

HMM 을 이용한 얼굴 검출과 인식

박호식* · 차영석* · 최현수* · 배철수* · 권오홍* · 최철재** · 나상동***

*관동대학교 전자통신공학과, **동우대학 전자계산학과, ***조선대학교 컴퓨터공학과

Face Detection And Recognition using Hidden Markov Models

Ho-sik Park* · Young-seok Cha* · Hyun-soo Choi* · Cheol-soo Bae*

Oh-heung Kwon* · Chul-jae Choi** · Sang-dong Ra***

*Kwandong University, **Dong-U College, ***Chosun University

요 약

Hidden Markov Model(HMM)을 기반으로 한 얼굴 검출과 얼굴 인식에 대한 프레임작업에 대한 것이다. 관찰 벡터는 Karhunen-Loves Transform(KLT)의 상관관계를 이용하여 얻은 HMM의 정역학 특성을 사용하였으며, 본 연구에서 보여준 얼굴인식 방법은 이전의 HMM 기반의 얼굴인식 방법에서 인식률을 약간 개선함으로써 컴퓨터 연산을 훨씬 간단히 할 수 있음을 보여준다.

1. 서 론

비디오 영상으로부터 얼굴을 인식하고 검출하는 것은 여러 가지의 상업, 법적인 이유로 인하여 최근에 활발히 연구되고 있는 신종 연구분야 이다.

얼굴 검출, 인식 시스템은 자동현금인출기 혹은 제한 영역의 출입제한, 특별구역(은행, 가게 등)의 출입허가, 특정 DB(경찰 DB)의 관리 목적으로 컴퓨터 등에 사용자 허가를 위해 사용되어진다. 얼굴을 인식하여 동일인으로 판정하는 시스템은 다양한 조명과 배경 같은 여러 상황에서, 정면이 아닌 영상과 나이에 따른 변화에 있다하더라도 특정한 얼굴을 두 개 혹은 여러 개의 영상에서 추려낼 수가 있어야 한다.

Chellappa 는 얼굴 인식에 방법을 연구하였으며, Beymer, Turk, Belhumeur 등은 기하학적 특징을 기반으로 한 방법과 형판을 기반으로 한 방법으로 연구하였다. 최근에 Samaria 와 Nefian은 모델을 기반으로 한 방법으로 얼굴인식 시스템을 개발하였으며, Samaria는 HMM 기반의 얼굴인식 방법에서 정면얼굴의 머리, 이마, 눈, 코, 그리고 입 같은 중요한 얼굴 영역을 위로부터 아래로 나타낼 수 있는 방법을 연구하였다. 그러므로 이러한 얼굴의 부분은 좌우 1차원적인 연속적인 HMM을 상태를 할당한다.

본 논문에서는 HMM 방식을 이용하여 얼굴을 인식하고 검출하는 시스템을 연구하며, KLT 계수 같은 관찰 벡터 추출 같은 효과적인 방법을 나타내며, 일관된 얼굴 인식의 HMM 프레임작업은 새로운 얼굴 인식 방법을 제공한다. 전통적인 평면 기반의 방법과 비교하여 HMM 기반의 방법이 훨씬 쉬운 검출과 인식 작업을 제공한다. 여러 시스템에서 능률적으로 사용이 가능하다.

2. Hidden Markov Model

Hidden Markov Model(HMM)은 신호의 통계학적 특성을 사용한 통계학적 모델의 조합이다.

HMM 은 유한한 상태수를 갖는 Markov 연

결에서 밑에 있거나 관찰되지 않는, 확률 모형과 최초의 확률 기능의 연결로 구성되어 있다.

HMM의 요소는:

- 모델의 상태의 수가 N , 만약 상태의 조합인 S 이면 $S = \{ S_1, S_2, \dots, S_N \}$ 이고, t 시간에 모델이 갖는 상태 $q_t \in S$, $1 \leq t \leq T$ 이다. T 는 관찰 결과의 길이 이다. (프레임의 수)

- Π , 첫 번째 상태 분산, 그러므로 $\Pi = \{ \pi_i \}$ 에서:

$$\pi_i = P[q_1 = S_i], 1 \leq i \leq N \quad (1)$$

- A , 확률모형의 상태변화
 그러므로 $A = \{ a_{ij} \}$ 에서:

$$a_{ij} = P[q_t = S_j | q_{t-1} = S_i], 1 \leq i, j \leq N, \quad (2)$$

$$0 \leq a_{ij} \leq 1,$$

그리고,

$$\sum_{k=1}^N a_{ik} = 1, 1 \leq i \leq N$$

- B , 확률모형 상태
 그러므로, $B = \{ b_j(O_t) \}$.

HMM의 연속 밀도의 상태는 연속 관찰 밀도 함수로 특성 지어진다. 대부분 일반적인 표시는 확률 밀도 함수(pdf)는 한정된 형태의 혼합이다.:

$$b_i = \sum_{k=1}^N c_{ik} N(O_t, \mu_{ik}, U_{ik}), 1 \leq i \leq N \quad (3)$$

c_{ik} 가 i 상태에서 k 를 혼합한 혼합상관관계일 때, N 의 손실 없이 의미 벡터 μ_{ik} 공

분산 모형 μ_{ik} 과 함께 가우시안 pdf 로 가정한다.

속기 표기를 사용하여, HMM 은 세쌍으로 규정된다.

$$\lambda = (A, B, \Pi). \quad (4)$$

3. HMM 얼굴 영상

HMM 은 화자 인식과 1차원 이상의 데이터의 활동 인식에 최근 성공적으로 사용되었고 본 연구에서는 흑백영상의 1차원 HMM의 인식과 검출기능에 대하여 연구하였다. 또한 얼굴의 정면 영상에서 중요한 얼굴영역(얼굴, 이마, 눈, 코, 입)은 심지어 영상이 약간 기울어져 있거나 수직으로 회전되어 있어도 위에서 아래로 나타낼 수 있으며 각각의 얼굴영역은 1차원 연속 HMM 으로 좌측으로부터 상태를 할당받고, 얼굴모델의 상태구조와 는 0이 아닌 변천 확률 a_{ij} 는 그림1과 같이 나타난다.

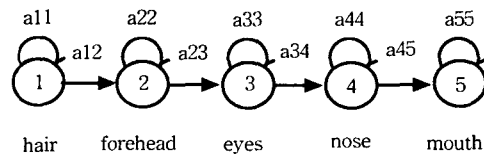


그림 1 : 얼굴 인식을 위한 좌측에서 우측으로의 HMM

4. 기능 추출

Wand 높이 H 의 각각의 얼굴 영상은 높이 L 과 너비 W 의 블록으로 겹쳐지면서 나누어진다. 연속적인 블록의 겹치는 양은 P 이다. (그림 2)

매개 변수 P 와 L 과 함께 인식 기능의 변화량은 광범위하게 연구되어왔다. 또한 시스템의 기능은 변수 L 의 변화량에 비하여 덜 민감하고, P 값은 큰 값을 나타낸다. ($P \leq L-1$).

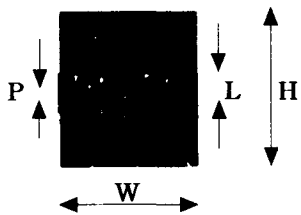


그림 2 : 얼굴영상의 매개변수와 블록 추출

F. Samaria는 관찰벡터 각 블록의 픽셀 값으로 구성되므로 관찰벡터의 넓이는 $L \times W$ 이다. 관찰벡터에서의 픽셀 값을 사용하여 두가지 손실이 있었다: 첫째, 픽셀 값이 영상의 회전이나 조명이 변함에 따라 매우 민감하므로 특징을 제대로 재현하지 못한다. 둘째, 관찰벡터가 큰 범위에서 검출과 인식 시스템의 훈련과정에서 높은 연산능력을 필요로 한다. 이것은 실시간으로 영상을 검출, 인식하는데 있어서 매우 중요한 문제이다.

A. V. Nefian은 각 블록의 2DCT 상관관계는 관찰벡터의 효과적인 조합으로 얻을 수 있었으며, 블록 에너지가 집중되는 DCT 영역에서 낮은 주파수의 직사각형 창에서의 상관관계는 관찰벡터를 사용하였다. 2D-DCT 상관관계의 선택으로 시스템의 복잡성을 극적으로 감소시켰다.

이 논문에서, 관찰벡터는 Karhunen Loeve Transform (KLT) 계수로 구성되어지며, 비상관관계 특성과 마찬가지로 KLT 압축 기능은 관찰벡터의 압축방법을 이루어낼 수 있었고, 훈련과정조합의 영상으로부터, 폼 벡터로의 컬럼-와이즈를 통해 블럭을 복원하고, KLT를 기본으로 한 벡터 형태의 공분산 모형의 큰 고정값은 고정 벡터에 상응하였다. 공분산 모형에서 계산된 벡터의 의미를 μ 라하고, 관찰 벡터를 얻어 의미 벡터 μ 를 영상의 블록과 상응하는 각각의 벡터값으로 나누어주고 결과 벡터는 공분산모형의 고정 벡터와 관찰벡터로부터의 결과 계수 위에 나타난다. 그림 3에서 일반적인 블럭의 관

찰벡터와 주요 얼굴영역과의 상관관계를 나타낸다.

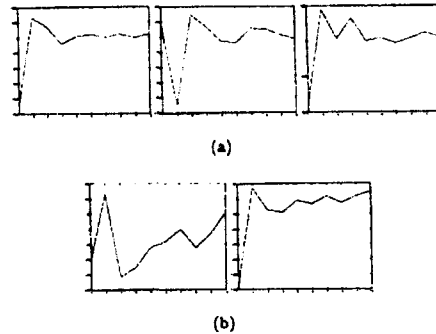


그림 3 : 일반적인 KLT 계수 for: a-hair(left), forehead(middle) and eyes(right), b-nose(left) and mouth(right).

5. 얼굴모델의 훈련

얼굴이미지의 조합에서 얼굴을 검출하기 위하여 하나의 HMM의 훈련을 이용하였다. 이미지 훈련 중에 다른 조명 하에서 다른 사람의 정면얼굴을 취하여 나타내었으며, 얼굴 인식에서 HMM 얼굴 모델에 의하여 데이터베이스에 각각 개인의 값을 나타내며 조합은 각 HMM 훈련을 이용하여 같은 얼굴을 나타내게 된다.

훈련과정조합에서 각 영상의 블럭을 복원한 후, 관찰벡터(KLT 계수)를 얻고 HMM의 각각 훈련하였다. 첫째, HMM $\lambda = (A, B, \Pi)$ 을 다음과 같이 초기화한다. 훈련 데이터를 $N=6$ 으로 위에서 아래로 한결같이 구분하고, 각 상태와 연결된 관찰벡터 관찰 확률모형 B 의 초기 예상값을 얻는다.

A 와 Π 의 초기 값을 얼굴모형의 좌측부터 우측으로 배열한다. 그림 4는 훈련연속 과정을 Viterbi 분절에 의하여 한결같이 분절을 바꾸고 다음 단계를 반복하며, 반복이 끝난 후, HMM을 초기화하고, Viterbi 분절이 임계값보다 작게 되도록 연속적으로

반복한다. HMM의 최종 변수를 Baum-Welch의 순환진행을 이용하여 얻는다.

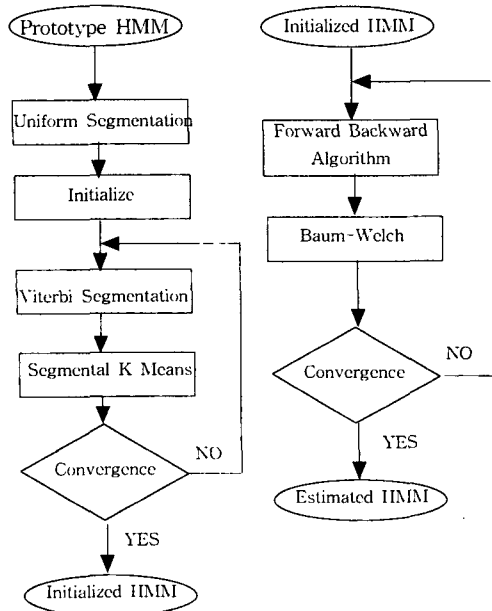


그림 4 : 훈련과정의 순서도

6. 얼굴 검출

실험 영상은 난잡한 배경의 하나 혹은 다수의 영상을 담고 있다, 실험 영상에서 각각의 직사각형 창안에서 얼굴 검출이 시작된다. 관찰벡터를 복원하고 얼굴모델에서 얻은 각 창의 데이터의 확률성을 계산한다. 얼굴 영상에서 인접한 창은 수직과 수평방향 모두 상당부분이 겹치게 된다. 간편화하기 위해서 Viterbi 알고리즘을 통해서 확률 연산을 수행하게 된다. 현재 창 위치에서 근접한 최대로 가능성 있는 창의 얼굴이 후보 얼굴로 선택된다. 이 접근은 짧은 관찰 벡터로부터의 얼굴 모형의 파괴되어지는 빠른 완성을 이룬다. 약간의 새 관찰 벡터의 연산이 요구되어지는 창의 가능한 얼굴이라면 연산은 연속적인 창사이의 이동량은 비교적으로 저렴한 방법인 Viterbi 분절과 일치한다. 또한

그림 5에 얼굴 검출 블록 다이어그램을 나타내었다.

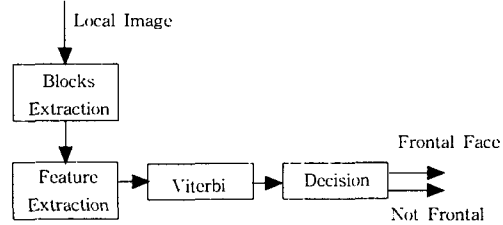


그림 5 : 얼굴 검출

그림 6은 얼굴 검출 결과를 보여주며 수평선은 Viterbi 상태 분절을 보여준다. 얼굴 검출 시스템은 MIT의 데이터베이스(배경 없이 16명의 48개의 60x90 픽셀의 영상)를 가지고 다른 조명 하에서 시험하였다. 9개의 수동으로 분류한 얼굴은 훈련을 위하여 사용하였다. 다른 39개의 영상을 시험하였다. 우리의 실험결과는 이 데이터베이스에서 90%의 검출을 보였다.

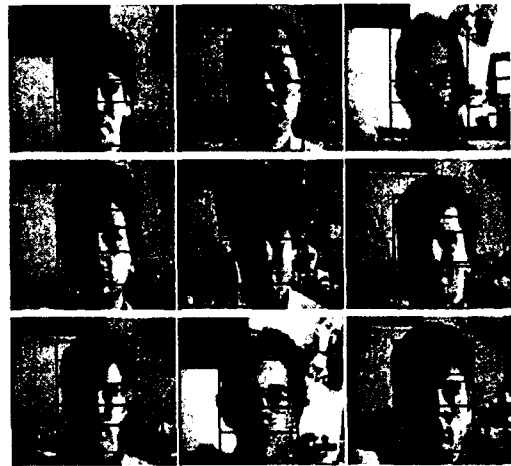


그림 6 : 얼굴 검출 결과

7. 얼굴 인식

HMM 얼굴 모델은 훈련을 통하여 관찰 벡터를 복원하고, 관찰결과의 확률성은 각각의 HMM 얼굴 모델을 Viterbi 인식기를 통하여 계산하였다. 높은 가능성을 지닌 모델을 선

택하고 이 모델은 모르는 얼굴에서 동일성을 드러냈다.

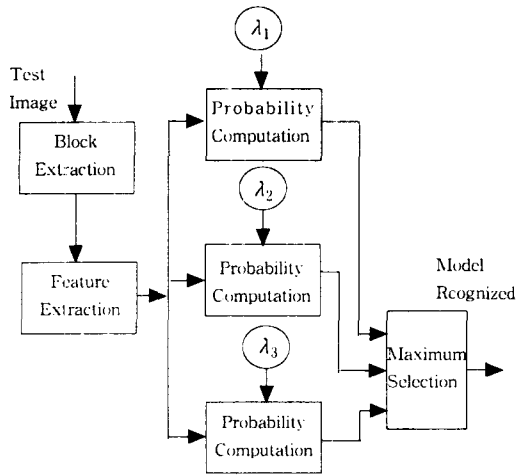


그림 7 : HMM 인식 개요

얼굴 인식 시스템의 Olivetti Research Ltd의 데이터베이스(10명의 40개의 다른 총 400개의 92x112 픽셀의 영상)를 가지고 시험하였다. 데이터베이스는 다른 얼굴 표현, 머리카락, 스타일, 안경 등의 영상으로 구성되었다. 방법[5]와 방법[6]의 인식 기능은 같은 데이터베이스 하에서 84% 이었다. 큰 용량의 관찰벡터 때문에 [5]의 컴퓨터 연산이 매우 복잡하게 되었다. 이 논문에서 보여준 시스템은 [6]에서 인식시간을 250ms에서 700 ms와 비교하였을 때 정확도를 86%까지 조금 향상 시켰다.

그림 8은 인식 결과를 나타낸다. 바르게 인식되고 남은 교차된 영상은 잘못 분류되었다.

8. 결 론

이 연구에서는 KLT 상관관계로부터 추출된 능률적인 관찰 벡터들을 이용하여 HMM을 기반으로 한 얼굴 인식과 검출에 방식에 대하여 설명하고 새로운 HMM을 기반으로 한 얼굴 검출 방식을 소개하였으며, 조명 같

은 자체내의 복잡한 변화에 주목하면서 이 방식이 정확도 등 얼굴 검출에 있어서 기존의 방식보다 뛰어난 것을 알 수가 있다.

인간 얼굴의 HMM 모델은 보다 넓은 영상의 범위를 표현을 할 수 있도록 하는 얼굴 인식, 검출방법으로서 매우 유망하다고 사료된다.



그림 8 : 얼굴 인식 결과

9. 참고문헌

- [1] R. Chellappa, C. Wilson, and S. Sirohey, "Human and machine recognition of faces: A survey," *Proceedings of IEEE*, vol. 83, May 1995.
- [2] D. Beymer, "Face recognition under varying pose," in *Proceedings of 23rd Image Understanding Workshop*, vol. 2, pp. 837-842, 1994.
- [3] M. Turk and A. Pentland, "Face recognition using eigenfaces," in *Proceedings of International Conference on Pattern Recognition*, pp. 586-591, 1991.
- [4] P. Belhumeur, J. Hespanha, and D.

Kriegman, "Eigenface vs Fisherfaces: Recognition using class specific linear projection," in Proceedings of fourth European Conference on Computer Vision, ECCV'96, pp. 45-56, April 1996.

[5] F. Samaria and S. Young, "HMM based architecture for face identification," Image and Computer Vision, vol. 12, pp. 537-583, October 1994.

[6] A. V. Nefian and M. H. Hayes, "Hidden markov models for face recognition," in ICASSP98, pp. 2721-2724, 98.

[7] L. Rabiner and B. Huang, Fundamentals of Speech Recognition, Englewood Cliffs, NJ: Prentice-Hall, 1993.

[8] E. Levin and R. Pieraccini, "Dynamic planar warping for optical character recognition," in ICASSP, pp. 149-152, 1992.

[9] S. Kuo and O. Agazzi, "Keyword spotting in poorly printed documents using pseudo 2-d Hidden Markov Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994.