The Third AFSS(1998), 731－739

An Efficient Learning Rule of Simple PR systems

Alan M. N. Fu and Hong Yan
Department of Electrical Engineering
University of Sydney, NSW 2006, Australia
phone: +61 2 9351-4824
fax: +61 2 9351-3847
e-mail: alanf@ee.usyd.edu.au
Gi Y. Lim
Department of Control & Instrumentation Engineering
Taejon National University of Technology, Korea
fax: +82 42-625-1485
e-mail: limgy@hyunam.tnut.ac.kr

Abstract

The probabilistic relaxation (PR) scheme based on the conditional probability and probability space partition has the important property that when its compatibility coefficient matrix (CCM) has uniform components it can classify m-dimensional probabilistic distribution vectors into different classes. When consistency or inconsistency measures have been defined, the properties of PRs are completely determined by the compatibility coefficients among labels of labeled objects and influence weights among labeled objects. In this paper we study the properties of PR in which both compatibility coefficients and influence weights are uniform, and then a learning rule for such PR system is derived. Experiments have been performed to verify the effectiveness of the learning rule.

Key words: Probabilistic relaxation, Compatibility coefficient matrix, Uniform components

1 Introduction

The relaxation labeling technique was first proposed by Rosenfeld *et al.* [1] to deal with ambiguity and noise in vision systems. Relaxation methods are iterative parallel procedures for producing reasonable labels of objects based on their incomplete information. Probabilistic relaxation (PR) methods have been successfully applied to many image processing tasks, such as scene labeling [1-3], pixel labeling [4-5], shape matching [6-7], line and curve enhancement [8-9], handwritten character recognition [10-12], breaking substitution ciphers [13], optical flow - template matching [14] and image segmentation [15]. Efforts have also been made towards the understanding of the properties of the method from mathematical analysis [16-24]. The essential issues in a PR system include the following four aspects: the measures of consistency or inconsistency among initial certainty measures and compatibility coefficients, the compatibility coefficients among labels of labeled objects, the influence weights among labeled objects, and an updating equation of the certainty measures. When measures of consistency or inconsistency and updating equations are given, the properties of the PR system are completely determined by the compatibility coefficients and influence weights. The measures of consistency in the PR [23] are defined exactly based on Bayes' formula and a probability space partition. So the PR in [23] overcomes the flaws that arise in traditional PRs, such as those in [1] [21]. The dynamics of the PR system in a general case have been studied theoretically in [23]. However, dynamics of the PR system in some special cases, such as both influence weights and compatibility coefficients being uniform components, are important in practical applications because a procedure for setting up a learning rule of the system is usually begun with compatibility coefficients being uniform components. In this paper, we study the dynamics theoretically of the PR system in which both influence weights and compatibility coefficients are uniform. For convenience such systems are referred to as uniform systems.

The paper is organized as follows. A brief review of PR systems and their some main properties in general cases is presented in Section 2. In Section 3 the dynamics and learning rule of simple PR systems are studied. Experiments to verify the learning rule is shown in Section 4 and the conclusion is made in Section 5.

2 The Elemental Properties of PR Systems

The PR system considered in this paper is one given in [23], in which the consistency or inconsistency measures between initial certainty measures and compatibility coefficients are defined exactly based on Bayes' formula and a probability space partition. The PR system shows good dynamics when its CCM has uniform components. Some main properties of the dynamics will be reviewed below. The notations used here are the same as those in [23].

Let n denote the total number of objects to be labelled, and $\mathbf{p}_i^{(0)}$ be the m-dimensional probabilistic vector associated with the ith object, where m is the total number of labels that each labeled object can be assigned. The λth component $p_i^{(0)}(\lambda)$ of $\mathbf{p}_i^{(0)}$ is the probability that the ith object is assigned the labeling value being λ initially. Let $\mathbf{p}_i^{(t)}$ be the m-dimensional probabilistic vector associated with the ith object at the tth step of an iterative process. The PR system is given by the following updating equation:

$$p_i^{(t+1)}(\lambda) = \frac{p_i^{(t)}(\lambda) + p_i^{(t)}(\lambda)v_i^{(t)}(\lambda)}{R_i^{(t)}} \tag{1}$$

where

$$R_i^{(t)} = \sum_{\lambda'=1}^{m} \left(p_i^{(t)}(\lambda') + p_i^{(t)}(\lambda')v_i^{(t)}(\lambda')\right) \tag{2}$$

and $v_i^{(t)}(z) = p_i^{(t)}(z) - s_i^{(t)}(z)$ is the controlling variable of the system, while

$$s_i^{(t)}(z) = \sum_{j=1}^{n} d_{ij}\left(\sum_{\lambda'=1}^{m} c_{ij}(z, \lambda')p_j^{(t)}(\lambda')\right), \tag{3}$$

where d_{ij} is the weight of the influence on the ith object label from the jth object label and satisfys: $0 \le d_{ij} \le 1$ and $\sum_{j=1}^{n} d_{ij} = 1$, $c_{ij}(q, r)$ represents the compatibility measure of object i having label q when object j having label r and satisfys: $0 \le c_{ij}(q, r) \le 1$ and $\sum_{q=1}^{m} c_{ij}(q, r) = 1$.

The following two theorems are the main theoretical results in [23].

Theorem 1 The updating rule in Eq. (1) can be rewritten as

$$p_i^{(k+1)}(\lambda) = p_i^{(k)}(\lambda) + p_i^{(k)}(\lambda)\frac{b_i^{(k)}(\lambda)}{R_i^{(k)}} \tag{4}$$

where

$$b_i^{(k)}(\lambda) = q_i^{(k)}(\lambda)v_i^{(k)}(\lambda) - \sum_{z\neq\lambda}^{m} p_i^{(k)}(z)v_i^{(k)}(z), \tag{5}$$

$$q_i^{(k)}(\lambda) = \sum_{z\neq\lambda}^{m} p_i^{(k)}(z). \tag{6}$$

Theorem 2 Assume that quantity N_i denotes the total number of labels λ which satisfy $v_i^{(0)}(\lambda) > 0$, i.e., $N_i = \sum_\lambda Y_i(\lambda)$ where

$$Y_i(\lambda) = \begin{cases} 1 & \text{if } v_i^{(0)}(\lambda) > 0 \\ 0 & \text{otherwise}. \end{cases} \tag{7}$$

Obviously, $0 \leq N_i < m$. For a given i,

(1) if N_i=0, then $p_i^{(t)}(\lambda) = p_i^{(0)}(\lambda)$, $\lambda = 1, \ldots, m$, $k \geq 0$;

(2) if N_i=1, i.e., there is only one label denoted by λ^* such that $v_i^{(0)}(\lambda^*) > 0$, and if

$$\mathrm{Min}\{c_{ij}(\lambda, \lambda^*) : \lambda \neq \lambda^*\} \geq \mathrm{Max}\{c_{ij}(\lambda, \lambda') : \lambda' \neq \lambda*\} \tag{8}$$

then
(i) $p_i^{(t)}(\lambda^*)$ converges to 1 as k increases infinitely,
(ii) $p_i^{(t)}(\lambda)$ converges to 0, for $\lambda \neq \lambda^*$ as k increases infinitely.

If labeled objects are independent, i.e., $d_{ii} = 1$ and $d_{ij} = 0$ as $i \neq j$, then Eq. (3) can be simplified and rewritten as

$$s_i^{(t)}(z) = \left(\sum_{\lambda'=1}^{m} c_{ii}(z, \lambda') p_i^{(t)}(\lambda') \right). \tag{9}$$

When the CCM of the system is selected to have uniform components, then Eq. (9) is further simplified and we have

$$s_i^{(t)}(z) = \frac{1}{m} \sum_{\lambda'=1}^{m} p_i^{(t)}(\lambda') = \frac{1}{m}. \tag{10}$$

Note that to derive the last equal relation in Eq.(10) we have used the relationship $\sum_{\lambda=1}^{m} p_i^{(t)}(\lambda) = 1$. Two important properties of dynamics of the PR system are [23]:

(I) If the PR system is given by Eqs. (1), (2) and (10), then m-dimensional probabilistic distribution vectors except a few are separated into m different groups. In order to show results graphically, the experimental results with $m = 3$ are given in Fig. 1 in which 450 3-dimensional probabilistic distribution vectors except few denoted by "?" have been classified into three classes denoted by "1", "2" and "3", respectively.

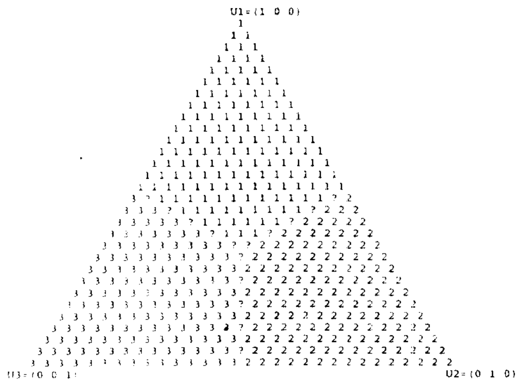

Figure 1: The initial probabilistic distribution vectors except a few are classified into three classes, i.e, U1, U2 and U3, under the PR system when the CCM has uniform components. Where U1=(1 0 0), U2=(0 1 0) and U3=(0 0 1) represent three basic unit vectors of 3-dimensional space, respectively.

(II) If an identical matrix is used instead of the CCM with uniform components in the case (I), then the PR system loses entirely the ability of classification. The experimental results are shown in Fig. 2 in which "*" means that the initial probabilistic distribution vector never changes under the PR system in this case.

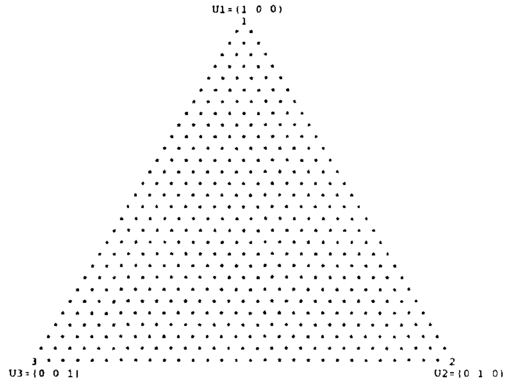

Figure 2: None of initial probabilistic distribution vectors is changed under the PR process when the CCM is an identical matrix.

3 Dynamics and Learning Rule of Simple PR Systems

A PR system with both the influence coefficients and compatibility coefficients being uniform is referred to as an uniform PR system. For uniform PR systems, the neighborhood support measure of the ith object at tth iteration defined by Eq. (3) now is replaced with

$$s_i^{(t)}(z) = \frac{1}{nm}\sum_{j=1}^{n}\sum_{\lambda'=1}^{m} p_j^{(t)}(\lambda') = \frac{1}{m}. \tag{11}$$

Based on the updating equation (1), the dynamics of PR systems is entirely ditermined the controlling variable $\nu_i(\lambda)^{(t)} = p_i^{(t)}(\lambda) - s_i^{(t)}(\lambda)$. From Eqs.(10) and (11) we have the following theorem

Theorem 3 Uniform PR systems and the PR systems with the labeled objects being independent and compatibility coefficients being uniform possess the same dynamics.

For convenience, an uniform PR and a PR with the labeled objects being independent and compatibility coefficients being uniform is known as a simple system. Because the condition given by Eq. (8) is true for simple systems, based on Theorem 2 and Eqs. (10) and (11) we obtain the following result.

Theorem 4 Suppose that the PR is a simple system and let $\mathrm{p}_i^{(0)}$ be an m-dimensional probabilistic vector. If the initial feature-factor $N_i = 1$, we assume $p_i^{(0)}(\lambda^*) > \frac{1}{m}$, i.e., $\nu_i^{(0)}(\lambda^* > 0)$, then

(i) $p_i^{(k)}(\lambda^*)$ converges to 1 as k tends to infinity,

(ii) $p_i^{(k)}(\lambda)$ converges to 0, for $\lambda \neq \lambda^*$ as k tends to infinity. In other words, p_i converges to the basic unit vector U_{λ^*} with λ^*th component being 1.

If $N_i > 1$ then the stable state $\bar{\mathrm{p}}_i$ of $\mathrm{p}_i^{(0)}$ may be a basic unit vector or not under the simple PR system. We are interested in the last case. We focus on how to adjust the compatibility coefficients to make $\mathrm{p}_i^{(t)}$ converge to a specific basic unit vector. Suppose that there are J nonzero components of $\bar{\mathrm{p}}_i$, say that $\bar{p}_i(\lambda_j) \neq 0$, $j = 1, \cdots, J$. For $\mathrm{p}_i^{(0)}$ we assume that the simple PR system arrive at stable state at kth steps, so $\mathrm{p}_i^{(l)} = \bar{\mathrm{p}}_i$ when $l \geq k$, i.e., $p_i^{(l)}(\lambda_j) = \bar{p}_i(\lambda_j)$, $j = 1, \cdots, J$ and $p_i^{(l)}(\lambda) = 0$ for $\lambda \neq \lambda_j$ when $l \geq k$. Based on Theorem 1, we have

$b_i^{(l)}(z)=0,\ z=1,\cdots,m$ when $l\geq k$. So

$$b_i^{(l)}(\lambda_j)=q_i^{(l)}(\lambda_j)v_i^{(l)}(\lambda_j)-\sum_{z\neq\lambda_j}^{m}p_i^{(l)}(z)v_i^{(l)}(z)=0,\ j=1,\cdots,J,\ \text{ as }\ l\geq k. \tag{12}$$

Namely,

$$\begin{aligned} b_i^{(l)}(\lambda_j) &= q_i^{(l)}(\lambda_j)[p_i^{(l)}(\lambda_j)-s_i^{(l)}(\lambda_j)]-\sum_{z\neq\lambda_j}^{m}p_i^{(l)}(z)[p_i^{(l)}(z)-s_i^{(l)}(z)] \\ &= q_i^{(l)}(\lambda_j)[p_i^{(l)}(\lambda_j)-s_i^{(l)}(\lambda_j)]-\sum_{z\neq\lambda_j}^{m}(p_i^{(l)}(z))^2+\sum_{z\neq\lambda_j}^{m}p_i^{(l)}(z)s_i^{(l)}(z) \\ &= 0,\ j=1,\cdots,J,\ \text{ as } l\geq k \end{aligned} \tag{13}$$

where

$$q_i^{(k)}(\lambda)=\sum_{z\neq\lambda}^{m}p_i^{(k)}(z). \tag{14}$$

What follow is to find a rule for adjusting the uniform compatibility coefficients to achieve that there is only one label, say λ_s such that $p_i^{(l)}(\lambda_s)$ converges to 1 while $p_i^{(l)}(\lambda_j)$ converges to 0, for $j=1,\cdots,m,\ j\neq s$. Let C_{ij} denote the compatibility coefficient matrix (CCM) of ith and jth labeled objects. We separate n^2 CCMs into two groups. One includes $C_{ii},\ i=1,\cdots,n$ and another consists of $C_{ij},\ i\neq j,\ i,j=1,\cdots,n$. Note that the CCMs in second group are not necessary for simple PR systems. The n CCMs in first group can be selected as a same CCM denoted by C with elements $c(\lambda,\lambda'),\ \lambda,\lambda'=1,\cdots,m$. We will present a learning rule in which we only need to adjust C. For simple PR system the neighborhood support measure $s_i^{(l)}(z)$ of object i at lth iteration is equal to $\frac{1}{m}$ (see Eq. (11)). In order to derive a learning rule for the PR systems, $s_i^{(l)}(z)$ can be represented as

$$s_i^{(l)}(\lambda_j)=\frac{1}{n}c(\lambda_j,\lambda_j)p_i^{(l)}(\lambda_j)+\frac{1}{n}\sum_{\lambda_j'\neq\lambda_j}^{m}c(\lambda_j,\lambda_j')p_i^{(l)}(\lambda_j')+\frac{n-1}{mn}. \tag{15}$$

Thus, $s_i^{(l)}(\lambda_j)$ reduces and some $s_i^{(l)}(\lambda),\ \lambda\neq\lambda_j$ increase if $c(\lambda_j,\lambda_j)$ is set to be zero rather than $\frac{1}{m}$. This is because $\sum_{\lambda=1}^{m}s_i^{(l)}(\lambda)=1$. Based on Eq. (13) $b_i^{(l)}(\lambda_j)>0$ if $c(\lambda_j,\lambda_j)=0$, while other $b_i^{(l)}(\lambda_i)\leq 0$, as $i\neq j$. Hence, if $\mathrm{p}_i^{(l)}$ is considered to be initial certainty support then its initial factor N_i is 1 under the adjusted PR system. Using Theorem 4 we know that $\mathrm{p}_i^{(l)}$ converges to basic unit vector U_{λ_j} when l tends to infinity. Note that $\sum_{z=1}^{m}c(z,\lambda)=1$, some $c(z,\lambda),\ z\neq\lambda_j$ and $\lambda\neq\lambda_j$ must be increase when $c(\lambda_j,\lambda_j)$ is set to be zero. Based on the above results we have the following learning rule for uniform PR systems.

The Learning Rule If an initial certainty support, i.e., a m-dimensional probabilistic vector $\mathrm{p}_i^{(0)}$ should converge to a basic unit vector Uk under the uniform PR system, but it does not, then

(a) set $c(k,k)=0$;
(b) set $c(j,k)=\frac{2}{m}$, where the index j satisfies $p_i^{(0)}(j)=\mathrm{Min}\{p_i^{(0)}(z):\ z\neq k,\ z=1,\cdots,m\}$;
(c) other components of CCM remain being $\frac{1}{m}$.

If there is another initial certainty support $\mathrm{p}_l^{(0)},\ l\neq i$, that should converge to a basic unit vector $Uq,\ q\neq k$ under the uniform PR system, but also, it does not converge to expected U_q, then to follow the same procedures, i.e.,

(a') set $c(q,q)=0$;
(b') set $c(r,q)=\frac{2}{m}$, where the index r satisfies $p_i^{(0)}(r)=\mathrm{Min}\{p_i^{(0)}(z):\ z\neq q,\ z=1,\cdots,m\}$;
(c') other components of CCM keep being $\frac{1}{m}$ except $c(k,k)=0$ and $c(j,k)=\frac{2}{m}$.

Table 1: The 7-dimensional initial probability vectors of models

	$\mathbf{p}_1^{(0)}$	$\mathbf{p}_2^{(0)}$	$\mathbf{p}_3^{(0)}$	$\mathbf{p}_4^{(0)}$
1	0.100	0.200	0.091	0.091
2	0.100	0.100	0.364	0.273
3	0.000	0.000	0.000	0.000
4	0.300	0.300	0.091	0.091
5	0.200	0.100	0.000	0.000
6	0.200	0.100	0.273	0.273
7	0.100	0.200	0.182	0.182

Table 2: The stable states of initial probability vectors of models under the uniform PR

	$\mathbf{p}_1^{(*)}$	$\mathbf{p}_2^{(*)}$	$\mathbf{p}_3^{(*)}$	$\mathbf{p}_4^{(*)}$
1	0.00	0.50	0.00	0.00
2	0.00	0.00	1.00	0.50
3	0.00	0.00	0.00	0.00
4	1.00	0.50	0.00	0.00
5	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.50
7	0.00	0.00	0.00	0.00

Because the adjustments of CCM for $\mathrm{p}_i^{(0)}$ and $\mathrm{p}_l^{(0)}$ are carried on different columns of CCM, so the later adjustment does not change the stable state of $\mathrm{p}_i^{(0)}$, i.e., $\mathrm{p}_i^{(0)}$ converges to Uk. Thus, we can adjust a CCM to achieve that more than two m-dimensional probabilistic vectors which are different in essence converge to different basic unit vectors.

4 Experiments

Experiments have been performed to verify the proposed learning rule for training uniform PR systems. The four 7-dimensional probabilistic vectors shown in Table 1 are associated with the four model contours shown in Figure 3.

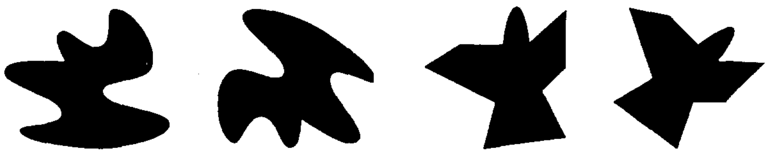

Figure 3: The model contours used in experiments.

Let $\mathbf{p}_i^{(*)}$ denote the stable state of $\mathbf{p}_i^{(k)}$, i.e., $\mathbf{p}_i^{(*)}$ is the limit of $\mathbf{p}_i^{(k)}$ when k tends to infinity. $\mathbf{p}_2^{(*)}$ should be basic unit vector $U4$ as $\mathbf{p}_1^{(*)}$ is and $\mathbf{p}_4^{(*)}$ should be basic unit vector $U2$ as $\mathbf{p}_3^{(*)}$ is, but they are not (see Table 2). Based on the learning rule we set $c(2, 2) = 0$, $c(3, 2) = 2/7$, $c(4, 4) = 0$ and $c(3,, 4) = 2/7$. The desired compatibility coefficient matrix denoted by $\mathbf{C}^*$ is

Table 3: The stable states of initial probability vectors of models under the PR with C^*

	$\mathbf{p}_1^{(*)}$	$\mathbf{p}_2^{(*)}$	$\mathbf{p}_3^{(*)}$	$\mathbf{p}_4^{(*)}$
1	0.00	0.00	0.00	0.00
2	0.00	0.00	1.00	1.00
3	0.00	0.00	0.00	0.00
4	1.00	1.00	0.00	0.00
5	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00

yielded

$$\mathbf{C}^* = \begin{pmatrix} 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/7 & 0.0 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/7 & 2/7 & 1/7 & 2/7 & 1/7 & 1/7 & 1/7 \\ 1/7 & 1/7 & 1/7 & 0.0 & 1/7 & 1/7 & 1/7 \\ 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \\ 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 & 1/7 \end{pmatrix}$$

The stable states of the four initial probabilistic vectors related to the four model contous under the PR with C^* are given in Table 3. The experiments show that the fresh contours shown in Fig. 4 are classified correctly based on their initial probabilistic vectors under the PR system with the adjusted compatibility coefficient matrix $\mathbf{C}^*$.

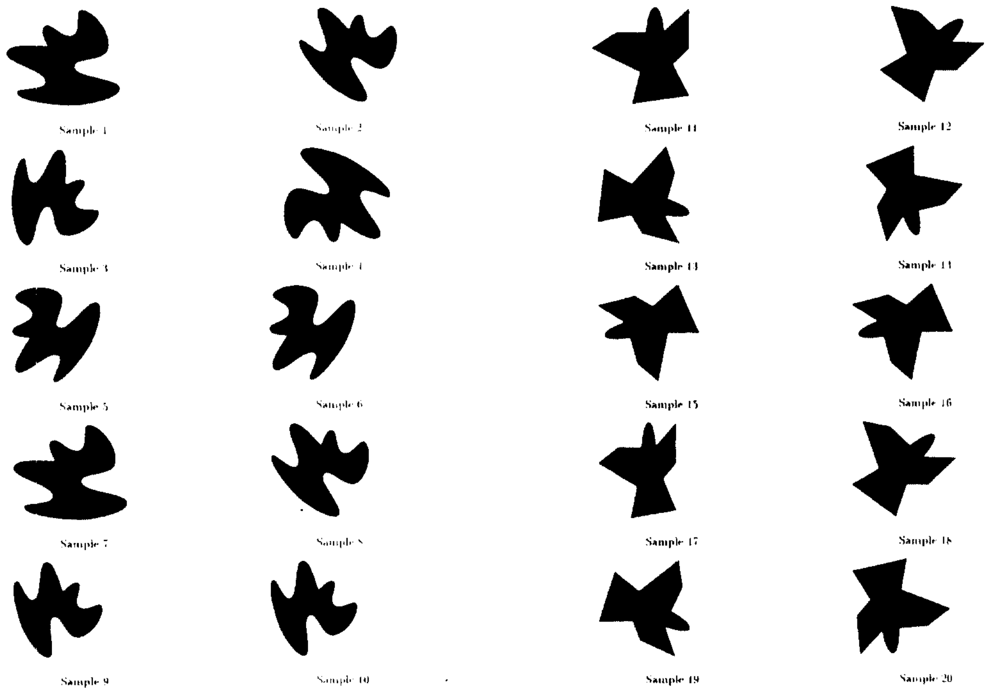

Figure 4: The sample contours used in experiments.

5 Conclusions

In this paper we have shown that the PR systems with uniform compatibility coefficients when labeled objects being independent and uniform PR systems have same dynamics, and therefore they have same learning rule. A simple learning rule for the two kinds of PR systems has also been proposed. The experiments have been shown that learning rule are effective and robust. By the learning rule the modified simple PR system can classified all m-dimensional probabilistic vectors into m different classes. If the stable states of simple PR systems which are not basic unit vectors are considered to be natural boundaries among the m classes, then with the learning rule the modified simple PR system is able to classify the boundaries to some specific classes. Thus, the essence of the learning rule is a procedure which modifies a simple PR to one which has the ability of treating the natural boundaries.

References

[1] A. Rosenfeld, R. A. Hummel and S. W. Zucker, Scene labeling by relaxation operations, *IEEE Trans. Syst. Man Cybern.* **SMC-6**(6), 420-433, (1976).

[2] S. W. Zucker, E. V. Krishnamurthy and R. L. Haar, Relaxation processes for scene labeling: convergence, speed, and stability, *IEEE Trans. Syst. Man Cybern.* **8**,41-48, (1978).

[3] R. A. Hummel and A. Rosenfeld, Relaxation processes for scene labeling, *IEEE Trans. Syst. Man Cybern.* **SMC-8**, 765-768, Oct. (1978).

[4] J. O. Eklundh, H. Yamamoto and A. Rosenfeld, A relaxation method for multispectral pixel classification, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-2**, 72-75, Jan. (1980).

[5] J. A. Richards, D. A. Landerebe and P. H. Swain, Pixel labeling by supervised probabilistic relation, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, 188-191, March (1981).

[6] B. Bhanu and O. D. Faugeras, Shape matching of 2-d object using a hierarchical stochastic labeling technique, *Proc. Conf. Pattern Recognition Image Process.* Las Vegas, NV, pp. 688-690, June (1982).

[7] W. S. Rutkowski, S. Peleg and A. Rosenfeld, Shape segmentation using relaxation, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-3**, 368-375, July (1981).

[8] S. W. Zucker, R. A. Hummel and A. Rosenfeld, An application of relaxation labeling to line and curve enhancement, *IEEE Trans. Comput.* **26**, 394-403, (1977).

[9] S. Peleg and A. Rosenfeld, Determining compatibility coefficients for curve enhancement relaxation processes, *IEEE Trans. Syst. Man Cybern.* **8**, 548-555, (1978).

[10] H. Ogawa and K. Taniguchi, On matching recognition of hand-printed Chinese characters by feature relaxation, *Pattern Recognition* **21**, 9-17, (1988).

[11] I. Sekita, K. Toraichi, R. Mori, K. Yamamoto and H. Yamada, Feature extraction of handwritten Japanese characters by sline function for relaxation, *Pattern Recognition* **21**, 9-17, (1988).

[12] S. Peleg, Ambiguity reduction in handwriting with ambiguous segmentation and uncertain interpretation, *Comput. Graphics Image Processing* vol. **10**, 235-245, July (1979).

[13] S. Peleg and A. Rosenfeld, Breaking substitution ciphers using a relaxation algorithm, *Commun. Ass. Comput. Mach.* vol. **22**, 598-605, Nov. (1979).

[14] Q. X. Wu, A correlation-relaxation-labeling framework for computing optical flow – template matching from a new perspective, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. **17**, No. 8, 843-853, 1995.

[15] M. W. Hansen and W. E. Higgins, Relaxation methods for supervised image segmentation, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. **19**, No. 9, 949-962, 1997.

[16] S. Peleg, A new probabilistic relaxation scheme, *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-2** 362-369, July (1980).

[17] S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images, *IEEE Trans. Pattern Anal. Mach. Intell.* **PAMI-6**, 721-741, Nov. (1984).

[18] R. A. Hummel and S. W. Zucker, On the foundations of relaxation labeling processes, *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-5**, 267-287, Nov. (1984).

[19] L. Pelkowitz, A continuous relaxation labeling algorithm for Markov random fields, *IEEE Trans. Syst. Man Cybern.* **SMC-20**, 709-715, May/June (1990).

[20] Q. Chen and J. Y. S. Luh, Ambiguity reduction by relaxation labeling," *Pattern Recognition* **27**, 165-180, (1994).

[21] Q. Chen and J. Y. S. Luh, Relaxation labeling algorithm for information integration and its convergence, *Pattern Recognition* **28**, 1705-1722, (1995).

[22] M. Pelillo and M. Refice, Learning compatibility coefficients for relaxation labeling processes, *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. **16**, No. 9, 933-945, 1994.

[23] A. M. N. Fu and Hong Yan, A new probabilistic relaxation method based on probability space partition, *Pattern Recognition* vol. **30**, No. 11, 1905-1917, 1997.

[24] S. Li, W. Y. C. Soh and E. K. Teoh, Relaxation labeling using augemented Lagrange-Hopfield method, *Pattern Recognition* vol. **31**, no. 1, 73-81, 1998.