

웨이블릿 계수와 Hidden Markov Model를 이용한 얼굴인식 기법

The Method for Face Recognition using Wavelet Coefficients and Hidden Markov Model

이경아*, 이대중*, 박장환**, 전명근*

*충북대학교 전기전자 및 컴퓨터 공학부

** 충주대학교 정보 제어 공학과

Kyong-Ah Lee, Dae-jong Lee, Jang-Hwan Park, Myung-Geun Chun

* Chungbuk National University School of Electrical & Computer Engineering

** Chungju National University School of information & control Engineering

E-mail : mgchun@chungbuk.ac.kr

요 약

본 논문에서는 웨이블릿 계수와 Hidden Markov Model(HMM)이용한 얼굴인식 알고리즘을 제안 한다. 입력 영상은 이산웨이블릿을 기반으로 한 다해상도 분석기법을 사용하여 데이터 수를 압축한 후, 각각의 해상도에서 얻어진 웨이블릿 계수를 특징벡터로 사용하여 HMM의 모델을 생성한다. 인식단계 에서는 웨이블릿 변환에 의해 생성된 개별대역의 인식값을 더하여 상호 보완함으로써 인식률을 높일 수 있었다. 제안된 알고리즘의 타당성을 검증하기 위하여 기본적 알고리즘인 벡터 양자화(VQ) 기법을 적용한 경우와 기존 얼굴인식에 제안된 DCT-HMM을 이용한 기법과의 인식률 비교를 한 결과, 제안된 방법이 우수한 성능을 보임을 알 수 있었다.

1. 서론

최근 정보통신 기술이 급속히 발달함에 따라 인간의 삶의 질은 향상되어 가고 있지만, 컴퓨터 간 정보의 불법 복제 및 삭제, 불법 정보유출 등에 의한 사회적 손실도 증가하고 있다. 이러한 문제점을 해결하기 위하여 해킹, 누출에 의해 정보가 도용될 수 없고, 또한 변경되거나 분실할 위험성이 없는 신분 검증 기법인 생체인식 기술이 각광을 받고 있다[1]. 이러한 생체인식 기술은 인터넷 뱅킹 및 은행거래, 정확한 본인 여부 판별, 통신상에서의 비밀번호 유출 및 망각에 의한 문제를 해결할 수 있을 것으로 기대되고 있을 뿐만 아니라, 공항, 항만, 건물 로비 등 불특정 다수가 이동하는 공간에서 특정인을 찾아내는 시스템으로도 적합하며, 테러 용의자, 범죄자 등의

접근을 차단하는 최첨단 감시시스템으로서도 주목받고 있다.

이와 같은 생체를 이용하는 방법 중, 지문인식 시스템과 홍채인식 시스템 등이 현재 상용화되어 활발히 사용되고 있는 추세이다. 또 다른 생체인식 기법인, 얼굴인식 방법은 지문을 이용한 방법에서 나타나는 접촉식 이라는 단점을 해결하고 홍채인식 과정에서 발생하는 사용자의 거부감을 최소화하여 개인을 식별할 수 있는 방법이다. 물론 얼굴인식을 하기 위해선 아직 해결해야 할 많은 문제점을 가지고 있다. 사람의 얼굴이 각 개인을 식별할 수 있는 많은 정보를 가지고 있다는 장점을 가지고 있지만 얼굴을 검출하고 인식하는 과정에서 외부 환경에 많은 영향을 받는 단점을 가지고 있다.

얼굴인식을 위해선 우선 정확한 얼굴의 특징을 추출하는 과정이 중요하다. 특히, 얼굴의 특징은 정확하게 찾기 힘들고 사람의 얼굴 특징은 대부분 비슷하기 때문에 특징 추출 시 발생하는 작은

감사의 글 : 본 연구는 한국과학기술단 목적 기초 연구(R01-2001-000-00315-0(2002) 지원으로 수행되었음.

오차가 인식에 큰 영향을 미치게 된다. 이러한 특징을 추출하는 기법 중, 주성분 분석(Principal Component Analysis)[2]을 이용한 방법은 대표적인 알고리즘이긴 하나, 조명이나 환경변화에 민감하고 학습영상 수가 증가함에 따라 계산의 양이 많아지는 단점을 가지고 있다.

이러한 문제를 해결하기 위해 새롭게 제안되고 있는 알고리즘 중에 하나가 HMM[3]을 이용한 얼굴인식기법 이다. HMM을 이용하여 기존에 제안된 얼굴인식기법 으로는 DCT계수를 이용한 HMM 얼굴인식기법이 있다[3][4]. DCT는 영상 데이터를 저주파 성분과 고주파 성분으로 나누는 특징을 가지고 있다. DCT로 변환된 영상은 푸리에 변환처럼 영상의 에너지가 낮은 공간주파수 쪽에 집중 되는데 여기에서 저주파 성분은 부호화 되고 고주파 성분은 제거되어 저주파 성분 만으로 영상을 표현한다. 이에 반해 웨이블릿은 다 해상도 분석을 통하여 저주파의 성분과 고주파의 성분을 모두 포함한다.

본 논문에서는 웨이블릿 변환에 의해 저주파의 성분과 고주파의 성분을 모두 포함한 4개 대역의 특징벡터를 산출한 후 HMM 모델 생성 및 최종 인식하는 기법을 제안하고자 한다.

2. 웨이블릿 변환과 HMM의 이론적 고찰

2.1 웨이블릿 변환

웨이블릿 변환(Wavelet Transform)은 비 주기적인 신호분리가 가능한 기저함수를 사용하여 신호를 해석하는 것으로 신호를 형성하고 있는 주파수가 다른 두 개의 사인함수와 하나의 델타 함수를 "시간-스케일" 공간에 정확하게 분리해 낸다[6]. 또한, 웨이블릿 변환은 직교변환의 일종으로서 식(1)과 같이 정의 할 수 있으며, 시평면 신호 $x(t)$ 에 대하여 다중윈도우(multi window) 기능을 제공함으로써 다중분해능 해석을 가능하게 한다.

$$CWT_x(\tau, a) = \frac{1}{\sqrt{a}} \int x(t)h^*\left(\frac{t-\tau}{a}\right)dt \quad (1)$$

$$x(t) = c \int_{a>0} \int CWT(\tau, a) h_{a,\tau}(t) \frac{dad\tau}{a^2} \quad (2)$$

2차원의 경우 웨이블릿은 아주 작은 비트 율로 정보를 표현함에도 불구하고 영상의 전체적인 정보뿐만 아니라 에지와 같은 미세한 정보도 스케일 계수로 모두 유지 시킬 수 있다. 따라서 계수들이 변환 전 영상의 위치정보를 포함하기 때문에 사용자가 원하는 영상정보를 변환 후에도 유지시킬 수 있다. 그리고, DWT를 영상신호에 적용하는 것은 영상을 공간상의 x축과 y축방향으로 저대역 통과필터(LPF)와 고대역 통과 필터(HPF)를 사용하여 신호를 추출하는 것을 의미하

는 것으로, DWT를 거친 신호는 총 네 개로 분리될 수 있다.

이때, 분리된 영역은 그림 1과 같이 영상에 적용된 필터에 따라 구분된다. 위 그림에서 LL영역은 영상에 저대역 통과 필터를 적용한 것으로, 영상에서 고주파 성분이 배제된 계수들로 이루어져 있다. 또한, HH영역은 영상에 고대역 통과 필터를 사용한 것으로 LL과는 반대의 고주파 성분만 나타나게 된다. 따라서 LL영역의 주파수 대역은 $0 < \omega < \pi/2$ 이고, HH영역의 주파수 대역은 반대로 $\pi/2 < \omega < \pi$ 이다. 또한, HL영역과 LH영역의 주파수 대역은 공간상에서 x축과 y축 방향으로 서로 다른 필터를 사용했기 때문에 공간적인 방향에 따라 달라질 수 있다.

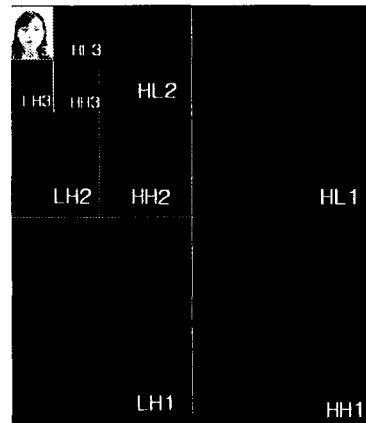


그림 1. 웨이블릿 변환후 4개의 밴드로 분리된 영상

2.2 Hidden Markov Model

HMM은 어떤 관측할 수 있는 시그널 과정에는 상태가 있다는 가정을 통하여 새로운 상태가 바로 전 상태에 의존하게 되는 상태전이 확률과 각각의 전이가 일어난 수 관측된 심벌이 현재의 상태에 의존하는 관찰확률을 구하게 된다.

HMM은 세가지 요소로 구성되어 있는데 상태(state)의 개수, 시간에 따른 상태의 변화를 결정하는 상태전이 확률분포, 그리고 각 상태에서의 출력심벌의 확률분포이다.

(1) 상태전이 확률 분포 $A = a_{ij}$

$$a_{ij} = P[q_{t+1} = j | q_t = i], \quad 1 \leq i, j \leq N$$

a_{ij} 는 임의 시각에 상태 q_i 에 있을 때 다음 순간에는 상태 q_j 로 천이할 확률을 나타낸다.

(2) 관찰심벌확률분포 $B = b_j(k)$

$$b_j(k) = P[v_t = k | q_t = j]$$

은닉 상태들 간의 전이 확률, 이전 상태에서 다음 상태로의 전이하는 확률 $b_j(k)$ 는 t시각에 상태 q_j 에 있을 때 k 번째 심벌 v_k 를 관찰할 확률을 나타낸다.

(3) 초기상태전이 확률분포 $\Pi = \pi_i$ 은닉 상태들의 초기 확률을 나타내는 벡터 $\pi_i = P(q_1 = i)$ 이것은 파라미터 A의 특별한 경우로 볼 수도 있는데 초기상태에서 상태의 확률분포를 나타낸다. 이상의 정의를 이용하여 HMM은 $\lambda = (A, B, \pi)$ 나타 낼 수 있다.

3. 웨이블릿 변환과 HMM을 이용한 얼굴 인식 시스템

본 논문에서 제안한 얼굴인식 시스템은 그림 2에서 보는 바와 같이 특징추출부, 코드북 형성부, HMM 모델 생성부, 인식부로 구성되어 있다.

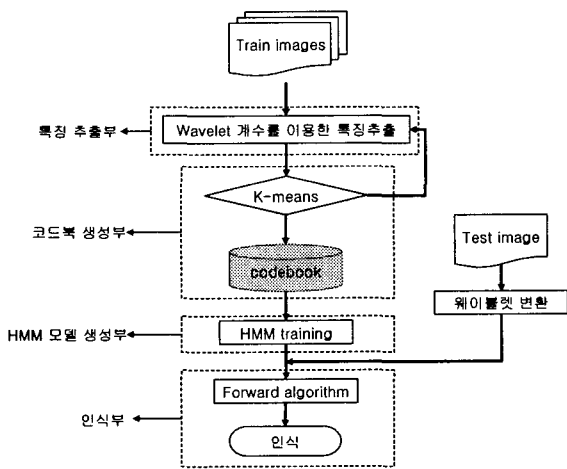


그림 2. 얼굴인식 시스템 개요도

특징 추출부에서는 입력영상을 웨이블릿 변환한 후 각각의 대역별로 특징벡터를 독립적으로 추출한다. 즉, 160×120의 학습 영상을 3번의 웨이블릿 과정을 통해서 20×15해상도를 가지는 네 개의 대역을 선택하여 특징 벡터로 사용한다. 이때 선택된 영상은 20×15해상도의 얼굴 영상 중에서 저대역 밴드 LL 영역과 고대역 부밴드 LH, HL, HH의 방향 성분을 가지게 된다. 네 개의 영상을 사용하기 때문에 데이터의 크기는 커졌지만 서로 다른 방향성분을 가진 네 개의 이미지를 얼굴인식에 사용하기 때문에 더 좋은 인식률을 보이게 된다. 코드북 형성부에서는 웨이블릿 변환에 의해 얻어진 특징벡터들을 K-means 클러스터링 알고리즘을 이용하여 입력영상을 대표하는 코드북을 생성한다.

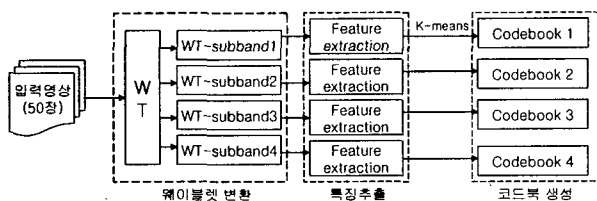


그림 3. 특징추출 및 코드북 생성 과정

HMM 모델 생성 및 인식과정을 그림 4에 나타냈다. 우선 모델 생성과정에서는 유클리디안 방법을 이용하여 HMM의 관측벡터인 입력영상의 특징벡터와 기준벡터인 코드북과 비교하여 벡터 공간상의 거리가 가장 가까운 코드북의 인덱스를 산출한 후, Baum-Welch 알고리즘에 의하여 각각의 얼굴 영상별 HMM의 모델변수인 A, B, π 를 결정한다. 인식부에서는 모델 생성과정과 동일하게 Test Image에 대해서 웨이블릿 변환을 통하여 특징벡터를 추출한 후 코드북과의 비교를 통하여 관측벡터인 인덱스를 산출하고, 각각의 영상별 HMM 모델에 적용하여 웨이블릿 대역별 매칭도를 산출하고, 최종적으로 개별적인 대역별 매칭값의 합산에 의해 최종 산출된 매칭도가 가장 높은 모델을 선택하여 인식한다.

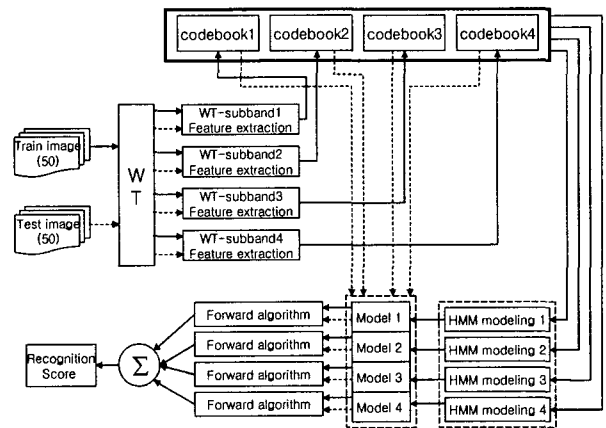
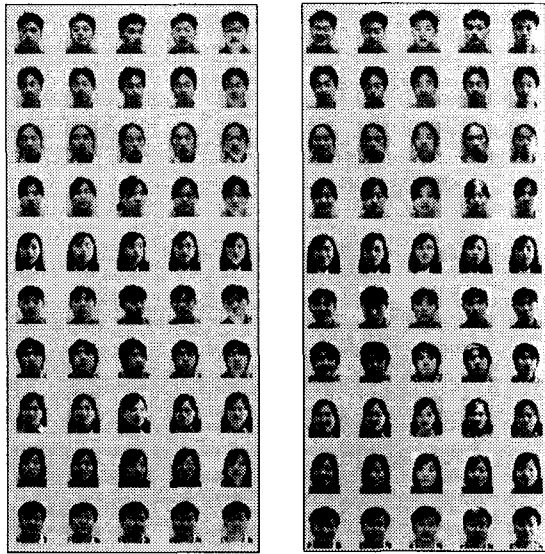


그림 4. HMM 모델 생성 및 인식과정

4. 얼굴인식 실험 및 결과

본 논문에서 제안한 방법의 우수한 성능을 보이기 위하여 CBNU라는 얼굴 데이터베이스를 구축하여 사용하였다. 실험에 사용된 CBNU 영상은 10명의 사람을 대상으로 각각 10장의 영상, 즉, 총 100장의 얼굴영상으로 구성되어 있다. 본 실험에서 홀수는 학습영상, 짝수는 테스트 영상으로 사용하여 각각 50장의 학습영상과 검증영상으로 구분하였다. 그림5-(a)는 학습영상으로 각 영상의 크기는 640×480 이다. 그림5-(b)는 검증영상으로 영상의 크기는 학습영상과 동일하다. 여기서, 웨이블릿 계수의 차수를 15차로 설정하기 위하여 원영상(640×480)을 160×120으로 재조정하였다. 즉, 다해상도 분석을 위해서 학습영상에 대해 이산 웨이블릿방법을 3번 적용함으로써, 20×15의 4개의 해상도 영상을 얻을 수 있다. 여기에서 4개의 해상도 영역은 다해상도 분할 방법의 마지막 단계로 이전대역의 LL(저주파영역)영역의 정보만을 연속으로 분해하여 얻어진

영역들이다. 마지막으로 사영된 영상은 각각의 대역별로 저장한다. 저장된 영상은 각각 K-means 알고리즘을 통하여 영상별 기준 코드북으로서 저장이 된다. 이때 코드북의 크기는 64로 설계 하였다.



(a) 학습영상 (b) 검증영상

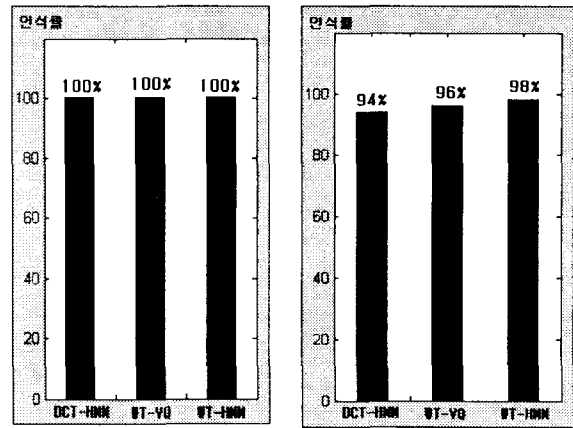
그림 5 사용된 영상

표 1에서는 본 논문에서 제안한 웨이블렛 계수와 HMM을 이용한 얼굴인식 기법을 적용한 경우 각각의 대역에 대한 인식률과 최종 인식률을 나타내었다. 표 1에서 보는 바와 같이 영상의 독립적인 대역별로 나온 인식에 대한 확률값을 산술적으로 더해 줌으로써 각각의 대역이 상호 보완하는 방식이 어느 특정 대역을 인식으로 쓰는 것보다 더 우수한 인식 값이 나오는 것을 보여 주고 있다.

표 1. WT-HMM을 사용한 인식률

구 분	1대역	2대역	3대역	4대역	total
인식률	94	68	72	46	98

본 논문에서 제안한 웨이블렛계수와 HMM을 이용한 얼굴인식의 우수한 성능을 비교하기 위해 HMM 과정 없이 단순한 유클리디안 비교만을 수행하는 일반적인 VQ기법과 현재 얼굴인식에 많이 사용되는 DCT-HMM을 비교 분석하여 그림 6에 나타냈다. 그림 6 에서 보는 바와 같이 실험 결과 WT-HMM은 DCT-HMM 보다는 4% VQ 보다는 2% 향상된 98%의 인식률을 보여 우수한 성능을 보임을 알 수 있다.



(a) 학습영상 (b) 검증영상

그림6 학습영상과 검증영상에 대한 인식률 비교

5. 결 론

본 논문에서는 웨이블렛 계수와 HMM을 이용한 얼굴인식을 제안하였다. 제안된 얼굴인식 시스템은 입력된 영상을 웨이블렛으로 다해상도로 분석하여 최종 확률값의 합이 가장 큰 모델을 인식하는 것으로, 단일 영역에서 인식하는 것 보다 인식률이 우수함을 볼 수 있었다. 또한 이전에 제안되었던 알고리즘과 비교해 볼 때 인식률이 상승 했음을 볼 수 있었다.

6. 참고문헌

[1] 전명근, 생체인식(Biometric) 총론, 한국정보통신교육원, pp. 8-1~21, 2002.
 [2]M. Turk, A. Pentland, "Eigenfaces for recognition", Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, 1991.
 [3] Rabiner,LR, "a tutorial on hidden markov models and selected applications in speech recognition" proceedings of the IEEE, Volume:77Issue:2,Page(s) :257-286 Feb.1989
 [4] Nefian, A.V.; Hayes, M.H., III"Hidden Markov models for face recognition" Acoustics, Speech, and Signal Processing, 1998. ICASSP '98. Proceedings of the 1998 IEEE International Conference on , Volume: 5 , 12-15 May 1998
 [5]Kohir, V.V.; Desai, U.B."Face recognition using a DCT-HMM approach" Applications of Computer Vision, 1998. WACV '98. Proceedings., Fourth IEEE Workshop on , 19-21 Oct. 1998
 [6] 이대중, 광근창, 유정용, 전명근, "웨이블렛 필터뱅크를 이용한 자동차 소음에 강인한 고립단어 음성인식", 한국퍼지 및 지능시스템학회논문지, Vol.12, No.2, pp.115-122, 2002