International Journal of Image and Graphics

CONSTRUCTIVE RELAXATION MATCHING INVOLVING DYNAMICAL MODEL SWITCHING AND ITS APPLICATION TO SHAPE MATCHING

KEISUKE KAMEYAMA, KAZUO TORAICHI
Tsukuba Advanced Research Alliance, University of Tsukuba,
1-1-1 Tennodai, Tsukuba-shi, Ibaraki-ken 305-8577, Japan
E-mail: [kame, toraichi]@tara.tsukuba.ac.jp

YUKIO KOSUGI
Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku, Yokohama 226-8502, Japan
E-mail: kosugi@pms.titech.ac.jp

Received (received date)
Revised (revised date)

This paper introduces a novel approach for contour-based shape matching named as Constructive Relaxation Matching (CRM). Relaxation Matching (RM) introduced by Rosenfeld et al. is one of the standard method for quasi-optimally solving the local correspondences of the template and input images. RM relies on an energy-minimizing nature of the dynamical system to update the label assignment to the input objects. In image matching relying on a particular modeling method, apparently similar images can be judged as being quite distant, according to the nature of the modeling process and its outcome. In the proposed CRM, the modeling stage of a novel input image contour, conventionally done in the same procedure used for modeling the templates, will be included in the procedure of iterative relaxation matching. The model of the input will be dynamically constructed during relaxation, by unifying the pairs of objects having similar template label assignment probabilities. After describing the CRM procedures, the method is applied to simple shape matching problems demonstrating the ability to adaptively model the input image during relaxation. It is shown that the proposed CRM improves the object-label correspondence for evaluation of the image similarities in the following stages of shape matching applications.

Keywords: Probabilistic Relaxation; Shape Matching; Labeling; Model Switching; Dynamic Systems

1. Introduction

General shape matching task by contour characterization commonly consists of three major processing stages. They are namely (Fig. 1),

- Image (shape) modeling,
- Matching,
- Distance evaluation.

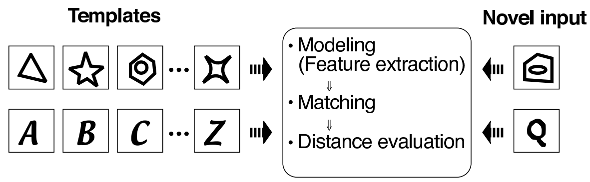


Fig. 1. General image/shape matching scheme.

In this work, we will assume that the images are binarized bitmaps. In the modeling stage, the pixels at the edge of levels (contour pixels) are extracted. Each closed contour is sectioned at points called *nodes*, sectioning the contour and making a collection of *contour segments*.

The contour segments generated by this process are called *labels* for the registered template images, and *objects* for the novel input images. The contour segments will represent the local feature of the image, and as a whole, they constitute the edge structure of the image. The segments may be further characterized by their coordinates, lengths, curvature and piecewise approximation functions.

The matching stage involves searching for the best correspondence of labels and objects. Relaxation Matching (RM) algorithm [1] is a variant of deterministic relaxation algorithms used for solving optimization problems, among which Hopfield neural networks [2] also belong to. Due to the common energy functional minimizing dynamics, RM is known to be very robust against noise, deformations and partial lack of information, where in image matching applications, all of such cases are common. RM have been used in various applications such as scene labeling [1] and handwritten character recognition [3] [4].

In this work, the set of objects that represent an image will be referred to as the *model*.

Although the contours of the template and the input images are usually processed by a common modeling operation, the outcomes are the descriptions of the contours in various models, namely models having different number of segments. These data are fed into the matching process of relaxation which relies heavily on the robustness of the operation (Fig. 2). Nevertheless, this can result in a very poor matching producing discrepancy measures that do not reflect the apparent similarity/discrepancy of the template and the input images.

In the top row of Fig. 3, three images of character “A” are shown. They are, (a) the reference “A”, (b) “A” in a different font (Slanted “A”), and (c) the reference “A” with noisy contours (Noisy “A”). The contours of these images were modeled by polygonal approximation used in [3] and [5]. The segments and the nodes of the modeled contours of the three images are shown in the bottom row ((d) - (f)).

Assume that the slanted and noisy “A” are matched against the reference “A” via the obtained models. In this case, it is highly likely that the slanted “A” will

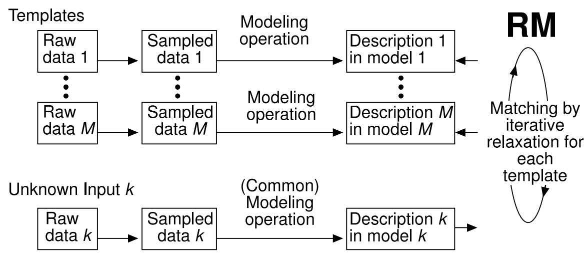


Fig. 2. Matching strategy of the conventional relaxation matching.

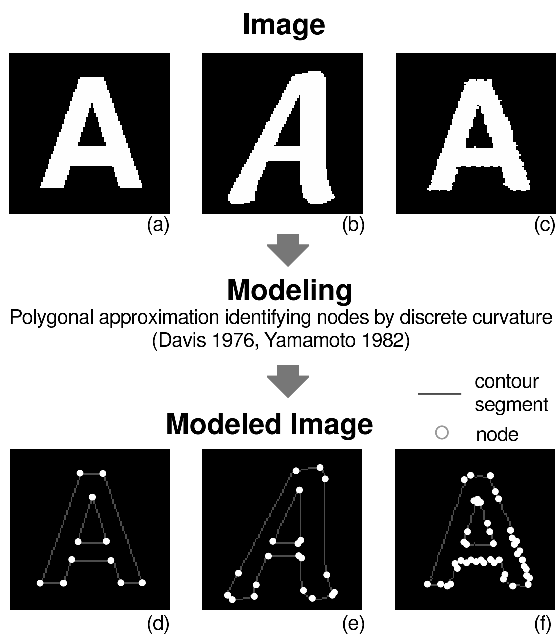


Fig. 3. Modeling example of three simple shapes. (a) The reference "A". (b) "A" in a different font (Slanted "A"). (c) The reference "A" with noisy contours (Noisy "A"). (d)(e)(f) The modeled images in contour segments and nodes.

be evaluated closer to the reference in comparison with the noisy "A", since it has better local correspondences of labels and objects.

There are several directions to solve this kind of problem. One way would be to enrich the segment types, incorporating piecewise arcs and polynomial curves in addition to line segments [4 6]. This enhancement is effective for characterizing smooth contours. However, for jagged noisy contours, its contributions are limited. Considering not only 1 to 1 object matchings, but 1 to n matchings can work to recover those segments that were modeled into two or more segments due to noise. Direct consideration of 1 to 2 segments are found in [3]. Also, use of higher-order compatibility in addition to the second-order measures, can contribute to find correspondences of three or more object-label subset [7]. However, sufficiently large n is not known in beforehand and the amount of required computation can be extremely large due to *the curse of dimensionality.*

Instead of defining a search domain of high dimensionality and trying to cover the whole, it can be more efficient *not* to define any search domain in the beginning, and try to determine the search space according to the data. In this work, an approach which adaptively re-models the input image according to the template, in the process of matching, is introduced. The model of the input image will be dynamically constructed in the process of iterative relaxation, where objects having similar responses to the labels of the existing template will be unified. By using this approach, the final matching score of the input image to the existing templates, will be evaluated via a different modeling, template by template, enabling a flexible consistency evaluation.

The structure of this paper is as follows. In Sec. 2, the conventional Relaxation Matching algorithm will be reviewed. In Sec. 3, the proposed Constructive Relaxation Matching algorithm will be introduced, and its application examples will be presented in Sec. 4. Finally, Sec. 5 concludes the paper with some discussions.

2. Relaxation Matching

In the conventional RM, a set of N objects $B = \{b_1, ..., b_N\}$ and a set of M labels $\Lambda = \{\lambda_1, ..., \lambda_M\}$ is initially assumed. The assignment of the labels to object i will be expressed in a vector of assignment probabilities as, $\boldsymbol{p}_i = [p_i(1), ..., p_i(M)]^T \in R^M$, where "$T$" denotes the transpose. The elements of $\boldsymbol{p}_i$ meet the usual condition of $\sum_{\lambda=1}^{M} p_i(\lambda) = 1$. The set of *object state vector* $\boldsymbol{p}_i$ for all the N objects will make *state vector* $\boldsymbol{p} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_N]^T \in R^{MN}$, which is the state vector of the whole dynamical system.

For updating the state $\boldsymbol{p}$, a symmetric matrix $R = [r_{ij}(\lambda, \mu)] \in R^{MN \times MN}$ is defined. The state vector will be updated according to,

$$\boldsymbol{p}(t+1) = \boldsymbol{f}(\boldsymbol{p}(t), \boldsymbol{q}(t)), \tag{1}$$

$$\boldsymbol{f}(\boldsymbol{p}, \boldsymbol{q}) = [f_1(1), \ldots, f_N(M)]^T, \tag{2}$$

$$f_i(\lambda) = q_i(\lambda)p_i(\lambda) / \sum_{\lambda'=1}^{M} q_i(\lambda')p_i(\lambda'), \tag{3}$$

and

$$\boldsymbol{q}(t) = R\boldsymbol{p}(t), \tag{4}$$

where t denotes the time. It has been shown [8 9] that the RM system with symmetric R will have a Liapunov (energy) function defined as,

$$-A(\boldsymbol{p}(t)) = -\boldsymbol{p}(t)^T R\boldsymbol{p}(t), \tag{5}$$

which is guaranteed to decrease by the iterative transition of $\boldsymbol{p}$ until one of the local minima of $-A(\boldsymbol{p})$ is reached.

When the above RM algorithm is applied to image shape matching, the template image will be modeled as a set of objects in particular spatial positioning. Each object in the template will be assigned a unique label to form the label set Λ above. The novel input image whose matching to the template is to be evaluated, will be modeled by the same process as in the template case to form the set of objects B. Matrix R is called the *compatibility coefficient matrix* whose element $r_{ij}(\lambda, \mu)$ reflects the *compatibility* of the situation when labels λ and μ are assigned to objects i and j, respectively. By using RM to the sets Λ and B, with the compatibility matrix R, the state vector $\boldsymbol{p}$ is expected to converge to one of the label assignments (states) that is at a local minimum of $-A(\boldsymbol{p})$. The value of function $A(\boldsymbol{p})$ can be a measure of the total consistency of the label assignments [7].

Upon using the RM procedure for classification, the value of $A(\boldsymbol{p})$ at the stable (converged) state of $\boldsymbol{p}$ can be used as the matching score, and the input will be classified to the class of the template image that resulted in the maximum $A(\boldsymbol{p})$ after convergence. As pointed out in the previous section, the modeling process of the input image can greatly affect the local label assignments to the objects, and also the final matching score ranking.

3. Constructive Relaxation Matching (CRM)

In this section, the proposed matching algorithm with dynamic modeling maneuvers will be presented. Instead of the conventional two-step procedure of modeling and relaxation matching, both stages will take place simultaneously, gradually constructing the model during the relaxation process as illustrated in Fig. 4. The proposed matching scheme will be referred to as Constructive Relaxation Matching (CRM) in the following.

The processing stage flow of the proposed algorithm is shown in Fig. 5. When compared with the conventional RM, three stages are added in the discrete-time global relaxation loop. Each processing stage will be detailed in the following. For the convenience of explanation, a simple matching example shown in Fig. 6 will be used. By using the polygonal contour modeling, the two shapes will be characterized

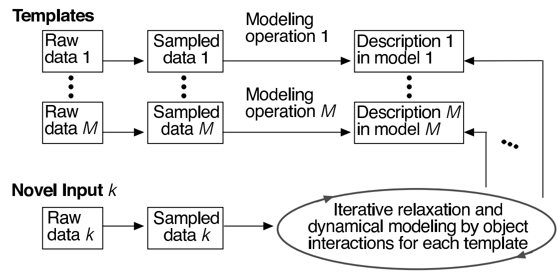


Fig. 4. Matching strategy of the proposed constructive relaxation matching.

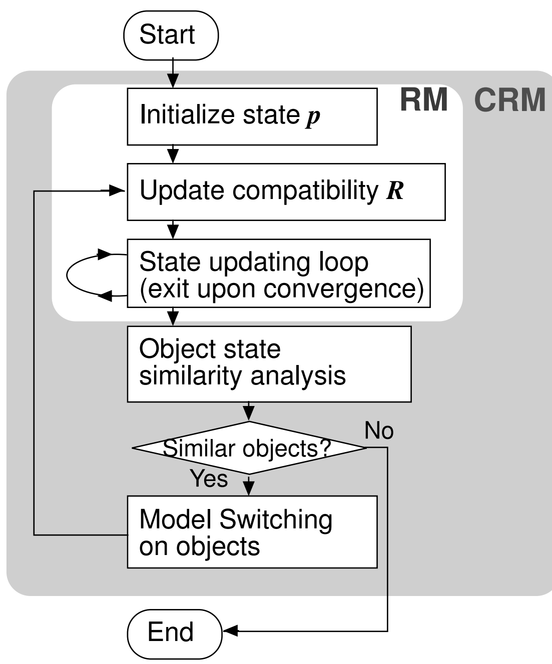


Fig. 5. The procedural flow of the proposed constructive relaxation matching.

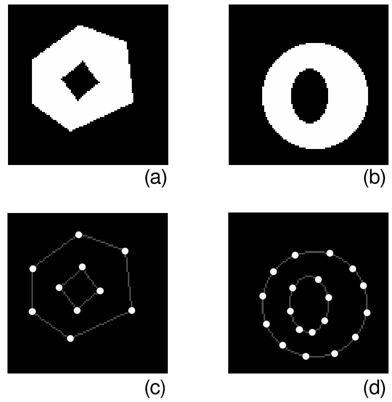


Fig. 6. Simple template-input pair. (a) Template. (b) Input. (c) Modeled template with 10 labels. (d) Modeled input with 19 objects.

in different models, as shown in Fig. 6(c) and (d), which can be an obstacle for evaluation of the apparent image similarity.

3.1. *Initial state and compatibility coefficient matrix*

Since all RM systems are deterministic dynamic systems, the initial state will determine the state of convergence. Therefore, it is important to choose a good initial state, which should be a reasonable labeling probability assignment obtained from the observations of the label and object sets.

Here, two elementary discrepancy factors between a label and an object are defined. First is the *distance* factor,

$$d_{ik} = min(d_{iks}, d_{ike}, d_{ikp}), \quad (i = 1, 2, \ldots, M, \quad k = 1, 2, \ldots, N) \tag{6}$$

which is a composite function of three Euclidean distances shown in Fig. 7. Subscripts s, e and p denote the start point, end point, and projection, respectively. Distance d_{ikp} is considered only when the normal projection of the object center falls on the label. Second is the *angle* factor θ_{ik} $(-\pi < \theta_{ik} \leq \pi)$ which is a signed measure of the differences in the directions of label and object edges.

Using the above measures, an element of the initial state $\boldsymbol{p}(0)$ is calculated as,

$$p_{ik}(0) = G(d_{ik} + \frac{\alpha}{\pi}|\theta_{ik}|; \sigma), \qquad (\alpha, \sigma : constant) \tag{7}$$

with $G(x;\sigma) = \exp(-x^2/\sigma^2)$. Also, an element of the compatibility coefficient matrix R is calculated as,

$$\begin{aligned} r_{ij}(k,l) \quad &= \quad G(|d_{ik} - d_{jl}| + \frac{\alpha}{\pi}|\theta_{ik} - \theta_{jl}|; \sigma). \qquad (8) \\ & \qquad (i, j = 1, 2, \ldots M, \quad k, l = 1, 2, \ldots N) \end{aligned}$$

Initial state vector $\boldsymbol{p}(0)$ and matrix R for the matching example are shown in Fig. 8.

Note however, that the calculations of the initial state and compatibility are defined here as one example, and that the idea of CRM is independent of these definitions.

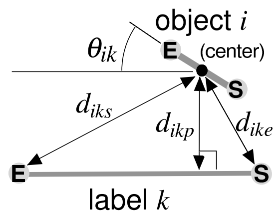


Fig. 7. Elementary discrepancy factors for calculating initial state and compatibility.

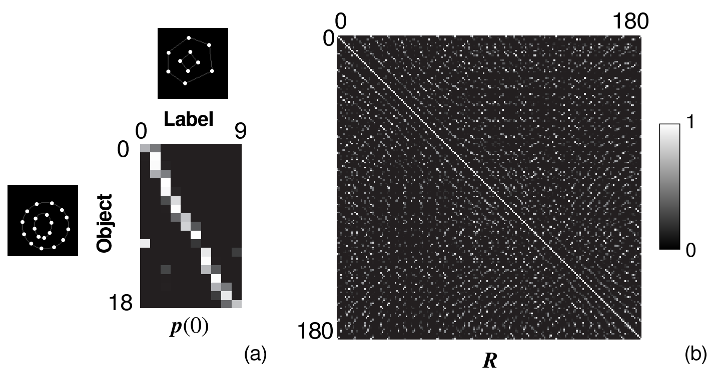


Fig. 8. (a) Initial state vector and (b) compatibility coefficient matrix for the matching case shown in Fig. 6, whose element's magnitudes shown in grayscale.

3.2. *State updating*

After initialization, the state vector $\boldsymbol{p}$ will be iteratively updated according to the formulas in Eqs. (1) - (4). The progress, together with the energy function $-A(\boldsymbol{p})$ for the matching example, is shown in Fig. 9.

3.3. *Model Switching*

In addition to the stages also found in the conventional RM, the possibilities of model switching (model size reduction of the input image), in accordance with the template, by means of object local unification are sought.

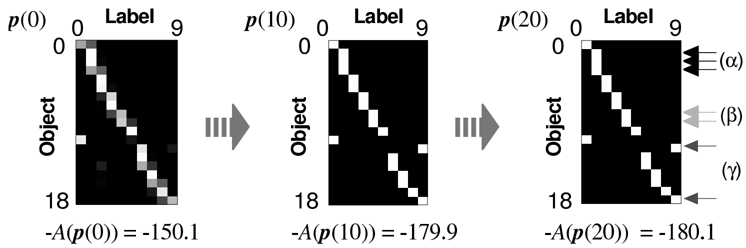


Fig. 9. The state vector and the energy function for the matching case in Fig. 6, at initial state, 10-th and 20-th iterations.Object sets shown by (α),(β) and (γ) are those with high object state similarities. Objects further satisfying neighboring conditions will be unified by fusion, a pair at a time.

Upon switching the model during relaxation, it is important that the progress of the relaxation process will be inherited by the new model. In [10] and [11], a scheme of layered neural network training involving a dynamic model alteration was introduced for efficient feature extraction and network model selection. This scheme named Model Switching (MS), is based on backpropagation training [12], but additionally enables an input-output map search in a way that the search domain is extended to multiple network models. The model crossover is done by switching the function (network) model at the instant which is possible to preserve the function acquired by the training up to that point. The local operation for switching the model in MS, is called *unit fusion*. The similarities of the unit response to the training set are evaluated, and pairs of units that have significant response correlation are unified to one unit. Also, It has been proved that this local operation effectively preserves the global progress of training [11].

A similar strategy can be taken to unify the objects in the input image, by evaluating their assignment probability vectors with the existing labels. If there are object pairs that have similar label probability assignments, further meeting the conditions such as being spatially adjacent objects in the input image, then, fusion of objects can take place.

In the object state similarity analysis stage of Fig. 5, the similarities of the object state vectors $\boldsymbol{p}_i (i = 1, \ldots, N)$ are evaluated as,

$$s_{ij} = \boldsymbol{p}_i^T \boldsymbol{p}_j. \tag{9}$$

Spatially adjacent objects i and j having similarity of $s_{ij} > s_{thres}$ will be fused to make a new object as shown in Fig. 10, where s_{thres} is a predefined constant. See object sets (α), (β) and (γ) in Fig. 9 for possible unification candidates in the example matching problem.

The states of the fused objects i and j are inherited by the new object i' as,

$$\boldsymbol{p}_{i'} = \frac{1}{2}(\boldsymbol{p}_i + \boldsymbol{p}_j) \tag{10}$$

This scheme of constructive process of nonparametric to parametric image matching also owes much to a biologically-inspired nonparametric method for image matching, which evaluate the local pixel correlations of the two bitmap images, for deriving a nonlinear mapping to match two similar images with nonlinear deformations [13].

After object fusion, the compatibility coefficient matrix R is to be recalculated for the new object set, and the relaxation process is to be restarted.

If the changes in the state vector converges, and if there are no more objects to be fused, the matching process will be stopped.

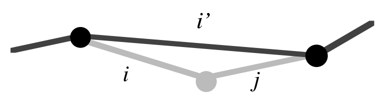


Fig. 10. Unification occurs for two adjacent objects with high object state similarity at a time.

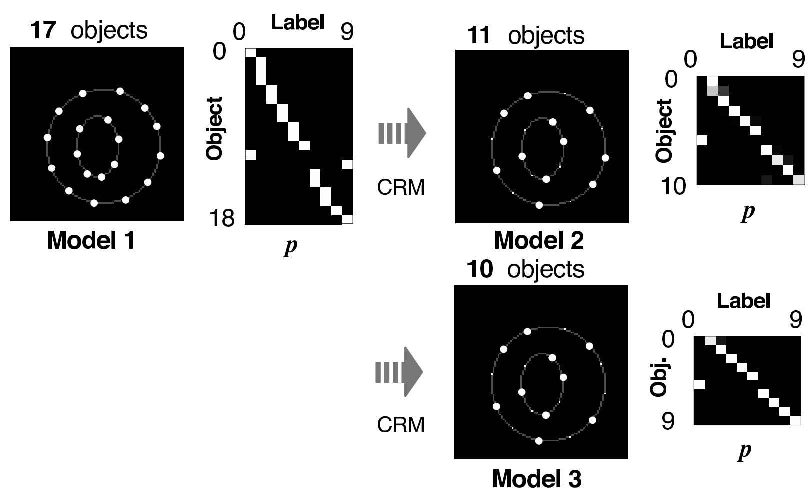


Fig. 11. Progress of model switching for the matching case shown in Fig. 6

4. Experimental Results

In this section, several results of applying the CRM algorithm will be presented focusing on the adaptation of the input image models. All results share the experimental conditions of, $\alpha = 30.0$, $\sigma = 10.0$ and $s_{thres} = 0.5$. All images were in 100×100 pixel dimensions. The templates were modeled via polygonal contour approximation.

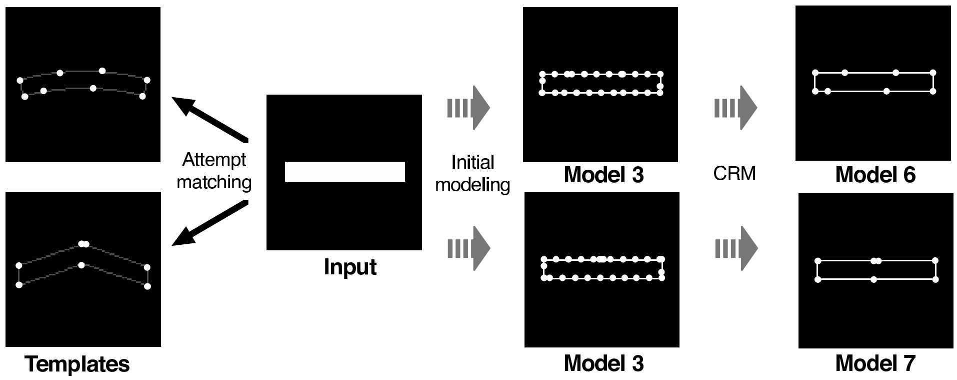


Fig. 12. A horizontal bar-shaped input matched against two different templates.

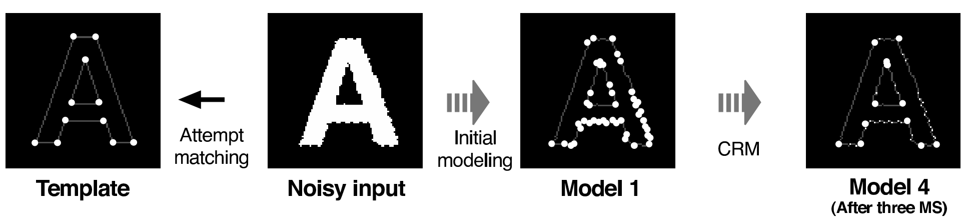


Fig. 13. The “A”-shape with noise in Fig 3 matched against the noiseless “A”.

4.1. *The simple matching example*

The result of the example used in previous sections is shown in Fig. 11. Starting from a 17-object-model, the input objects were gradually fused to a 10-object-model. Note that the objects are modeled to have better correspondences with the labels shown in Fig. 6(c).

4.2. *Difference in the modeling according to the template*

A bar-like input was matched against two different templates shown in the left hand side of Fig. 12. The input started the matching from a *nonparametric* status, namely with all contour pixels identified as nodes of minimal-length contour edges. In the 6-th and 7-th generations, the input was modeled as shown in the rightmost. It is observed that the modeling maneuver of the input will differ according to the model of the template.

4.3. *The noisy "A"*

The noisy "A" shown in Fig. 3(c) was matched against the reference "A" in Fig. 3(a). Note that in the 4-th model, most of the minimal sections produced due to contour noise are fused to make large contour sections. Also note that the sections and nodes have very similar positioning as those in the reference "A".

5. Discussion and Conclusion

In this paper, a method for image shape matching named Constructive Relaxation Matching (CRM) was introduced. Spurious discrepancies in image matching caused by the modeling process was mentioned, and the possibility of modeling the input image in accordance with the template image, was pointed out. Inspired by the Model Switching learning which is a neural network learning scheme with dynamical model alteration, the proposed CRM introduced to the relaxation matching algorithm, a dynamic model constructing maneuver applied to the input image. Through the experiments, it was shown that by using the CRM, the correspondence of the input image contours to those of the template could be resolved dynamically. This process enabled a modeling of the input image *customized* for each template, for better evaluation of the local image feature correspondences.

The CRM method introduced in this paper has a merit that it can be incorporated in the conventional RM method with simple addition of Model Switching procedures. This provides a way to override the limitations of searching the correspondences within the predetermined models. Moreover, additional paths of model alterations can be provided via label unification maneuvers applied to the template images, and also object and label *splittings*.

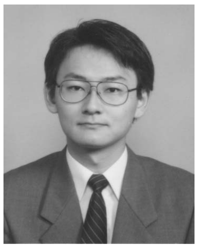
Acknowledgement

This work was supported by Grant-in-Aid for Encouragement of Young Scientists, Japan Society for the Promotion of Science (JSPS), project number 12750362.

References

1. A. Rosenfeld, R. A. Hummel, and S. W. Zucker. Scene labeling by relaxation operations. *IEEE Trans. Sys., Man and Cybern.*, SMC-6(6):420–433, 1976.
2. J. J. Hopfield and D. W. Tank. "neural" computation of decisions in optimization problems. *Biological Cybernetics*, 52:141–152, 1985.
3. K. Yamamoto. Recognition of handprinted kanji characters by relaxation matching. *IECE Trans.*, J65-D(9):1167–1174, 1982.
4. K. Toraichi, S. Ishiuchi, T. Horiuchi, K. Yamamoto, and H. Yamada. Recognition of handwritten kanji and hiragana characters by relaxation matching. *IEICE Trans.*, J73-D-II(9):1448–1457, 1990.
5. L. S. Davis. Shape matching usnig relaxation techniques. *IEEE Trans. Pattern Analysis and Machine Intelligence*, PAMI-1(1):60–72, 1979.
6. Paul Wing Hing Kwan, Keisuke Kameyama, and Kazuo Toraichi. Trademark retrieval by relaxation matching on fluency function approximated image contours. *Proc. IEEE Pacific Rim Conf. on Comm., Comp. and Sig. Pro.*, pages 255–258, 2001.
7. R. A. Hummel and S. W. Zucker. On the foundations of relaxation labeling processes. *IEEE Trans. of Pattern Analysis and Machine Intelligence*, PAMI-5(3):267–287, 1983.
8. Marcello Pelillo. On the dynamics of relaxation labeling processes. *Proc. IEEE Int'l. Conf. on Neural Networks*, 2:1006–1011, 1994.
9. Marcello Pelillo. Dynamics of nonlinear relaxation labeling processes. *Journal of Mathematical Imaging and Vision*, 7(4):309–323, 1997.
10. K. Kameyama and Y. Kosugi. Model switching by channel fusion for network pruning and efficient feature extraction. *Proc. Int'l Joint Conf. on Neural Networks 1998*, pages 1861–1866, 1998.
11. K. Kameyama and Y. Kosugi. Neural network model switching for efficient feature extraction. *IEICE Trans. on Inf. and Sys.*, E82-D(10):1372–1383, 1999.
12. D.E. Rumelhart, J. L. McClelland, and the PDP Research Group. *Parallel distributed processing*. MIT Press, 1986.
13. Y. Kosugi, M. Sase, H. Kuwatani, N. Kinoshita, T. Momose, J. Nishikawa, and T. Watanabe. Neural network mapping for nonlinear stereotactic normalization of brain mr images. *Journal of Computer Assisted Tomography*, 17(3):455–460, 1993.

Photo and Bibliography



Keisuke Kameyama received the B.E., M.E., and Dr. Eng. degrees in electrical engineering (1989), precision machinery engineering (1991), and computer science (1999), respectively, all from Tokyo Institute of Technology, Tokyo, Japan. From 1992 to 2000 he worked as a Research Associate at the Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology. In 2000, he joined the Tsukuba Advanced Research Alliance (TARA), University of Tsukuba, where he currently works as an assistant professor. His research interests include pattern recognition, signal processing using higher-order statistics, and neural networks. Dr. Kameyama is a member of IEEE and IEICE.

Kazuo Toraichi received the M.E. and Ph.D. degrees in electronics engineering from Hokkaido University, Japan, in 1968 and 1971, respectively. From 1971 to 1980, he worked in the Electro Technical Laboratory of MITI. In 1980, he became an Associate Professor at the Institute of Information Sciences and Electronics, University of Tsukuba, Ibaraki, Japan, where he is currently a Multimedia Aspect Professor in the Center for Tsukuba Advanced Research Alliance, University of Tsukuba. His main fields of research are wisdom systems and medical electronics. Dr. Toraichi was the recipient of the "Fluency Multimedia Project" awarded by The Ministry of Education, Culture, Sports, Science and Technology of Japan, in 2002 for the promotion of the industry-academic-government cooperation. Dr. Toraichi is a member of the Institute of Electrical Engineers of Japan.

Yukio Kosugi received the B.E. degree from Shizuoka University in 1970, and the M.S. and Dr. Eng. degrees from Tokyo Institute of Technology in 1972 and 1975, respectively, both in electronics. From 1975 to 1985 he was with the Research Laboratory of Precision Machinery and Electronics, at Tokyo Institute of Technology. During 1985 to 1999, he served as an associate professor at the Interdisciplinary Graduate School of Science and Engineering, of the same institute. During fiscal years 1998-2000, he conducted the project of "R&D of Brain-like Information Systems" at the Frontier Collaborative Research Center, Tokyo Institute of Technology. He is currently a Professor at the Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology and also serving as a visiting professor at Tsukuba Advanced Research Alliance of Tsukuba University. His research interests include information processing in the nervous system and neural network-aided image processing in medical and geographic fields. Dr. Kosugi is a member of IEEE and The Institute of Electronics, Information and Communication Engineers Japan.