Journal of Korea Multimedia Society Vol. 14, No. 5, May 2011(pp. 642-649)

Mass-Spring-Damper Model for Offline Handwritten Character Distortion Analysis

Beom-Joon Cho†

ABSTRACT

Among the various aspects of offline handwritten character patterns, it is the great variety of writing styles and variations that renders the task of computer recognition very hard. The immense variety of character shape has been recognized but rarely studied during the past decades of numerous research efforts. This paper tries to address the problem of measuring image distortions and handwritten character patterns with respect to reference patterns. This work is based on mass-spring mesh model with the introduction of simulated electric charge as a source of the external force that can aid decoding the shape distortion. Given an input image and a reference image, the charge is defined, and then the relaxation procedure goes to find the optimum configuration of shape or patterns of least potential. The relaxation process is based on the fourth order Runge-Kutta algorithm, well-known for numerical integration. The proposed method of modeling is rigorous mathematically and leads to interesting results. Additional feature of the method is the global affine transformation that helps analyzing distortion and finding a good match by removing a large scale linear disparity between two images.

Key words: Character Recognition, Distortion, Mass-Spring-Damper, Mesh

1. INTRODUCTION

Offline handwritten character recognition is perhaps the last task in optical character recognition field. This is largely due to the great variety of shape variations including style differences and dynamic variations. Hence, unlike other problems and tasks, little progress in offline character recognition technology has been made during the past two or three decades to say nothing of mature technologies and commercial products [1].

Among the various aspects of offline handwritten character patterns, it is the great variety of writing styles and variations that renders the task of computer recognition very hard. The immense variety of character shape has been recognized but rarely studied during the past decades of numerous research efforts. This paper tries to address the problem of measuring image distortions and handwritten character patterns with respect to reference patterns.

There have long been research efforts that involve the measurement of image distortion or shape difference between two images. For instance stereo correspondence and optical flow estimation are among the commonest problems. Typical methods start with computing the discrepancies between two frames. Most solutions to the problems are based on the assumption of at least local linearity and smooth constraints [2]. On the other hand there is a problem in computer graphics of rendering, distorting and animating nonrigid objects effectively. One of most popular technique is mass-spring damper model in human cloth simulation [3]. Under some physical constraints such a mesh model represents a two dimensional sheet

※ Corresponding Author: Beom-Joon Cho, Address: (501-759) Department of Computer Engineering, Chosun University, 375 Seosuk-dong, Dong-gu, Gwangju, Korea, TEL: +82-62-230-7103, FAX: +82-62-233-6896, E-mail: bjcho@chosun.ac.kr

Receipt date: Apr. 14, 2011, Revision date: May 11, 2011
Approval date: May 16, 2011

† Department of Computer Engineering, Chosun University, Korea

and is deformed in 3D space. There are explicitly three-dimensional mesh models, too, for the simulation of 3D objects in facial animation.

In this paper we will introduce the mass-spring mesh model with simulated electric charge for analyzing the distortion of handwritten character patterns. In the literature there exist a number of studies on character shape deformation analysis. Among them Sin [4] work is most similar to ours in basic idea but differs in his relaxation algorithm based on highly unstable Euler algorithm and the tendency to fall into local optima of potential wells. The inclusion of affine transformation in this work is somewhat reminiscent of Wakahara's local and global affine transform for skeleton matching and shape normalization [5,6]. There are several other reports concerning distortion of spatial patterns but they limited to line structures or their methods are remote or irrelevant to the method to be proposed here [7-9].

In this paper we propose a robust method of rectangular mesh model consisting of mass-spring-dampers elements for the analysis and measurement of nonlinear shape variations of handwritten digit characters. The rest of the paper is organized as follows: Section 2 describes the basic idea of the mass-spring model. Section 3 discusses an interesting idea of simulated electric charge from an intensity image. Then Section 4 presents the description of the physics, the three types of basic forces acting on the mesh nodes. Section 5 describes the relaxation algorithm based on a numerical integration method called Runge-Kutta algorithm. Section 6 is about global affine transformation that supports the nonlinear distortion of the mesh model. Section 7 explains the test environments and test results step by step. Section 8 concludes the paper.

2. MASS-SPRING MODEL

A mass-spring model consists of a collection of point masses arranged into a two dimensional array and their interconnection to form a two dimensional mesh structure as shown in Figure 1. In this model each mass represents a local feature for the corresponding location in relation to neighboring masses. Here the location is defined only in the context of immediate neighbors. Hence, given a neighborhood, the current mass is independent of other masses beyond the neighborhood. This is the so-called Markovian property.

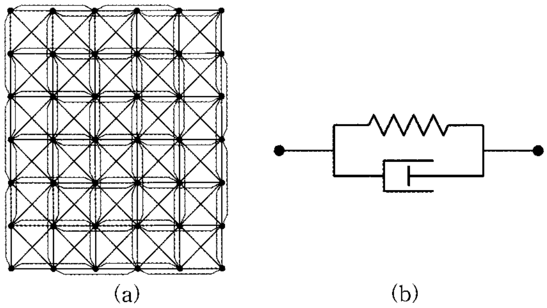


Fig. 1. Mesh model of mass-spring-damper. (a) mesh, (b) spring-damper model.

The detailed structure of the mesh is defined by three types of springs, each representing or constraining a specific type of mesh distortion. Figure 2 gives labeling of the springs around a mass at the center. The structural spring is a link that connects to four nearest neighbors, up, down, left, and right. The structural springs describes the two dimensional layout, but they can not maintain planar sheet structure because of shearing or row-wise sliding distortion. Thus we need the second type of shear springs. The third type of springs prevents the mesh from folding by connecting to a mass two-steps away. The set of all masses connected

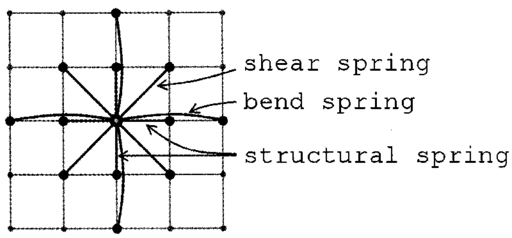


Fig. 2. Three types of springs.

to a central mass is defined as the neighborhood. It affects or conditions the displacement and motion of the central mass.

3. ELECTRIC CHARGE

Unlike cloth simulation which considers wind and air drag as well as gravity as the source of distortion, this study introduces the concept of local attraction and repulsion between input pixels and reference image pixels for the analysis of shape variation of hand-printed characters. Thus one distinguishing feature of the model is the definition of positive and negative electric charge for a grayscale or binary valued pixel, just like protons and electrons in atomic physics.

Let us consider a pixel I_{ij}, $1 \le i \le m$, $1 \le j \le n$ of a grayscale image $I \in [0,1]^{m \times n}$. We define the electric charge corresponding to the intensity I_{ij} as

$$q_{ij} = 2I_{ij} - 1 \tag{1}$$

This simulated charge is different from individual electron's charge in that it takes fractional charge. But the analogy holds when the image is binary. Given two images where one is superposed over the other with a sufficiently small distance, there occurs interaction between two pixels each belonging to a different image. Here we define attraction between similarly colored pixels, and repulsion between differently colored pixels. The strength depends on the strength of the charges as well as the distance between them. It will be defined in the next section.

4. FORCES IN AND ON MESH

Once the mesh is under the influence of external force or field, it begins to change shape nonlinearly. And due to physical constraints of masses and interconnection springs, it is no easy to predict the resulting shape. But we can at least describe the physics in the following equation

$$-c_i \dot{\mathbf{x}}_i + \sum_j \mathbf{g}_{ij} + \mathbf{f}_i = 0 \tag{2}$$

where $\mathbf{x}_i \in \Re^3$ is the position of mass i, and the three terms on the right-hand side are in the respective order, (1) velocity-dependent damping force working in the opposite direction of the motion, (2) internal force exerted on mass i by the spring between i and all others j, and (3) sum of all external forces (gravity or user-applied) acting on mass i.

The internal force given a displacement vector $\mathbf{x}_{ij}$ from mass i to j is

$$\mathbf{g}_{ij} = k\mathbf{x}_{ij} \tag{3}$$

where k is the elastic constant. The sum of them applied to mass i is

$$\mathbf{g}_i = \sum_j \mathbf{g}_{ij} \tag{4}$$

External force includes only electrostatic force caused by the electric charges between two pixels, each from different image, described by

$$\mathbf{f}_{ij} = k_Q \frac{q_i q_j}{\| \mathbf{r}_{ij} \|^3} \mathbf{r}_{ij} \tag{5}$$

where k_Q is the Coulomb constant that determines the scale of the resulting force. Then the equation of motion for entire system is described by concatenating N equations, each for a three dimensional vector x. Then the resulting equation describes a motion for a single $3N$ dimensional vector.

$$C\dot{U} + KU = F \tag{6}$$

where C and K are $3N \times 3N$ matrices corresponding to damping force, internal/stiffness force matrices, which are sparse albeit very large. The first one is diagonal while the second matrix is banded. Note that spring work is strictly localized with 12 connections to other masses. U is the composite vector of displacement of all masses, and U is the time derivative of U. Finally F is the composite vector of external forces. Typically the system equation evolves by the following first-order differential equation

$$\dot{\mathbf{v}} = m^{-1}(-C\mathbf{v} - K\mathbf{x} + \mathbf{f}) \tag{7}$$

$$\dot{\mathbf{x}} = \mathbf{v} \tag{8}$$

where $\mathbf{v}$ is the velocity vector of and $\dot{\mathbf{v}}$ the acceleration applied to a mass of m. A variety of numerical integration techniques are available to compute $\mathbf{x}$ and $\mathbf{v}$ as functions of time. In this study we take Runge-Kutta algorithm which highly stable and more accurate than the simple Euler algorithm [10].

5. RELAXATION

The formula for Euler method in numerical integration for solving ordinary differential equations is based on the linear approximation of a system function f at x_n. It has been applied to character shape analysis in [4]. The problem with this method is that it uses derivative information only at the beginning of the interval h of temporal step. This means that the step error is one power of h smaller than the correction of $O(h^2)$. Therefore the error accumulates quickly, which is especially true when h is not sufficiently small, and frequently shows very a unstable behavior, disrupting the mesh out of control.

By far the most popular method is the classical fourth-order Runge-Kutta formula [10]. It requires four evaluations of the system function per step h to make an accurate estimation at the next step x_{n+1}.

$$\begin{aligned} k_1 &= hf(x_n, y_n) \\ k_2 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_1}{2}) \\ k_3 &= hf(x_n + \frac{h}{2}, y_n + \frac{k_2}{2}) \\ k_4 &= hf(x_n + h, y_n + k_3) \\ y_{n+1} &= y_n + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) + O(h^5) \end{aligned} \tag{9}$$

6. AFFINE TRANSFORMATION

Handwritten characters are highly variable depending on the style and writing conditions like pen and paper. Often the distortion is too great involving simple scaling, rotation, translation, or even shearing. They are none other than a constrained form of linear transformation, or affine transformation. Of course the correct distortion of handwritten characters is nonlinear. But in this research the whole nonlinear distortion is modeled by a combination of affine transformation and nonlinear distortion modeling using a mesh. With this approach, the distortion of a character shape is analyzed by global affine transformation and then nonlinear analysis fined tuned to local features between the mesh and the input image.

Affine transformation can be determined by a set of feature points in both mesh (reference image) and the input image. Let us consider a set of feature points $X = \{\mathbf{x}_i = (x_{i1}, x_{i2}) : i = 1, \ldots, N\}$ from a mesh and $Y = \{\mathbf{y}_i = (y_{i1}, y_{i2}) : i = 1, \ldots, N\}$ from an input image. The points in the ordered pair of points from each set corresponds to each other so that we can write their relation by employing the homogeneous coordinate system as

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{i1} \\ x_{i2} \\ 1 \end{bmatrix} = \begin{bmatrix} y_{i1} \\ y_{i2} \\ 1 \end{bmatrix} \tag{10}$$

By changing the role of unknowns and using more than three points we can determine the system coefficients a_{ij}, $i = 1, 2$, $j = 1, 2, 3$, approximately based on the least squares error method. With the introduction of the global affine transformation we can improve the accuracy of the distortion analysis or settle down to a better configuration and less potential energy of the mesh.

7. EXPERIMENT

7.1 Mesh models

For the experiment we tested the improved relaxation method on the data set in [4]. The basic dataset includes isolated digit images as shown in

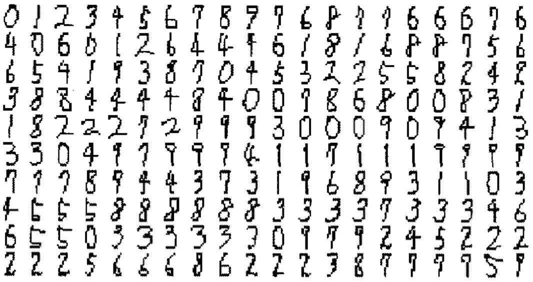


Fig. 3. Training data samples.

Figure 3. From these we computed the average images to design the mesh models as shown in Figure 4. Here each node is rendered into a small dot with its size or radius visualizing its darkness. Note that those big dark dots correspond to negative charge while the grid points with no dots have positive charge. These meshes will be put afloat above input test images and then start to distort their shape.

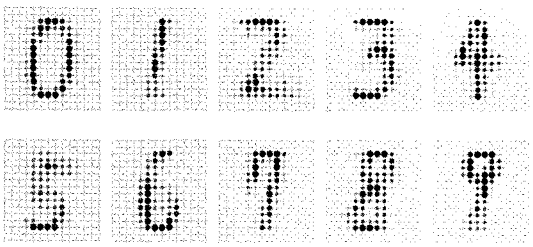


Fig. 4. Digit model meshes

7.2 Test data

The test data is different from those used in preparing the mesh models. For the following experiments, we used part of the numeral images from the famous CENPARMI dataset of Concordia University [11]. Figure 5 shows sample images of digit zero only. They differ not only in shaped but is size as well. As can be seen here handprint patterns in general are so variable that it is not easy to analyze the whole distortion in simple terms.

7.3 Preliminary test

The first test of the proposed method involves an image shown in Figure 6(a). It is overlaid by a mesh of digit zero as shown in Figure 6(b). Note that their shape of dark pixels does not match well to that of the underlying image pixels at first and that they have different size as well. The scale difference has been resolved by extending the springs of the reference mesh appropriately as shown in Figure 6(b). The mesh undergoes relaxation according to the relaxation algorithm described in Section 5. The final result of the step is shown in Figure 6(c) which shows not only the distortion of the mesh but also the displacement of each mass from the original location by a thick solid line segment. Note that there are dark nodes at the top that could not move downward to the dark pixels. This is can be attributed to partly the over distortion of the handwritten character and partly to the spring constraints settling down to a local optimum in the potential landscape.

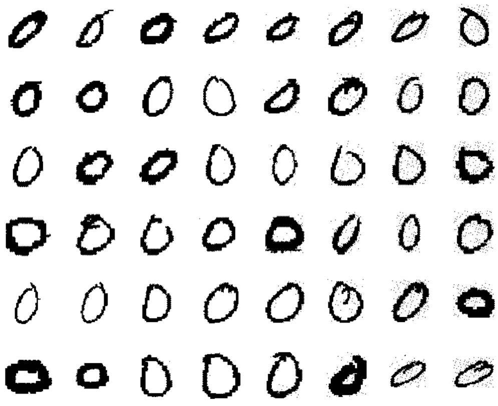
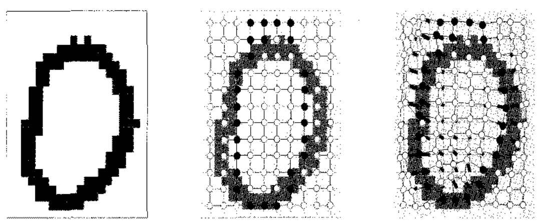


Fig. 5. Test images.

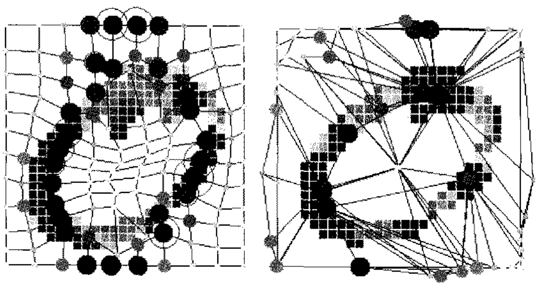
It is generally well-known that Euler method of



(a) (b) (c)

Fig. 6. (a) a test image, and a mesh (b) before relaxation and (c) after relaxation.

numerical integration is very simple but often unstable and the evaluation may go rampant and eventually out of control. Figure 7 shows two cases. Figure 7(a) shows a result similar to that Figure 6(c). It was obtained from a conservative and slow setting of constants and step sizes. On the other hand Figure 7(b) shows an unstable mesh that has masses oscillating wildly. In general, Euler method is sensitive to the values of model parameters and step sizes. Therefore it is not considered a preferred method.



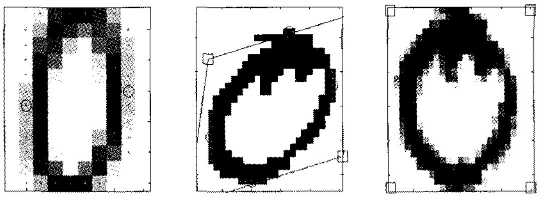
(a) (b)

Fig. 7. Sample results by Euler method. (a) normal, (b) unstable.

7.4 Global affine transformation

Affine transformation is a type of linear transformation which restricted to a combination of rotation, scaling, shearing, and translation. It is often used for image registration of superresolution studies. In the proposed method it is used as an approximate matching of images with a small number of distortion parameters. The residual distortions are nonlinear and highly local is analyzed by mesh relaxation. The affine transformation helps greatly in interpreting the large scale distortion and often indispensable for overall handwriting shape analysis.

Let us consider a reference image corresponding to a mesh such as Figure 8(a). Here we define a set of simple and robust features: the vertical maximum and the minimum points and the leftmost and the rightmost points from the dark stroke pixels as marked by small circles. Then the corresponding points in an input image are found by similar methods. See Figure 8(b). Using the four pairs of points, the affine transformation can be determined. Then we can determine the quadrilateral corresponding to the rectangle of Figure 8(a) in the transformed space as shown in Figure 8(b). By the inverse transformation, one can easily eliminate the effect of global affine distortion as shown in Figure 8(c). It is the closest form possible by a linear transformation.



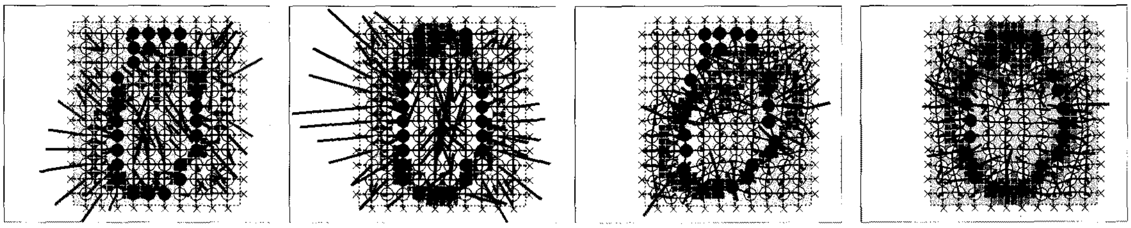
(a) (b) (c)

Fig. 8. Affine transformation. (a) mesh, (b) affine distortion, (c) correction result.

7.5 Relaxation

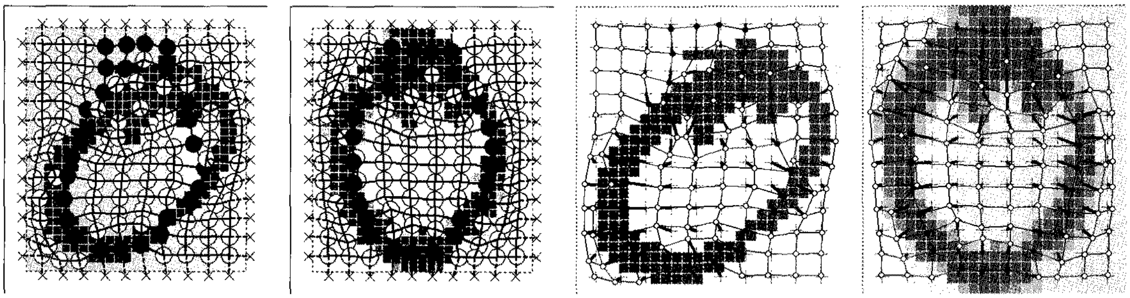
With or without affine transformation, the input mesh is put on top of the de-transformed image. Prior to this each pixel's intensity is converted into an electric charge. When the charged mesh is first applied, the mesh will receive a huge amount of stress from the image charges at the background. It is graphically shown in Figure 9(a) (without affine transformation) and 9(b) (with affine motion). The nodes with long bars are those under great pressure for motion away from the current location in the direction of the bars. Figures 9(c) and (d) show the resulting meshes after relaxation of 100 iterations. In these figures the bars became much shorter but with the introduction of the second bars in the opposite direction. The new bars represent the elastic force from the mesh springs. Note that the electric force and the elastic force are in good balance across the entire mesh, which means a near equilibrium.

When the force bars are removed from the fires, detailed nonlinear distortion is evident as shown



(a) (b) (c) (d)

Fig. 9. A mesh (a), (b) before and (c), (d) after relaxation.



(a) (b) (c) (d)

Fig. 10. Comparison of results by methods of (a),(c) relaxation only, and (b),(d) relaxation with affine transformation.

in Figure 10(a) and (b). When we forgo the affine transformation, Figure 10(a) follows. Note that there are several black nodes which are not aligned to dark stroke pixels at the top and right in the background image. With an affine transformation of the background image, the resulting nonlinear distortion is next to perfect. Refer to Figure 10(b). Figures 10(c) and (d) show the displacement vectors of the mesh nodes. With the displacement vectors, you can tell the distortion field of Figure 10(d) is much smoother and look natural, too. Note, however, that there are a few dark masses that are not aligned to dark underlying pixels. This implies that the introduction of affine transformation does not eliminate but reduce the chance of unacceptable local optimum matching.

The final analysis into the distortion is about the quantitative evolution of the meshes. Once a mesh is applied to an input image, the optimal mesh shape is found through a relaxation process with low potential energy. Figure 11 shows the change clearly. At first the electric potential is very high,

Fig. 11. Temporal evolution of per-mass potential energy relaxation.

but over time reduces and some of the loss is converted to elastic potential.

8. CONCLUSION

In this paper we proposed a new idea of mass-spring mesh relaxation with simulated electric charge. Global affine transformation for approximate global distortion analysis is another interesting feature of the proposed approach. In a word the proposed method is about interpreting image distortion via the combination of global affine transformation and local nonlinear distortion. It is quite interesting to view the nonlinear distortion vectors as obtained from the proposed method. This method can be applied to several other fields requiring detailed analysis and quantization of image distortion. Handwritten character analysis and recognition could be one of the most appropriate tasks for this method.

REFERENCES

[1] A. Vinciarelli, "A Survey On Off-Line Cursive Script Word Recognition," *Pattern Recognition*, Vol.35, pp. 1433-1446, 2002.

[2] B. Born and B Schunck, "Determining Optical flow," *Artificial Intelligence*, Vol.17, pp. 185-203, 1981.

[3] A. Baraff and D. Witkin, "Large Steps in Cloth Simulation," in Proc. SIGGRAPH, pp. 43-54, 1998.

[4] Bong-Kee Sin, "2D Pattern Deformation Analysis using Particle and Spring-Damper Mesh", *Journal of Korea Institute of Information Science and Engineers: Software and Applications*, Vol.32, No.8, pp. 769-780, 2005.

[5] T. Wakahara, "Shape Matching using LAT and Its Application to Handwritten Numeral Recognition," *IEEE TPAMI*, Vol.16, No.6, pp. 618-629, 1994.

[6] T. Wakahara and K. Odaka, "Adaptive Normalization of Handwritten Characters using Global/Local Affine Transformation," *IEEE TPAMI*, Vol.20, No.12, 1998.

[7] M. Revow, C. Williams, and G. Hinton, "Using Generative Models for Handwritten Digit Recognition," *IEEE TPAMI*, Vol.18, No.6, pp. 592-606, 1996.

[8] K.-W. Cheung, D.-Y. Yeung, and R. Chin, "A Bayesian Framework for Deformable Pattern Recognition with Application to Handwritten Character Recognition," *IEEE TPAMI*, Vol. 20, No.12, 1998.

[9] R. Webster, T. Nagasaki, T. Teramura, and M. Nakagawa, "Several Possibilities for Character Recognition Based on a Dynamic Model," in *The 4th IWFHR*, Taipei, pp. 423-430, 1994.

[10] M. Heath, *Scientific Computing*, p. 563, 2002.

[11] http://www.cenparmi.concordia.edu/

Beom-Joon Cho

received B.S. and M.S. degrees from Chosun University in 1980 and 1982, respectively. He received his Ph.D degrees in Dept. of Electrical Engineering from Hanyang University in 1988 and Electrical Engineering & Computer Science from KAIST in 2004. He is currently a professor in the Department of Computer Engineering at Chosun University. His research interests include the Pattern Recognition, Neural Network, Character Recognition, and Computer Vision.