# 2-D Conditional Moment for Recognition of Deformed Letters†

Myoong-Young Yoon*

**요 약** 본 논문에서는 화상자료의 특성인 이웃 화소간의 종속성을 표현하는데 적합한 깁스분포를 바탕으로 특징벡터를 추출하여 변형된 글자를 인식하는 새로운 방법을 제안하였다. 추출된 특징벡터는 이미지의 크기, 위치, 회전에 대하여 불변한 특성을 갖는 2차원 조건부 모멘트로 구성된다. 변형된 글자 인식을 위한 알고리즘은 특징벡터 추출하는 과정과 패턴을 인식하는 과정으로 구성하였다. (i) 특징벡터는 하나의 이미지에 대하여 추정된 조건부 깁스분포를 바탕으로 2차원 조건부 모멘트를 계산하여 추출한다. (ii) 변형된 문자인식은 제안된 판별거리함수를 계산하여 최소거리를 산출한 미지의 변형된 문자를 원형문자로 인식한다. 제안된 방법에 대한 성능평가를 위하여, 생성된 훈련 데이터를 만들어 Workstation에서 실험 한 결과 96%이상의 인식성능이 있음을 밝혔다.

**Abstract** In this paper we propose a new scheme for recognition of deformed letters by extracting feature vectors based on Gibbs distributions which are well suited for representing the spatial continuity. The extracted feature vectors are comprised of 2-D conditional moments which are invariant under translation, rotation, and scale of an image. The Algorithm for pattern recognition of deformed letters contains two parts: the extraction of feature vector and the recognition process. (i) We extract feature vector which consists of an improved 2-D conditional moments on the basis of estimated conditional Gibbs distribution for an image. (ii) In the recognition phase, the minimization of the discrimination cost function for a deformed letters determines the corresponding template pattern. In order to evaluate the performance of the proposed scheme, recognition experiments with a generated document was conducted on Workstation. Experiment results reveal that the proposed scheme has high recognition rate over 96%.

## 1. Introduction

An essential issue in the field of pattern analysis is the recognition of objects and characters regardless of their positions, size, and orientations. In the recent computer vision literature there has been increasing interest in use of statistical techniques for recognition and processing image data. Statistical image analysis concerns the measurement of quantitative information from an image to produce a probabilistic description. A feature-based recognition of objects or patterns independent of their position, size, orientation and the other variations has been the goal of ongoing research. Finding efficient invariant features is the key to solving this problem. In other words, selection of "good" features is a crucial step in the process. "Good" features are those satisfying the following requirements: (i) small interclass invariance, and (ii) large interclass separation. These features (or shape descriptors) may be divided into five groups as follows[1, 2]:

- Visual features(edges, texture and shape);
- Transform coefficient features(Fourier descriptors or Hadamard coefficients);
- Algebraic features(based on matrix decomposition of an image);
- Differential invariant features(used especially for curved objects)
- Statistical features(moments invariants).

Moments and functions of moments have been extensively employed as the invariant global features of an image in pattern recognition, image classification,

† 본 연구는 2000년 충청대학 학술연구진흥비의 지원으로 연구된 논문임.

* 충청대학 컴퓨터학부 부교수

target identification, and scene analysis[3, 4]. Generally, these features are invariant under image translation, rotation, scale change, and rotation only when they are computed from the original non-distorted analog two dimensional image. The moment invariants are very useful features for pattern recognition. Tsirikolias, Shen, and Arezki proposed moments that provide features for recognition of patterns have been used for a number of image classifying applications[1, 5, 6]. However, their performance for pattern recognition is poor since the moments did not included spatial information which is the characteristic of the most images.

In order to overcome the drawback of their pattern recognition methods, we propose a new scheme for moment-based recognition of a deformed letters using the spatial information based on Gibbs random field (GRF). Gibbs random fields are well suited for representing statistical dependence(or spatial continuity) of the pixel value at a lattice point on the those of its neighbors[5, 6]. The proposed a new scheme contains two parts: feature extraction and pattern recognition. First of all, we estimate the parameters of Gibbs random fields to model a pattern image. And then we extract feature vector which consists of the calculated 2-D conditional moments. In the recognition phase, the minimization of the discrimination cost function for a deformed letter determines the corresponding template pattern.

## 2. Gibbs Distributed Image

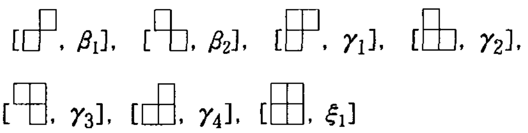
In this section, we present a particular class of Gibbs distribution(GD) which is suited for describing the deformed letters and estimate the parameters of Gibbs distributed for a deformed letter image. We focus our attention on discrete 2-D random fields defined over a finite $N_1 \times N_2$ rectangular lattice of points(pixels) defined as $L = \{(x, y) : 1 \le x \le N_1, 1 \le y \le N_2\}$.

Suppose $Q = \{q_{xy}\}$ represents an image, where $q_{xy}$ measures the grey-level(or intensity) of the pixel in the $x$-th row and $y$-th column. Let $\eta$ be neighborhood system defined over the finite L. A random field $Q = \{Q_{xy}\}$ on L has Gibbs distribution or equivalently is a Gibbs Random Field(GRF) with respect to $\eta$ if and only if its joint distribution is of the form[7, 8]

$$P(Q=q) = \frac{\exp\{-Energy\ function\}}{\{Partition\ function\}} = \frac{\sum_q \exp\{-E(q)\}}{\exp\{-\sum_{c \in C} V_c(q)\}} \quad (1)$$

where $c$ is a clique, and $C$ is the set of all cliques of a lattice-neighborhood pair $(L, \eta)$ and $V_c(q)$ is the potential associated with clique $c$, arbitrary except for the fact that it depends only on the restriction of $q$ to $c$. Let $\eta^m$ be the $m$th order neighborhood system

The GD characterization in some applications provides a more workable spatial model. We assume that the random field $Q$ consists of binary-valued discrete random variables $\{Q_{xy}\}$ taking values in $\Omega = \{\omega_1, \omega_2\}$ To define GD it suffices to specify the neighborhood system $\eta$, the associated cliques, and the clique potentials $V_c(q)$'s. Assume that the random field is homogeneous.




<Figure 1> The parameters associated with clique types.

The distribution is specified in terms of the second order neighborhood system $\eta^2$. Figure 1 shows the parameters associated with clique types, except for the single pixel clique. The clique potentials associated with $\eta^2$ are defined as follows

$$V_c(q_{xy}) = \begin{cases} -\zeta & \text{if all } q_{xy}\text{'s in c are equal} \\ \zeta & \text{otherwise} \end{cases} \quad (2)$$

where $\zeta$ is the parameter specified for the clique type c. For the single pixel cliques, the clique potential is defined as

$V_c(q_{xy}) = \alpha_k$ for $q_{xy} = \omega_k$ (3)

The parameters $\alpha_k$ control the percentage of pixels in each site, that is the marginal distribution of the single random variables $Q_{xy}$'s, while the other parameters control the size and direction of clustering.

## 3. Conditional Moment and Classification

In this section, we describe a method for estimating parameter of Gibbs distributed image. And then we extract feature vector which consists of the calculated 2-D conditional moments. These feature vectors are invariant under image translation, rotation, size, and rotation. Finally, We propose a discrimination distance function which recognize a deformed letter.

### 3.1 Parameter Estimation

In this subsection our aim is to estimate the parameters of Gibbs distributed image. The most commonly used parameter estimation method to date is the so-called "coding method," first presented by Besag[9]. It requires the solution of a set of nonlinear equations. Therefore, it is cumbersome and difficult to use reliably. In view of the practical difficulties involved in using the coding method [10, 11], we describe an alternative parameter estimation scheme for finite range space GRF, which consists of histogramming and a standard, linear, least squares estimation as its components. Suppose $Q$ is a GD with a discrete range space of $\Omega = \{\omega_1, \omega_2\}$ A realization $q$ of this random field is available to be used in estimating the parameters of the distribution. For convenience of notation, let $s$ represent $q_{xy}$ and $t'$ represent the vector of the neighboring values of $q_{xy}$, that is, $t' = [u_1, u_2, u_3, u_4, v_1, v_2, v_3, v_4]^T$ Where the location of $u_i$'s and $v_i$'s with respect to $s$ are shown in Figure 2.

| | | |
|---|---|---|
| $v_1$ | $u_2$ | $v_2$ |
| $u_1$ | ■ | $u_3$ |
| $v_4$ | $u_4$ | $v_3$ |

<Figure 2> $q_{xy}$ and $\eta_{xy}$.

We define indicator functions

$$I(z_1, z_2, \ldots, z_k) = \begin{cases} -1 & \text{if } z_1 = z_2 = \cdots = z_k \\ 1 & \text{otherwise} \end{cases} \quad (4)$$

and

$$J_m(s) = \begin{cases} -1 & s = \omega_m \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

We can express the potential functions of the GD in terms of these quantities. Let $V(s, t', \theta)$ be the sum of the potential functions of all the cliques that contain $(x, y)$, the site of s. That is $V(s, t', \theta) = \sum_{c s \in C} V_c(q)$ where $\theta$ is the parameter vector $\theta = (\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2, \gamma_3, \gamma_4, \xi_1)$. Using the clique potentials for this class of GD we can write $V(s, t', \theta)$ as $V(s, t', \theta) = \rho^T(s, t')\,\theta$ where

$$\begin{aligned} \rho(s, t) = [&J_1(s), J_2(s), \\ &(I(s, v_2) + I(s, v_4)), (I(s, v_1) + I(s, v_3)), \\ &(I(s, u_2, v_2) + I(s, u_4, u_3) + I(s, u_1, v_4)), \\ &(I(s, u_4, u_3) + I(s, u_2, u_3) + I(s, u_1, v_1)), \\ &(I(s, u_2, v_1) + I(s, u_1, u_4) + I(s, u_3, v_3)), \\ &(I(s, u_1, u_2) + I(s, u_4, v_4) + I(s, u_3, v_2)), \\ &(I(s, u_1, v_1, u_2) + I(s, u_2, v_3, u_3) \\ &+ I(s, u_3, v_3, u_4) + I(s, u_4, v_4, u_1))]^T. \end{aligned} \quad (6)$$

Now Suppose $P(s|t')$ is the joint distribution of the random variables on the $3 \times 3$ window centered at $(x, y)$ and $P(t')$ is the joint distribution of the random variables on $\eta_{xy}$ only. Then the conditional distribution $P(s|t')$ is given by the ratio of $P(s, t')$

to $P(t')$. It follows from the GRF-MRF equivalence and the resulting local characteristic that

$$\frac{e^{-V(s,t',\theta)}}{P(s,t')} = \frac{Z(t',\theta)}{P(t')} \tag{7}$$

where $Z(t',\theta)$ is the appropriate normalizing constant.. Note that the right-hand side of (7) is independent of $s$. Considering the left-hand side of (7) for any two distinct values of $s$, e.g., $s=j$ and $s=k$ we have

$$(\rho(k,t')-\rho(j,t'))^T\theta = \ln\frac{P(j,t')}{P(k,t')} \tag{8}$$

where $\rho^T(k, t')\theta = V(k, t',\theta)$
Consideration of all possible triplets $(j,k,t')$, $j<k$ generates from equation (8) a large set of linear equations, which may be solved for $\theta$ by least squares procedures. The question that remains to be answered , now, is how to determine or estimate $P(s,t')$ for all $(s,t')$ combinations using a single or a few realizations. We will calculate the probability $P(s,t')$ using histogram techniques.

### 3.2 Calculation of 2-D Conditional Moments

The basic and classical moment, a regular 2-D moment of order $(k+l)$ is defined by[1, 8]

$$m_{kl} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^k\, y^l\, f(x,y), \tag{9}$$

where $f(x,y)$ is the intensity at a point $(x,y)$ in the image and $k,\ l=0,1,2,\cdots$The moments proposed by many researchers[1, 6, 8, 12] have not included spatial information which is the characteristic of most images. An an alternative to cope with the drawback of the moments, we propose 2-D conditional moments which include spatial information by using the estimated conditional Gibbs distribution. The corresponding 2-D conditional moments are given by the following steps.

● Step1) Calculate the centroids $\bar{x}$, $\bar{y}$ of the considered shape as follows.

$$\bar{x} = \sum_{x=1}^{N_1}\sum_{y=1}^{N_2} x\cdot \hat{P}(Q_{xy}=q_{xy}|\eta_{xy})$$

and

$$\bar{y} = \sum_{y=1}^{N_2}\sum_{x=1}^{N_1} y\hat{P}(Q_{xy}=q_{xy}|\eta_{xy}) \tag{10}$$

where $\hat{P}(Q_{xy}=q_{xy}\ |\eta_{xy})$ is estimated conditional probability of the site $(x,y)$

● Step2) Calculate the standard deviation $\sigma_x$ and $\sigma_y$

$$\sigma_x = \sqrt{\{\sum_{x=1}^{N_1}\sum_{y=1}^{N_2}(x-\bar{x})^2\cdot \hat{P}(Q_{xy}=q_{xy}|\eta_{xy})\}}$$
$$\sigma_y = \sqrt{\{\sum_{y=1}^{N_2}\sum_{x=1}^{N_1}(y-\bar{y})^2\cdot \hat{P}(Q_{xy}=q_{xy}|\eta_{xy})\}}. \tag{11}$$

● Step3) Calculate and Store the 2-D conditional moments for $(k,l)$

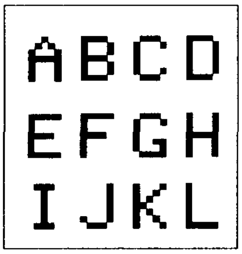
$$m_{kl} = \sum_{x=1}^{N_1}\sum_{y=1}^{N_2}(x-\bar{x}/\sigma_x)^k\,(y-\bar{y}/\sigma_y)^l \times \hat{P}(Q_{xy}=q_{xy}\ |\eta_{xy}). \tag{12}$$

The required number of moments depends on: (i) the level of the existing noise on the application and (ii) the form of the considered shapes. The above moments are invariant under translation, magnification, and rotation of the image, but not under rotation. Thus, in order to use them as feature vector in recognition phase, we have to normalize respect to the rotation. The normalization is a simple operation, since only a multiplication of the coordinates of the image by $e^{-j\phi}$, where $\phi$ is the rotation change of the object. Table 1 shows the normalized 2-D conditional moments for template letter A and D of Figure 3.

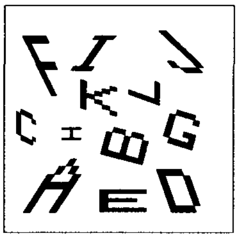
<Table 1> The proposed moments of each letters

| Moments | $m_{30}$ | $m_{40}$ | $m_{50}$ | $m_{60}$ | $m_{70}$ | $m_{80}$ |
|---|---|---|---|---|---|---|
| Template 'A' | 1.02 | 2.65 | 1.29 | 4.11 | 3.97 | 7.08 |
| Template 'D' | 1.01 | 2.44 | 1.01 | 2.04 | 1.08 | 5.72 |

| Moments | $m_{03}$ | $m_{04}$ | $m_{05}$ | $m_{06}$ | $m_{07}$ | $m_{08}$ |
|---|---|---|---|---|---|---|
| Template 'A' | 1.11 | 2.92 | 1.02 | 3.00 | 1.59 | 3.09 |
| Template 'D' | -1.10 | 3.20 | -1.93 | 4.20 | -2.91 | 7.94 |




(a) The template



(b) the letters to be recognized

<Figure 3> The template and the deformed letters

### 3.3 Classification

In order to classify deformed letters, we define a the discrimination cost function (DCF) $F(i, v)$ which is defined by

$$F(i,v)=\sum_{j=1}^{d}\,[T_{vj}-U_{ij}]^2 \qquad (13)$$

where $T_{vj}$ denotes the $j-th$ feature of the $v-th$ template, $U_{ij}$ denotes the $j-th$feature of the $i-th$ shape under consideration and $d$ is the dimension of the feature vectors. The minimization of the index $F(i,v)$, $v=1,2,...$for a specific deformed letter $i$ determines the corresponding template $v$.

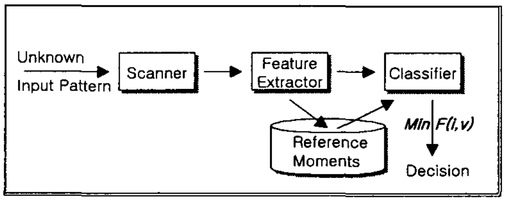
The proposed DCF is a kinds of Euclidean distance between an arbitrary pattern vector $U_{ij}$ and the $v-th$ prototype vector $T_{vj}$. Furthermore, since the proposed DCF only require some simple analytic algebraic calculations, it is characterized by low computation cost. The ideal discrimination of a deformed letter corresponding exactly to a template, without any noise and computational error, the index $F(i, v)$ should be zero. However, in practice, the discrimination is clear if $F(i, v)$ is sufficiently smaller in comparison with the other templates, as well as small enough itself. Table 2 shows the discrimination functions of the letters of Fig. 4. It is seen that $F(i, v)$ is sufficiently smaller for the deformed letter "A" in Table 2.

<Table 2> The DCF for deformed letter A

| | |
|---|---|
| F(template "A", deformed Letter "A") = | 4.07 |
| F(template "A", deformed letter "F") = | 79.28 |

## 4. Experimental Results

In order to illustrate the performance of the proposed moment for deformed letter recognition, we carried out the following experiments was carried out. In this setting, the generated training document consists of 10 lines of 52 alphabetic characters each. Two documents were created for testing the performance of the proposed classification method on the basis of the extracted feature vector. Each document consists of 24 lines 52 characters each. Figure 4 shows the overall block diagram of the proposed method for classification of patterns, where it is shown that a document to be processed is at first scanned. Then the classification feature vectors are extracted by formulae (10) through (12). These features are sent to a classifier, which is described by formula (13), for a decision in order to identify the input character.




<Figure 4> Overview of the proposed scheme for classifying of the deformed letters

The gross structural features of the shape can be better characterized by the proposed moments derived from the silhouette. In our experiments we use only silhouette moments since these moments are less sensitive to noise. The used feature vector $F_v$ for the templates is considered to be $F_v=[m_{03}, m_{04}, m_{05}, m_{06}, m_{07}, m_{08}, m_{30}, m_{40}, m_{50}, m_{60}, m_{70}, m_{80}]^T$

A classification simulation was run six times. The first simulation used a library set of 52 feature vectors derived from the first line of characters of the training document. The second simulation used two library sets derived from the first two lines of the training document. The third, fourth, fifth and sixth simulations used four, six, eight and ten library sets, respectively. The classification rates resulting from these simulations are presented in Table 3. As Table 3 reveals, we can achieve better than 96% increase in classification rates when we use eight or ten library sets. Since the proposed the 2-D improved moments have properties of the affine or geometric moments, as well as spatial information which describe dependance between pixels, our proposed method was superior to otner methods using the affine moments[12] and the geometric moments[8], respectively.

<Table 3> The recognition rates

| No of library sets | Flusser's method using the affine moments (%) [8] | Tsirikolias's method using the geometric moments (%) [12] | Proposed method using the 2-D moments (%) |
|---|---|---|---|
| 1 | 73 | 72 | 80 |
| 2 | 82.5 | 81 | 86 |
| 4 | 84 | 85 | 88 |
| 6 | 90.5 | 89 | 95 |
| 8 | 93 | 91 | 96 |
| 10 | 95 | 94.5 | 98.5 |

In our method, the incorrected classification of the deformed letters is caused by the insufficient clique function $V(s, t', \theta)$ described in equation (6). Since the clique parameter vector $\hat{\Theta}$ is a measure which is the strength of interaction between pixels, the clique potentials $V(s, t', \theta)$ affect the Gibbs distribution $\hat{P}(Q_{xy}=q_{xy} \mid \eta_{xy})$ which is used in calculation the proposed conditional moments depends only the clique function. In other words, The success of classification depends on how good the used the clique parameter $\hat{\theta}$ fits characteristic of the image. So, we will focus our efforts on further development of the clique functions in order to improve recognition of the deformed letters.

## 5. Concluding Remarks

In this paper we propose a new algorithm for pattern recognition for the deformed letters using an improved 2-D conditional moments based on GD. Experiment results reveal that the proposed scheme has high classification rate over 96%. The proposed method appears to be efficient with respect to the existing ones, since it shares the following advantages. (i) The discrimination process is invariant under translation, scaling and rotation of the considered shape. (ii) Fast processing, since calculations of the moment are simple. (iii) Each shape is uniquely described.

The success of pattern classification depends on how good the used clique parameter $\hat{\theta}$ fits characteristic of the image. Upon completion of the pattern classification, we will focus our efforts on further development of the clique functions.

## REFERENCES


[1] K. Tsirikolias and B. G. Mertzios, "Statistical Pattern Recognition Using Efficient Two-dimensional Moments with Applications to Character Recognition," Pattern Recognition, Vol. 26, No. 6, pp. 877-882, 1993.

[2] S. C. Pei and J. H. Hong, "A Moment-based Approach for Deskewing Rotationally Symmetric Shapes," IEEE Trans. on Image Processing, Vol. 8, No.12, pp. 1831-1834, 1999.

[3] Torres-Mendez, Ruiz-Suarez, Sucar, Gomez, "Translation, Rotation, and Scale-invariant Object Recognition," IEEE Trans. on Systems, Man & Cybernetics Part C: Applications & Reviews, Vol. 30, No. 1, 2000

[4] Yzhu, De Silva, "Using Moment Invariants and HMM in Facial Expression Recognition," Prceedings of the 4th IEEE Southwest Symposium on Image Analysis and Interpretation, pp. 305-309, 2000.

[5] B. C. Li and J. Shen, "Fast computation of moment invariants," Pattern Recognition Vol. 24, pp. 807-813, 1991.

[6] H. Arezki, "A modified moment Inavariants Technique for Classification of Noisy Objects," Conference on Electrical and Computer Engineering, Canada, pp. 25-27, 1991.

[7] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," IEEE Transactions on Pattern Analysis Machine Intelligence, Vol. 6, pp. 721-741, 1984.

[8] W. Chung and E. M. Tzanakou, "A Simulation for Different Moment sets," Proceedings in SPIE, Vol. 2181: Document Recognition, San Jose, California, pp. 378-389, 1994.

[9] J. Besag, "Spatial Interaction and the Statistical Analysis of Lattice Systems," Journal of Royal Statistics Society B, Vol. 36, pp. 192-236, 1974.

[10] C. H. Wu and P. C. Doerschuk, "Cluster Expansions for the Deterministic Computation of Bayesian Estimators Based on Markov Random Fields," IEEE Transactions on Pattern Analysis Machine Intelligence, Vol. 17, No. 3, pp. 275-293, 1995.

[11] A. J. Gray, J. K. Kay, and D. M. Titterington, "On the Estimation of Noisy Binary Markov Random Fields," Pattern Recognition, Vol. 25, No. 7, pp. 749-768, 1992.

[12] J. Flusser and T. Suk, "Classification Objects and Patterns by Affine Moments Invariants," Image Processing: Theory and Applications, pp. 345-348, 1993.

**윤 명 영**

1984년 충북대학교 계산통계학과 졸업(이 학사)
1987년 충북대학교 대학원 계산통계학과(이학석사)
1997년 충북대학교 대학원 전자계산학과(이학박사)
1987년~1990년 육군 전산장교 근무
1997년~1999년 충청대학 정보센터원장
1992년~현재 충청대학 컴퓨터학부 부교수
관심분야: 화상처리, 패턴인식, 퍼지이론, 한국어정보처리