

다선형 모델을 이용한 얼굴 및 표정 인식

박용찬*, 이성오**, 박귀태**
고려대학교

A Case Study on Face and Expression Recognition using AAMs and Multilinear Analysis

Yong-Chan Park*, Seong-Oh Lee**, Gwi-Tae Park**
Korea University

Abstract - 얼굴 인식은 얼굴의 특징적인 패턴을 이용하지만, 이러한 패턴은 표정, 포즈, 조명의 변화에 민감하여 인식에 어려움이 있다. 본 논문은 표정 변화에 강한 인식 모델을 개발하기 위해 Cohn-Kanade 표정 데이터베이스와 AAM을 이용하여 다양한 데이터를 추출하였고, 추출된 데이터를 다선형 분석을 이용하여 분석하였다. 이를 적용한 인식 실험에서 PCA보다 표정에 좀 더 강한 인식 성능을 나타내었다.

1. 서 론

얼굴 인식에 사용되는 얼굴의 특징적인 패턴은 표정, 포즈, 조명 등 다양한 요소들에 영향을 받는다. 따라서 기존의 PCA와 같은 단일한 요소의 특징적인 패턴을 이용하는 방법은 다양한 요소들의 상호 작용을 분석하는데 어려움이 따르며, 각각의 요소에 개별적인 모델을 구성해야 한다.

최근 얼굴 인식 연구에서는 다선형 분석(Multilinear analysis)을 이용하여 데이터 형성에 영향을 미치는 요소들의 상호 작용을 분석하고, 얼굴 및 표정 인식에 적용하고 있다[1],[6].

한편 얼굴 인식에 사용하기 위한 영상에 배경이 포함될 경우 인식과 인식 성능에 영향을 미치게 된다. AAM은 영상에서 얼굴을 추적하여 정확한 얼굴 영상을 추출하는데 있어 좋은 성능을 나타내고 있다[2].

본 논문은 정확한 얼굴 영상의 다양한 데이터를 추출하기 위해 AAM을 이용하였고, 표정 변화에 강한 성능을 나타내기 위해 다선형 분석으로 데이터를 분석하였다. 다양한 데이터에 따라 분석한 결과를 개별적으로 얼굴 인식에 적용하여 PCA와 비교 실험하였다.

2. 본 론

본 논문의 알고리즘은 AAM과 다선형 분석으로 구성되고, 본 논문에서 이 알고리즘을 다선형 모델이라고 표기한다. 다선형 모델은 AAM을 이용하여 정확한 얼굴 영상의 다양한 데이터를 추출하여 정규화한 후 다선형 분석을 적용하여 학습 모델을 구현한다.

2.1 Active Appearance Models

AAM은 통계학적 기반의 알고리즘으로, 영상에서 얼굴을 추적하고 정확한 얼굴 영상을 얻을 수 있다. 이러한 AAM은 현재 다양한 종류가 있으며, 본 논문에서는 독립적인 shape 모델과 텍스처(texture) 모델을 적용한 AAM을 이용하였다[2]. shape 모델은 가변적인 얼굴의 형태를 나타내며, 텍스처 모델은 shape 영역에 속한 화소들의 값을 나타내고, 이 두 가지 모델을 결합하여 가변적인 얼굴 영상을 나타낼 수 있다.

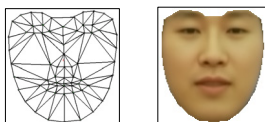
AAM의 shape s 는 mesh를 구성하는 점들의 집합으로 정의되며, 평균 shape s_0 와 n 개의 shape 벡터 s_i 의 선형 결합의 합으로 표현할 수 있다.

$$s = (x_1, y_1, x_2, y_2, \dots, x_n, y_n)^T$$

$$s = s_0 + \sum_{i=1}^n p_i s_i$$

AAM의 텍스처는 평균 shape s_0 내의 영상으로 정의되며, 동일한 방법으로 표현할 수 있다.

$$A(x) = A_0(x) + \sum_{i=1}^n \lambda_i A_i(x) \quad \forall x \in s_0$$



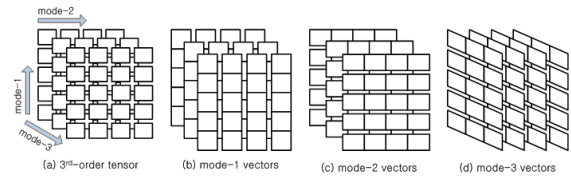
<그림 1> AAM shape과 appearance

2.2 다선형 분석(Multilinear Analysis)

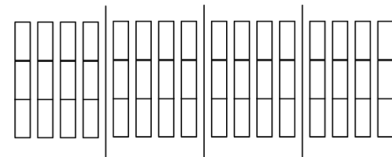
다선형 분석은 데이터 형성에 영향을 주는 다양한 요소들의 상호 작용을 단일한 구조에서 분석할 수 있다. 다선형 분석을 하기 위해 데이터는 벡터와 행렬의 고차원 일반화된 다차원 행렬(tensor)로 구성되어야 한다. 이러한 다차원 행렬은 N-mode SVD를 이용하여 분해할 수 있다.

$$T_{(i)} = U_i \Sigma V_i^T$$

행렬 $T_{(i)}$ 의 열벡터는 mode- n 벡터이며, 모드 행렬 U_n 의 열벡터는 mode- n 벡터들로 형성되는 mode space의 orthonormal 기저 벡터이다.



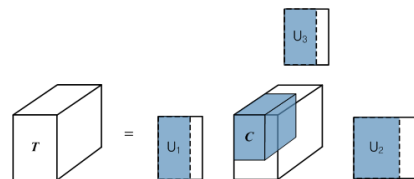
<그림 2> 3차 다차원 행렬과 mode 벡터



<그림 3> 3차 다차원 행렬의 $T_{(3)}$

n 차 다차원 행렬은 다음과 같이 분해되며, 다차원 행렬 C (core tensor)는 모드 행렬 사이의 상호작용을 나타낸다.

$$T = C \times_1 U_1 \times_2 U_2 \times \dots \times_n U_n$$



<그림 4> 3차 다차원 행렬의 N-mode SVD

2.3 다선형 모델

다선형 모델은 AAM을 이용한 데이터 획득과 다선형 분석을 통한 학습 두 단계로 구성된다. 우선 AAM을 이용하여 학습 영상에서 얼굴 영상을 추출하여 정규화 한다. 그 다음 정규화된 데이터를 다차원 행렬로 구성하여 다선형 분석을 적용한다. 본 논문에서는 인물과 표정 요소를 분리하기 위해 3차 다차원 행렬을 다음과 같이 구성하였다.

$$D = C \times_1 U_{\text{people}} \times_2 U_{\text{express}} \times_3 U_{\text{data}}$$

다차원 행렬 D 를 분해한 후 다차원 행렬 P 와 E 를 생성할 수 있다.

$$P = C \times_2 U_{\text{express}} \times_3 U_{\text{data}}$$

$$E = C \times_1 U_{\text{people}} \times_3 U_{\text{data}}$$

P 는 얼굴 영상을 인물 파라미터 공간으로 투영하는 기저 벡터의 집합이며, E 는 얼굴 영상을 표정 파라미터 공간으로 투영하는 기저 벡터의 집합이다. 이 두 다차원 행렬을 이용하여 얼굴 영상으로 구성된 다차원 행렬을 다음과 같이 나타낼 수 있다.

$$D = P \times U_{\text{people}}$$

$$D_{(\text{people})} = U_{\text{people}} P_{(\text{people})}$$

최종적으로 특정한 표정과 인물에 따른 얼굴 영상을 다음과 같이 나타낼 수 있다.

$$d_{p,e} = P_e^T c_p$$

U_{people} 의 행벡터 c_p 는 각각의 인물에 대한 파라미터이며, 다차원 행렬 $P_{(\text{people})}$ 의 일부분인 행렬 P_e^T 의 열벡터는 특정한 표정 e 에 대응하는 기저 벡터이다. 이러한 대응 관계에 따라 다양한 표정에 상관 없이 동일한 인물에 대해 같은 파라미터를 얻을 수 있다.

$$c_p = P_e^{-T} d_{p,e}$$

인식 과정은 학습 과정과 동일한 방법으로 AAM을 이용하여 입력 영상의 정규화된 얼굴 영상 데이터를 추출한다. 모든 표정에 대한 기저 벡터를 이용하여 입력된 얼굴 영상의 인물 파라미터 c_p 를 구한 후, 학습된 모든 파라미터 c_p 와 비교한다. 이때 가장 작은 오차로써 특정 인물 p 를 판단한다.

$$p = \text{argmin} \| c_p - c_e \|$$

표정 인식 역시 얼굴 인식과 동일한 과정으로 진행된다.

$$d_{p,e} = P_{p(\text{express})}^T c_e$$

$$c_e = P_{p(\text{express})}^{-T} d_{p,e}$$

3. 실험

다선형 모델의 학습과 얼굴 및 표정 인식 실험에 Cohn-Kanade 표정 데이터베이스를 이용하였다[7]. 이 데이터베이스에서 얼굴 인식과 표정 인식에 동일한 학습 영상과 실험 영상을 사용하였다. 한편 동일한 4가지 표정이 있는 인물의 수가 적어 학습과 실험에 적은 수의 데이터를 사용하였다. 학습에는 12명 대한 무표정, 놀람, 행복, 역겨움 4가지 표정의 영상 48장을 사용하였다. 실험은 동일한 4가지 표정의 10명에 대한 영상 31장을 사용하였으며, 몇 명에 대해서는 일부 표정만 사용하였다.

학습과 실험 영상에서 AAM을 이용하여 얼굴을 자동으로 찾은 후 3가지 데이터를 추출하여 실험에 사용하였으며 다음과 같다.

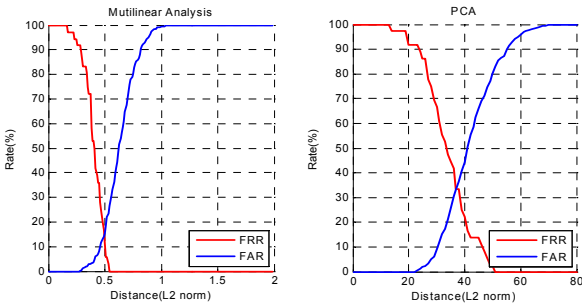
C-APP: AAM에서 찾은 얼굴 영상을 평균 shape으로 정규화한 영상

N-APP: AAM에서 찾은 얼굴 영상을 shape 정보를 이용해 회전,

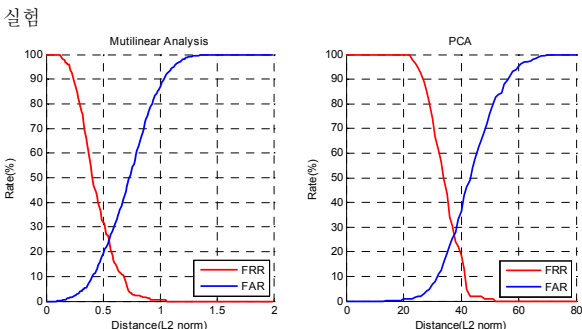
크기를 예상하여 정규화한 영상

N-S: 정규화된 shape 데이터

실험 결과는 FRR(False Rejection Rate)와 FAR(False Acceptance Rate)의 ERR(Equal Error Rate) 수치를 사용하였다.



<그림 5> C-APP를 이용한 얼굴 인식



<그림 7> C-APP를 이용한 표정 인식

<표 1> 얼굴 인식에 대한 ERR

	다선형 분석	PCA
C-APP	14.4%	33.3%
N-APP	15.2%	19.6%
N-S	46.7%	42.2%

<표 2> 표정 인식에 대한 ERR

	다선형 분석	PCA
C-APP	24.9%	26.7%
N-APP	31.2%	44.9%
N-S	27.6%	27%

실험 결과 얼굴 인식과 표정 인식에 가장 큰 영향을 미치는 얼굴 영상 데이터에 대하여 다선형 분석이 PCA 보다 더 낮은 ERR 수치를 나타내었다. 특히 C-APP를 이용한 얼굴 인식과 N-APP를 이용한 표정 인식에서 더욱 낮은 수치를 보였다.

얼굴 인식에서 본 논문의 다선형 모델은 표정에 대한 요소를 추출하여 인식에 적용하였기 때문에 PCA보다 표정 변화에 강인한 결과를 나타내고 있다. 또한 표정 인식에서도 인물에 따른 요소를 고려할 수 있기에 본 논문의 결과가 PCA보다 더 높은 성능을 보이고 있다.

N-S에 이용한 얼굴 인식의 결과는 PCA가 좀 더 낮은 수치를 나타내고 있다. 표정의 경우 N-S는 표정에 따라 큰 특징을 나타내는 반면에 인물의 경우 인물에 따른 특징을 나타내는 것이 부족하기 때문인 것으로 생각된다.

4. 결론

AAM과 다선형 분석을 이용한 다선형 모델을 구성하여 얼굴 및 표정 인식에 적용하였으며, 그 결과를 기존의 PCA와 비교하였다. 얼굴 인식 실험의 결과 다선형 모델이 PCA보다 표정 변화에 강인한 결과를 나타내었고, 표정 인식에서도 좀 더 좋은 성능을 나타내었다.

[참고 문헌]

[1] M.A.O. Vasilescu and D. Terzopoulos, "Multilinear analysis for facial image recognition", Proceedings of the International Conference on Pattern Recognition, Vol. 3, pp. 511-514, 2002.

[2] I. Matthews and S. Baker, "Active appearance models revisited", International Journal of Computer Vision, Vol. 60(2), pp. 135-164, 2004.

[3] M. Turk and A. Pentland, "Eigenfaces for Recognition", Journal of Cognitive Neuroscience, Vol. 3, pp. 72-86, 1991.

[4] T.F. Cootes, G.J. Edwards and C.J. Taylor, "Active Appearance Models", Proceedings European Conference on Computer Vision, Vol. 2, pp. 484-498, 1998.

[5] Ives Macedo, Emilio Vital Brazil, Luiz Velho, "Expression Transfer between Photographs through Multilinear AAM's", XIX Brazilian Symposium on Computer Graphics and Image Processing, pp. 239-246, 2006

[6] H. Wang and N. Ahuja, "Facial expression decomposition", In ICCV, Vol. 2, pp. 958-965, 2003.

[7] T. Kanade, J.F. Cohn & Y. Tian, "Comprehensive database for facial expression analysis", Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, pp. 46-54, 2000.