

표정 HMM과 사후 확률을 이용한 얼굴 표정 인식 프레임워크

(A Recognition Framework for Facial Expression by Expression HMM and Posterior Probability)

김진옥^{*}

(Jin Ok Kim)

요약 본 연구에서는 학습한 표정 패턴을 기반으로 비디오에서 사람의 얼굴을 검출하고 표정을 분석하여 분류하는 프레임워크를 제안한다. 제안 프레임워크는 얼굴 표정을 인식하는데 있어 공간적 정보 외 시간에 따라 변하는 표정의 패턴을 표현하기 위해 표정 특성을 공간적으로 분석한 PCA와 시공간적으로 분석한 Hidden Markov Model(HMM) 기반의 표정 HMM을 이용한다. 표정의 공간적 특징 추출은 시간적 분석 과정과 밀접하게 연관되어 있기 때문에 다양하게 변화하는 표정을 검출하여 추적하고 분류하는데 HMM의 시공간적 접근 방식을 적용하면 효과적이기 때문이다. 제안 인식 프레임워크는 현재의 시각적 관측치와 이전 시각적 결과간의 사후 확률 방법에 의해 완성된다. 결과적으로 제안 프레임워크는 대표적인 6개 표정뿐만 아니라 표정의 정도가 약한 프레임에 대해서도 정확하고 강건한 표정 인식 결과를 보인다. 제안 프레임 워크를 이용하면 표정 인식, HCI, 키프레임 추출과 같은 응용 분야 구현에 효과적이다.

키워드 : 표정 인식, 표정 분류, HMM, 얼굴 검출

Abstract I propose a framework for detecting, recognizing and classifying facial features based on learned expression patterns. The framework recognizes facial expressions by using PCA and expression HMM(EHMM) which is Hidden Markov Model (HMM) approach to represent the spatial information and the temporal dynamics of the time varying visual expression patterns. Because the low level spatial feature extraction is fused with the temporal analysis, a unified spatio-temporal approach of HMM to common detection, tracking and classification problems is effective. The proposed recognition framework is accomplished by applying posterior probability between current visual observations and previous visual evidences. Consequently, the framework shows accurate and robust results of recognition on as well simple expressions as basic 6 facial feature patterns. The method allows us to perform a set of important tasks such as facial-expression recognition, HCI and key-frame extraction.

Key words : facial expression recognition, facial expression-classification, HMM, face detection

1. 서론

언어는 물론 몸짓, 표정과 같은 비언어적인 방법 역시 사람들 간의 의사소통에서 중요한 역할을 담당하고 있는데 특히 얼굴의 표정이 그 중에서도 가장 대표적이다. 이 때문에 컴퓨터 비전, 감성 컴퓨팅, 컴퓨터 그래픽스, HCI, 심리학 등의 학문 분야에서 실시간으로 얼굴 표정을 인식하고 분석하는 시스템을 개발하여 이를 여러 용

용 분야에 적용하려는 노력이 계속되고 있다.

얼굴 표정의 시간적 변화는 의사소통에 따른 주요 정보의 추론에 도움이 되기도 하고 감정 상태를 이해할 수 있게 하지만 시간에 따라 아주 복잡하게 변화하는 행동 패턴을 보인다. 따라서 자동화된 얼굴 표정 인식 시스템은 기계가 사람과 자연스럽게 상호작용을 할 수 있도록 공간상에서 사람의 얼굴 동작을 인식하고 얼굴 동작의 시간적 변화 정도를 감지하여 인간 표정이 변화되어 가는 여러 가지 상태를 능동적으로 해석할 수 있도록 구현되어야 한다. 본 연구는 얼굴 표정 인식 시스템이 시간에 따라 다양하게 변하는 얼굴 표정을 분석하여 자동적으로 사람의 표정을 인식하는 프레임워크를

^{*} 이 논문은 2004년도 대구한의대학교 기린연구비 지원에 의한 것이다

[†] 정 회 원 : 대구한의대학교 멀티미디어학부 교수

bit@dhu.ac.kr

논문접수 : 2004년 9월 24일

심사완료 : 2005년 3월 15일

제안하고자 한다.

1.1 관련 연구

사람의 얼굴은 감정뿐만 아니라 현재 상태에 대한 중요한 정보를 담고 있기 때문에 감정 상태를 인식하기 위해 기쁨, 슬픔, 화남, 혐오, 두려움, 놀람 등 전 문화권에서 동일하게 인지되는 6가지 대표 얼굴 표정을 중심으로 표정을 인식하는 연구가 진행되어 왔다.

얼굴 표정 분석에 대한 연구 방법은 대표적으로 공간적 분석과 시공간적 분석 두가지로 나뉘는데 공간적 분석의 첫번째 방법은 신경망을 이용한 방법으로 신경망 네트워크를 얼굴 표정 인식에 직접 적용하거나 주요성분 분석, 독립성분 분석, 가버 웨이블릿 필터와 연계하여 적용된다. Rosenblum[1]은 복잡한 얼굴 표정 분석을 감정표현, 얼굴 특징의 움직임, 움직임 방향 등 3가지 층으로 나눈 네트워크 시스템을 이용했다. Fasel[2]은 크기 변환과 무관한 다중 크기의 단순한 특징 추출기 층을 이용한 컨볼루션 신경망 기반의 시스템을 개발했다. Dailey[3]가 제안한 신경망 방법은 첫번째 단계에서 2D 가버 필터를 겹친 그리드를 적용하여 이미지를 필터링하고 두번째 단계에서는 PCA를 적용하여 차원을 줄인 다음 이 데이터를 6 가지 대표 감정에 대해 각각 6개의 출력을 처리하는 신경망의 입력 값으로 처리한다. 이외 가버 웨이블릿을 이용한 Lyons et al. [4]과 규칙 기반의 신경망 네트워크를 이용한 얼굴 표정분석방법[5]도 제안되었으며 Schulze[6]는 얼굴 움직임 분류를 처리하기 위해 SVM(Support Vector Machine)을 이용했다.

시공간적인 방법으로는 Essa & Pentland[7]가 얼굴 표정의 시공간적 움직임 에너지를 추출하는 방법을 제안했다. Otsuka & Ohya[8]는 얼굴 표정을 분류하는데 HMM(Hidden Markov Model)을 이용, 6개 기본 표정의 모델링에 15 차원의 특징 벡터를 적용하여 시간상의 시퀀스를 정합시켰다. Lien, Kanade, Cohn와 Li[9]는 얼굴 표정을 움직임 단위로 분류한 Ekman과 Friesen의 FACS(Facial Action Coding System)[10]에서 제안한 표정 특징을 HMM에 적용한 연구를 진행했다.

의사 2D HMM을 확장한 3D HMM은 Muller, Eickerler, Rigoll[11]이 제안하여 동적인 얼굴 표정 인식에 적용했다. Hoey[12]는 혼합 마코프 체인으로 이루어진 베이저안 네트워크를 제안했다. 이외 가버 웨이블릿과 HMM을 이용한 표정 인식 방법도 제안되었다[13].

이러한 기존 연구의 공간적 분석 방법은 얼굴 표정의 변화로 나타나는 시간적 행동을 포착하기 어렵다. 또한 시공간적 분석 방법은 불확실한 표정 상태와 시간적인 변화를 모델링하기에는 적절하지만 이전의 표정 상태에 종속적인 현재의 얼굴 표정을 표현하기는 어렵다.

얼굴 표정은 시간적 정보뿐만 아니라 표정을 결정하는 특징들의 일시적인 조합으로도 형성되기 때문이다.

또한 얼굴의 시간적 정보만을 이용한 경우 일련의 표정에서 표정의 시작, 최고 상태, 종료 이 세 가지 이산 정보 상태만을 고려하며 표정의 움직임 그 자체는 고려하지 않는다. 이 때문에 인간의 행동 분석에서 더 많은 정보를 담고 있는 관찰된 표정의 순간적 의도와 시간적 진전은 감안하지 않는다는 단점이 있다.

따라서 본 연구에서는 얼굴 표정 분석에 공간적 분석 방법과 시공간적 분석 방법을 동시에 적용함으로써 얼굴 표정의 공간적 특성과 시간적 특성을 적절히 인식하고 표정의 순간적 의도를 분석하도록 한다. 이를 위해 영상 데이터의 공간적 분석 방법인 PCA(principal component analysis)와 순차적 데이터를 인식하는데 널리 사용되어 온 HMM을 시공간적 방법으로 병용하여 연구를 진행했다. 특히 움직임 단위의 표정 데이터에 단순히 HMM을 적용한 기존 연구방법[9]대신 PCA기반의 공간적 분석 방법에서 추출한 특징을 이용하여 얼굴 표정 패턴을 분류하는 표정HMM을 제안함으로써 관찰된 표정의 순간적 의도와 표정의 시간적 변화를 감지하여 제안 프레임워크가 시간적 진전에 민감한 얼굴 표정 변화에 적절하게 대응하게 한다. 또한 표정HMM이 분류한 표정 분류의 신뢰도를 높이기 위해 표정 분류 결과에 표정 특징 벡터를 적용하여 나타나는 사후 확률을 이용함으로써 표정 인식의 정확도를 높인다.

1.2 연구의 개요

본 연구에서는 기존 시간적, 시공간적 얼굴 표정 분석 방법의 문제를 해결하기 위해 학습한 시공간 패턴에 기반하여 비디오 시퀀스에서 얼굴 표정을 감지, 추적, 해석하기 위한 프레임워크를 제안한다. 얼굴 표정의 시공간적 성질에서 공간적 정보량이 적어도 시간적 특성은 유지되기 때문에 본 연구에서는 공간적 정보는 고정 크기로 표현하고 시간적 변화의 학습과 모델링을 더 강조하여 단순한 시각적 인식뿐만 아니라 현재 표정 상태를 정확히 해석하는 방안을 제시한다.

본 연구에서는 음성과 움직임 인식 분야[14]뿐만 아니라 시각적 추적이 적용되고 있는 HMM[15]을 이용하여 얼굴 표정의 시간적 패턴을 모델링한다. HMM은 시간 변화에 따라 복잡하게 변하는 행동을 나타내는 신호를 인식하고 학습하는데 효과적인 확률적 프레임워크를 제공하기 때문이다. 또한 공간적 정보를 나타내기 위해 얼굴 표정 에너지를 PCA로 표현한다.

이를 위해 먼저 주요 표정을 취하는 사람들의 표정에서 주요 시퀀스를 수집하고 이 시퀀스에서 얼굴 영역을 검출한다. 검출한 이미지에는 PCA를 적용하여 현재의 얼굴 특징을 인식한다. PCA를 통해 찾은 고유 벡터를

이미지의 고유공간 투영에 이용하면 각 이미지 시퀀스는 궤적을 형성하게 된다. 이 궤적들은 연속적인 관측 밀도의 HMM을 학습하는데 이용된다. 이와 같은 방법으로 각 표정별로 PCA와 HMM 학습을 수행한다.

실제로 인식해야 할 표정 데이터는 학습과 같은 과정을 거친 다음 HMM 학습 데이터와의 확률 비교를 통해 표정 인식과 분류가 이루어진다. 이 때 6가지 대표 표정보다 그 정도가 약한 표정은 인식과 분류가 쉽지 않으므로 사후 확률을 이용하여 표정 범주를 결정하도록 한다.

2. 시스템 구조

본 연구의 표정 분석 시스템은 그림 1과 같이 3 가지 대표적인 단계인 얼굴 검출(추적), 얼굴 특징 추출, 얼굴 표정 분류로 이루어진다. 얼굴 검출 단계에서는 입력된 비디오 이미지에서 얼굴 영역을 검출하며 검출된 얼굴에서 얼굴 표정의 공간적 분석을 통해 PCA를 적용, 얼굴의 표정 특징을 추출 한 다음 표정 HMM을 통해 표정을 분류한다.

2.1 얼굴 검출

본 연구에서는 얼굴 피부 색조를 이용하여 얼굴 영역을 검출하는 선행 연구[16]를 바탕으로 얼굴 검출을 처리한다. 이미지시퀀스에서 얼굴 영역 부분을 검출하는데는 빛, 조명 등 환경적 영향이 크기 때문에 색상의 노출

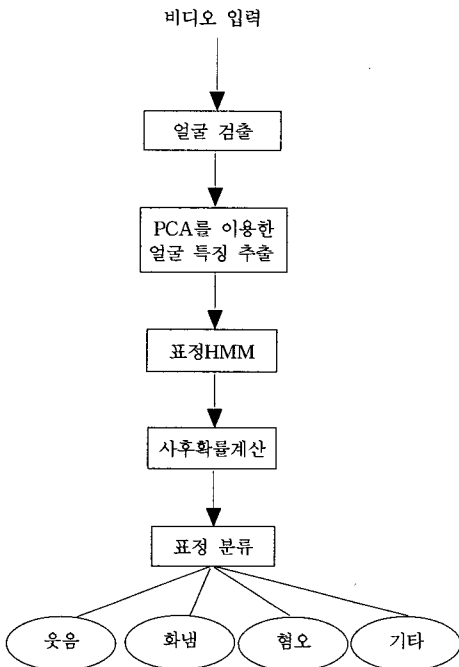


그림 1 얼굴 표정 인식 프레임워크 구조

을 정규화하기 위해 “참조 흰색”을 이용하여 빛의 조건을 보정하는 방법으로 이미지를 전처리한다. 얼굴의 각 부분은 색상과 모양 정보를 기반으로 결정되기 때문에 피부와 유사한 색도와 휘도를 가진 영역을 먼저 추출하고 그 다음 입과 눈 영역을 포착하는 최적의 타원을 계산한다. 이를 위해 피부 색조의 RGB 피부 색상을 YC, C, α 값으로 변환하고 이 YC, C, α 값을 이용하여 피부 영역을 찾아 내기 위해 피부 색상을 정의함으로써 얼굴의 다른 영역에서 피부 색조를 구별하는데 적절한 갈라 이미지의 값을 찾아내어 이미지에서 얼굴 피부 영역만을 검출한다.

눈과 입은 색도에 대해 타 영역과 다른 반응을 보인다는 점을 고려하여 수평, 수직적 투영과 윤곽선을 이용, 눈과 입 영역을 타원으로 표시한 다음 그림 2의 (a)와 같이 눈 영역의 위와 입 영역의 바로 아래에 선을 그어 선이 그어진 영역만을 대상으로 눈과 입을 이용한 기하학적 거리 값을 통해 얼굴 특성의 위치를 결정한다.

얼굴의 검출 영역은 눈과 입이 포함된 이미지를 이용하여 얼굴이 전면에 보일 때 그 이미지를 고정시켜 그림 2의 (b)와 같이 검출을 처리한다.

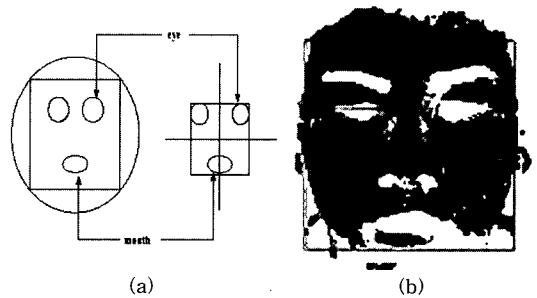


그림 2 얼굴 검출 결과

2.2 표정 특징 추출

얼굴 표정의 특징은 자세, 빛의 영향, 감정 상태 등에 따라 다양하게 나타나며 이 특징은 계측이나 사람의 개입 없이 실시간으로 추적하여 포착해야 한다[17]. 또한 표정 분석을 위해서는 표정 특징을 적절하게 추출해야 하는데 얼굴 특징의 위치보다는 표정 특징을 결정하는 파라미터를 찾아내는 것이 중요한 일이다.

본 연구의 선행 연구[18]에서는 PCA를 이용, 이미지의 고유벡터를 예측하여 얼굴 특징을 계산한다. 이 방법을 적용하면 얼굴 분포도의 주요 성분을 찾거나 얼굴 이미지 집합의 공분산 행렬의 고유벡터를 찾을 수 있다. 고유벡터들은 얼굴 이미지간의 변화를 특성화하는 특징 집합으로 간주된다. 따라서 본 연구에서는 여러 환경적 조건하에서 얼굴의 특징을 표현하는 고유 얼굴 효과를

이용하여 표정 특징을 추출하기 위해서는 검출된 얼굴 영역의 PCA를 이용한다.

본 연구에서는 그림 3과 같이 표정을 결정하는 5개의 고유 벡터 값으로 양 눈 중심점과 눈썹 시작점간의 거리, 양 눈 중심점과 입의 끝점간의 거리, 입의 양 끝점간의 거리를 이용한다. 표정 특징 값을 찾기 위해서는 먼저 얼굴 피부 경계값을 얼굴 경계 영역의 피부 색상에 적용하여 얼굴 경계선을 찾은 후 검출된 얼굴 영역의 눈과 입의 윤곽선을 이용하여 두 눈 바로 위와 입 바로 아래에 선을 그린 다음 128 x 128 크기로 얼굴 경계 지도를 만든다. 그 다음 얼굴 대칭 축을 중심으로 눈, 코, 입의 기하학적 위치 값을 이용하여 피부 색상과 다른 어두운 영역인 눈의 중심점과 눈썹의 머리점을 반복하여 찾아 낸다. 눈의 위치를 찾은 다음 눈의 위치를 이용하여 입의 위치는 적절하게 예측하여 찾는다. 얼굴 대칭 축을 이용하기 때문에 한 쪽만 찾으면 동일한 거리 값을 다른 쪽에 적용하여 원하는 거리 값을 구할 수 있다.



그림 3 표정 인식을 위한 특징 파라미터

2.3 표정 학습과 분류

사람의 표정은 여러 상태로 나눌 수 있다. 무표정한 얼굴에서 웃거나 화를 내거나 싫어하는 표정으로 변하는 움직임 패턴은 시간의 흐름에 따라 초기 상태에서 최고조의 상태 그리고 종료 등의 상태로 나뉜다. 표정 패턴을 여러 상태로 세분화하면 더 복잡해진다. 또한 잡음이 표정 변화 신호에 섞일 수도 있다. 이러한 상황은 음성, 움직임 인식[14,15]과 유사하기 때문에 이 분야 연구에 다양하게 적용된 HMM을 표정 학습과 분류에 적용한다.

2.3.1 은닉 마코프 모델(HMM)

HMM은 시간에 따라 변하는 신호를 분석하는데 적당한 통계적 동작 모델로 두 가지 상태의 통계 과정을 생성한다. 그 하나는 관측자에게는 보이지 않는 과정으로 시스템 내에서 상태 변환을 담당한다. 다른 하나는 시스템 종속적인 확률 분포에 따라 관측치를 발산하는 일을 담당한다. HMM은 시간 급수에 대한 확률 분포로

$\theta = \{S, P_i, P_{j \rightarrow i}, P_{i(x)}\}$ 으로 구성되어 있다.

$S = \{s_1, \dots, s_n\}$ 은 이산 상태의 집합이고 확률행렬 $P_{j \rightarrow i}$ 은 j 상태에서 i 상태로 전이하는 확률이며 확률 벡터 P_i 는 i 상태에서 시작한 시퀀스의 확률이다. 확률 분포 $P_i(x)$ 는 i 상태에서 X 를 관측하는 확률로 식 (1)과 같이 평균 μ_i 과 편차 k_i 의 가우스 혼합 분포 모델을 따른다.

$$p_i(x) = N(x; \mu_i, k_i) = \frac{1}{\sqrt{2\pi k_i}} e^{-\frac{1}{2}(x-\mu_i)^T k_i^{-1}(x-\mu_i)} \quad (1)$$

HMM 이론은 학습 수행, 인식, 상태 예측을 수행하는데 효과적인 확률적 절차를 제공하는 상태공간 모델로 HMM에서는 시간 급수 X 를 배치 공간을 경유하는 경로로 간주하여 배치 공간을 확률 분포에 따라 특정 은닉 상태가 점유하는 영역으로 분할한다.

은닉 상태 $S = \{s_1, s_2, \dots, s_T\}$ 의 특정 시퀀스에 해당하는 경로 가능성도 $X = \{x_1, x_2, \dots, x_T\}$ 는 현재 은닉 상태에 해당하는 경로의 각 포인트별 확률이다. 이 확률 간격은 상태 시퀀스 자체 확률의 시간 간격에 따라 정한다. 모든 은닉 상태 시퀀스를 합하면 전체 HMM에 관한 경로 가능성이 된다. HMM은 동작간의 변화와 모든 학습 동작을 결합하여 개별적인 동작 클래스를 더 큰 모델 형태로 합성하는 특성을 가지고 있기 때문이다.

2.3.2 표정 은닉 마코프 모델(Expression Hidden Markov Model)

표정 인식 시스템이 분석해야 하는 이미지 시퀀스의 표정 패턴은 여러 학습 표정 패턴 중의 하나이므로 많은 상황 중에 한 가지를 가정하여 처리하여야 한다. 따라서 다양한 표정의 패턴을 학습하는 표정 HMM을 구성하는 것이 필요하다. 표정 인식 시스템이 최초에 어떤 표정 HMM을 선택할 것인지를 결정하기 위해 본 연구에서는 기쁨, 슬픔, 화냄, 혐오, 두려움, 놀람등의 표정을 나타내는 대표 표정 변수를 v 로 설정한다. 시스템은 자동적으로 이미지 시퀀스의 대상 표정을 분류해야 하기 때문에 확률적으로 v 값을 결정해야 한다.

표정 HMM은 표정 특징 파라미터가 기능적으로 대표 표정 변수 v 에 따르는 HMM으로 발산 확률 함수 $p_i(x_i)$ 는 평균과 편차가 v 에 따라 변하는 가우스 분포이다.

사람의 표정은 화냄, 웃음, 두려움, 지루함 등 주요 표정에 따라 다른 표정 패턴을 보인다. 따라서 각 표정을 결정하는 특징을 나타내는 변수 v 에 따라 다른 표정 HMM을 선택하고 표정 HMM 내에서 표정간의 전이가 발생하면 전이된 다른 표정 모델의 HMM을 적용한다.

표정 HMM은 $\theta = \{S, P_i, P_{j \rightarrow i}, \mu_i, K_i, P, D, v\}$ 로 정의

된다.

평균벡터 μ_i , 편차 행렬 K_i , 분산 행렬 P, \bar{D} 그리고 대표 표정 변수 v 는 식(2)와 같이 i 상태에서 관측하는 데이터 $P_i(x_i)$ 의 가우시안 확률 $p_i(x_i)$ 의 매개변수를 정한다.

$$p(x_i|s_i) = N(x_i; \mu_i + Pv, K_i + \bar{D}v) \quad (2)$$

여기서 표정 편차 행렬 $K_i + \bar{D}v$ 는 절대값에 고유값을 사소한 양의 유한값이다.

파라미터 $\{P_i, P_{i \rightarrow i}, \mu_i, K_i\}$ 는 예측을 통해 얻어지고 P, \bar{D} 는 PCA의 결과치인 2.2에서 추출한 5개의 얼굴 표정 고유 벡터와 벡터의 평균값이다. v 는 시스템이 인식한 사용자의 얼굴 표정 특징 데이터에서 예측한다. v 값을 고정시키면 모델은 표준 이산 상태의 가우시안 HMM이 된다. 표정 HMM은 그림 4와 같은 형태로 구성된다.

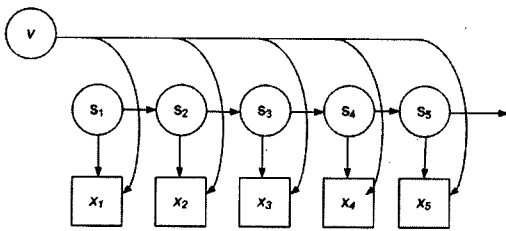


그림 4 표정 은닉 마코프 모델

그림 4에서 관측된 신호 x_i 는 시간에 따라 변하는 이산수치의 은닉상태 변수 S_i 에 의해 결정되며 v 와 S_i 는 은닉 상태이므로 확률적으로 추론해야 한다. 그림 4의 호는 표정 변수간의 조건적 종속 관계를 표시하는데 계수화된 호환 함수 형태를 취한다.

2.3.3 표정 패턴의 학습

본 연구에서는 6개의 주요 얼굴 표정이 담긴 이미지 시퀀스를 이용하여 학습을 한다. 먼저 학습 시퀀스에서 얼굴 영역을 검출하면 모든 표정 클래스 $n=1, \dots, N$ 에서 시퀀스 집합 $\{I_i[t], i=1, \dots, N, t=1, \dots, T_m\}$ 이 생성된다.

표정 특징 파라미터를 추출하기 위해 이미지 간 공간적 차이에 식(3)과 같이 PCA를 수행한다.

$$D_i[t] = a_i[t] \cdot I_i[t+1] - I_i[t] \quad (3)$$

스케일 계수 a_i 는 분산치에 따라 적용된다. PCA는 평균 이미지 \bar{D} 와 주요 성분 $P=[I_i], i=1, \dots, N_c$ 을 생성하는데 N_c 는 적용된 주요 성분의 수이다. 각 이미지

시퀀스 $\{D_i[t]\}$ 는 식(4)에 따라 고유공간에 사영된다.

$$Q_i[t] = P \cdot (D_i[t] - \bar{D}) \quad (4)$$

시간 시퀀스 집합 $\{Q_i[t]\}$ 은 N_c 차원의 고유공간에서의 궤적이다. 주요 성분 P 와 평균 이미지 \bar{D} 와 더불어 시간 시퀀스 $\{Q_i[t]\}$ 는 학습 시퀀스의 시공간적 특징을 나타낸다. 시퀀스 $\{Q_i[t]\}$ 는 HMM을 학습하는데 이용한다.

HMM은 학습(Baum-Welch 예측), 인식(forward-backward 알고리즘), 상태 예측(Viterbi 알고리즘)[19]으로 이루어진 확률 방법론으로 상태 변화 $p(\{Q_i[t]\}|\theta)$ 수를 제한하여 표정을 분류한다. 표정 클래스당 한개인 HMM에서 N 개의 클래스를 학습하기 위해 시간 시퀀스 $\{Q_i[t]\}$ 를 이용하는데 먼저 표정 학습 시퀀스가 주어지면 다른 표정 클래스의 사전 균등치 추측에 의해 식 (5), (6)과 같이 선행행 (forward-backward) 알고리즘과 베이즈 규칙을 적용한다.

$$n_{max} = \arg \max_n p(\theta(Q_i[t])) \quad (5)$$

$$= \arg \max_n p(Q_i[t]|\theta) \quad (6)$$

식 (5)와 (6)에 의해 주어진 관측치가 가장 가능성이 높은 표정 클래스 $\theta^{(n_{max})}$ 가 생성되고 그에 따라 표정의 분류가 이루어진다.

만약 HMM에서 최적의 상태 시퀀스를 찾기 위해 분석 대상인 이미지 시퀀스에 Viterbi 알고리즘[19]을 적용하면 관측치가 가장 가능성이 높은 상태 시퀀스 $Q=\{q_1, \dots, q_n\}$ 와 표정 모델 $\theta^{(n_{max})}$ 가 생성된다. 이 방법으로 표정 주기와 이벤트, 키 프레임 표정과 같은 정보들을 추출할 수 있다.

2.3.4 사후 확률을 이용한 약한 표정 분류

미소나 찌푸림 등의 그 정도가 약한 얼굴 표정이 대표 얼굴 표정보다 그 표현빈도가 높은 반면 표현 방식은 대표 감정보다 훨씬 복잡하고 다양하기 때문에 인식하기에 상당히 어렵다. 또한 약한 얼굴 표정은 표정을 결정하는 각 얼굴 부위에서 동시에 다양한 표정들의 파라미터가 포착된다.

이런 문제점은 표정 분류에 오류를 일으키는 요인이 되므로 약한 표정 분류의 오류를 줄이기 위해서 본 연구에서는 Zhang과 Ji[20]가 제안한 다양한 시각적 인식 정보를 이용한 확률로써 인식율을 높이는 방법을 제안 표정 인식 프레임워크에 적용한다.

주어진 이미지시퀀스에 표정HMM을 적용하여 표정 상태 $X=(x_1, \dots, x_n)$ 중의 한 표정이 결정될 때 해당 얼굴 표정은 m 개의 PCA를 이용한 표정 특징 벡터 정보 $E=e_1, \dots, e_m$ 를 가지고 있으므로 이렇게 주어진 표정 벡터를 이용하여 표정 상태의 사후 확률을 예측함으로써 인식율을 높인다.

사후 확률 p 는 벡터 정보 E 가 주어졌을 때 $X=x_i$ 로 인식되는 확률 즉, $p=(X=x_i|E)$ 이다. 이 때 인식한 표정의 벡터값들을 평균하여 사후 확률 예측에 적용하면 E 가 주어졌을 때 표정 상태 X 의 확실성을 높이는 데 도움이 된다.

사후 확률 $p=(X=x_i|e \in E)$ 가 1 또는 0에 가깝게 되면 어느 대표 표정으로 분류 가능한지 아닌지가 판단되어 정도가 약한 얼굴 표정을 대표 표정 중 하나로 분류하기 쉽다.

사후 확률이 적용된 정도가 약한 표정 특징 벡터 E 에는 E_1, E_2 의 각각 다른 표정 벡터가 섞여 있을 수 있으므로 표정 인식의 불확정 데이터 감소치를 이용하여 약한 표정을 대표 표정으로 정확히 분류하도록 한다. 이를 위해 식(7)과 같이 사논의 엔트로피 계측[21]을 확장하여 불확정 감소치 $R(X|E_1, E_2)$ 를 구한다.

구한 불확정 감소치를 분류하고자 하는 해당 표정 특징 벡터에서 빼면 서로 다른 표정 벡터 중 더 큰 값만 남게 되어 표정 인식의 신뢰도가 높아진다.

$$R(X|E_1, E_2) = ENT(X) - ENT(X|E_1, E_2)$$

$$= - \sum_{x \in X} p(x) \log(p(x)) + \sum_{x \in X} \sum_{e_1 \in E_1} \sum_{e_2 \in E_2} p(e_1, e_2) p(x|e_1, e_2) \log p(x|e_1, e_2)$$

(7)

x 와 e_i 는 X 와 E 에서 파생되는 값이다. $p(x|e_1, e_2)$ 는 E_1, E_2 가 조건 독립이라는 가정 아래 식 (8)과 같이 베이즈 규칙을 적용하여 예측한다. 식 (8)의 $p(x|e_1)$ 과 $p(x|e_2)$ 는 표정HMM에서 얻는다.

$$p(x|e_1, e_2) = \frac{p(x|e_1)p(x|e_2)}{p(x) \sum_{x \in X} \frac{p(x|e_1)p(x|e_2)}{p(x)}} \quad (8)$$

이 과정을 거쳐 표정 HMM을 통해 분류된 표정 상태 $p(X)$ 의 인식 결과는 거쳐 더 정확하게 수정된다.

3. 실험

실험을 위한 데이터로는 그림 5와 같이 10명으로부터 찍은 6개의 대표 표정, 즉, 화남 표정, 혐오스러운 표정, 기쁨 표정, 놀란 표정, 슬픈 표정, 두려운 표정 외에 대표 표정과 유사하나 정도가 약한 미소 짓는 얼굴, 불쾌한 표정, 당황스러운 표정이 담긴 이미지 시퀀스를 이용했다. 표정을 포착하기 위해 실험 대상자는 실험 주체자로부터 훈련을 받았으며 실험 대상자의 얼굴 움직임은 정면 카메라로 실내에서 찍었다.

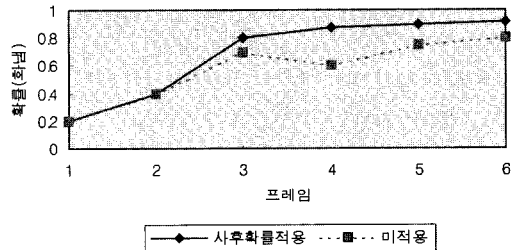
정도가 약한 얼굴 표정의 인식 확률은 낮기 때문에 특정 시점에서 잡은 표정은 어느 표정 클래스에 속하는지 분류하기 어려워 인식 오류를 보인다. 따라서 정도가



그림 5 실험 대상 표정 데이터



(a) 화내는 표정의 이미지 시퀀스



(b) 사후 확률 적용 및 미적용 확률의 비교 결과

그림 6 사후 확률을 이미지시퀀스에 적용한 결과

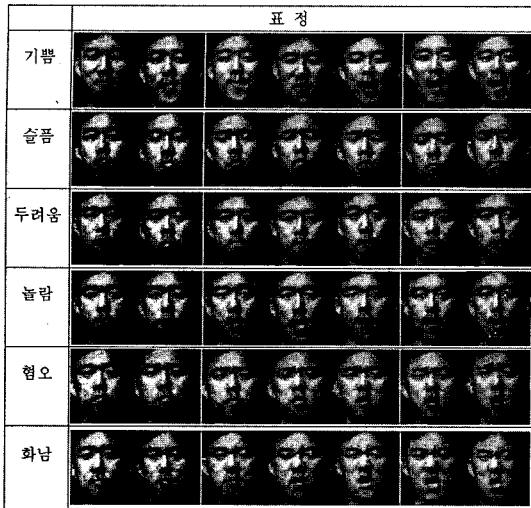
약한 표정을 대표 표정으로 분류하기 위해 사후 확률을 적용한 표정 분류 결과는 그림 6과 같다.

그림 6의 (a)는 정도가 약한 표정의 이미지 시퀀스이다. 이 표정을 대표 표정으로 분류한 확률은 6 (b)의 실선과 같다. 그림 6 (b)의 점선은 사후확률을 이용하지 않은 상태의 표정 분류 결과이다.

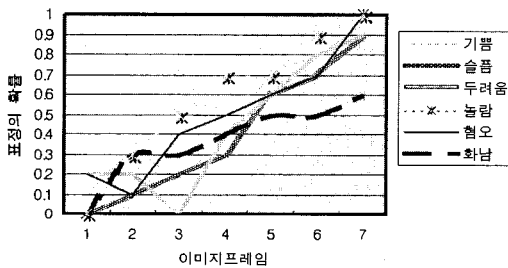
그림 7은 실험 이미지 시퀀스에서 얼굴 표정을 인식한 결과를 보여 준다

얼굴 표정의 원 이미지 시퀀스는 대표 표정과 정도가 약한 표정을 담고 있다. 원 이미지 시퀀스에서 적절한 표정을 담고 있는 프레임을 키프레임으로 한 이미지 시퀀스를 그림 7(a)과 같이 정리한 후 그 결과를 본 논문에서 제안한 얼굴 표정 인식 프레임워크에 적용했다. 그림 7(a)의 표정 확률은 그림 7(b)로 얼굴 표정의 정도가 약한 상태에서 대표 표정으로 시간적으로 변화하는 과정을 확률로 나타냈다. 얼굴 표정은 시간 슬라이스에서 한 슬라이스 정도만 나타나는 것이 아니라 시간 슬라이스 앞 뒤로 가우시안 분포를 이루기 때문에 표정 특징의 시공간적으로 변화 정도를 포착하면 대표 표정의 예측이 가능하다.

표 1은 제안 프레임워크의 인식결과를 보여 준다. 실험 결과 사후 확률을 적용한 제안 프레임워크는 대표



(a)얼굴 표정 시퀀스



(b) 표정 분류 결과

그림 7 대표 표정과 정도가 약한 표정 인식 결과

표정의 그보다 정도가 약한 표정도 대표 표정으로 분류, 인식함을 알 수 있다. 또한, 두려움과 같은 다양한 특징을 보여주는 표정보다는 기쁨과 놀람 등 얼굴의 움직임이 큰 표정 인식에 강건함을 알 수 있다.

4. 결론 및 향후 연구

본 연구에서는 자동화된 표정 인식 프레임워크를 구현하여 이미지 시퀀스에서 인물의 얼굴 검출, 얼굴 표정 특징 추출을 통해 다양한 표정을 대표 표정중 하나로 인식하는 과정에서 공간적 분석방법으로는 PCA를 이용하고 시공간적 분석 방법으로는 표정 HMM을 적용했다.

제안 프레임워크에 적용된 표정 HMM은 표정 특징 벡터를 시간 변화에 종속적인 표정 패턴을 학습하여 대표 표정 중의 하나로 분류하여 인식한다. 이 방법은 기존 연구의 움직임 단위 얼굴 표정에 적용된 HMM을 개선함으로써 표정 패턴의 시공간적 변화를 포착, 이미지 시퀀스 내에서 순간순간 다른 표정으로 전이하는 것을 감지하여 시간의 변화에 따른 표정의 상태를 인식하는데 강점을 보인다.

또한 제안 프레임워크에서는 기본 표정보다 정도가 약한 표정의 인식 오류를 줄이기 위해 표정 HMM에 의해 인식, 분류된 표정 상태가 사전 표정 특징 벡터에 의해 결정된다는 가정하에 표정 특징 벡터가 주어졌을 때 나타나는 대표 표정의 사후 발생 확률을 이용했다.

사후 확률을 이용하면 대표적인 6개 표정 뿐 아니라 웃음, 슬픔 등의 표정보다 정도가 약한 미소, 서운함과 같은 표정의 인식률을 높여 제안 프레임워크는 더 정확한 표정 분류가 가능하다.

본 연구에 적용한 실험 데이터 관련하여 이미지시퀀스 상의 다양한 인물의 얼굴 포즈 및 얼굴 각도, 얼굴이 다른 대상과 겹친 경우 등이 자동 얼굴 검출 및 표정 인식 과정에서 중요한 변수로 작용할 수 있으나 본 연구에서는 고려하지 않았으므로 향후 연구에서 이를 개선해 나갈 예정이며 제안 프레임워크의 신뢰도를 높이기 위해 표정 특징 파라미터의 다양화와 기존 표정 인식 알고리즘과의 비교 연구도 계속할 계획이다.

본 연구는 표정 인식을 통한 HCI 에 응용 가능 할 뿐 아니라 애니메이션의 얼굴 표정 구현 및 합성에도

표 1 표정 인식 결과

대표 표정	약한 표정	화남	혐오	두려움	기쁨	놀람	슬픔	인식율
화남		14	2	1	0	0	1	78%
	불쾌	4	1	0	0	0	1	67%
혐오		2	15	0	1	0	0	83%
	싫어함	1	4	0	0	0	1	67%
두려움		1	1	13	0	2	1	72%
	겁냄	0	0	3	0	2	1	50%
기쁨		0	0	0	18	0	0	100%
	미소	0	1	0	6	0	0	86%
놀람		0	0	1	0	17	0	94%
	당황	0	0	1	0	6	0	86%
슬픔		1	1	2	0	0	15	79%
	서운함	1	1	0	0	0	4	67%

적용 가능하다.

참고 문헌

- [1] M. Rosenblum, Y. Yacoob, and L. Davis, "Human Expression Recognition from Motion Using a Radial Basis Function Network Architecture" *IEEE Trans. on Neural Network*, vol. 7, no. 5, pp. 1121-1138, 2002.
- [2] B. Fasel, "Multiscale Facial Expression Recognition using Convolutional Neural Networks," In *Proc. of the Third Indian Conference on Computer Vision, Graphics and Image Processing(ICVGIP'2002)*, 2002.
- [3] M. N. Daile, W. C. Cottrell, C. Padgett and R. Adlophs, "EMPATH: A Neural Network that Categorizes Facial Expressions," *Journal of Cognitive Science*, vol. 14, no. 8, pp. 1158-1173, 2002.
- [4] M. J. Lyons, J. Budynek and S. Akamatsu, "Automatic Classification of Single Facial Images," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 21, no. 12, pp. 1357-1362, 1999.
- [5] M. Pantic and L. Rothkrantz, "Expert System for Automatic Analysis of Facial Expression," *J. Image and Vision Computing*, vol. 18, no. 11, pp. 881-905, 2000.
- [6] M. Schulze, K. Scheffeller and C. W. Omlin, "Recognizing Facial Actions with Support Vector Machines," In *Proc. PRASA 2002*, pp. 93-96, 2002.
- [7] I. Essa and A. Pentland, "Coding, Analysis, Interpretation, Recognition of Facial Expressions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol 10, no. 7, pp. 757-763, 1999.
- [8] T. Otsuka and J. Ohya, "Spotting Segments Displaying Facial Expression from Image Sequences Using HMM," In *Proc. Inter. Conf. on Automatic Face and Gesture Recognition 98*, pp. 442-477, 1998.
- [9] J. J. Lien, H. Kanade, T. Kitamura, J. F. Cohn and C. C. Li, "Detection, Tracking and Classification of Action Units in Facial Expression," In *Journal of Robotics and Autonomous Systems*, pp. 432-329, 1998.
- [10] P. Ekman and W. V. Friesen, "The Facial Action Coding System: A Technique for Measurement of Facial Movement," *Consulting Psychologists Press*, San Francisco, CA, 1978.
- [11] S. Muller, S. Eickerler and G. Rigoll, "Pseudo 3D HMMs for Image Sequence Recognition," in *IEEE Proc. Inter. Conf. on Image Processing*. 1999, pp. 237-241, 1999.
- [12] J. Hoey, "Hierarchical Unsupervised Learning of Facial Expression Categories," *IEEE Workshop on Detection and Recognition of Events in Video (EVENT'01)*, pp. 92-99, 2001.
- [13] 이경아, "웨이블렛 계수와 Hidden Markov Model를

이용한 얼굴인식 기법", 한국 퍼지 및 지능시스템학회 03년 추계 학술대회 학술발표 논문집, 162-165쪽, 2003.

- [14] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Pro. IEEE*, vol 77, no. 2, pp. 257-286, 1989.
- [15] C. Bregler, "Learning and recognizing human dynamics in video sequences," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 568-574, 1998.
- [16] Jin Ok Kim, Sung Jin Seo and Chin Hyun Chung, "Real-Time Face Recognition by the PCA with Color Images," *Lecture Notes in Computer Science*, LNCS 3036, pp. 687-690, June 2004.
- [17] 최명근, 정현숙, 신영숙, 이일병, "표정 변화에 따른 얼굴표정에서의 특징점 추적", *한국정보과학회 논문집 7권*, 2호, 425-427쪽, 2000.
- [18] Jin Ok Kim, Sung Jin Seo, Chin Hyun Chung, Jun Hwnag and WoongJae Lee, "Face Detection by Facial Features with Color Images and Face Recognition using PCA," *Lecture Notes in Computer Science*, LNCS 3043, pp. 1-8, May 2004.
- [19] A. J. Viterbi, "Error Bounds for Convolution codes and an Asymptotically Optimum Decoding Algorithm," *IEEE Trans. Information Theory*, vol. 13, pp. 260-269, 1967.
- [20] Y. Zhang, Q. Ji, "Facial Expression Understanding in Image Sequences Using Dynamic and Active Visual Information Fusion," *IEEE Inter. Conf. on Computer Vision(ICCV2003)*, vol. 2, 2003.
- [21] C. E. Shannon, "A mathematical theory of communication," *Bell System Technical Journal*, vol. 27, pp. 379-423 and 623-656, 1948.



김진옥

1989년 성균관대학교 졸업(학사). 1998년 성균관대학교 정보통신공학과 졸업(석사) 2002년 성균관대학교 전기전자 및 컴퓨터 공학과 졸업(박사). 2003~2004 세종대학교 컴퓨터공학과 초빙교수. 2004~ 현재 대구한의대학교 멀티미디어학부 전임강사. 1992년~1994년 (주)현대전자산업 정보통신사업본부 1994년~1999년 (주)현대정보기술 인터넷사업본부 과장. 1999년~2000년 (주)은세통신 온라인사업 팀장. 2000년~2001년 (주)유료코넷 기술담당 이사. 관심분야는 Multimedia Processing, Image Processing, Biometrics, Pattern Recognition, Multimedia Contents