

PCA을 이용한 얼굴표정의 감정인식 방법

Emotion Recognition Method of Facial Image using PCA

김호덕¹, 양현창¹, 박창현², 심귀보¹

¹서울시 동작구 흑석동 221, 중앙대학교 전자전기공학부
E-mail: kbsim@cau.ac.kr

²대전광역시 유성구 가정동 161, 한국전자통신연구원 전파방송연구단 전파기술연구그룹
E-mail: 3rr0r@etri.re.kr

요 약

얼굴 표정인식에 관한 연구에서 인식 대상은 대부분 얼굴의 정면 정지 화상을 가지고 연구를 한다. 얼굴 표정인식에 큰 영향을 미치는 대표적인 부위는 눈과 입이다. 그래서 표정 인식 연구자들은 얼굴 표정인식 연구에 있어서 눈, 눈썹, 입을 중심으로 표정 인식이나 표현 연구를 해왔다. 그러나 일상생활에서 카메라 앞에서 서는 대부분의 사람들은 눈동자의 빠른 변화의 인지가 어렵고, 많은 사람들이 안경을 쓰고 있다. 그래서 본 연구에서는 눈이 가려진 경우의 표정 인식을 Principal Component Analysis (PCA)를 이용하여 시도하였다.

Key Words : 표정인식, 템플릿 정합기법, PCA

1. 서 론

얼굴인식에 관한 연구들은 예전부터 많은 연구자들에 의해서 이루어지고 있다. 얼굴 인식 뿐만 아니라 얼굴 표정인식에 관해서도 여러 가지 방법을 이용한 연구들이 진행되고 있다 [1-3]. 최근에는 얼굴 표정 인식에서 PCA를 이용한 연구도 진행되고 있는데, PCA는 1990년 처음으로 얼굴 인식을 하는 인식 알고리즘으로 사용되었다. 특히, PCA의 특징 중에 차원을 감소시키는 장점 때문에 많은 얼굴 인식 연구자들이 많이 사용했고 지금도 사용되고 있다 [4-5]. 얼굴 표정인식에 관한 연구내용을 간단히 살펴보면 다음과 같다. 주영훈 등[1]은 개인에 대한 식별과 감정인식은 사용된 특징벡터들의 추출로 인한 Eigenface의 가중치와 상관관계를 통해서 인식을 했으며, Maja Pantic 와 Leon Rothkrantz [3]는 얼굴의 특징점들의 변화에 따라 그 특징점들의 배열 모양에 따라 얼굴의 감정인식 모델을 하였다. 이와 같이 대부분의 연구자들은 표정 인식에서 두 가지 방법을 사용하는데, 그것들은 기준점들의 지리적 위치를 사용하는 방법과 기준점에서의 Gabor filter coefficient를 사용하는 방법이다. 본 논문에서는 기존의 얼굴전체 영역을 대상으로 감성을 인식하지 않고 얼굴에서 눈을 가리고 입의 표정만으로 PCA를 이용한 감정인식을 하

였다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 표정인식에 관해서 다룬다. 3장에서는 눈이 가려진 경우 사람의 인식 실험 및 결과를 보여주고, 4장에서는 눈이 가려진 경우의 PCA를 이용한 표정 인식 실험 및 결과를 보여준다. 그리고 5장에서 결론으로 마무리 짓는다.

2. 표정인식

2.1 표정인식 소개

의사소통에는 음성, 얼굴 표정, 몸짓 등을 이용한 방법들을 주로 사용한다. 표정인식 연구는 크게 두 가지로 나눌 수 있는데, 첫째는 정지 얼굴 영상에 대한 인식으로써 템플릿 매칭 (Template Matching), 특징(Feature), 3D 얼굴 모델을 이용한 방법들이 있다. 동영상에 비해 처리량은 적지만, 얼굴의 비강체적 변형이 큰 경우 매칭이나 정합에 의한 인식이 힘들고, 얼굴 표정영상을 자동으로 지정하기 어렵다는 단점이 있다. 둘째는 동영상에서 사용자의 얼굴 변화를 분석하여 표정을 인식하는 연구로, 표정의 변화를 연속적으로 처리할 수 있어 실시간 응용제품에 유용하게 사용될 수 있다. 주로 사용되는 알고리즘으로는 신경망을 이용한 학습에 의한 인식, 얼굴 전체의 움직임 패턴에

주목하여 생성시킨 모션필드의 데이터베이스와 이용자의 모션필드에 부합되는 3차원 모델과의 상관관계에 의해 인식하는 방법 등이 있다.

2.2 템플리트 정합기법

기하학적 특징에 기반 한 순수 얼굴영역 검출기법에 의해 회전과 이동 변화를 보정한 얼굴 영역을 산출 할 수 있으므로 간단한 템플리트 정합으로도 높은 정확도를 보장할 수 있는 얼굴 인식 기법을 구현할 수 있다. 얼굴 표정의 다양함을 고려하여 템플리트 영상의 자동 생성 기법을 다음과 같이 제안한다. k번째 클래스에 속한 학습영상의 집합을 $I^k = \{I_1^k, I_2^k, \dots, I_n^k\}$ 로 표기한다.

이 기법은 다음의 과정으로 진행된다.

Step 1: $m=1$ (m : 템플리트의 개수)

Step 2: 모든 화소의 좌표(x, y)에 대해 화소 값들 $\{I_1^k(x,y), I_2^k(x,y), \dots, I_n^k(x,y)\}$ 의 median 값을 산출하여 m번째 템플리트 영상 T_m 을 생성

Step 3: T_m 과 각 영상 I^k 의 거리를 다음과 같이 계산한다.

$$\delta_i = \frac{1}{N_\Delta} \min_{\Delta} \sum_{(x,y)} |T_m(x,y) - I_i(x,y + \Delta)|$$

여기서 N_Δ 는 T_m 과 I_i 영상에서의 중첩된 영역에 속한 화소수를 의미한다.

Step 4: 템플리트 영상의 생성에 참여한 모든 영상의 δ 값이 Q보다 작다면 템플리트 생성 과정을 마친다.

Step 5: 일부 영상의 δ 값이 Q보다 작다면, 영상의 다음 단계의 템플리트 생성과정에서 m을 증가 시키고 Step 2의 과정을 되풀이한다.

Step 6: δ 값이 Q보다 작은 영상이 존재하지 않을 경우, δ 값이 가장 큰 영상으로 T_m 을 대체한다. m을 증가시키고 Step 2의 과정을 되풀이 한다. 얼굴영상의 표정 변화에 따른 템플리트의 자동생성을 위해 Step 2에서는 median을 사용하였다. 식에서는 수직 좌표축의 변화의 가능성을 고려하여 δ_i 의 계산과정에서 수직 축으로 일정 구간만큼 이동시키면서 최소의 δ_i 값을 계산한다. 또한 Q값이 작을수록 템플리트의 개수가 증가한다. 적절한 Q값은 전체 얼굴영상의 표정 변화와 유사도에 의해 결정되므로 이론적인 최적치를 설정하기 매우 어렵다.

3. 눈이 가려진 경우 사람의 인식 실험 및 결과

3.1 실험 개요

눈은 시각 도구이기도 하지만 몸의 어떤 부

분보다도 인간의 감정을 잘 나타내는 중요한 감정도구이다. 이처럼 눈은 감정을 인식하는데 있어서 매우 중요한 역할을 수행하지만 현대의 비전 시스템은 사람이 눈을 인식하는 만큼 눈의 변화를 잘 검출하지 못한다. 눈은 끊임없이 움직이며 변화하지만 변화의 규모가 미세한 반면, 입은 가만히 있는 경우도 있지만, 일단 움직이면 셀 수도 없이 많은 방식과 모양으로 움직인다는 면에서 눈보다 더욱 다양한 감정을 전달 할 수 있다. 즉, 입만으로도 일정 수준까지 감정을 인식 할 수 있다고 가정할 수 있다. 그래서 본 논문에서는 입의 모양만으로 표정을 인식하는 실험을 하기 전에 눈이 가려진 사진을 보고 피험자들이 각 사진들이 나타내는 표정을 어느 정도까지 인식할 수 있는지를 설문조사를 통해 통계를 내보았다. 본 실험을 통해 입의 모양만으로 하는 표정인식의 한계와 목표치를 설정 할 수 있다.

3.2 실험 방법

본 실험을 위해 6명의 남성(26~37세의 다양한 지역 출신의 대학원생)으로부터 5가지 표정(무표정, 화, 웃음, 슬픔, 놀람)을 연기하도록 하고 눈을 제외한 부분만을 사진으로 찍었다. 이렇게 수집 된 30장의 사진을 20명의 사람들에게 보여주고 사진 속의 사람이 어떤 감정을 갖고 있는지를 물어 보았다.



그림 1. 감정별 대표 표정

그림 1은 30개의 사진 중 감정별 대표 사진을 보여준다. 설문 대상자들에게 보인 문항은 사진의 사람들이 의도한 감정들 5가지와 애매하게 느껴질 수 있는 경우 4가지로 이루어져 있다

3.2 실험 결과

앞에서 설정한 방법대로 30개의 사진으로 이루어진 설문지를 설문 대상자 20명에게 보여주고 답하는 방법에 대한 간단한 설명을 약 3분간 해준 뒤에 문항 당 10초씩 답을 하도록 하였다. 이는 너무 오랜 시간동안 표정을 보는

것보다 짧은 시간만 보는 것이 감정 인식에서 더욱 의미 있기 때문이다. 이와 같은 실험의 결과로 각 감정에 대한 사람들의 인식률을 살펴보면 다음의 그래프와 같다.

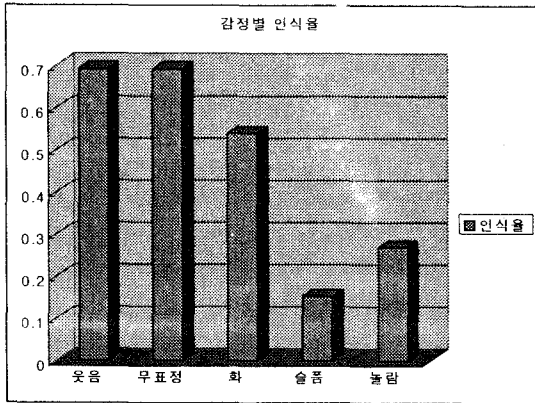


그림 2. 감정별 사람들의 인식률

그림 2 에서 보이는 바와 같이, 웃음과 무표정에 대한 사람들의 인식률은 거의 70%정도로 정확한 반면, 슬픔과 놀람은 30%도 안 되는 매우 저조한 인식률을 보인다. 이는 슬픔과 놀람의 경우 눈의 모양을 통한 인식 의존도가 매우 높음을 역설하는 것이다. 화는 표정의 경우는 약53%의 인식률을 보였는데 이 또한 높은 인식률이라고 할 수 없다.

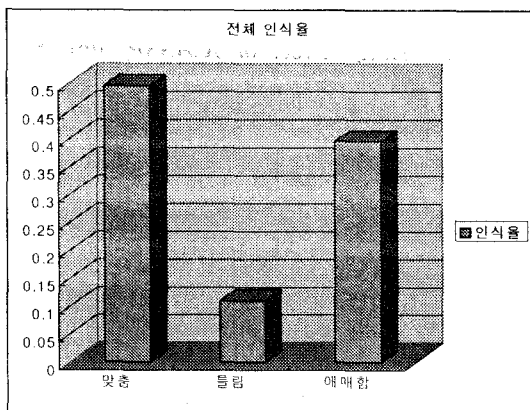


그림 3. 표정에 대한 인식률

그림 3의 그래프는 전체 사진들에 대해서 질문에 대한 답이 옳은 개수와 특정 감정으로 답했으나 의도한 감정과 다른 개수, 그리고 애매하다고 답한 개수를 나타내었다. 그 결과로 전체의 인식률은 약49%이고 오답의 확률은 약 10%, 약39%는 애매함으로 답을 하였다. 애매함에 속한 사진들은 화, 슬픔, 놀람의 감정을 의도한 것들로서, 앞의 결과와 마찬가지로 3가지 표정들의 경우 눈이 가려졌을 때 인식에서의 어려움이 있다는 것을 보여준다.

4. 눈이 가려진 경우의 PCA를 이용한 표정 인식 실험

3.1 실험 개요

표정 인식은 얼굴 전체의 영역을 대상으로 하는 것이 일반적이다. 대부분의 연구자들이 얼굴 전체에서 주요 부위에 대한 연구를 진행하는 반면, 일부에서는 얼굴의 일부분이 가려졌을 경우에 대한 연구를 진행해 왔다[7]. 본 논문에서는 눈을 가린 사진에 대해 표정 인식을 시도하였다. 또한, 분류 알고리즘으로는 Principal Component Analysis를 사용하였다.

4.2 실험 환경 및 결과

본 실험을 위해 사용된 사진의 사이즈는 768 × 1024 이고 학습용 사진으로는 5가지 감정을 위해 5장의 사진을 사용하였다. 768 × 1024의 행렬을 786432 × 1의 행벡터로 인식후보 얼굴 벡터 집합을 구성한다. 그 후 발생하는 에러를 줄이기 위해 미리 설정된 평균과 분산을 기준으로 모든 이미지를 정규화 하였다

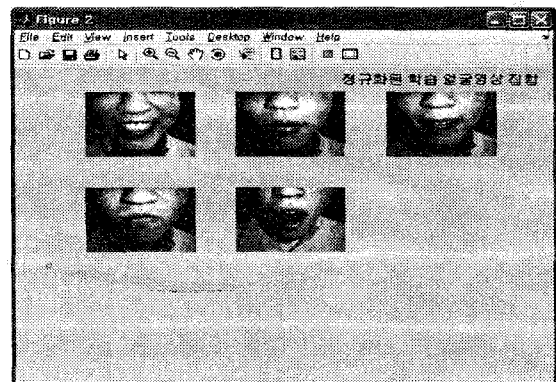


그림 4. 정규화된 학습 얼굴영상 집합

그리고 인식후보 얼굴 벡터 집합으로부터 평균 얼굴 벡터를 계산한다. 각 인식후보 얼굴 벡터와 평균 얼굴 벡터의 차를 계산 한 뒤 그 차 벡터의 공분산 행렬을 계산하고 그로부터 고유값을 계산한다. 이때의 고유값은 평균 얼굴 영상에 대한 분산의 정도를 나타내는 것이며, 이를 고유 얼굴이라고 한다. 학습단계를 마치고, 새로운 얼굴에 대해 인식을 수행하게 되는데, 새로운 얼굴이 입력되면 고유 얼굴에 대한 사영을 취하여 성분값을 구한다. 이 성분값은 퓨리에 급수에서의 사인과 코사인 같은 성분의 가중합으로 주기신호가 표현 가능하고 퓨리에 계수를 구하는 방식과 같다. 이 값이 구해지면, 후보 얼굴영상들의 고유얼굴상에서의 가중치와 유클리디안 거리를 비교하여 그 거리가 최소가 되는 표정이 입력과 가장 유사

한 표정이므로 이 후보를 인식 결과로 결정하는 것이다.

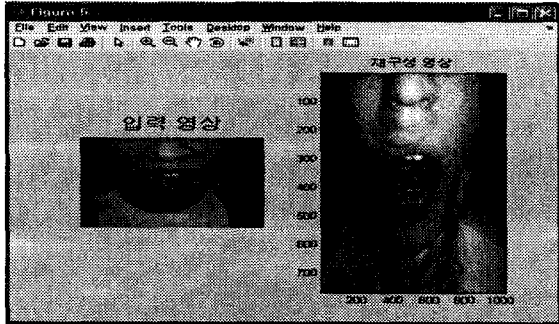


그림 6. 테스트용으로 입력된 영상

입력된 영상은 놀라는 표정인데, 학습된 영상들과 비교하여 유클리디안 거리가 제일 작은 것을 찾아보았더니 5번 영상임을 나타낸다. 학습시 사용된 5번 영상 역시 놀람을 나타내고 있으므로 이 경우는 올바른 인식을 한 것이다.

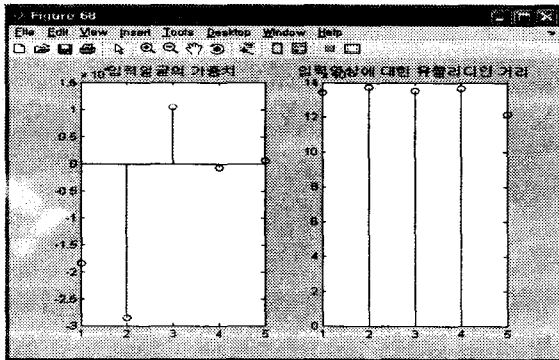


그림 7. 결과 화면

이와 같은 방식으로 학습용 영상을 바꿔가면서 5번 실험을 하였는데, 그 결과는 다음의 그래프와 같다.

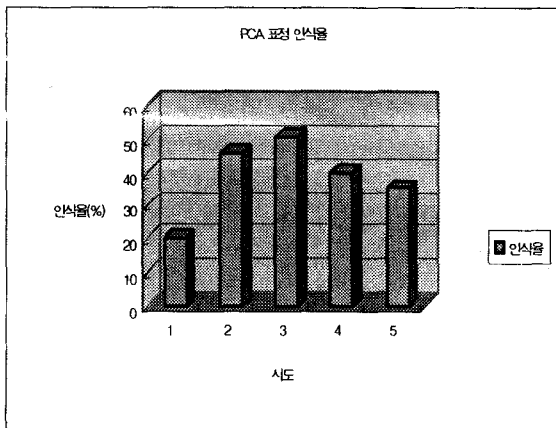


그림 8. 실험 결과 그래프

5. 결 과

본 논문은 눈이 가려진 사람의 얼굴 표정 사진을 사람이 인식한 것과 PCA를 이용하여 얼굴표정에서의 감성인식을 비교하였다. PCA를 이용해서 얼굴표정에서의 감성인식이 최고로 높을 때의 인식률은 51%로써 절대적인 수치로는 그다지 좋은 인식률은 아니지만, 사람의 인식률도 49%밖에 안 되었던 것을 감안한다면 그리 나쁜 성능을 보이는 것이 아님을 알 수 있다. 앞으로 입모양뿐만 아니라 코 옆과 입근처의 근육의 미세한 움직임도 인식해서 인식률을 높이고 감성 인식에서의 학습에 있어서도 다른 많은 유전 알고리즘을 이용해서 인식률 높이는 연구가 진행 될 것이다.

감사의 글 : 본 논문은 서울시 산학연 협력사업(과제번호 : 106876)에 의해 수행되었습니다. 연구비 지원에 감사드립니다.

참 고 문 헌

- [1] 주영훈, 이상윤, 심귀보, "Eigenface를 이용한 인간의 감성인식 시스템," *한국퍼지 및 지능시스템학회 논문지*, 제13권, 제2호, pp. 216-221, 2003.
- [2] I. Badiu, I. Kotsia, and I. Pitas, "Facial expression analysis under partial conclusion," *Proc. of ICASSP*, vol. 5, pp. 453-356, March 2005.
- [3] Maja Pantic and Leon Rothkrantz "Case-Based Reasoning for User Profiled Recognition of Emotions from Face Images." *IEEE International Conference on Multimedia and Expo*, 2004.
- [4] Vo Dinh Minh Nhat and Sung Young Lee, "Two-dimensional Weighted PCA algorithm for Face Recognition," *Proc. of 2005 IEEE International Symposium on Computational Intelligence in Robotics and Automation*, 2005.
- [5] M. Anouar Mellakh, Dijana Petrovska-Delacretaz, Bernadette Dorizzi "Using Signal/Residual Information of Eigenfaces for PCA Face Space Dimensionality Characteristics," *The 18th International Conference on Pattern Recognition*, 2006.