2003년도 대한전자공학회 하계종합학술대회 제26권 제1호

평탄도 측정을 이용한 GMM 얼굴인식기 구현 및 성능향상

천 영 하, 고 대 영, 김 진 영, 백 성 준
전남대학교대학원 전자공학과
전화 : (062) 530-0472 / 팩스 : (062) 530-0472

Implementation and Enhancement of GMM Face Recognition System using Flatness Measure

Young Ha Chun, Dae Young Ko, Jin Young Kim, Sung Jun Back
Dept. of Electronics Engineering, Chonnam National University
catsby1004@empal.com

Abstract

This paper describes a method of performance enhancement using Flatness Mesure(FM) for the Gaussian Mixture Model(GMM) face recognition systems. Using this measure we discard the frames having low information before training and test. As the result, the performance increases about 9% in the lower mixtures and calculation burden is decreased. As well, the recognition error rate is decreased under the illumination change surroundings. We use the 2D DCT coefficients for face feature vectors and experiments are carried out on the Olivetti Research Laboratory (ORL) face database.

Ⅰ. 서론

사회가 점차 고도화되고 복잡화되면서 보안의 중요성이 커지면서, 생체인식 기술이 주목되고 있다. 얼굴인식은 그 한 분야로 오래 전부터 다양한 연구가 활발하게 진행중이며, 그 활용분야는 출입통제장치, 개인용 컴퓨터에 대한 접근 제어, 모바일 금융거래, 사이버대학 등의 온·오프라인에서 다양하다.

지금까지 얼굴인식을 위한 많은 방법들이 연구, 개발되었다[1]. 많이 사용되는 방법들로 Eigenface방법, LDA 방법 등이 있으며, 음성과 문자 인식에 성공적으로 사용된 HMM 방식도 얼굴인식을 위한 통계적 모델링 방법으로 사용되어졌다. HMM 방식[2][3]은 좋은 인식성능을 나타내고 있지만, 그만큼 파라미터 수와 계산량이 많고 모델링에 있어서 다소 복잡하다.

본 논문에서는 HMM보다 모델링이 간단한 GMM 기반 얼굴인식 시스템을[4] 구현하는 방법과, 국부적 특성을 모델링 하는 GMM에 평탄도 측정법을[5] 적용하여, 좀더 효율적인 모델링을 통해 인식 성능을 개선하는 방법에 대해서 기술한다.

Ⅱ. GMM 기반 얼굴인식

2.1 특징벡터 추출

본 논문에서 특징벡터는 2D DCT 계수를 사용한다. 픽셀값 대신 DCT계수를 이용함으로써 특징벡터의 크기와 인식 시스템의 복잡성을 줄일 수 있기 때문이다.
이는 한 장의 얼굴 영상을 8*8픽셀 윈도우를 중첩하여 나누는데, 블록간의 중첩은 6*6픽셀씩 이루게 한다. 각 이미지 블록당 추출한 특징벡터들 중 중요성분을 포함

하는 6개의 계수만을 사용한다. 이미지 한 장의 크기는 112*92이므로 중첩을 이루어 생성된 블록들은 총 2279 프레임이다. 따라서 한 장의 이미지에서는 6차원의 2279 개의 특징벡터들이 추출되는 셈이다.

2.2 GMM

가우시안 혼합모델(GMM)은 통계적인 모델링 방법으로, 영상에서 추출한 특징벡터들을 몇 개의 가우시안 성분(가우시안 함수)의 혼합으로 하나의 모델을 정의한다. 그 파라미터는 혼합계수, 평균, 공분산으로 식(3)과 같이 정의된다. 식(2)는 이렇게 정의된 혼합 모델과 다른 영상으로부터 추출한 특징벡터와 비교확률을 나타내고 있다.

$$L(X \mid \lambda_C) = \frac{1}{N_V} \sum_{i=1}^{N_V} \log p(\vec{x}_i \mid \lambda_C) \quad (1)$$

$$(\vec{x} \mid \lambda) = \sum_{j=1}^{N_M} m_j N(\vec{x}, \vec{\mu}_j, \Sigma_j) \quad (2)$$

$$\lambda = \{m_j, \vec{\mu}_j, \Sigma_j\}_{j=1}^{N_M} \quad (3)$$

λ_C : C라는 사람의 모델

N_V : 특징벡터들의 개수

N_M : 가우시안 혼합 성분 개수

m_j : Mixture j에 대한 혼합 가중치 계수

$N(x, \mu, \sum)$: 다변수 가우시안 함수

2.3 학습화 과정

위의 모델 파라미터들을 학습화하기 이전에 특징벡터들을 균등하게 나누어 초기화 모델을 만든다. 초기화 된 모델은 expectation maximization(EM)[6] 알고리즘을 통해 학습화 과정을 거치게 된다. 이는 학습화 데이터들이 M개의 가우시안 분포를 따르는 최적의 부분을 찾아주는 알고리즘으로, 학습화 데이터를 이용하여 식1과 같이 구한 사후 평균로그확률이 최대가 될 때까지 다음 과정을 계속 반복하면서 최종 모델을 생성한다.

· Expectation step :

$$l_{j,i} = \frac{m_j N(\vec{x}_i, \vec{\mu}_j, \Sigma_j)}{\sum_{n=1}^{N_M} m_n N(\vec{x}_i, \vec{\mu}_n, \Sigma_n)} \quad (4)$$

· Maximization step :

$$\hat{m}_j \Leftarrow \frac{1}{N_V} \sum_{i=1}^{N_V} l_{j,i} \quad (5)$$

$$\hat{\vec{\mu}}_j \Leftarrow \frac{\sum_{i=1}^{N_V} \vec{x}_i l_{j,i}}{\sum_{i=1}^{N_V} l_{j,i}} \quad (6)$$

$$\hat{\Sigma}_j \Leftarrow \frac{\sum_{i=1}^{N_V} l_{j,i} (\vec{x}_i - \hat{\vec{\mu}}_j)(\vec{x}_i - \hat{\vec{\mu}}_j)^T}{\sum_{i=1}^{N_V} l_{j,i}} \quad (7)$$

실제 인식하는 과정은 새로운 입력 테스트 영상의 특징벡터들과 학습화된 모델들간 평균로그확률을 식(1)을 통해 구하고 그중 최대확률이 나오는 모델로 인식하게 된다.

Ⅲ. 평탄도를 이용한 인식 성능 개선 방법

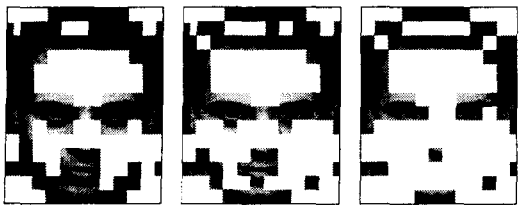
각각 가우시안 성분은 유사한 국부적 특성을 나타내는 것이므로, 모델을 세우기 전에 얼굴에 대한 정보가 적은 불필요한 영역을 제거하여 좀더 효율적인 모델링을 하고자 평탄도 측정법을 도입한다.

이는 원래 spectral flatness measure (SFM)라고 한다. 음성의 스펙트럼 공간에서 유성음 구간을 추출해 내는 방법으로 사용되어졌으며, magnitude의 산술평균과 기하평균의 비로 정의된다. 이를 얼굴인식에서 적용시에는 영상공간에서 픽셀값의 산술평균과 기하평균의 비로 구한다. 그렇기 때문에 우리는 그냥 Flatness Measure라고 명명하기로 한다.

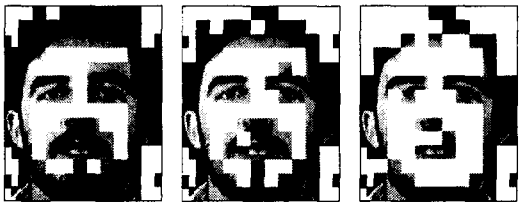
$$FM_{dB} = 10 \log_{10} \frac{G_m}{A_m} \quad (8)$$

G_m : 기하평균 A_m : 산술평균

영상에서 평탄도라는 것은 그 영역이 얼마나 저주파 성분을 많이 포함하고 있느냐를 나타내는 척도이다. 이는 곧 정보량이 적다는 것과 동일하다. 이 값이 큰 프레임들은 그림1 에서 보이는 바와 같이 얼굴에 대한 정보가 별로 없는 배경이라든지 이마, 뺨 등의 영역에 해당한다. 이런 영역들을 미리 잘라내어 유의미한 영역만을 남겨 효율적인 모델링을 하고자 했다.



(a) th = -0.05 (b) th = -0.08 (c) th = -0.15



(d) th = -0.05 (e) th = -0.08 (f) th = -0.15

그림 1. 평탄도 측정법을 적용한 후 얼굴 영상들

영역을 잘라내는 방법은 기준치(threshold)를 정해 이 보다 큰 값을 갖는 프레임들을 불필요한 영역으로 간주하고 잘라내게 된다. 이는 학습화와 테스트 전과정의 모든 영상에 적용된다. 기준치를 낮게 하면 그림 1의 (c)나 (f)와 같이 너무 많은 프레임들이 잘려나가 계산시간은 단축되지만, 오히려 성능개선에 도움이 되지 않는다. 여러 가지로 기준치를 변화시켜 실험해본 결과 -0.05와 -0.08에서 비교적 우수한 성능을 보였으며, 이는 전체 2279개 프레임 중 대략 42% ~54%의 프레임만을 사용하는 것이다. 이에 따라 계산량도 훨씬 줄어드는 효과를 볼 수 있다.

Ⅳ. 실험구성

4.1 데이터 베이스

실험은 ORL(Olivetti Research Laboratory) 데이터 베이스를 사용하였다. 이는 총 400장의 얼굴영상으로 구성되어 있으며, 40명에 대해 각각 10장의 이미지를 포함하고 있다. 우리는 개인 당 10장의 이미지 중에 5장을(총 200장) 학습화모델을 만드는데 사용하였으며, 나머지 5장들은(총200장) 테스트에 사용한다.

4.2 실험 방법

원래 평탄도는 스펙트럼 공간에서 측정하는 것이므로 우선은 주파수 공간에 적용시에는 어떤 결과를 보이는지 영상 공간에 적용했을 때와 비교실험을 한다. 그리고

delta 20 delta 40 delta 60 delta 80

그림 2. 조명변화에 따른 얼굴 영상

영상공간에서 평탄도를 적용한 것과 하지 않은 것을 비교 실험하였다. 보통의 이미지와 인위적으로 조명변화를 준 이미지에 두 가지에 대해 실험하여, 실제와 유사한 환경에서도 견고한 성능을 보여주는지를 테스트하였다. 인위적으로 조명변화를 주는 방법은 아래와 같다. 식(9)를 통해 영상을 변환시키면 그림 2에 보이는 바와 같이 좌측에서 조명을 주는 효과가 나타나게 된다. 델타 값이 커질수록 더 강한 조명을 나타낸다. 본 논문에서는 델타 값을 0 ~ 80까지 20간격으로 변화시키면서 실험하였다.

$$v(y,x) = w(y,x) + mx + \delta \tag{9}$$

여기서, $$m = \frac{-\delta}{xSize/2} \tag{10}$$

Ⅴ. 실험결과

주파수 공간과 영상공간에 평탄도 적용 비교실험을 한 결과, 그림 3에 보이는바와 같이 영상공간보다 주파수 공간에서 성능이 더 저하되었다. 이로써 영상에서 평탄도 측정은 주파수공간이 아닌 영상공간에서 더 의미 있다고 판단되어, 영상공간에서만 평탄도 적용실험을 하였다. 그냥 보통의 이미지에 대해 평탄도를 적용하지 않았을 때 인식성능은 가우시안 성분개수에 따라서 70%~96% 의 인식률로 많은 차이를 보인다(그림 4). 여기에 평탄도를 적용하여 실험한 결과 30 mixture 미만에서는 인식성능이 향상됨을 볼 수 있었다. FM1은 기준치를 -0.05로 두고 잘라내어 평균 54%의 프레임만을 이용하였고, FM2는 기준치를 -0.08로 두고 잘라내어 평균 42%의 프레임만을 이용하여 학습화와 테스트를 한 결과이다. 전체프레임의 절반정도만을 이용하였지만, 결과는 최고 9%까지 향상되었다.

그림 5는 인위적으로 조명변화를 시킨 이미지에 대한 25 mixture에서 인식결과를 나타내고 있다. 평탄도를 적용하지 않았을 때는 약한 조명 변화 환경에서도(δ=20) 에러율은 급격하게 높아짐을 볼 수 있다. 하지만 여기에

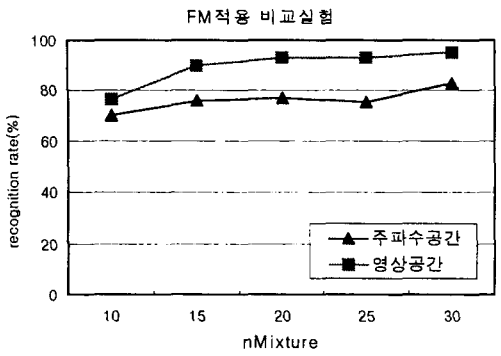


그림 3. 주파수공간과 영상공간에서 FM적용 인식결과

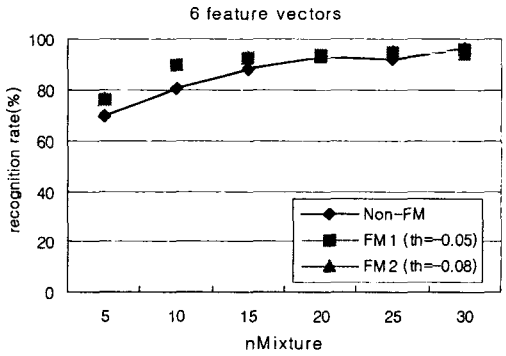


그림 4. 보통의 이미지에 대한 인식실험 결과

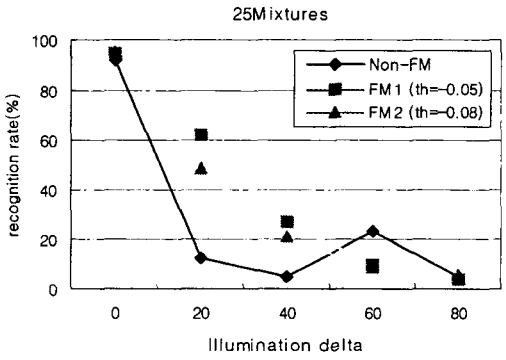


그림 5. 조명변화 이미지에 대한 인식실험 결과

평탄도를 적용하여 실험한 결과, 약한 조명변화 환경에서는 적용전보다 최고 40%이상 성능이 개선되었다.
평탄도를 적용하여 실험한 경우 원래 데이터의 절반 정도만 사용하기 때문에, 너무 많은 가우시안 성분을 이용할 경우 최적의 모델링과 거리가 멀어지게 되어 성능개선에는 도움이 되지 않았다.

VI. 결론

본 논문에서는 가우시안 혼합모델을 이용하여 얼굴인식 시스템을 구현하고, 평탄도 측정법을 도입하여 인식성능을 개선하고자 하였다.
평탄도 측정법을 적용한 결과 적용 안 했을 때보다 30 mixture 미만에서는 인식성능이 개선됨을 볼 수 있었고, 약한 조명변화 환경에서도 역시 더 강인함을 나타냈다.
결과적으로 평탄도 측정법이 국부적 특성들을 모델링하는 GMM기반 얼굴인식 시스템의 유용한 전처리 과정임을 입증하였고, 계산시간과 성능을 동시에 개선하면서 GMM 기반 실시간 얼굴인식의 가능성을 제시하였다.

참고 문헌

[1] R. Chellappa, C. Wison, and S. Sirohey, "Human and Machine Recognition of Faces: A Survey," *Prodeedings of IEEE, vol. 83,* pp. 705-740, May 1995.

[2] A. Nefian, *A Hidden Markov Model Based Approach for Face Detection and Recognition, PhD Thesis*, University of Georgia Institute of Techonology, August 1999.

[3] F. Samaria, *Face Recognition Using Hidden Markov Models, PhD Thesis*, University of Cambridge, 1994.

[4] C. Sanderson and K. K. Paliwal, "Likelihood Normalization for Face Authentication in Variable Recording Conditions," *Proc. Internaltional Conf. on Image Processing, Rochester*, New York, 2002.

[5] Robert E. Yantorno, Kasturi Rangan Krishnamachari and Jereme M. Lovekin, "The Spectral Autocorrelation Peak Valley Ratio," *IEEE International Workshop on Intelligent Signal Processing*, May 2001.

[6] L. Xu and M. I. Jordan. "On Convergence Properties of the EM algorithm for Gaussian mixtures," *Neural Computation,* 8(1): 129-151, 1996.