated GAT correlation with PAT correlation [8] so as to absorb nonlinear distortion. Also, use of cutting-edge processing schemes such as GPUs to reduce the processing time of the proposed method is most interesting.

Acknowledgment

The authors would like to thank Dr. Yozo Tamura for his fruitful discussion regarding this work.

A part of this work was supported by Grants-in-Aid for Scientific Research No. 23500225 from JSPS.

References

[1] C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa. "Handwritten digit recognition: benchmarking of state-of-the-art techniques". *Pattern Recognition*, 36:2271–2285, 2003.

[2] M. Rebow, C. K. I. Williams, and G. E. Hinton. "Using generative models for handwritten digit recognition". *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-18:592–606, 1996.

[3] A. K. Jain and D. Zongker. "Representation and recognition of handwritten digits using deformable templates". *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-19:1386–1390, 1997.

[4] M. Ronee, S. Uchida, and H. Sakoe. "Handwritten character recognition using piecewise linear two-dimensional warping". *Proc. of Sixth Int. Conf. on Document Analysis and Recognition*, pages 39–43, Seattle, Sept. 2001.

[5] T. M. Ha and H. Bunke. "Off-line, handwritten numeral recognition by perturbation method". *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-19:535–539, 1997.

[6] P. Simard, Y. LeCun, and J. Denker. "Efficient pattern recognition using a new transformation distance". *Advances in Neural Information Processing Systems*, 5:50–58, 1993.

[7] T. Wakahara, Y. Kimura, and A. Tomono. "Affine-invariant recognition of gray-scale characters using global affine transformation correlation". *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-23:384–395, 2001.

[8] T. Wakahara and Y. Yamashita. "Multi-template GAT/PAT correlation for character recognition with a limited quantity of data". *Proc. of Twentieth Int. Conf. on Pattern Recognition*, pages 2873–2876, Istanbul, Aug. 2010.

[9] T. Wakahara and K. Odaka. "On-line cursive Kanji character recognition using stroke-based affine transformation". *IEEE Trans. Pattern Anal. Machine Intell.*, PAMI-19:1381–1385, 1997.

[10] J. Flusser and T. Suk. "Character recognition by affine moment invariants". *Proc. of 5th Int. Conf. on Computer Analysis of Images and Patterns*, pages 572–577, Budapest, Sept. 1993.

[11] O. Golubitsky, V. Mazalov, and S. M. Watt. "Toward affine recognition of handwritten mathematical characters". *Proc. of 9th IAPR Int. Workshop on Document Analysis Systems*, pages 35–42, Boston, June 2010.

[12] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Second Edition, Springer, 2009.

[13] T. Iijima. *Pattern Recognition*. Chap. 6, Corona, 1973 (in Japanese).

[14] Mathematical Society of Japan and K. Ito. *Encyclopedic Dictionary of Mathematics*. Second Edition, the MIT Press, 1987.

[15] 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 (in Japanese).

[16] M. Shi, Y. Fujisawa, T. Wakabayashi, and F. Kimura. "Handwritten numeral recognition using gradient and curvature of gray scale image". *Pattern Recognition*, 35:2051–2059, 2002.

[17] R. G. Casey. "Moment normalization of handprinted characters". *IBM J. Res. Develop.*, 14:548–557, 1970.