

계층적인 가버 특징들과 베이지안 망을 이용한 필기체 숫자인식

(Hierarchical Gabor Feature and Bayesian Network for Handwritten Digit Recognition)

성재모^{*} 방승양^{**}
(JaeMo Sung) (Sung-Yang Bang)

요약 본 논문에서는 필기체 숫자인식을 위해서 계층적으로 서로 다른 레벨의 정보를 표현할 수 있는 구조화된 특징들의 추출 방법과 특징들 사이에 의존도를 이용하여 분류하는 베이지안 망을 제안한다. 이러한 계층적 특징들을 추출하기 위해서 레벨 단위로 가버 필터들을 정의하고, FLD(Fisher Linear Discriminant) 척도를 이용하여 최적화된 가버 필터들을 선택한다. 계층적 가버 특징들은 최적화된 가버 필터들을 이용하여 추출되며, 하위 레벨일수록 더욱 국부적인 정보를 표현한다. 추출된 계층적 가버 특징들의 분류성능 향상을 위해서 가버 특징들 사이의 계층적 의존도를 이용하는 베이지안 망을 생성한다. 본 논문에서 제안하는 방법은 naive Bayesian 분류기, k-nearest neighbor 분류기, 그리고 신경망 분류기들과 함께 필기체 숫자인식에 적용되어 계층적 가버 특징들의 효율성과 계층적 의존도를 이용하는 베이지안 망은 분류성능을 향상시킬 수 있다는 것을 보여준다.

키워드 : 베이지안 망, 가버 필터, 필기체 숫자인식, 계층적인 모델

Abstract For the handwritten digit recognition, this paper proposes a hierarchical Gabor features extraction method and a Bayesian network for them. Proposed Gabor features are able to represent hierarchically different level information and Bayesian network is constructed to represent hierarchically structured dependencies among these Gabor features. In order to extract such features, we define Gabor filters level by level and choose optimal Gabor filters by using Fisher's Linear Discriminant measure. Hierarchical Gabor features are extracted by optimal Gabor filters and represent more localized information in the lower level. Proposed methods were successfully applied to handwritten digit recognition with well-known naive Bayesian classifier, k-nearest neighbor classifier, and backpropagation neural network and showed good performance.

Key words : Bayesian network, Gabor filter, handwritten digit recognition, hierarchical model

1. 서론

지금까지 패턴인식을 위해서 가버 필터[1], PCA (Principle Component Analysis)[2], 그리고 ICA(Independent Component Analysis)[3] 등을 이용하는 특징추출에 관한 연구가 진행되어 왔다.

가버 필터는 컴퓨터 비전[4] 분야에서 널리 적용되어 왔으며, 인간의 시각과정에 있는 단순세포를 모델링할 수 있다는 것도 알려져 있다[5]. 이러한 사실에 근거하

여 가버 필터를 이용해 추출된 특징을 사용하여 필기체 숫자인식[6,7]에 적용되기도 하였다. 또한 문자인식분야에서 사용된 특징들의 종류에 따라서 성능에 많은 영향을 준다는 것도 알려져 있다. 대부분의 문자인식에 관한 연구는 특징추출에 관한 연구에 치우쳐 있었으며, 분류기로는 널리 알려진 k-nearest neighbor[2]나 신경망[2]에 기반한 분류기 등을 사용하여 성능평가가 되었다. 그러나 대부분의 특징추출 방법들은 서로 다른 레벨의 정보를 표현하는 구조화된 특징들을 추출하도록 고안되지 않았다.

본 논문에서는 효율적인 인식 시스템을 위해서 구조화된 특징추출 방법과 추출된 특징들 사이의 구조적인 의존도를 이용하여 분류하는 것이 숫자인식을 위해서 더 효율적일 수 있다는 것을 보여준다. 이를 위해서 계

^{*} 비회원 : 포항공과대학교 컴퓨터공학과
emtidi@postech.ac.kr

^{**} 종신회원 : 포항공과대학교 컴퓨터공학과 교수
sybang@postech.ac.kr

논문접수 : 2002년 11월 7일

심사완료 : 2003년 10월 6일

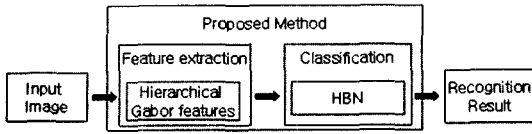


그림 1 제안된 방법의 흐름도

층적으로 구조화된 특징들을 추출할 수 있는 가버 필터들을 레벨 단위로 정의하고, FLD(Fisher's Linear Discriminant) 척도를 이용하여 최적화된 가버 필터들을 선택한다. 최적화된 가버 필터들에 의해서 추출된 특징들은 계층적으로 서로 다른 레벨의 정보를 표현할 수 있도록 구조화된다. 즉 본 논문에서는 더 낮은 레벨의 특징들은 더 국부적인 정보들을 표현할 수 있는 구조화된 특징추출 방법을 제안한다. 또한 분류기로서 추출된 특징들 사이의 계층적인 구조적 의존도를 고려할 수 있는 계층적 베이저안 망(Hierarchical Bayesian Network)을 제안한다. 널리 알려진 naive Bayesian 분류기, k-nearest neighbor 분류기 그리고 신경망 분류기들과 함께 수행된 필기체 숫자인식에 대한 성능평가는 본 논문에서 제안하는 특징추출 방법과 분류방법의 우수성을 보여준다.

2. 계층적 가버 특징들의 추출

가버 필터들은 공간과 주파수(spatial-frequency) 도메인에서 정의되며, 본 논문에서 사용된 가버 필터[8]는 식 (1)과 같이 정의된다.

$$G(x, y, \omega, \sigma, r, \theta) = \frac{1}{\sqrt{\pi} r \sigma} e^{-\frac{1}{2} \left[\frac{(rR_1)^2 + R_2^2}{(r\sigma)^2} \right]} e^{i\omega R_1},$$

where $R_1 = x \cos \theta + y \sin \theta$, $R_2 = -x \sin \theta + y \cos \theta$ (1)

여기서 ω 는 레이디얼 주파수이며, θ 는 라디안 단위의 방향이다. σ 는 x축에 대한 Gaussian envelope의 표준편차이며, $r = (\sigma_x / \sigma_y)$ 는 x축과 y축의 표준편차의 비율이다. 또한 위의 가버 필터는 공간 도메인에서 ($x=0, y=0$)을 중앙으로 위치한다.

주어진 이미지에서 계층적으로 구조화된 정보를 표현하기 위해서는 추출된 특징들은 서로 다른 레벨의 정보를 표현할 수 있어야 한다. 즉, 낮은 레벨에서 추출된 특징일수록 더 국부적인 정보를 표현해야 한다. 본 논문에서는 가버 필터들을 이용하여 서로 다른 레벨의 정보를 표현하는 특징들의 추출방법을 제안한다. 이를 위해서 먼저 레벨마다 주어진 방향에 대해서 가버 필터 집합들을 정의하고, 각 가버 필터 집합마다 FLD 척도를 이용해서 하나의 최적화된 가버 필터를 선택한다. 이렇게 각 레벨마다 선택된 최적화된 가버 필터들에 의해서

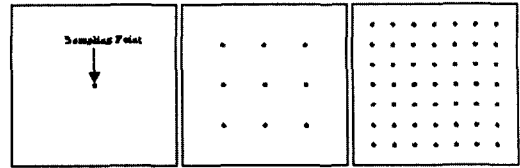


그림 2 각 레벨의 샘플링 포인트(왼쪽은 최 상위 레벨, 오른쪽은 최 하위 레벨)

주어진 이미지의 계층적 가버 특징들을 추출하게 된다.

p^k 을 l 번째 레벨에 있는 s 번째 샘플링 포인트(sampling point)라 하자. 또한 첫 번째 레벨 ($l=1$)은 최 상위 레벨이라고 하자. 최 상위 레벨에는 이미지의 중앙에 하나의 샘플링 포인트가 있게 된다. 최 상위 레벨에서부터 계속적으로 각 샘플링 포인트는 하위 레벨에서 9개의 샘플링 포인트들로 분할된다. 그림 2는 각 레벨마다 분할된 샘플링 포인트들을 보여준다.

$\Theta = (\theta_j; j \in [1, 2, \dots, N_\theta])$ 을 N_θ 개의 방향 θ_j 을 원소로 가지는 주어진 방향 집합이라고 하고, $\Omega = (\omega_i; i \in [1, 2, \dots, N_\omega])$ 을 N_ω 개의 주파수 ω_i 을 원소로 가지는 주어진 주파수 집합이라고 하자. 각 샘플링 포인트 p^k 와 방향 θ_j 에 대해서 주파수 집합 Ω 에 대응하는 가버 필터 집합 GF_j^k 은 식 (2)와 같이 정의된다.

$$GF_j^k = \{ G_{R_1}^k, G_{R_2}^k, \dots, G_{N_\omega}^k \},$$

$$G_{R_i}^k(x, y) = G(x^k - x, y^k - y, \omega_i, \sigma^k, r, \theta_j) \quad (2)$$

(x^k, y^k) 는 이미지 좌표상의 샘플링 포인트 p^k 의 좌표이다. 가버 필터 $G_{R_i}^k(x, y)$ 는 이미지 좌표에서 (x^k, y^k) 을 중앙으로 위치하며, 주파수 ω_i , x축 표준편차 σ^k , 그리고 방향 θ_j 를 파라미터로 가진다. r 은 가버 필터의 타원성에 대한 정도를 나타내며 $r=1$ 인 경우 가버 필터는 원 모양이 된다. 본 논문에서는 $r=2$ 로 고정하여 타원 모양의 가버 필터를 사용하였다.

하위 레벨일수록 더욱 국부적인 정보를 표현하기 위해서 각 레벨마다 가버 필터의 표준편차 σ^k 을 식 (3)과 같이 정의하여 하위 레벨에서 가버 필터가 더 작은 영역을 커버하게 만든다(그림 3의 (a)).

$$\sigma^k = \frac{m(d_1, d_2, d_3, d_4)}{\sqrt{2 \ln 2}} \quad (3)$$

d_1, d_2, d_3, d_4 은 p^k 과 이웃한 샘플링 포인트들 사이의 거리이며(그림 3의 (b)), $m(d_1, d_2, d_3, d_4)$ 은 이들 사이의 평균거리이다. 이렇게 정의된 σ^k 을 가지는 원형 가버 필터 ($r=1$)의 Gaussian envelop는 최대값의 반에 해당하는 외형(contour)의 반지름이 $m(d_1, d_2, d_3, d_4)$ 에 놓이게 된다.

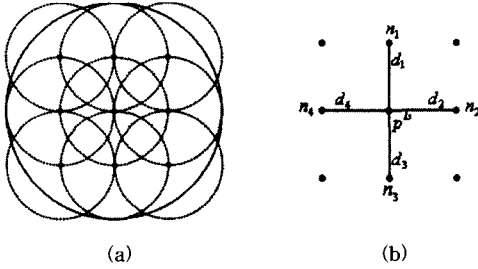


그림 3 (a) 큰 원은 상위 레벨의 샘플링 포인트에서 정의된 가버 필터의 영역을 나타내며, 작은 원들은 하위 레벨에서 분할된 9개의 샘플링 포인트들에서 정의된 가버 필터들의 영역들을 나타낸다. (b) 샘플링 포인트 p^k 의 이웃한 4개의 샘플링 포인트 n_1, n_2, n_3, n_4 와 그들과의 거리 d_1, d_2, d_3, d_4 를 보여준다.

위와 같이 레벨마다 각 샘플링 포인트와 방향에 대하여 가버 필터 집합 GF_j^k 을 정의한 후 FLD(Fisher's Linear Discriminant) 척도를 이용하여 GF_j^k 로부터 최적화된 가버 필터 $G_{j^*}^k$ 을 선택한다. 필기체 숫자인식을 위해서 특징들의 추출 후 최종 목표는 추출된 특징들의 분류이다. 그러므로 FLD 척도를 사용하여 보다 더 쉽게 분류될 수 있는 특징들을 추출할 수 있는 주파수를 가지는 가버 필터를 선택하는 것은 합리적이다. $C = c$ 을 분류하고자 하는 클래스의 집합이라고 하고, I 를 주어진 학습 이미지의 집합이라고 하자. $h_d \in C$ 이고 $I_d \in I$ 이면, (h_d, I_d) 는 사전 분류된 하나의 학습 데이터가 된다. 그러면 최적화된 가버 필터 $G_{j^*}^k$ 는 식 (4)에 의해서 선택된다.

$$G_{j^*}^k \in GF_j^k, i=1, \dots, N_w,$$

$$X_{j^*}^k = (h_1, g_1), \dots, (h_{N_s}, g_{N_s}),$$

$$g_d = \sum_x \sum_y I_d(x, y) G_{j^*}^k(x, y), G_{j^*}^k = G_{j^*}^{k^*},$$

$$i^* = \arg \max_i (DM_{j^*}^k), DM_{j^*}^k = FLD(X_{j^*}^k) \quad (4)$$

$X_{j^*}^k$ 은 사전 분류된 이미지들과 가버 필터 $G_{j^*}^k$ 와의 response들의 집합이며, N_s 는 주어진 학습데이터 이미지의 수이다. $FLD(X)$ 는 데이터 집합 X 에 대한 Fisher's Linear Discriminant 척도를 위한 함수이다. $FLD(X)$ 에 대해서는 부록을 참고하기 바란다. 식 (4)에서 각 가버 필터 $G_{j^*}^k \in GF_j^k$ 에 대해서 먼저 학습 이미지들의 response들이 계산된 후 가장 큰 FLD 척도

를 주는 주파수의 가버 필터는 최적화된 가버 필터 $G_{j^*}^k$ 로 선택된다.

각 샘플링 포인트 p^k 에 대해서 N_σ 개의 최적화된 가버 필터 $G_{j^*}^k$ 을 선택한 후 이미지 I 에 대한 가버 특징을 식 (5)와 같이 정의하자.

$$a^k = [a_1^k, a_2^k, \dots, a_{N_\sigma}^k]^T,$$

$$a_j^k = \sum_x \sum_y I(x, y) G_{j^*}^k(x, y) \quad (5)$$

a_j^k 는 $G_{j^*}^k$ 에 대한 이미지 I 의 response이다. 그러면 a^k 는 샘플링 포인트 p^k 에서 정의된 모든 방향에 대한 최적화된 가버 필터들에 의해서 추출된 가버 특징이며, N_σ -차원의 벡터가 된다. 또한 정의된 모든 샘플링 포인트에서 추출된 가버 특징들의 집합인 $a = (a^k)$ 을 계층적 가버 특징들로서 식 (6)과 같이 정의한다.

$$a = a^1 \cup a^2 \cup \dots \cup a^{N_L},$$

$$a^l = a^{l_1}, a^{l_2}, \dots, a^{l_{N_l}} \quad (6)$$

N_L 은 총 레벨 수이며, N_l 은 레벨 l 에서의 총 샘플링 포인트의 수이다. a^l 은 레벨 l 에서의 가버 특징들의 집합을 나타낸다.

3. 계층적 베이지안 망

3.1 베이지안 망

베이지안 망(Bayesian Network)[9-11]은 결합 확률 분포의 랜덤 변수들 사이의 의존도 구조를 표현하기 위한 하나의 그래프 표현방법이다. 유한한 랜덤 변수 집합 $V = A_1, \dots, A_n$ 에 대해서, 일반적으로 베이지안 망은 $\langle DAG, CP \rangle$ 로 정의된다. 사이클이 없는 방향그래프(Directed Acyclic Graph) $DAG = (V, E)$ 는 베이지안 망의 구조를 정의한다. $V = \{A_1, \dots, A_n\}$ 는 노드들의 집합을 나타내며, 랜덤 변수들과 일대일 대응관계가 있다. $E = \{(A_i, A_j) : A_i, A_j \in V, \text{ where } i \neq j\}$ 는 방향성 있는 간선들의 집합을 나타내며 랜덤 변수들 사이에 직접적인 확률적 의존도를 표현한다. (A_i, A_j) 는 노드 A_i 로부터 노드 A_j 로 방향성 있는 간선을 나타내며, DAG 에 의해서 정의된 베이지안 망이 표현하고 있는 결합 확률분포에서 랜덤 변수 A_i 는 랜덤 변수 A_j 에 대해서 직접적인 확률적 의존도가 있음을 나타낸다. CP 는 노드들의 조건부 확률분포들의 집합이며, 노드 A_i 의 조건부 확률분포는 $P(A_i | \mathcal{I}_i)$ 로 정의된다. \mathcal{I}_i 는 DAG 에서 노드 A_i 의 부모노드들의 집합을 나타낸다. 이것은 만약 부모 노드들의 값이 주어지면 노드 A_i 는 다른 조상 노드에 대해서 조건부 독립임을 나타낸다.

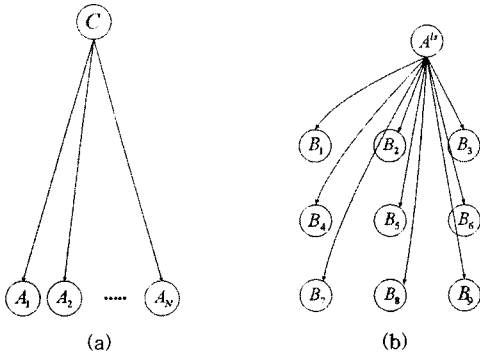


그림 4 (a) naive Bayesian 분류기 구조 (b) 계층적 베이지안 망의 부분 구조

노드의 조건부 독립을 이용하면 베이지안 망에 의해서 설명되는 결합 확률분포는 다음과 같이 각 노드의 조건부 확률분포들에 의해서 식 (7)과 같이 인수분해 될 수 있다.

$$P(A_1, \dots, A_N) = \prod_{i=1}^N P(A_i | \Pi_i) \quad (7)$$

예로써 그림 4의 (a)는 naive Bayesian 분류기의 구조를 보여준다. 또한 naive Bayesian 분류기에 의해서 설명되는 결합 확률분포는 식 (8)과 같이 나타낼 수 있다.

$$P(A_1, \dots, A_N, C) = \prod_{i=1}^N P(A_i | C) P(C) \quad (8)$$

더 깊은 베이지안 망의 이론은 본 논문의 범위를 벗어나므로 참고문헌을 참고하기 바란다.

3.2 계층적 베이지안 망

앞에서 계층적인 정보를 표현할 수 있도록 구조화된 가버 특징들을 추출하였다. 본 논문에서는 보다 좋은 인식 성능을 위해서 가버 특징들의 의존도를 표현하는 베이지안 망으로써 계층적 베이지안 망을 제안한다. 제안하는 계층적인 베이지안 망은 임의의 한 샘플링 포인트에서 추출된 가버 특징과 하위 레벨에서 분할된 9개의 샘플링 포인트들에서 추출된 가버 특징들 사이의 직접적인 의존도를 표현한다. 즉 계층적 베이지안 망은 임의의 특징과 그것보다 더 국부적인 특징들 사이의 의존도를 표현하며, 인식을 위해서 분류기로 이용된다.

클래스 노드 C를 포함하지 않은 계층적 베이지안 망의 구조는 DAG = <V, E>로 정의된다. A^k을 가버 특징들을 위한 노드 집합 V의 한 노드라 하자. 또한 앞서 언급한 샘플링 포인트 p^k와 노드 A^k는 일대일 대응 관계가 있다고 하자. 랜덤 변수 A^k는 앞서 언급한 가버 특징 a^k를 값으로 가진다. 그러면 노드 집합 V는 다음과 같이 정의된다.

$$V = \bigcup_{l=1}^{N_L} A^l, \quad A^l = A^1, \dots, A^k, \dots, A^{N_l} \quad (9)$$

N_l은 레벨 l에서의 샘플링 포인트들의 총 수이며, A^l은 레벨 l에서의 노드들의 집합이다. 노드 집합 Φ^k 를 하위 레벨에서 분할된 p^k의 9개 샘플링 포인트들에 대응하는 노드들의 집합이라고 하자. 그러면 노드 집합 Φ^{k^s} 는 식 (10)으로 정의된다.

$$\Phi^{k^s} = B_1^k, \dots, B_i^k, \dots, B_9^k,$$

$$B_i^k \in A^{l+1}, \quad l=1, \dots, N_L - 1 \quad (10)$$

여기서 N_L은 총 레벨 수이다. 그러면 가버 특징들 사이의 계층적인 의존도를 표현하기 위한 간선들의 집합 E는 식 (11)과 같이 정의된다.

$$E = \bigcup_{l=1}^{N_L-1} E^l, \quad E^l = E^1 \cup \dots \cup E^k \cup \dots \cup E^{N_l},$$

$$E^k = (A^k, B_1^k), \dots, (A^k, B_i^k), \dots, (A^k, B_9^k) \quad (11)$$

여기서 B_i^k ∈ Φ^{k^s} 이고 l=1, ..., N_L-1이다.

계층적 베이지안 망에서 직접적인 의존도는 노드 A^k와 하위 레벨의 9개의 노드들(B_i^k ∈ Φ^{k^s}) 사이로 제한되지만, 전체적으로 계층적인 의존도를 표현하게 된다. 그림 4의 (b)는 이 제한된 직접적 의존도 구조를 보여준다.

계층적 베이지안 망을 분류기로써 사용하기 위해서는 클래스 노드 C가 V의 모든 노드들의 부모노드로서 포함하도록 수정되어야 한다[13]. 수정된 계층적 베이지안 망의 구조 DAG' = <V', E'>는 식 (12)로 정의된다.

$$V' = V \cup \{C\}, \quad E' = E \cup E_c, \quad E_c = \{(C, A^k) : \text{for all } A^k \in V\} \quad (12)$$

DAG'를 구조로 가지는 계층적 베이지안 망을 완전히 정의하기 위해서는 조건부 확률분포 집합 CP를 정의해야만 한다. 계층적 베이지안 망은 연속 랜덤 변수와 이산 랜덤 변수 모두를 가지고 있다. 계층적 가버 특징들을 위한 V의 변수들은 연속 랜덤 변수이며 클래스 변수 C는 이산 랜덤 변수이다. Π_k^s 는 DAG'_H에서의 A^k의 부모 노드의 집합이라 하면, Π_k^s 는 연속 랜덤 변수와 이산 랜덤 변수 모두를 포함하고 있다. 그러므로 각 연속 랜덤 변수 A^{ks}의 조건부 확률분포 P(A^{ks} | Π_k^{ks})는 조건부 multivariate Gaussian 분포[11]로 정의된다. 또한 부모 노드를 가지고 있지 않은 클래스 노드 C의 조건부 확률분포 P(C)는 multinomial 분포[10,11]로 정의된다. 계층적 베이지안 망은 은의 랜덤 변수를 가지고 있지 않기 때문에 각 조건부 확률분포의 파라미터들은 학습 데이터로부터 ML(Maximum Likelihood) 방법으로 학습될 수 있으며, 본 논문에서도 성능평가를 위해서 학습 이미지로부터 계층적 가버 특징들을 추출한 후 ML 방법을 사용하여 파라미터들을 학습하였다.

α 를 추출된 계층적 가버 특징들의 집합이라고 할 때,

계층적 베이지안 망을 이용하여 분류하기 위해서 \mathbf{a} 가 주어진 후 클래스 변수에 대한 $P(C | \mathbf{V} = \mathbf{a})$ 을 추론하여야 한다. 그래서 가장 큰 $P(C = c | \mathbf{V} = \mathbf{a})$ 을 주는 클래스 c 로 \mathbf{a} 를 분류한다. $P(C = c | \mathbf{V} = \mathbf{a})$ 는 베이지안 망의 추론 알고리즘[9]을 이용하여 구할 수 있으며 본 논문에서는 성능평가 시 junction tree 알고리즘[12]을 사용하였다.

4. 성능평가

제한한 계층적인 가버 특징들의 효율성을 간접적으로 보여주기 위해서 참고논문 [7]에서 성능평가에 사용된 학습 이미지 데이터를 사용하였다. 참고논문 [7]에서는 본 논문에서 제안하는 가버 특징과 다른 특징을 사용하여 tolerant rough set, k-nearest neighbor, 그리고 신경망으로 분류하였다. 그러므로 본 논문의 성능평가에서는 제한한 계층적 베이지안 망의 분류 성능평가에 주안점을 두었다.

성능평가를 위해서 UCI(University of California, Irvine) 데이터베이스의 이진 필기체 숫자 이미지 데이터[14]들을 사용하였다. 그림 5는 사용된 일부 데이터를 보여준다. 가버 특징들의 계층적인 의존도가 성능을 향상시킬 수 있다는 것을 보여주기 위해서 제한한 계층적 베이지안 망은 널리 알려진 naive Bayesian 분류기, k-nearest neighbor 분류기, 그리고 신경망 분류기들과 함께 동일한 계층적 가버 특징들에 대해서 평가되었다. 또한 성능평가에 사용된 모든 방법들의 학습에 대해서도 동일한 계층적 가버 특징을 그들의 입력으로 사용하였다.

성능평가는 다음과 같은 조건에서 수행되었다. 학습 데이터를 위해서 무작위로 클래스 당 25, 50, 300의 숫

자 이미지 데이터를 선택하였고, 평가 데이터를 위해서 남아 있는 숫자 이미지 데이터들 중 무작위로 1,934개를 선택하였다. 또한 학습 데이터와 평가 데이터는 중복되지 않았다. 표 1은 클래스 당 평가 데이터 수를 보여준다. 계층적 가버 특징들을 추출하기 위해서 가버 필터의 허수 부분을 사용하였다. 가버 필터 집합을 위해서 주파수 집합 $\Omega = \{0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.4, 0.6, 0.8, 1\}$, 방향 집합 $\Theta = 0, \frac{1}{4}\pi, \frac{1}{2}\pi, \frac{3}{4}\pi, \pi$, $r=2$ 을 사용하였으며, 총 레벨은 3으로 하였다.

주어진 학습 이미지 데이터로부터 각 레벨마다 최적화된 가버 필터들은 앞에서 언급한 방법으로 FLD 척도에 의해서 선택되었다. 그림 6은 선택된 최적화된 가버 필터들을 보여준다. 계층적 베이지안 망과 naive Bayesian 분류기의 조건부 확률 분포들의 파라미터들은 학습 이미지 데이터의 계층적 가버 특징들을 가지고 ML 방법을 사용하여 학습되었다.

계층적 베이지안 망과 naive Bayesian 분류기는 59개의 가버 특징들을 위한 노드들을 가졌으며, 각각의 가버 특징들은 4개의 방향들에 대한 4차원 벡터였다. 계층적 베이지안 망과 naive Bayesian 분류기의 가장 큰 차이점은 naive Bayesian 분류기는 가버 특징들 사이에 어떠한 의존도도 표현하지 않는 것이다. k-nearest neighbor 분류기와 신경망 분류기 역시 특징들 사이의 어떠한 의존도도 표현하지 않았으며, 그들의 입력 형식에 맞게 계층적 가버 특징들은 변형되었다. 즉 k-nearest neighbor 분류기는 하나의 236(=59×4) 차원 벡터를 사용하였으며, 신경망 분류기는 236개의 입력 노드들을 사용하였다. 또한 신경망 분류기는 150개의 은의 노드, 학습 비율 $\eta=0.01$, momentum rate $\alpha=0.5$, 그리고 총 학습 반복수=10,000으로 파라미터들을 고정시켰다. 그림 7과 표 2는 30번의 반복 수행의 평균으로 분류 성능평가의 결과를 보여준다. k-nearest neighbor

표 1 클래스 당 평가 데이터 수

Class	0	1	2	3	4	5	6	7	8	9	Total
#	189	198	195	199	186	187	195	201	180	204	1934

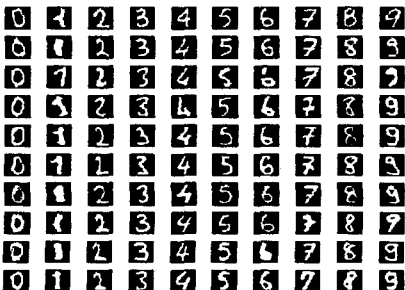
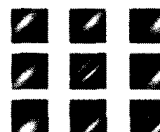


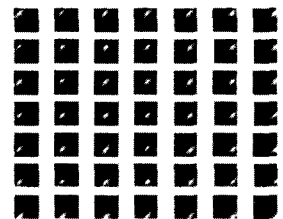
그림 5 성능평가에 사용된 이진 필기체 숫자 이미지 데이터



(a) 레벨 1



(b) 레벨 2



(c) 레벨 3

그림 6 방향 $\theta = \pi/4$ 을 가지는 최적화된 가버 필터들

분류기의 평가에서 $k=1, 3, 5$ 인 경우들을 수행하였으며 k 가 1일 때 가장 좋은 성능을 보여주었다.

같은 데이터에 대해서 500개의 학습 데이터(클래스 당 50)와 1434개의 평가 데이터에 대해서 참고문헌 [7]에서 보여 주었던 결과는 tolerant rough set에서 0.892, $k=1$ 일 때 k -nearest neighbor 분류기에서 0.873 그리고 신경망 분류기에서 0.889의 성능을 보여주었다. 본 논문에서 수행된 계층적 가버 특징들을 이용하는 모든 분류기는 500개의 학습 데이터에 대해서 위의 결과에서 보여 주었던 것보다 우수한 성능을 보여주었다. 위의 성능 평가 비교는 간접적으로 제안된 계층적 가버 특징들의 효율성을 보여준다. 총 레벨 2의 계층적 가버 특징들의 평가에도 모두 참고문헌 [7]의 결과보다 우수한 성능을 보였다. 또한 본 논문에서 수행된 4개의 분류기들의 성능평가에서 제안된 계층적 베이지안 망이 가장 우수한 성능을 보였다. 이는 분류 시 가버 특징들 사이의 계층적인 의존도를 고려하는 것이 성능의 향상을 준다는 것을 보여준다. 그러나 클래스 당 20개의 학습데이터의 경우 naive Bayesian 분류기가 더 나은 성능을 보였는데, 이는 계층적 베이지안 망은 의존도를 표현하기 위해서 naive Bayesian 분류기보다 더 많은 파라미터를 가지고 있기 때문에 클래스 당 20개의 학습데이터는 제안된 계층적 베이지안 망의 파라미터 학습에는 부족함을 보여준다.

표 2 클래스 당 300개의 학습 데이터일 때 인식율의 90% confidence interval. HBN : 계층적 베이지안 망, NBC : naive Bayesian 분류기, KNN : k -nearest neighbor 분류기, NN : 신경망 분류기

HBN	NBC	KNN with $k=1$	NN
0.9831 ± 0.0051	0.9666 ± 0.0061	0.9804 ± 0.0047	0.9755 ± 0.0055

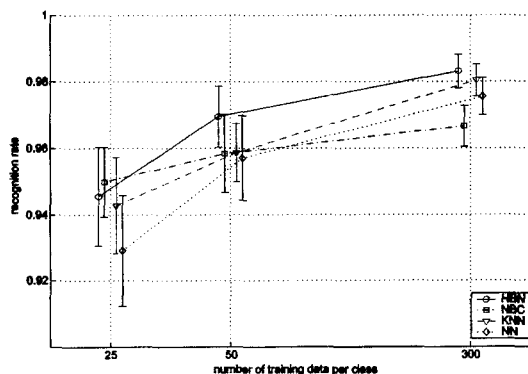


그림 7 성능평가 결과(세로 줄은 90% confidence interval을 나타냄)

5. 결론

본 논문에서는 서로 다른 레벨의 정보를 표현하는 계층적 가버 특징 추출방법과 분류를 위한 계층적 베이지안 망을 제안하였다. 계층적으로 구조화된 특징들을 추출하기 위해서 가버 필터를 고려하였다. 하위 레벨에서 9개로 분할되는 샘플링 포인트를 정의하였고, 가버 필터들은 각 샘플링 포인트를 중앙으로 하여 정의되었다. 최적화된 가버 필터는 FLD 척도를 이용하여 선택되었다. 계층적 베이지안 망은 추출된 가버 특징들 사이의 의존도를 표현하도록 생성하여 더 나은 인식성능을 가능하게 하였다.

제안한 방법들은 널리 알려진 naive Bayesian 분류기, k -nearest neighbor 분류기 그리고 신경망 분류기들과 함께 필기체 숫자인식에 성공적으로 적용되었다. 성능평가 결과로부터 계층적 가버 특징들의 효율성을 보여 주었으며, 가버 특징들의 계층적 의존도를 고려하는 것이 더 좋은 분류성능을 준다는 것을 보여주었다.

본 논문에서는 계층적 의존도를 표현하기 위해서 미리 정해진 9개의 노드들로 제한하였는데, 향후 더 좋은 인식성능을 위해서 레벨 간 가버 특징들 사이의 의존도 구조를 찾는 연구를 할 계획이며, 필기체 숫자 인식의 다른 응용분야에 적용을 할 계획이다.

참고 문헌

- [1] Gabor, D. : Theory of communication. J. Inst. Electr. Engng., vol. 93, pp. 429~459, 1946.
- [2] Bishop, C. M. : Neural Networks for Pattern Recognition. Oxford University Press, 1995.
- [3] Hoyer, P. O. and Hyärinen, A. : Independent component analysis applied to feature extraction from colour and stereo images. Network : Computation in Neural Systems, vol. 11, no. 3, pp. 191~210, 2000.
- [4] Porat, M. and Zeevi, Y. Y. : The generalized Gabor scheme of image representation in biological and machine vision. IEEE trans. PAMI, vol. 10, pp. 452~468, 1988.
- [5] Daugman, J. G. : Two-dimensional spectral analysis of cortical receptive field profiles. Vision Res, vol. 20, pp. 847~856, 1980.
- [6] Hamamoto, Y., Uchimura, S., Watanabe, M., Yasuda, T., Mitani, Y., and Tomita, S. : A Gabor Filter-Based Method for Recognizing Handwritten Numerals. Pattern Recognition, vol. 31, no. 4, pp. 395~400, 1998.
- [7] Daijin Kim and Sung-Yang Bang. : A Handwritten Numeral Character Classification Using Tolerant Rough Set. IEEE trans. PAMI, vol. 22, no. 9, pp. 923~937, 2000.
- [8] Tai Sing Lee. : Image Representation Using 2D

Gabor Wavelets. IEEE trans. PAMI, vol. 18, no. 10, pp. 959~971, 1996.

- [9] Pearl, J. : Probabilistic Inference in Intelligent Systems. Morgan Kaufmann, San Mateo, California, 1988.
- [10] Heckerman, D. and Geiger, D. : A tutorial on learning with Bayesian networks. In M.I.Jordan (Ed.), Learning in graphical models, Kluwer, 1998.
- [11] Cowell, R. G., Dawid, A. P., Lauritzen, S. L., and Spiegelhalter, D. J. : Probabilistic Networks and Expert Systems. Springer, 1999.
- [12] Lauritzen, S. L. and Spiegelhalter, D. J. : Local computations with probabilities on graphical structures and their applications to expert systems. J. Royal Statist. Soc, B, vol. 50, no. 2, pp. 154~227, 1988.
- [13] Friedman, N., Geiger, D., and Goldszmidt, M. : Bayesian network classifiers. Mach. Learn, vol. 29, pp. 131~163, 1997.
- [14] Blake, C. L. and Merz, C. J. : UCI Repository of Machine Learning Databases. Dept. of Information and Computer Science, University of California, Irvine, 1998.

부 록(Fisher's Linear Discriminant(FLD) me

FLD 척도는 주어진 정보가 분류를 위해서 얼마나 효율적인가를 평가하는 방법의 하나이다. FLD 척도는 within-class matrix와 between-class matrix에 의해서 정의되어진다. 즉 한 클래스 내에서는 정보의 분산이 작고, 클래스들 간에는 정보의 분산이 클수록 FLD 척도는 좋은 값을 가진다.

c 클래스 문제에 대해서, n개의 d차원의 예제 데이터, $x_1 \dots, x_n$, 들로 구성된 데이터 집합 χ 가 있다고 하자. 클래스 c_i 에 속한 데이터들을 χ 의 부분 집합 χ_i 의 원소라 하고, 원소의 개수를 n_i 라 하자. 그러면 S_w (within-class scatter matrix)와 S_B (between-class scatter matrix)는 다음과 같이 정의된다.

$$S_i = \sum_{x \in \chi_i} (x - m_i)(x - m_i)^T,$$

$$m_i = \frac{1}{n_i} \sum_{x \in \chi_i} x, \quad S_w = \sum_{i=1}^c S_i$$

$$m = \frac{1}{n} \sum_{x \in \chi} x = \frac{1}{n} \sum_{i=1}^c n_i m_i$$

$$S_B = \sum_{i=1}^c n_i (m_i - m)(m_i - m)^T$$

여기서 T 는 matrix의 전치 연산이다. 위와 같은 분산 척도로부터 Fisher's Linear discriminant 척도는 다음과 같다.

$$FLD = \frac{|S_B|}{|S_w|}, \text{ where } | \cdot | \text{ denotes determinant.}$$

본 논문에서는 x 는 스칼라 값이다.



성 재 모

2001년 인하대학교 전자계산산공학과(공학사). 2003년 포항공과대학교 컴퓨터공학과(공학석사). 2003년~현재 포항공과대학교 컴퓨터공학과 박사과정. 관심분야는 인공지능, 데이터마이닝



방 승 양

1966년 일본 Kyoto대학 전기공학에서 학사. 1969년 서울대학교 전기공학에서 석사. 1974년 미국 University of Texas 전산학에서 박사를 받았음. 미국 Wayne State University, NCR, Bell 연구소 등에서 근무하다가 1981년 귀국. 한국전자기술연구소 시스템부 실장, 부장 역임, (주)유니온시스템 전무. 1986년부터 포항공대 컴퓨터공학과 교수. 현재 뇌연구센터 소장. 관심분야는 패턴인식, 신경회로망.