

얼굴 표정 인식 기술

얼굴 표정 인식은 인간 중심의 human-machine 인터페이스의 가장 중요한 요소 중 하나이다. 현재의 얼굴 표정 인식 기술은 주로 얼굴 영상을 이용하여 특징을 추출하고 이를 미리 학습시킨 인식 모델을 통하여 각 감정의 범주로 분류한다. 본 논문에서는 이러한 얼굴 표정 인식 기술에 사용되는 표정 특징 추출 기법과 표정 분류 기법을 설명하고, 각 기법에서 많이 사용되고 있는 방법들을 간략히 정리한다. 또한 각 기법의 특징들을 나열하였다. 또한 실제적 응용을 위해서 고려해야 할 사항들에 대하여 제시하였다. 얼굴 표정 인식 기술은 인간 중심의 human-machine 인터페이스를 제공할 뿐만 아니라 로봇 분야에서도 활용 가능할 것으로 전망한다.

■ 허경무, 강수민*
(단국대학교 전자공학과)

I. 서론

얼굴 표정(Face Expression)은 사람의 감정 뿐 만 아니라 마음의 상태, 사회적 상호작용, 생리학적인 신호 등과 같은 다양한 정보를 반영한다. 최근 컴퓨터 비전, 컴퓨터 그래픽스 그리고 HCI 분야에서 실시간 얼굴 표정 인식 및 분석 기술이 주요 관심사로 부각되고 있다. 특히 비전 기반 실시간 얼굴 모션 추적 및 표정 인식 기술은 효과적인 인간과 컴퓨터의 상화 작용을 위한 매력적인 입력 모드로 관심을 모으고 있다. 이처럼 얼굴 표정 인식에 관련된 연구의 중요성이 증대되고 있는 이유는 컴퓨터 성능의 향상에 따라 저렴한 비용으로 빠른 처리가 가능할 뿐만 아니라 얼굴 검출, 얼굴 추적, 얼굴 인식 등과 같은 영역에서의 연구와 밀접한 연관성이 존재하여 상호보완적으로 연구 수행을 향상시키기 때문이다. 얼굴 표정의 동적 변화를 실시간으로 분석하기 위해서는 얼굴 동작의 시간적 변이를 효과적으로 반영할 수 있는 최적의 표정 정보의 추출과 실시간 추적이 필요하며 특정 표정 사이의 변화를 능동적으로 설명할 수 있는 표정변화 모델에 기반한 얼굴 표정의 해석 방법이 필요하다. 그러나 실시간 분석을 위한 얼굴 표정 변화의 추적은 조명 및 배경의 변화에 따라 매우 어려운 작업으로 인식되고 있다. 본 논문에서는 이러한 문제점들을 해결하기 위해 현재 연구가 진행되고 있는 얼굴 표정 인식에 관한 여러 가지 방법들을 소개하고자 한다.

II. 얼굴 표정 인식 연구 동향

1. 표정 특징 추출 기법

표정 인식을 위한 특징 추출 방법은 크게 통계적 성분 분석 기반의 특징 추출 방법과 특징점 기반의 방법으로 나눌 수 있다. 우선, 성분 분석을 기반으로 특징을 추출하는 방법은 추출된 얼굴 영역의 전체 또는 일부분의 픽셀 정보를 이용하는 것으로 주로 고유 벡터(eigen vector)를 이용하여 특징을 추출한다. 반면에 특징 점 기반의 특징 추출 방법은 눈, 입과 같은 얼굴의 특정 부분에 해당하는 특징점의 위치 정보를 활용한 것으로 특징점 구성 방법에 따라 여러 가지 알고리즘이 존재한다.

1.1 성분 분석 기반의 특징 추출 기법

얼굴을 나타내는 픽셀 정보를 표정 인식에 그대로 사용하면 불필요한 정보가 많고 처리시간도 길어진다. 따라서 차원(dimension)을 축소하여 얼굴을 나타내는 대표적인 특징을 추출해 내는 과정인 PCA를 통해 고유 벡터 또는 고유 얼굴(eigen face)을 획득한다. 그러나 고유 벡터는 눈, 코, 입과 같은 세부적인 정보보다는 얼굴 전체에 대한 표현이기 때문에 표정에 의한 분류가 어렵다. 따라서 Buciu는 분류되지 않은 특징들 중에서 확률적으로 독립 성분을 충분히 포함하고 있는 새로운 특징을 추출하는 기법인 ICA(Independent Component Analysis)를 이용하여 특징을 추출하는 방법을 제시하였다. 그

림 1은 ICA를 이용하여 추출한 특징 벡터를 시각화 한 것으로 눈으로 특징을 구별하기 힘들지만 색상 등과 같이 불필요한 데이터는 사라지고 각 얼굴을 구별하는 주요 특징들만 추출된 모습이다.

ICA와는 달리 고차원 데이터들의 통계적인 특성과 기하학적 구조를 활용한 LLE(Locally Linear Embedding)를 사용한 방법도 제안되었다. LLE는 고차원에서의 인접 관계를 저차원에서도



그림 1. ICA를 이용하여 추출한 특징 벡터.

그대로 유지하면서 차원의 축소를 하는 방법으로 벡터의 이동, 회전 등의 변화 없이 특징 정보를 추출하는 방법이다. 이러한 통계적 성분 분석 기반의 특징 추출 기법은 학습 데이터의 의존도가 높아서 사람의 성별, 인종, 나이 등에 많은 영향을 받는 단점이 존재한다.^[12]

1.2 특징점 기반 특징 추출 기법

특징점 기반의 특징 추출 기법은 눈, 코, 입 등의 얼굴 구성체의 특정 위치를 특징점으로 정의하고 각 표정이 특징점의 조합으로 구성된다는 것을 이용한다. 표정 인식 연구의 초기에는 애니메이션에서 얼굴 구성을 위해 주로 사용되는 MPEG-4 face model을 사용하였다. MPEG-4 모델은 그림 2와 같이 84개의 특징점을 정의하였으며 이 특징점의 움직임에 관련된 FAP(Facial Animation Parameter)의 조합에 의해 표정을 나타내었다.^[1] MPEG-4 모델에 의해 정의된 특징점은 optical flow를 이용하여 추적되며 FAP의 활성화여부가 판단된다. 그러나 MPEG-4 모델은 특징점에 대한 정의만 있을 뿐 표정 인식을 위한 모델링 기

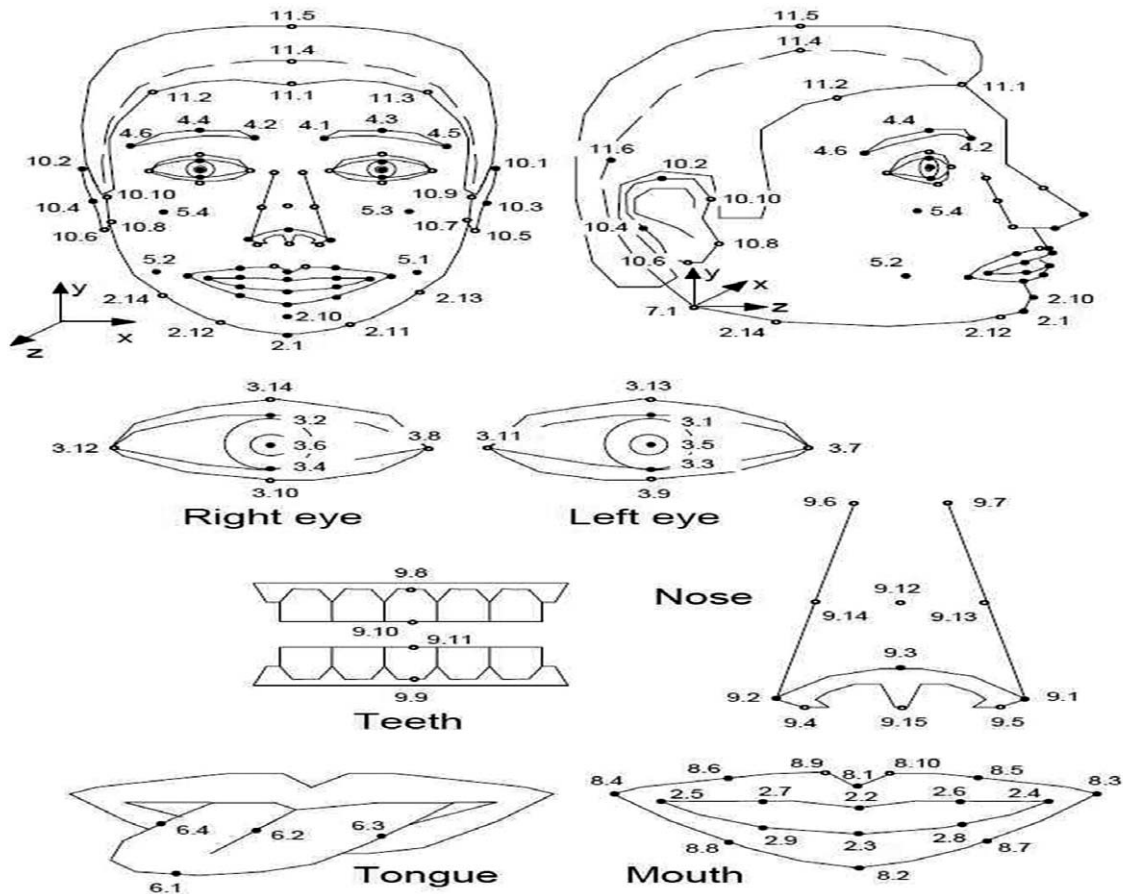


그림 2. MPEG-4 특징점.

법 등이 존재하지 않는다. 또 다른 특징점 기반의 특징 추출 방법으로 ASM(Active Shape Model)과 BSM(Bayesian Shape Model)이 있다. ASM은 통계학적 모델링 기법으로 미리 정의된 특징점의 위치를 벡터화하여 각 표정별 파라미터를 얻는 방법이다. 파라미터와 PCA로 추출한 고유 벡터와의 조합을 통해 다양한 표정을 생성할 수 있게 된다. BSM은 확률변수와 에너지 함수를 이용하여 특징점의 위치를 추정하는 방법으로 ASM 보다 나은 성능을 보여준다. ASM은 특징점의 위치 정보만 가지고 있기 때문에 표정이 변하거나 새로운 얼굴에 대한 적응력이 낮다. 이러한 단점을 보완한 것으로 특징 점의 위치뿐만 아니라 내부 픽셀 정보까지 함께 모델링을 한 AAM이 있다. AAM은 ASM의 파라미터와 특징점이 구성하는 내부 픽셀 에 대한 모델링을 결과를 결합하여 하나의 모델을 구성한 것으로써 모델의 적응력이 높아 표정 인식 분야에 활발히 연구되고 있다.^[45]

1.2.1 Feature invariant 접근 방법

이 기법은 표정 검출을 위해서 사람의 회전이나 크기 변화, 조명 변화에 무관한 불변 특징치를 추출한 후 이를 이용하여 표정 검출을 시도한다. 이 기법의 경우 facial features, texture, skin, multiple features 등을 이용하여 사전에 추출된 특징치들과 입력 영상에 존재하는 특징치들 간의 검색 및 분류 작업을 통해서 표정을 검출하게 된다. 하나의 예로 facial feature를 이용하는 경우 그림 3과 같은 특징치들을 추출 후 이 특징치들 간의 연결을 통해서 정상적인 표정 검출 여부를 판단하게 된다.^[6]

그림 4는 위와 같이 눈썹, 눈, 코, 입의 윤곽선 추출을 통해 얻

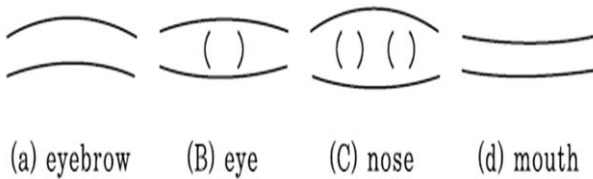


그림 3. 얼굴 특징들의 예.

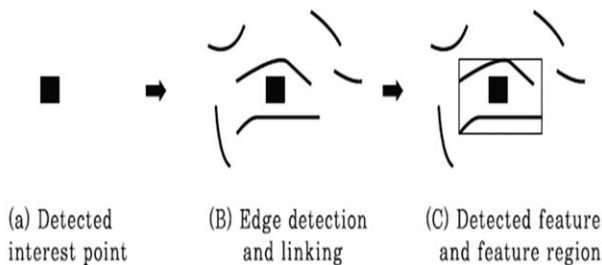


그림 4. 얼굴 특징들의 연결 및 검증을 통한 특징 추출.

어진 특징을 이용해서 추출된 각각의 특징들의 관계를 묘사하고 해당 특징들의 존재 여부를 검증함으로써 얼굴의 특징을 검출하는 예를 보여준다.

1.2.2 Template matching 방법

이 기법은 그림 5와 같이 표준 얼굴 패턴에 대한 정보를 사전에 추출한 후 입력 영상과 사전에 정해 놓은 template와의 상관 관계를 추출함으로써 표정 검출을 시도한다. 표준 얼굴 패턴은 일반적으로 얼굴 형태, 눈, 코, 입 등을 독립적으로 표준 패턴으로 등록하여 사용한다. 그리고 검출된 표준 얼굴 패턴 사이의 관계 요소로부터 얼굴 표정을 최종적으로 판단한다.^[6]

이 기법은 구현이 매우 쉬운 장점이 있으나 크기 변화, 회전 변화, 형태 변화 등에 민감하게 반응하여 실제 사람 검출에 적용하기에는 많은 문제점이 발견되었으며 이를 해결하기 위한 다해상도 기법, 변형 가능한 template를 이용한 연구들이 진행되고 있다.

2. 표정 분류 기법

추출된 표정 특징을 이용하여 여러 가지 표정으로 분류하는 방법은 다양하다. 단순히 변화되는 특징점의 조합을 통해 웃는 표정인지 화난 표정인지 판단할 수도 있으며, 머신 러닝(machine learning)이나 신경망(neural network)을 통하여 표정 특징을 분류할 수 있다. 기존의 방법들은 주로 HMM, SVM과 같은 기계 학습 방법을 사용하였으며 사용하는 특징 및 데이터베이스에 따라서 80~90%의 인식률을 보여주고 있다. 다음은 표정 분류 및 인식에 활용되는 주요 알고리즘을 설명한다.^[79]

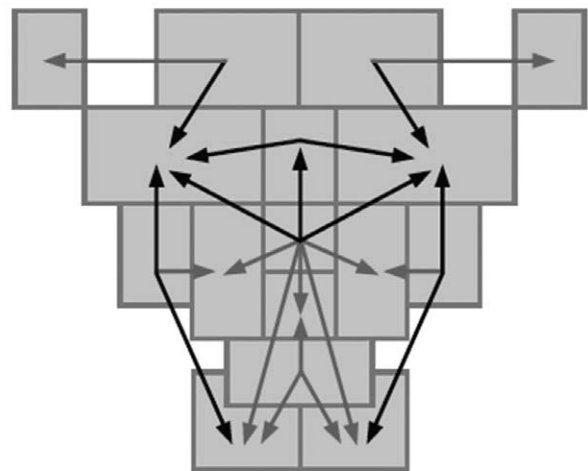


그림 5. 16개의 템플레이트와 23개의 각 템플레이트와의 관계 정보를 이용한 얼굴 표정.

2.1 Support vector machines

SVM은 입력공간에서 maximum-margin hyper plane을 이용하여 2개의 영역을 분류하는 지도학습 방법의 일종이다. 두 그룹으로 나눌 수 있는 벡터가 있을 때 각 벡터 그룹을 분류하는 선을 hyper plane 이라고 하며 이러한 hyper plane은 하나 이상 존재할 수 있다. 수많은 hyper plane 중에서 SVM은 각 그룹 사이를 가장 멀리 지나는 maximum-margin hyper plane을 이용하여 분류하는 방법을 제공한다. 벡터 그룹 중에서 가장 외곽에 위치한 벡터들을 convex hull이라고 하며 convex hull 내부에 위치하는 벡터들은 hyper plane의 결정에 크게 영향을 주지 않는다. 따라서 SVM은 convex hull에 위치한 벡터 중에서 두 그룹 사이의 거리가 가장 가까운 벡터들과 수직 거리가 가장 먼 hyper plane을 maximum-margin hyper plane으로 결정함으로써 두 벡터 그룹을 분류하게 되며, hyper plane과 가장 가까운 벡터들을 support vector라고 한다. 그림 6은 support vector와 maximum-margin hyper plane을 나타낸 것으로, 이 중 원으로 표시된 것이 support vector이며 maximum-margin hyper plane은 굵은 선으로 표시되어 있다.

SVM은 선형적인 hyper plane으로 분류할 수 있는 경우에 좋은 성능을 나타내며 비선형적인 경우라도 더 높은 차원에서 선형성을 나타낼 경우 차원을 확대하여 비선형 분류기로써 활용이 가능하다. 또한 그림 3과 같이 완벽하게 분류할 수 없는 경우에도 거리의 오차를 최소화 할 수 있는 maximum-margin hyper plane을 찾을 수 있기 때문에 패턴 분류에서 많이 사용되고 있다.^[7]

2.2 Hidden Markov 모델

HMM은 패턴 인식에 주로 사용되는 통계 학습 방법으로써 모

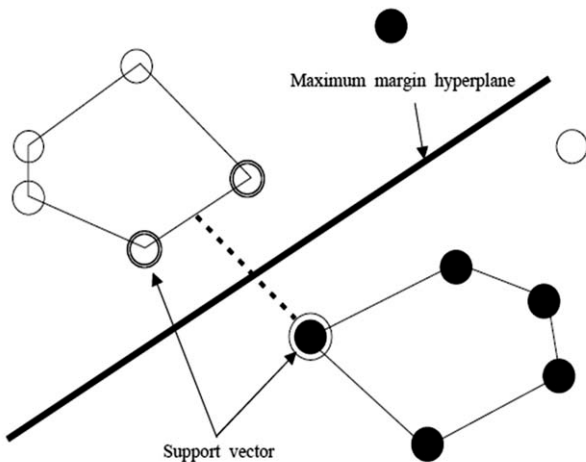


그림 6. SVM의 분류 방법.

델링하는 시스템이 임의의 파라미터를 가진 Markov process 라고 가정하고, 그 가정에 기초하여 관측된 파라미터로부터 숨겨진 파라미터를 추론하는 알고리즘이다. 일반적인 Markov 모델은 그 상태를 관찰자가 볼 수 있지만 HMM은 상태를 볼 수 없고 상태에 따라 출력되는 값만 알 수 있다. 따라서 HMM은 입력 값을 통해서 상태들의 순서에 대한 결과를 이용하여 추론하게 된다. 이 때 출력되는 값은 각 상태에 따라 특정 확률을 가지고 있으며, 상태가 전이할 때에도 확률을 가지고 전이가 된다. 그림 7은 HMM의 구조에 대한 예시이다. X는 상태를 나타내며 a는 각 상태를 전이할 때의 확률 값이고 b는 상태출력에 대한 확률 값이다. HMM은 출력 y를 이용하여 Bayesian network를 기반으로 상태 X를 추론하게 되며, 상태 전이에 대한 결과를 패턴인식에 사용한다.^[9]

2.3 역전파 알고리즘

역전파 알고리즘은 피드포워드(feedforward)구조를 가진 다층 신경회로망으로써 문자 인식 등의 패턴 인식분야에 많이

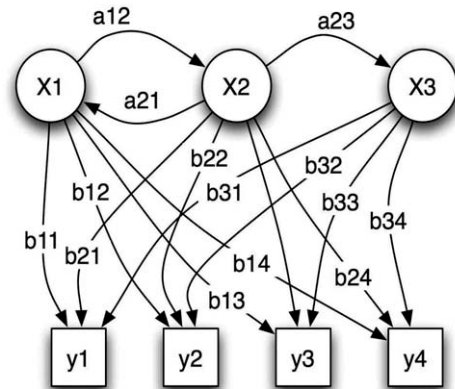


그림 7. HMM의 구조도 예시.

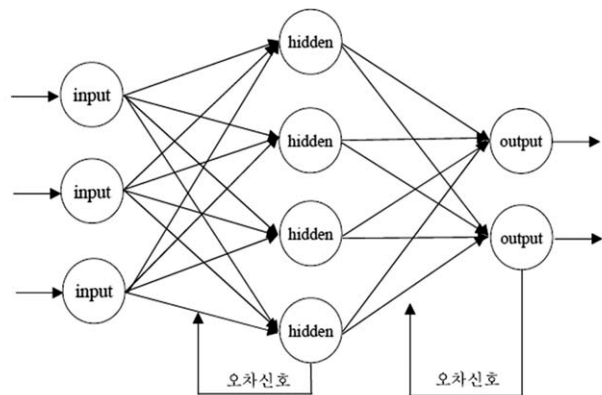


그림 8. 역전파 알고리즘의 구조도 예시.

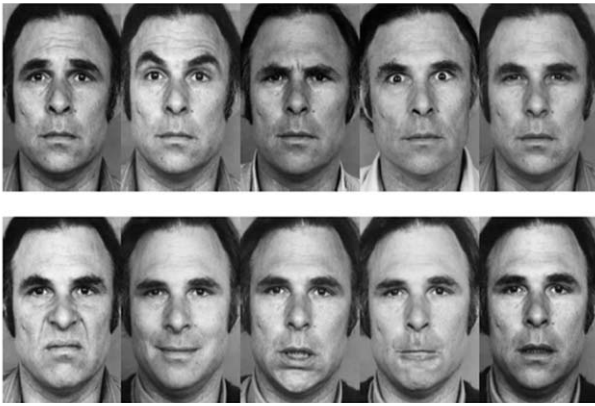


그림 9. FACS에 정의된 Action Unit의 표현.

표 1. Action Unit의 설명.

Action Units	동작 설명
AU1	안쪽 눈썹이 올라감
AU2	바깥쪽 눈썹이 올라감
AU4	눈썹이 내려감
AU5	윗눈꺼풀이 올라감
AU7	눈꺼풀이 얇아짐
AU10	윗입술이 올라감
AU12	입꼬리가 올라감
AU16	아랫입술이 내려감
AU23	입술이 가늘게 붙음
AU25	입이 벌어짐

사용되고 있다. 역전파 알고리즘은 그림 8과 같이 출력층의 오차신호를 이용하여 은닉층과 출력층 사이의 연결강도를 변경하고, 출력층의 오차신호를 은닉층에 역전파하여 입력층과 은닉층 사이의 연결강도를 변경하는 학습방법이다. 역전파 알고리즘은 *generalized delta rule*을 이용하여 학습을 하게 되며 인식률이 뛰어나고 잡음이 포함된 입력에도 좋은 성능을 보여준다. 그러나 초기 연결강도 선택에 따른 *local minima*에 빠지거나 학습률의 선택 등 학습단계에 고려할 사항이 많은 단점도 존재한다.

2.4 Facial Action 코딩 시스템(FACS)

FACS는 1976년 심리학자 Ekman에 의해서 제안된 것으로써 사람의 표정을 해부학적으로 분석하기 위해 개발되었으며, 표정은 얼굴 근육 움직임의 조합임을 과학적 연구를 통해 증명하였다. Ekman은 슬픔, 괴로움, 분노, 놀라움, 두려움, 역겨움 등의 8가지 감정을 분석하였으며, 얼굴 표정을 통해 이러한 감정을 읽는 방법을 연구한 결과가 FACS이다. FACS는 각 얼굴 근육의 움직임을 나타내기 위해 46개의 AU(Action Unit)을 정의하였으며 표정은 이 AU의 조합으로 나타내었다. 그림 9와 표 1은 FACS에 정의된 몇 가지 AU의 표현과 설명을 나타낸다.^[10]

만약 웃는 표정이라면 주로 눈꺼풀이 얇아지고 입꼬리가 올

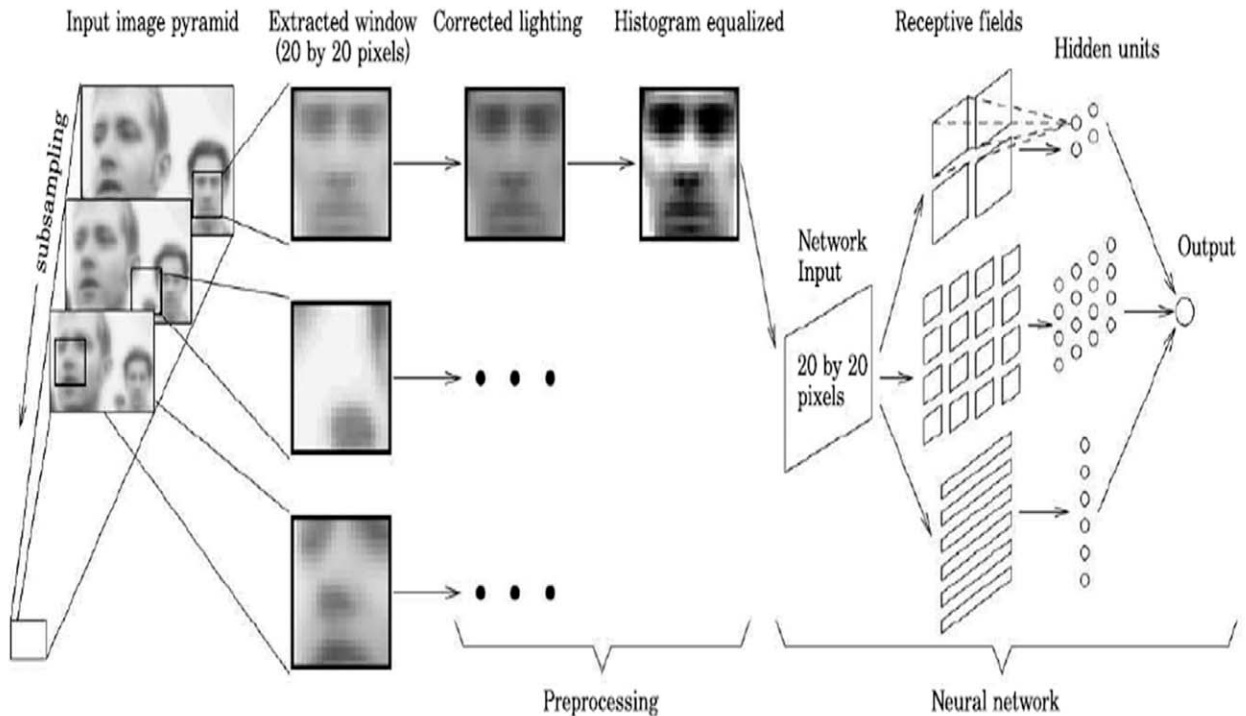


그림 10. Eigenface와 neural network를 이용한 얼굴 표정 인식의 예.

라가며 입 이 벌어지거나 턱이 내려갈 것이다. 따라서 AU7 + AU12 + AU25로 간단히 표현할 수 있다. 이와 같이 FACS는 다양한 표정에 대하여 AU의 조합을 통해 표현을 하였으며 하나의 표정에 대해서도 다양한 조합을 정의해 놓았다. 그러나 FACS가 표정에 대하여 세밀하게 정의를 해놓았다라도 표정 인식의 절대 적인 기준이나 방법은 될 수 없다. 각 AU의 활성화에 대한 기준이 명확하지 않으며 불완전한 조합일 경우에 판단이 모호해지는 경우가 생기기 때문이다.

2.5 신경망(Neural Networks)

신경망은 문자 인식, 물체 인식 등에서 성공적으로 적용되어 많은 연구 결과를 도출하였으며 이러한 기법들을 이용하여 얼굴 표정 검출에도 적용되고 있다. 이 기법의 장점은 매우 많고 다양한 얼굴 표정 패턴에 대한 학습 체계가 사전 연구로 잘 갖추어져 있는 점이고 단점은 좋은 성능을 얻기 위해서는 매우 많은 layer 수, node 수와 iteration 수를 요구한다는 점이다.^[11]

III. 결론

현재 다양한 얼굴 표정 인식을 위한 방법들이 존재하고 있다. 이러한 방법들은 독자적으로 쓰이거나 여러가지 방법들이 혼용되어 사용되기도 한다. 이러한 방법을 통한 현재 얼굴 표정 인식률을 약 95%이다. 하지만 인식률을 높이기 위한 연구는 현재도 진행되고 있으며, 얼굴 표정 검출 알고리즘의 성능이 향상됨으로써 실시간 처리가 필수적인 고급 응용 분야에까지 널리 활용되고 있는 것으로 판단된다. 특히 표정 인식 기술은 시시각각으로 변화하는 표정의 실시간 인식을 위해서는 고속의 얼굴 검출이 필수적이라고 판단된다. 또한, 하드웨어 구조의 설계를 통해 실시간 처리 가능하고 이렇게 설계된 하드웨어 구조가 주문형 반도체로 제작되어 저가로 공급되면 최첨단 디지털 카메라, 인공지능 로봇 등과 같은 다양한 시스템에서 부담 없이 활용할 수 있으므로 많은 응용 분야에 적용될 것으로 기대된다.

참고문헌

- [1] I. Buciu, C. Kotropoulos, I. Pitas, "Comparison of ICA approaches for facial expression recognition," *International Journal of Signal, Image and Video Processing*, Vol.3, No.4, pp.345-361, 2009.
- [2] Y. Shin, "Facial expression recognition based on emotion dimensions on manifold learning," *Proceedings of the 7th International Conference on Computational Science (ICCS 2007)*, LNCS 4488, pp.81-88, 2007.
- [3] A.M. Tekalp, J. Ostermann, "Face and 2-D mesh animation in MPEG-4," *International Journal of Signal Processing: Image Communication*, Vol.15, Issue 4-5, pp.387-421, Jan. 2000.
- [4] J. Bu, M. Song, Q. Wu, C. Chen, C. Jin, "Sketch based facial expression recognition using graphics hardware," *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction (ACII 2005)*, LNCS 3784, pp.72-79, 2005.
- [5] J.H. Kwon, G.Y. Kim, Y. Mun, "Efficient facial features warping using BSM(Bayesian Shape Model)," *Proceedings of the International Conference on Computational Science and Its Applications (ICCSA 2008)*, LNCS 5073, pp.56-66, 2008.
- [6] Yea-Shuan Huang and Wei-Cheng Lie, "Face detector with oriented multiple templates", *International MultiConference of Engineers and Computer Scientists*, Hong Kong, March, 2008.
- [7] S. Park, D. Kim, "Subtle facial expression recognition using motion magnification," *Pattern Recognition Letters*, Vol.30, Issue 7, pp.708-716, May 2009.
- [8] E. Cerezo, I. Hupont, "Emotional facial expression classification for multimodal user interfaces," *Proceedings of the 4th International Conference on Articulated Motion and Deformable Objects (AMDO 2006)*, LNCS 4069, pp.405-413, 2006.
- [9] Z. Hammal, L. Couvreur, A. Caplier, M. Rombaut, "Facial expression recognition based on the belief theory: Comparison with different classifiers," *Proceedings of the 13th International Conference on Image Analysis and Processing (ICIAP 2005)*, LNCS 3617, pp.743-752, 2005.
- [10] Michel F. Valstar, Bihan Jiang, Marc Mehu, Maja Pantic, and Klaus Scherer, "The first facial expression recognition and analysis challenge", *Automatic Face & Gesture Recognition and Workshops (FG 2011)*, pp 921 - 926, March 2011
- [11] 주영훈, 이상윤, 심귀보, "Eigenface를 이용한 인간의 감정인식 시스템", *퍼지 및 지능시스템학회 논문지* Vol. 13, No. 2, pp 216-221, 3, 2003

◎ 저 자 약 력



허경무

- 1979년 서울대학교 전자공학과 학사.
- 1981년 한국과학기술원 전기 및 전자공학과 석사.
- 1989년 한국과학기술원 전기 및 전자공학과 박사.
- 1993년~현재 단국대학교 전자공학과 교수.
- 2005년 2월~2011년 6월 단국대 전자부품 검사자동화 지역혁신센터(RIC) 소장.
- 2011년 1월~2012년 12월 대한전자공학회 시스템및제어 소사이 어티 회장.
- 관심분야 : 시스템제어, 머신비전, 검사자동화, 로봇제어, 학습제어.



강수민

- 2004년 단국대학교 전자공학과 학사.
- 2006년 단국대학교 전자공학과 석사.
- 2007년~현재 단국대학교 전자컴퓨터공학과 박사 수료.
- 관심분야 : 로봇트 제어, 검사 자동화, Machine Vision, 시스템 제어