

# 내적상태 차원모형에 근거한 얼굴표정 합성 시스템\*

## A system for facial expression synthesis based on a dimensional model of internal states

한 재 현\*\*      정 찬 섭\*\*  
(JaeHyun Han) (ChanSup Chung)

**요 약** 쾌-불쾌, 각성-수면의 2차원 감정상태 공간에서 임의의 좌표값을 지정하면 Parke와 Waters[1]의 근육 기반 얼굴모형에 근거하여 그것에 해당하는 표정을 자동 합성할 수 있는 시스템을 개발하였다. 얼굴 변형의 준거 설정을 위하여 정찬섭, 오경자, 이일병과 변해란[2]의 표정 DB에 수록된 표정 자료 및 내적감정상태 모형을 사용하였다. DB 내 표정 자료들은 감정 모형에 근거하여 차원값 및 범주로 기술된 내적상태 정보를 포함하고 있었다. 내적상태 변화에 따른 표정의 변형 규칙을 찾기 위해 표집된 표정들에 대해 21개 얼굴 근육의 변화량을 측정하였다. 내적감정상태와 얼굴 근육들의 관계 구조를 파악하기 위해 수집된 자료들을 대상으로 회귀분석을 실시하였다. 이러한 과정을 거쳐 얻어진 변형 규칙은 쾌-불쾌 및 각성-수면 차원값으로 주어지는 내적상태에 대한 자연스러운 표정을 합성해내었다. 이같은 결과는 방대한 양의 자료 및 개별 근육의 변화를 포착하는 회귀분석으로부터 도출된 규칙이 얼굴표정을 합성하는데 유용하고 강력한 도구가 될 수 있음을 시사한다.

**주제어** 표정, 합성 시스템, 차원모형, 내적상태

**Abstract** Parke and Waters' model[1] of muscle-based face deformation was used to develop a system that can synthesize facial expressions when the pleasure-displeasure and arousal-sleep coordinate values of internal states are specified. Facial expressions sampled from a database developed by Chung, Oh, Lee and Byun[2] and its underlying model of internal states were used to find rules for face deformation. The internal - state model included dimensional and categorical values of the sampled facial expressions. To find out deformation rules for each of the expressions, changes in the lengths of 21 facial muscles were measured. Then, a set of multiple regression analyses was performed to find out the relationship between the muscle lengths and internal states. The deformation rules obtained from the process turned out to produce natural-looking expressions when the internal states were specified by the pleasure-displeasure and arousal-sleep coordinate values. Such a result implies that the rules derived from a large scale database and regression analyses capturing the variations of individual muscles can be served as a useful and powerful tool for synthesizing facial expressions.

**Keywords** facial expression, synthesis system, dimensional model, internal states

### 1. 서론

표정은 내적감정상태에 따라 얼굴 근육의 조성 형태가 달라지고 그로 인해 얼굴의 표면 구조가 변화되어 나타나게 된다. 따라서, 표정을 해석하거나 합성하는 시스템

을 개발하려면 내적감정상태와 얼굴표정을 각기 체계적으로 기술할 수 있는 모형을 만들고 이 두 모형간의 상관 구조를 파악하는 작업이 필요하다. 즉, 내적감정상태의 변화에 의한 얼굴 구조의 변형 규칙을 알아야 한다는 것이다. 이와 같은 내적감정상태에 따른 얼굴의 변형 규칙으로서 지금까지 MAX(Maximally Discriminative Facial Movement Coding System) [3], EMFACS(Emotional Facial Action Coding System)[4], 그리고 AFFEX(A System for Identifying Affect Expressions by Holistic

\* 이 논문은 2000년 한국학술진흥재단의 지원에 의하여 연구되었음. (KRF-2000-908-C00027)

\*\* 연세대학교 인지과학 연구소  
e-mail : jaehyun@psylab.yonsei.ac.kr. cschung@yonsei.ac.kr

Judgements) [5] 등이 제안되어 왔다. 표정의 동영상을 관찰하여 표정을 구성하는 각 얼굴 움직임들을 정밀하게 추출해내고 이를 바탕으로 표정과 정서를 연결지을 수 있도록 개발된 이러한 체제들은 정서와 표정의 관계 구조에 대한 일반적인 연구뿐만 아니라 정서 장애 등 특수한 상황에서의 표정 및 정서 연구에 다양하게 사용되어 왔다. 그러나 이와 같은 표정 해석 시스템들은 기본 정서에 해당하는 표정이 범문화적으로(cross-culturally) 표준화되어 있다는 가정 하에 단편적인 표정 자료들을 기반으로 이 자료들에서의 얼굴 움직임과 몇 가지의 정서 범주들을 단순히 대응시켜 개발되었다는 제한점이 있다. 그 결과 감정의 연속적인 변화 과정이나 정도의 차이가 표정에 반영되는 양상을 효율적이고 섬세하게 포착하는 측면에는 소홀한 면이 없지 않다. 또한 시스템 개발자들은 사람들이 얼굴 움직임을 포착하는 훈련에 초점을 두고 있어 표정을 정량화 하고자 하는 시도가 부족하였다. 그 결과 얼굴의 개별적인 움직임에 대응되는 근육이 규정되어 있지 않았으므로 [6] 이 체제들에서의 정서와 표정의 관계 규칙들이 표정을 자동으로 구현하거나 인식하는 시스템에 직접 응용되기에는 무리가 있다.

내적감정상태의 변화가 얼굴표정에 반영되는 양상을 경험적으로 파악하고 이 결과가 표정 구현 시스템에 응용될 수 있기 위해서는 먼저 감정상태와 얼굴표정의 기저 구조를 확인할 필요가 있다. 체계화된 두 영역에 대해서 이들을 연결 지을 수 있는 타당한 해석 체계가 조사될 수 있으며 이 관계를 효과적으로 구현할 수 있는 구체적인 접근 방법들이 검토될 수 있기 때문이다. 이를 위해 본 연구에서는 먼저 사람의 내적감정상태 및 얼굴표정의 기저 구조에 대한 기본 가정을 점검하였다.

내적감정상태의 기저 구조에 대해서는 기본정서모형(또는 범주 모형)과 차원모형의 두 입장이 나뉘어 감정의 구성적 특징을 서로 다른 측면에서 접근하고 있다. 자기 보고된(self-reported) 정서 경험들을 요인 분석한 결과를 바탕으로 내적 정서는 여섯 개에서 열 두 개 사이의 독립적이고 단극적(monopolar)인 요인들로 구성되어 있다고 보는 입장이 기본정서모형이다. 이에 의하면 표정은 내적감정상태에 대한 일종의 신호 체계로 간주된다. 그러나 기본정서모형은 감정의 정도(intensity)나 중간점을 포함하지 않기 때문에 미세하게 변화하는 감정상태를 포괄적으로 다룰 수 없고 감정의 분류에서 과일반화의 문제를 야기할 수 있다. 뿐만 아니라 이 모형 자체는 연구 방법의 가공물(artifact)일 가능성이 높다는 견해가 제시되고 있다 [7]. 차원모형은 내적감정상태들은 서로 독립적이지 않으며 체계적인 방식으로 관련되

어 연속적인 양극 차원으로 이루어진 공간에 의해 더 잘 표상될 수 있다고 본다. 이 모형을 가정하면 표정으로부터 추론되는 다양한 감정상태들 간의 관계를 수리적으로 규정함으로써 경제적인 방법으로 내적감정상태를 표상할 수 있다. 그러나 감정 범주에 대한 언급이 없이 각 차원에 대한 정도만이 제시되는 경우 직관적인 의미가 없어지는 단점이 있다. 이러한 이유로 내적감정상태에 대해서는, 사람들이 일상적으로 '기쁨', '슬픔' 등과 같이 범주로서 감정을 판단하는 경향을 보이지만 심층적으로는 '쾌-불쾌', '각성-수면'의 두 기저 차원에 근거하여 처리하고 있는 것으로 간주할 수 있다. 이러한 내적감정상태의 기저 구조에 대한 가정은 사람의 미묘한 감정의 변화 및 정도를 표현할 수 있도록 한다는 점에서 이전 연구들의 기본 전제들과 차별화 된다.

내적감정상태와 연결지를 얼굴표정에 대해서는 단순히 얼굴의 움직임을 기술하지 않고 특징 요소들을 기반으로 수치화하고자 하였다. 이를 통해 내적상태의 변화가 얼굴의 특징 요소들에 반영되는 양상을 조사함으로써 표정과 정서에 관한 기존 연구 결과(예, [8])와의 비교가 가능하도록 하였다. 이와 동시에 몇 개의 변수를 설정하느냐, 어떤 두상 모형을 선정하느냐, 그리고 어떠한 알고리즘을 사용하여 변형하느냐 등의 문제를 고려하여 분석 결과를 적용, 표정을 구현함으로써 이전 연구들과의 차별화를 시도하였다. 구현을 위한 문제 해결의 핵심인 사람의 얼굴은 매우 정교하고 곡선적인 골격(skeletal) 기초를 갖고 있으며 골격은 근육에, 근육은 다시 여러 층의 피부 조직에 연결되어 있는 복잡한 구조를 지니고 있다 [9]. 이와 같이 복잡한 구조적 및 역동적 특성들을 모두 하나의 모형에 반영한다는 것은 가능하지 않다. 또 설사 그것들을 다 반영한다고 하여도 용량이나 속도 때문에 실용적 가치가 전혀 없을 수 있다. 그러므로 두상 모형의 열개는 사람과 유사한 골격 및 근육 구조를 갖지만 지나치게 복잡하지 않아 계산이 용이하도록 구성되어야 한다. 얼굴을 구성하는 근육은 200개가 넘지만 표정과 직결된 것만을 선별해서 변형하도록 하여야 한다. 변형 알고리즘은 근육의 움직임에 의한 자연스러운 텍스처 변형이 가능한 것으로 선정하여야 한다. 이러한 측면들을 모두 고려하여 본 연구에서는 Parke와 Waters [1]의 모형을 합성 시스템의 기저 모형으로서 사용하였다.

지금까지 논의한 것과 같이 표정 합성 시스템은 여러 가지 복합적인 작업을 필요로 한다. 내적감정상태의 구조적 특징이 규명되고 이에 대응될 수 있는 표정 변화를 수리적으로 표현할 수 있는 변수들이 선정될 필요가

있다. 이들을 바탕으로 한 감정과 표정의 관계 규칙은 모형과 변형 변수 및 변형 알고리즘에 적용됨으로써 시스템으로 응용될 수 있다. 이러한 단계적인 과정을 통해 개발된 표정 합성 시스템은 사람처럼 자연스럽게 유연한 표정을 합성할 수 있을 것이다.

## 2. 표정 합성의 준거 설정

표정과 감정의 대응 규칙은 다양한 표정 자료를 수집하여 사람들이 표정으로부터 감정을 추론하는 경향을 조사하고 그에 해당하는 표정의 특징적 구조들을 분석함으로써 파악될 수 있다. 정찬섭, 오경자, 이일병과 변혜란[2]의 표정 DB는 영화와 TV 드라마에서 표집된 표정 자료와 연극 배우들의 표정 연기 자료 및 평정 자료를 포함하고 있어 감정과 표정의 연결 관계를 찾아내기 위한 자료로서 적합하였다. 표정 자료를 합성 시스템의 설계에 반영하는 과정에서 영상 처리 작업이 필요하므로 이 DB에서 화질과 촬영각이 영상 처리에 적합한 것만을 분석 대상 자료로 선정하였다. 합성 시스템의 성능은 이들 선정된 설계 기반 자료들에 의해서 결정되므로 시스템 설계에 앞서 이 자료들에 대한 타당성과 신뢰성을 검증하여 시스템의 구현 기반을 강화하였다.

표정을 합성하기 위해서는 중립 표정으로부터 특정 감정에 해당되는 표정으로의 얼굴 변화에 대한 정보가 있어야 한다. 얼굴의 변화는 근육과 연결된 피부와 조직뿐만 아니라 표면의 주름, 선, 골 등 다양한 요소들의 복잡한 움직임을 수반한다. 따라서 시스템의 구현 속도와 성능을 동시에 높일 수 있는 방안의 하나로 감정을 잘 반영하는 주요 변수들만을 선정하고 이 변수들을 정교하게 측정하는 작업을 실시하였다.

### 2.1. 준거 설정을 위한 자료 확보 및 타당성 확인

시스템의 성능 향상을 위해 정찬섭 등의 표정 DB 내의 자료들을 점검하였다. 연구자들이 가정하고 있었던 것처럼 포즈된 표정들은 배우들이 의도한 내적감정상태로 인식되는가를 조사하였다. 이러한 과정을 통해 일차적으로는 DB에 포함되어 있는 감정상태 정보의 정확성을 확인하고자 하였다. 나아가서 배우들이 얼마나 주어진 스크립트 및 형용사 범주에 충실하게 연기하였는가를 평가하였으며 이를 통해 배우들의 표정 자료에 그들이 표현하고자 의도한 내적상태 범주를 명명하는 작업의 타당성을 검증하고자 하였다.

#### 방법

**평정자 및 대상자료** 연세대학교에 재학 중인 학부

생 192명이 평정자로 조사에 참가하였다. 평정 대상 자료는 정찬섭 등의 DB에 수록되어 있는 480×480 화소(pixel) 크기 및 24비트 컬러 비트맵 형식의 디지털 표정 영상 498장이었다. 이들은 연기자별로 구분되어 83장씩의 표정 사진을 포함하는 별도의 묶음으로 구성되었으며 각 묶음 내에서의 순서는 무선회되었다.

**평정절차** 평정자는 개별적으로 모니터에 제시되는 사진을 보고 마우스로 응답하도록 하였다. 평정자가 시작 버튼을 누르고 평정이 시작되면 17인치 모니터의 왼쪽에 평정 대상 사진이 제시되었다. 평정자는 제시된 표정으로부터 추론되는 내적감정상태를 먼저 쾌-불쾌, 각성-수면의 두 차원에 대한 9점 척도에 평정하였다. 쾌와 각성의 두 차원이 구성하는 좌표 평면(9×9) 위에 척도에 평정한 위치가 점으로 표시되었다. 정찬섭 등의 DB 내 감정 모형에 근거하여 83개의 감정 범주 목록 중 지정된 점에서 가장 가까운 내적상태 범주 다섯 개가 평정자에게 제시되었다. 평정자들은 제시된 형용사 다섯 개 중 표정으로부터 추론되는 가장 적합한 감정상태 범주를 선택하거나 적절한 범주가 없는 경우에는 마우스로 점의 위치를 변경하면서 각 점에서부터 가장 가까운 형용사 다섯 개를 지속적으로 탐색하여 가장 적합한 형용사를 찾아내도록 하였다. 내적감정상태의 평정 결과는 평정자가 선택한 형용사 범주의 쾌-불쾌 및 각성-수면 차원값으로 환산되어 평정자들의 총 평균 평정값으로 계산되었다. 배우에게 지시한 내적감정상태 범주의 차원값과 그로부터 표현한 표정에 대한 평균 평정 차원값간의 상관 구조 및 거리를 분석하였다.

#### 결과 및 논의

배우 여섯 명 각각이 포즈한 83장의 표정 자료에 대하여 배우에게 제시된 내적상태 범주의 차원값과 포즈된 표정 자료들이 평정된 평균 차원값과의 상관은 <표 1>과 같다. 상관 조사 결과 쾌-불쾌 차원에서는 각 배우별로 두 값이 .847에서 .947의 높은 상관 관계에 있음이 밝혀졌으며 각성-수면 차원에서도 두 값은 평균 .701의 상관을 가지는 것으로 조사되었다. 남자 배우들(A, B, C)의 평균 상관값과 여자 배우들(D, E, F)의 평균 상관값을 비교하면 여자 배우들의 표정에서 표현하고자 의도한 내적감정상태와 평정된 내적감정상태 사이의 상관이 다소 높음을 알 수 있다.

위에서 제시한 상관 구조 분석 결과는 수집된 두 자료의 분포 경향에 의해서 결정된다는 한계를 포함하고 있다. 이를 극복하고 지시된 감정상태와 평정된 감정상태의 관계를 직접적으로 파악할 수 있도록 둘 간의 거

<표 1> 배우별 지시된 내적상태 범주의 차원값과 평정자들의 평균 평정 차원값간의 상관

배우	평정된 내적상태						평균 (표준편차)
	A	B	C	D	E	F	
배-볼쾌 차원	0.847	0.924	0.923	0.924	0.928	0.947	0.915 (0.035)
지시된 내적상태 각성-수면 차원	0.611	0.666	0.743	0.740	0.676	0.768	0.701 (0.059)

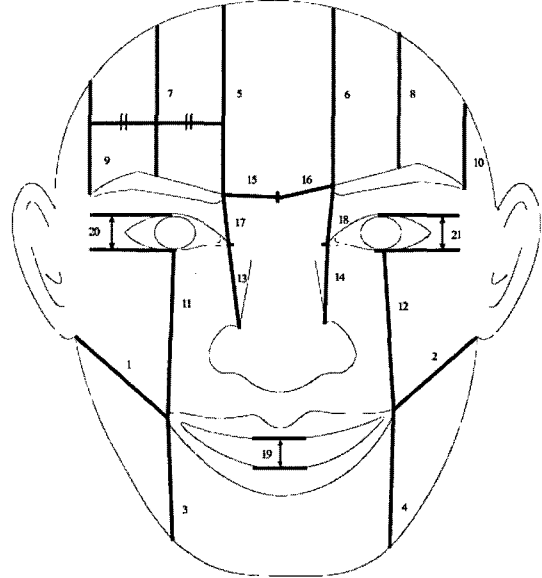
리를 조사하였다. 쾌와 각성에 대한 9점 척도의 두 차원으로 구성된 평면에서 두 값간의 평균 거리는 1.11이었다. 평정 결과로부터 배우가 포즈한 표정 자료들은 이들이 표현하고자 한 내적감정상태와 비교적 일관되게 사람들에게 인식되고 있음을 확인하였다. 이러한 분석 결과로부터 DB 내의 연극 배우 실연을 통한 표정 자료들을 배우들이 의도한 내적감정상태로 명명하는 것에 무리가 없음을 확인하였다.

## 2.2. 표정 변화를 포착하기 위한 변수 설정

표정과 감정의 관계에 대한 분석 결과가 지금까지 수행되어 온 연구 결과들과 비교될 수 있고 앞으로 연구될 응용 시스템에 반영될 수 있기 위해서는 표정의 특징적 구조를 기술하기 위한 표준이 필요하다. 이러한 표준을 마련하기 위하여 근육과 연결된 피부와 조직뿐만 아니라 표면의 주름, 선 골 등 다양한 요소들의 변화를 가장 효율적인 방법으로 반영하여 얼굴을 변형할 수 있는 변수의 확인 작업을 수행하였다. 이 과정에는 합성된 모형의 적절성, 정밀도뿐만 아니라 합성 속도 등 시스템의 수행력도 고려되었다. 특히 합성된 표정의 적절성과 합성 속도는 서로 상충적(trade-off) 관계에 있다는 점을 고려하여 이를 절충한 최적의 해결책을 찾도록 하였다.

시스템의 설계에 사용된 얼굴의 변형 변수는 21개의 변수였다. 이들 중 18개의 근육은 Parke 등이 FACS (Facial Action Coding System) [10]에 근거하여 얼굴 움직임을 반영하는 기본 근육들로 선택한 것이며 이후에 턱의 관절에 의한 입열림 정도를 추가하였다. 그런데 제시된 근육 목록에는 표정 인식 및 발현에 매우 중요한 것으로 알려진 [8] [11] 눈의 열림 정도가 제외되어 있다. 감정의 변화가 민감하게 반영되는 것으로 밝혀진 눈의 열림과 단합을 구현하기 위해 양쪽 눈의 눈꺼풀을 움직이는 2개의 근육을 변수로 추가하였다.

표정의 변화는 근육의 수축과 이완에 의해서 발현되



(그림 1) 표정 수치화를 위해 설정된 21개 변수. Parke 등이 제시한 근육 18개에 턱의 움직임에 의한 입열림 정도, 눈 주위 괄약근에 의한 눈열림 정도가 변수로 추가되었다.

지만 얼굴 표면에는 드러나지 않기 때문에 근육이 직접 관찰될 수 없다. 이러한 이유로 표정을 수치화 하는 작업은 근육을 지정하는 것만으로 완료될 수 없다. 시스템 설계의 기초 자료로서 수집할 근육의 변화 정도는 표정 영상에서 측정할 것이므로 개별 근육들의 끝점에 대응되는 영상에서의 위치를 결정할 수 있는 기준이 필요하다. 표정에 수반되는 모든 얼굴 움직임은 얼굴 표면의 표지물(landmarks) 또는 요소(organs)들 간의 거리 변화에 의해서만 정의될 수 있다[12]는 점에 착안하여 Parke 등이 제시한 근육들의 움직임이 가장 잘 반영되는 얼굴 특징 요소들이 무엇인지 확인하였다. 각 근육의 양 끝점에 대응되는 얼굴 특징 요소들의 위치를 지정하였고 이 21개 거리들을 얼굴 변수로서 규정하였다. 표정의 변화를 효과적으로 포착할 수 있으며 측정이 가능하도록 규정한 21개의 변수는 (그림 1)에 제시하였다.

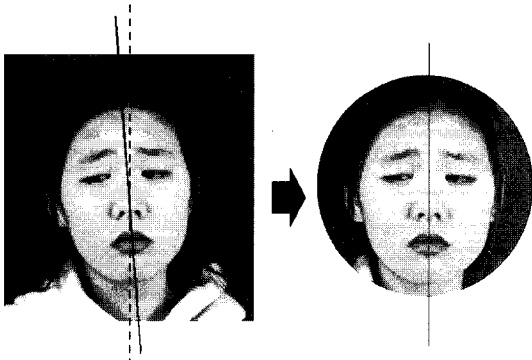
## 2.3. 얼굴표정 자료들의 변수 측정

얼굴표정의 특징적 구조를 수치화하기 위하여 수집된 표정 자료들에 대해 표정 변화가 민감하게 반영되는 변수들을 측정하였다. 측정의 정확성을 높이기 위하여 표정 자료들을 정비하였으며 측정 과정을 자동화하였다. 측정값들간의 비교가 용이하도록 각 얼굴에 대하여 표

준화하였으며 시스템에 적용하기 위해 표준화된 값들을 중립 얼굴에서의 측정값에 대한 비율로 환산하고 이들의 평균값을 계산하였다.

**2.3.1. 각도보정 및 잘라내기(crop)**

변수 측정 작업에 정확성을 기하기 위하여 표정 자료의 기울어진 각도를 보정하였다. 각 사진에서 코의 가장 높은 부분을 기준으로 수직선을 내리고 이 직선과 이마의 중간점과 턱의 중간점을 연결하는 직선의 기울기를 비교하였다. 두 직선이 일치하지 않는 경우에는 턱과 이마의 중간점을 연결하는 직선이 수직이 되도록 사진의 각도를 보정하였다. 보정된 사진 자료들에 대해서는 얼굴 변수들의 측정이 용이하도록 지름이 400화소인 원으로 사진을 잘라내었다. (그림 2)에 각도 보정 및 잘라내기 과정을 제시하였다.



(그림 2) 얼굴 변수 측정을 위한 표정 자료의 각도 보정 및 잘라내기(crop) 과정. 이마의 중간점과 턱의 중간점을 연결하는 직선이 코의 가장 높은 점에서 내린 수직선과 일치하도록 각도를 보정한 뒤 얼굴 변수들을 정확히 측정하기 위해 얼굴 부분만을 잘라내었다.

**2.3.2. 변수 측정을 위한 점찍기(pointing) 및 표준화(normalization)**

이마의 중간점과 턱의 중간점이 수직선상에 있도록 각도가 보정된 사진들에 대하여 각 변수들의 길이를 측정하였다. 규정된 각 변수들의 시작점과 끝점을 마우스로 순서대로 클릭하면 각 점들 간의 거리가 화소값으로 환산되어 나오도록 측정 과정을 자동화하였다. 점찍기 프로그램은 총 21개의 변수 및 표준화를 위한 참조값(reference value)들을 측정하기 위해서 34개점의 위치를 지정하도록 구성되어 있었으며 이들 중 위치가 분명하지 않은 점들에 대해서는 참조선이 표시되어 작업에

용이하도록 설계하였다.

표정은 얼굴 요소들간 거리의 미세한 변화에 의해 지각될 수 있으므로 카메라로부터의 거리 차이로부터 발생할 가능성이 있는 측정값들의 오차를 최소화하기 위해서 측정치들을 표준화하였다. 표준화를 위한 참조값으로는 비교적 표정 변화에 독립적인 얼굴 고유값[13]을 지정하였다. 가로 길이는 양 눈의 안쪽 끝점 사이의 거리로, 세로 길이는 이마 끝에서 코 끝점까지의 거리로 결정하여 두 길이를 곱한 면적으로 측정치들을 나누어 표준화하였다.

**2.3.3. 변화 비율 환산 및 평균값 계산**

표정은 중립얼굴로부터의 변형에 의해 인식되므로 특정 감정상태의 표정을 정량화 하기 위해서는 중립얼굴과의 차이의 형태로 기술될 필요가 있다. 이를 위해 표정을 연기한 여섯 명 각 배우의 중립 표정 자료에 대하여 표정 자료와 동일하게 21개 변수들 및 참조값들을 측정 후 표준화하였다. 표정 자료에서 측정하여 표준화 한 값들 중 Parke 등에 의해 제시된 18개의 근육 길이값들과 양쪽 눈이 열린 정도는 그에 대응되는 중립얼굴의 값으로 나누어 비율로 환산하였다. 입을 벌린 정도는 각 얼굴의 코 끝점에서부터 이마 끝점까지를 잇는 세로 참조값에 대한 비율로 환산하였다. 총 498장의 표정 자료에 대하여 조사된 21개 변수의 변화 비율에 대해서 표정에 명명된 내적감정상태 범주를 기준으로 여섯 명 배우의 자료에 대한 평균값을 계산하였다.

**3. 표정 합성 시스템의 설계**

모든 감정상태는 연구에서 선정된 21개 근육들의 변형에 의해 표정으로 합성될 수 있다는 가정 하에 내적 감정상태의 두 차원값으로부터 각 근육의 길이를 예측할 수 있도록 21개 회귀 방정식을 구하였다. 이 21개의 회귀식을 표정 합성의 집합적 장치 모음으로 사용하여 특정 감정이 주어지면 21개의 근육 모두가 회귀식이 규정하는 바에 따라 변형된 값을 갖도록 하였으며 이것에 근거하여 중성 얼굴의 텍스처가 변형되도록 하였다. 표정 합성이 모든 내적감정상태에 걸쳐 유연하게 구현될 수 있도록 내적감정상태 차원모형을 설계하고 그 차원 공간상의 한 점을 선정하면 그에 맞추어 표정이 자동 합성되도록 하였다. 사용자의 효율성을 높이기 위하여 시스템의 인터페이스를 고안하였다.

**3.1. 얼굴 변수별 변형 규칙 결정**

얼굴표정을 구성하는 각 변수들이 내적감정상태의 변

화에 따라 체계적으로 변화하는 양상을 조사하였다. 이를 위해 각 표정 자료에 대한 21가지의 측정값들 간의 상관 구조 및 내적상태 차원값 사이의 상관 구조를 분석하였다. 내적상태의 변화에 따른 각 얼굴 변수들의 변화 경향을 자세히 파악하기 위해 내적상태 쾌-불쾌 및 각성-수면 차원값을 독립변수로 하고 얼굴의 21개 측정값 각각을 종속변수로 하는 다중 회귀 분석(Multiple Regression Analysis)을 실시하여 21개의 회귀식을 도출하였다.

### 3.1.1. 얼굴 변수 측정값과 내적감정상태 차원값의 상관 조사

얼굴표정의 21가지 특징 변수들에 대한 측정값들 간의 상호 상관 및 내적감정상태 차원값과의 상관 구조를 조사하였다. 좌우 대칭인 변수들 간의 상관은 평균 .872로 이들은 매우 높은 상관 관계를 갖고 있는 것으로 조사되었다. 인위적으로 만들어진 표정이 자연스럽게 발생한 표정보다 좌우가 비대칭적이라는 연구 결과[14]를 참조하면 정찬섭 등이 구축한 DB에 수록된 연극 배우의 표정 자료들은 자연스러운 표정에 가깝게 표현된 것임이 확인된 것으로 볼 수 있다. 눈썹 바깥쪽 끝점에서 이마 끝점까지의 길이, 눈썹 안쪽 끝점에서 미간 중점까지의 길이만이 좌우 측정값간 상관이 .711과 .527로 좌우의 변화 정도간에 상대적으로 낮은 상관 관계를 갖고 있는 것으로 나타났다.

입꼬리에서 콧끝까지의 거리와 눈 아래쪽 중간점에서 입꼬리까지의 거리, 눈 안쪽 끝점에서 콧망을 윗점까지의 거리 사이에, 그리고 눈썹 안쪽 끝점에서 이마 끝까지의 거리와 눈썹 중간점에서 이마 끝까지의 거리, 눈썹 바깥쪽 끝점에서 이마 끝까지의 거리사이에는 .7 이상의 높은 상관 관계가 있는 것으로 조사되었다. 이로부터 눈과 입 사이 뺨 주위의 변수들, 이마 주위의 변수들 사이에는 높은 상관이 있어 움직임이 서로 독립적이지 않음을 알 수 있다.

얼굴의 21개 변수들과 내적감정상태 차원값간의 상관 관계를 조사한 결과는 내적감정상태 쾌-불쾌 차원과 높은 상관을 가지는 얼굴 변수들과 각성-수면 차원과 높은 상관 관계에 있는 얼굴 변수들이 서로 차이가 있음을 보여준다(표 2 참고). 눈썹 안쪽 끝에서 이마 끝까지의 거리와 얼굴 왼쪽의 눈 안쪽 끝점에서 콧망을 윗점까지의 거리, 그리고 눈썹 정도와 입을 벌린 정도만이 내적감정상태 각성-수면 차원값과 유의미한 상관을 가지는 것으로 조사되었다. 내적감정상태 쾌-불쾌 차원상의 변화와는 눈썹 안쪽 끝에서 이마 끝까지의 거리와

눈썹의 안쪽 끝점에서 눈의 안쪽 끝점까지의 거리, 그리고 눈썹 정도의 세 변수를 제외한 모든 변수들의 변화 비율과 유의미한 상관 관계에 있었다. 이러한 상관 분석 결과는 내적감정상태 기저 차원에 근거하여 표정을 합성하려면 각 차원의 변화에 따라 변형시켜야 할 얼굴 요소에 차이가 있음을 의미한다. 내적감정상태 기저 두 차원간에는 .132의 상관 관계가 있는 것으로 조사되어 두 차원이 비교적 서로 독립적임을 확인하였다.

### 3.1.2. 얼굴 변수 측정값과 내적감정상태 차원값의 회귀 분석

내적감정상태 쾌-불쾌, 각성-수면 차원 상의 변화가 얼굴표정에 반영되는 양상을 조사하기 위하여 표정 자료로부터 수집 및 측정된 내적감정상태 차원값과 얼굴의 21개 변수에 대한 평균 변화 비율을 대상으로 회귀 분석을 실시하였다. <표 2>에 제시한 바와 같이 분석을 위한 독립변수는 내적감정상태 쾌-불쾌 및 각성-수면 차원값으로 하였으며 각 값에 해당되는 표정의 21개 얼굴 변수의 변화 비율 각각을 종속변수로 하여 총 21개의 회귀 방정식을 도출하였다.

내적감정상태 두 기저 차원상의 변화는 21개 얼굴 변수들의 변화 비율을 평균 26.3% 정도 설명하는 것으로 조사되었다. 회귀 분석 결과 쾌한 상태일수록 입꼬리가 올라가고 눈썹의 가운데와 바깥쪽이 올라가며, 각성된 상태일수록 눈썹 안쪽이 올라가고 눈이 커지며 입이 많이 벌어지는 것으로 밝혀졌다. 이와 같은 분석 결과는 영화로부터 표집된 표정 자료들을 사용하고 Pilowsky, Thornton과 Stokes[15]가 제시한 열 두 차원을 측정된 결과로부터 도출된 한재현 등의 연구 결과 및 그 이전의 발견들[16][17]과 일관된다. 얼굴 변수 중 눈썹 안쪽 끝에서 이마 끝까지의 거리, 입열림 정도 및 눈열림 정도만이 내적감정상태의 각성 차원상의 변화에 의해 민감하게 변형되는 것으로 나타났다. 회귀 방정식으로부터 조사된 내적감정상태 차원상의 변화에 따른 얼굴표정 변수들의 변형 규칙은 차원모형을 기반으로 한 얼굴표정 합성 시스템의 기초 자료로서 사용하였다.

## 3.2. 표정 구현을 위한 모형 선정 및 시스템 개발

Parke 등이 개발한 얼굴모형은 사람 얼굴의 근육과 조직의 복잡한 구조를 모사하여 개발되었을 뿐만 아니라 근육 움직임에 의해 표정을 표현할 수 있으며 각 근육의 움직임에 영향을 받는 영역이 설정되어 있어 얼굴 움직임을 구현할 모형으로서 적합하였다. 내적감정상태의 쾌-불쾌 및 각성-수면 차원값을 지정하면 이 모형

<표 2> 내적상태 두 차원 상에서의 변화가 얼굴의 21개 변수의 변형 비율을 예언하는 정도에 대한 다중회귀분석 결과

얼굴변수	adjusted R <sup>2</sup>	내적상태	Beta	t	p
1	.682	패	-.830	-13.305	.000
<i>F</i> (2,81)=90.105, <i>p</i> <.001		각성	-.001	-.015	.988
2	.621	패	-.782	-11.468	.000
<i>F</i> (2,81)=68.956, <i>p</i> <.001		각성	-.068	-.996	.322
3	.244	패	.500	5.195	.000
<i>F</i> (2,81)=14.421, <i>p</i> <.001		각성	.064	.664	.509
4	.263	패	.510	5.368	.000
<i>F</i> (2,81)=15.843, <i>p</i> <.001		각성	.093	.973	.333
5	.062	패	-.152	-1.419	.160
<i>F</i> (2,81)=3.735, <i>p</i> <.05		각성	-.228	-2.128	.036
6	.059	패	-.173	-1.613	.111
<i>F</i> (2,81)=3.588, <i>p</i> <.05		각성	-.205	-1.908	.060
7	.329	패	-.576	-6.348	.000
<i>F</i> (2,81)=21.380, <i>p</i> <.001		각성	-.065	-.717	.475
8	.305	패	-.542	-5.866	.000
<i>F</i> (2,81)=19.176, <i>p</i> <.001		각성	-.110	-1.195	.236
9	.155	패	-.422	-4.147	.000
<i>F</i> (2,81)=8.599, <i>p</i> <.001		각성	.053	.518	.606
10	.160	패	-.411	-4.047	.000
<i>F</i> (2,81)=8.895, <i>p</i> <.001		각성	-.065	-.643	.522
11	.346	패	-.583	-6.513	.000
<i>F</i> (2,81)=22.998, <i>p</i> <.001		각성	-.091	-1.015	.313
12	.339	패	-.566	-6.281	.000
<i>F</i> (2,81)=22.276, <i>p</i> <.001		각성	-.127	-1.410	.162
13	.325	패	-.554	-6.095	.000
<i>F</i> (2,81)=21.020, <i>p</i> <.001		각성	-.126	-1.390	.168
14	.397	패	-.582	-6.772	.000
<i>F</i> (2,81)=28.324, <i>p</i> <.001		각성	-.203	-2.363	.021
15	.430	패	.660	7.890	.000
<i>F</i> (2,81)=32.309, <i>p</i> <.001		각성	-.214	-2.563	.012
16	.436	패	.676	8.128	.000
<i>F</i> (2,81)=33.044, <i>p</i> <.001		각성	-.103	-1.237	.220
17	.012	패	.183	1.665	.100
<i>F</i> (2,81)=1.488, <i>ns</i>		각성	-.074	-.668	.506
18	.026	패	.213	1.951	.054
<i>F</i> (2,81)=2.098, <i>ns</i>		각성	-.096	-.875	.384
19	.287	패	.271	2.901	.005
<i>F</i> (2,81)=17.681, <i>p</i> <.001		각성	.445	4.763	.000
20	.018	패	-.022	-.198	.844
<i>F</i> (2,81)=1.749, <i>ns</i>		각성	.205	1.870	.065
21	.031	패	-.023	-.214	.831
<i>F</i> (2,81)=2.337, <i>ns</i>		각성	.236	2.161	.034
평균 (표준편차)	.263 (.192)				

의 21개 변수가 도출된 회귀방정식에 근거하여 변형된 값을 가짐으로써 표정을 합성해내도록 하여 시스템을 설계하였다.

### 3.2.1. Parke 등의 얼굴모형 변형 코드의 응용

자연스러운 얼굴 움직임의 구현을 위하여 Parke 등이 개발한 코드는 얼굴모형을 만들고 근육에 기반하여 모형의 기하학적 구조를 조정하여 그 결과를 화면에 보여 줄 수 있도록 구성되어 있다. 이 프로그램에서는 사람의 해부학적 구조 및 FACS에 근거하여 19개의 얼굴 근육이 정의되고 각 근육의 길이가 변화할 때 영향을 받는 영역이 설정되어 있다. 이 모형에서 정의된 근육들을 점진적으로 늘이거나 줄이기 위해 키보드로 입력을 받도록 설계되어 있었는데 이를 기초로 프로그램의 일부를 수정, 보완함으로써 표정 합성 시스템으로 개발하였다.

Parke 등의 코드는 키보드의 c/C키의 입력에 의해 각 근육이 줄어들거나 늘어나도록 설계되어 있다. 프로그램 내부에서는 c/C 키가 입력된 횟수를 저장하여 그에 대응되는 각 근육의 수축값(contraction value :  $\pm 0.1$ )을 결정하고 그로부터 근육과 연결된 표면이 변화되도록 하는 구조를 갖고 있다. 그러나 이 때 주어지는 수축값이 원래 근육의 길이와 어떠한 관계를 가지는가에 대해서는 명백한 규정이 없다.

중립 표정일 때의 각 변수들의 길이에 대한 변화 비율을 기준으로 모형을 변형하기 위해 프로그램의 일부를 수정하였다. 근육의 수축값과 근육 길이의 관계를 밝혀내기 위해 모형에서 21개의 변수의 양 끝점(head, tail)에 해당하는 꼭지점들을 찾아내고 근육의 수축값에 따른 이 점들의 변화를 추적하여 길이의 변화로 환산하였다. 수축값의 증가 또는 감소에 따른 양 끝점간의 2차원 거리 및 3차원거리를 계산하여 비교하였다. 그 결과 두 값의 차이가 아주 작고 입력될 측정값들이 2차원을 기반으로 한다는 점을 고려하여 2차원 상의 거리를 변형 기준으로 사용하였다. 얼굴 변형을 위해 입력되는 21개 얼굴 변수의 변화 비율로부터 근육의 수축값을 역으로 찾아내어 변수들에 해당하는 꼭지점들의 위치를 변화시키도록 프로그램을 조정하였다.

### 3.2.2. 텍스처(texture) 제작

모형의 변형에 의해 나타나는 변화를 효과적으로 관찰하기 위해 모형에 씌울 표면의 텍스처를 구성하였다. 사람 같은 느낌이 나도록 눈썹의 솟과 입술의 질감을 만들어냈으며 프로그램 내의 조명 효과만 나타나도록 피부 표면은 균일하게(homogeneous) 제작하였다.

### 3.2.3. 합성 시스템 및 인터페이스 개발

내적감정상태의 쾌-불쾌 및 각성-수면 차원값을 입력하면 회귀 방정식에 근거하여 얼굴 21개의 변수들의 변화 정도가 결정되고 이 값들을 기초로 표정이 합성되도록 시스템을 구성하였다. 시스템을 정교하게 만들기 위하여 회귀 방정식의 일부를 조정하였다. 쾌-불쾌 차원과 각성-수면 차원이 만나는 증점은 중립얼굴이 되도록 조율하였다. 또한 모형의 각 근육들은 변형될 수 있는 최소값과 최대값이 결정되어 있어 이를 넘을 경우 텍스처가 깨지거나 구동 자체가 불가능하였으므로 근육들의 한계점을 계산하여 식을 일부 조정한 뒤 시스템에 적용하였다. 최종적으로 결정된 회귀식을 시스템에 대입한 후에 사용자가 쉽고 편리하게 이용할 수 있도록 슬라이드 바와 좌표평면 등 두 가지의 인터페이스를 설계하였다.

슬라이드 바에 의한 내적상태 차원값의 입력 화면은 그림 3에 제시하였다. 쾌-불쾌와 각성-수면의 두 차원값을 슬라이드 바를 움직여가며 조정할 수 있도록 구성하였으며 숫자로도 제시되어 확인할 수 있도록 하였다. 슬라이드 바의 양쪽 끝에는 화살표를 만들어 1에서 9까지의 차원값들을 0.01 단위로 변화시킬 수 있도록 제작하였다. 주어진 차원값에 의해 결정된 표정은 실시간으로 구현되었다.

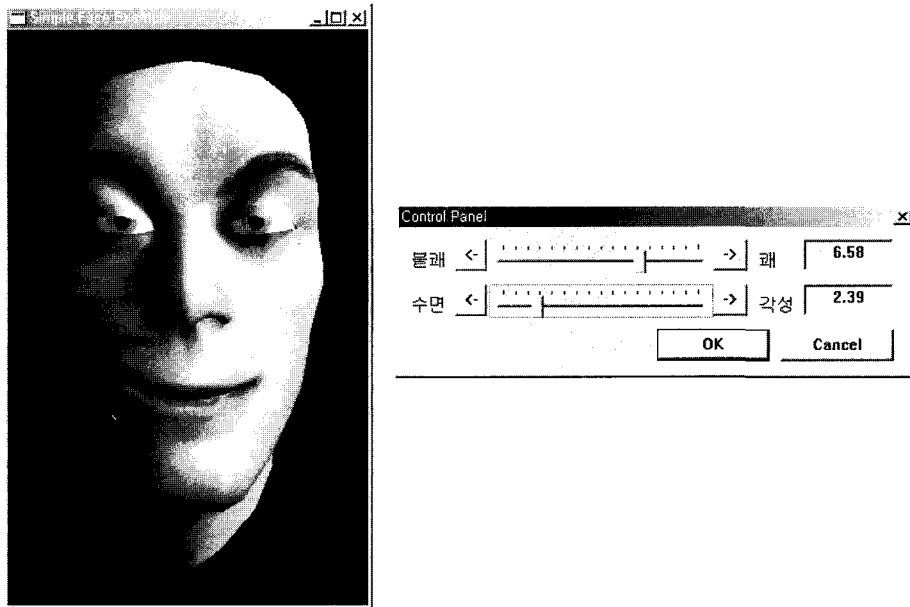
슬라이드 바에 의한 내적감정상태 차원값의 입력 방식은 감정상태 공간 내에서의 위치를 파악하기에 용이하지 않다. 이를 보완하기 위해 그림 4와 같이 쾌-불쾌 및 각성-수면 차원으로 구성된 좌표 평면 상에서 특정 위치를 지정하면 그 값에 근거하여 표정이 합성될 수 있는 인터페이스도 고안하였다. 정밀한 탐색이 가능하도록 필요한 경우 사용자는 격자(grid)를 볼 수 있도록 설계하였으며 평면 내의 어느 곳이든 마우스로 클릭하면 그 위치에 해당되는 표정이 합성되도록 구성하였다. 좌표 평면을 통한 표정의 합성은 실시간으로 진행되어 사용자는 변형의 효과를 즉각적으로 확인할 수 있었다.

## 4. 종합논의

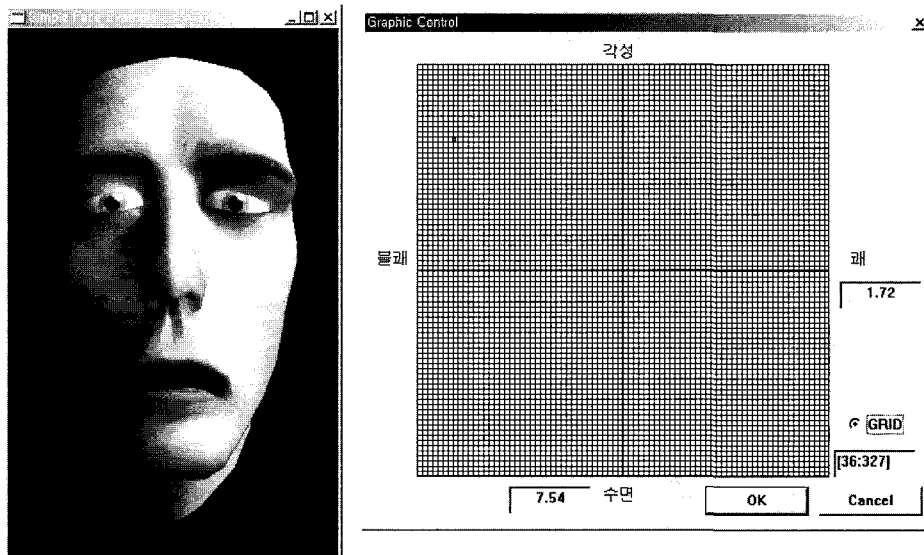
본 연구에서는 경험적인 표정 자료를 바탕으로 감정과 표정 사이의 관계 구조를 분석하였다. 이 조사 결과를 Parke 등의 근육 기반 얼굴모형에 적용하여 쾌-불쾌 및 각성-수면의 2차원 평면에 임의의 값을 지정하면 그 점에 해당하는 표정을 자동으로 합성할 수 있는 시스템을 개발하였다.

표정 합성 시스템의 기저 자료로서 감정과 표정간의





(그림 3) 자동 표정 합성 시스템에서 슬라이드 바에 의한 인터페이스. 시스템의 사용자는 볼래 및 각성 차원의 슬라이드 바를 움직이며 각 차원상의 변화가 표정에 반영되는 양상을 관찰할 수 있도록 구성하였다.



(그림 4) 자동 표정 합성 시스템의 좌표평면에 의한 인터페이스. 시스템 사용자가 볼래 및 각성 차원으로 구성되는 2차원의 좌표 평면에서 특정 위치를 지정하면 그에 대응되는 표정이 합성되어 관찰 가능하도록 구성하였다.

관계 모형은 표정 자료들의 내적감정상태 쾌-불쾌 및 각성-수면 차원의 평정값을 독립변수로 하고 21개의 각 얼굴 변수 측정값들을 종속변인으로 하는 회귀분석의 결과로부터 수립되었다. 이를 통해 내적 감정의 쾌한 상태로의 변화는 입꼬리가 올라가고 눈썹이 올라가는 것으로, 각성된 상태로의 변화는 입을 벌리고 눈을 크게 뜨는 것으로 얼굴표정에 반영됨이 확인되었다. 이 결과는 내적감정상태 차원모형을 가정하고 감정과 표정의 관계 구조를 밝히고자 했던 이전의 연구 결과들과 비교될 수 있다.

Snodgrass[18]는 FACS에서 규정하고 있는 얼굴 움직임 단위(Action Unit: AU)에 의해 전달되는 내적감정상태를 분석하고자 하였다. Snodgrass는 Ekman이 사람들에게 특정한 얼굴 움직임을 표현하도록 하고 그 장면을 촬영한 동영상 및 정지 영상을 분석 대상 자료로서 사용하였다. 이 표정 자료들은 일상 생활에서 감정이 표현되는 표정과 유사할 것으로 기대하기는 어렵지만 각 얼굴 움직임들을 정확히 표현하고 있다는 장점이 있다. 평정자들이 대상 표정 자료들을 내적감정상태 두 차원에 대하여 평정한 결과를 근거로 입꼬리를 당기거나 내리는 것(AU 12, AU15)이 쾌-불쾌 차원 평정에, 눈썹이나 눈꺼풀을 올리거나 내리는 것, 턱을 내리는 것, 그리고 눈을 가늘게 뜨거나 감는 것(AU1, AU2, AU5, AU26, AU41, AU42, AU43)이 각성-수면 차원 평정에 결정적인 요소임이 확인되었다.

감정과 표정의 연결 관계 모형을 조사하기 위한 한재현 등의 연구에서는 영화와 TV드라마로부터 표집된 자연스러운 표정들을 연구 대상 자료로 이용하였다. 표정 자료들에 대하여 Pilowsky 등이 제시한 열 두 얼굴 차원값을 측정하고 내적 감정의 기저 차원값을 조사하였다. 이들에 대하여 다중 회귀 분석 및 정준 상관 분석을 실시한 결과 내적감정상태 쾌-불쾌 차원상의 변화는 입의 가로의 길이에 의해, 각성-수면 차원상의 변화는 눈을 크게 뜨는 정도와 입을 많이 벌리는 정도에 의해 민감하게 반영됨을 확인하였다.

다양한 방법으로 수집된 표정 자료에 대하여 서로 다른 얼굴 변수들을 측정하여 표정과 감정의 관계 구조를 조사한 연구들로부터 이처럼 일관된 결과들이 조사되었다. 이와 같은 결과는 내적감정상태 차원모형에 근거하여 감정 정보를 수집하고 Parke 등의 얼굴 변수로 표정의 특징들을 수치화 한 뒤 이 두 값에 대해 회귀분석을 실시하여 얻은 둘 간의 관계 모형 및 이를 기반으로 한 합성 시스템의 개발 과정이 타당함을 확인해준다.

개발된 합성 시스템이 더 '그럴듯한' 표정을 만들어

내기 위해서 반드시 보완되어야 할 몇 가지 사항들이 있다. 시스템 내부의 기저 알고리즘의 기반으로 사용된 회귀방정식의 설명력을 향상시킬 필요가 있다. 도출된 회귀식에 의하면 내적감정상태 두 기저 차원 상의 변화는 각 근육 길이의 변량을 평균 26.3% 정도 설명하는 것으로 조사되었다. 이러한 수치는 내적감정상태의 변화와 이에 따른 얼굴표정 변형간의 비선형적 관계 구조에 대해서도 조사할 필요가 있음을 시사한다. 외부적으로는 시스템의 두상 모형의 위상이 외국인을 근거로 하여 제작되어 근본적으로 서구인의 얼굴 모습에 가까울 뿐만 아니라 인종의 길이가 지나치게 길어 여색하게 보였다. 모형의 조정을 위해서는 다양하게 개발되고 있는 모델러 등의 사용을 검토해 볼 필요가 있다. 표정 합성 시스템은 궁극적으로는 임의의 얼굴 영상을 적용하여 표정을 구현할 수 있어야 하므로 다양한 텍스처를 정합하여 표정을 구현할 수 있는 기술이 보완되어야 한다. 이와 더불어 포함하는 구강 구조와 눈동자의 움직임 등 세부 요소가 모형에 추가되면 훨씬 더 사람 같은 자연스러운 표정 구현에 접근할 수 있을 것이다.

## 참고문헌

- [1] Parke, F. I., & Waters, K. (1996). *Computer facial animation*, A K Peters, Wellesley.
- [2] 정찬섭, 오경자, 이일병, 변혜란. (1998). **표정/제스처에 의한 감정측정기술 및 DB개발**. 과학기술부 감성공학기술개발사업 최종보고서, G-17-01-09.
- [3] Izard, C. E. (1979). *The maximally Discriminative Facial Movement Coding System (Max)*, Instructional Resources Center, University of Dalaware, Newark.
- [4] Ekman, P., & Friesen, W. V. (1982). *Rationale and reliability for EMFACS Coders*, Unpublished.
- [5] Izard, C. E., Dougherty, L. M., & Hembree, E. A. (1983). *A system for Identifying Affect Expressions by Holistic Judgements(Affex)*, Instructional Resources Center, University of Dalaware, Newark.
- [6] Rosenberg, E. L. (1997). The study of spontaneous facial expressions in psychology. In P. Ekman & E. Rosenberg (Eds.), *What the face reveals*, Oxford University Press, New York.
- [7] Russell, J. A., & Bullock, M. (1986). Fuzzy concepts and the perception of emotion in facial expressions. *Social Cognition*, 4, 309-341.
- [8] 한재현, 정찬섭. (1998). 얼굴표정에 의한 내적상태

- 추정. *한국감성과학회지*, 1(1), 41-58.
- [9] Waters, K., & Terzopoulos, D. (1992). The computer synthesis of expressive faces. In V. Bruce, A. Cowey, A. W. Ellis & D. I. Perrett (Eds.), *Processing the Facial Image*, Oxford University Press, Oxford.
- [10] Ekman, P., & Friesen, W. V. (1978). *Facial Action Coding System(FACS)*, Consulting Psychologist Press, Palo Alto.
- [11] Smith, C. A., & Scott, H. S. (1997). A componential approach to the meaning of facial expressions. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression*, Cambridge university press, New York.
- [12] Yin, L., & Basu, A. (1998). Analysis and synthesis of facial expressions for MPEG-4 system. *Proceedings of IEEE International Conference on Systems, Man and Cybernetics*, 5, 4608-4613.
- [13] Tatsuno, Y., Suzuki, S., Yokoya, N., Iwasa, H. & Takemura, H. (1996). Analysis and synthesis of six primary facial expressions using range images. *Proceedings of the 13th International Conference on Pattern Recognition*, 3, 489-493.
- [14] Skinner, M., & Mullen, B. (1991). Facial asymmetry in emotional expressions: A meta-analysis of research. *British Journal of Social Psychology*, 30, 113-124.
- [15] Pilowsky, I., Thornton, M. & Stokes, B. B. (1986). Towards the quantification of facial expressions with the use of a mathematic model of the face. In H. D. Ellis, M. A. Jeeves FRSE, F. Newcombe & A. Young (Eds.), *Aspects of face processing*, Martinus Nijhoff Publishers, Dordrecht.
- [16] Boucher, J., & Ekman, P. (1975). Facial areas and emotional information. *Journal of Communication*, 25, 21-29.
- [17] Fraser, I. H., Craig, G. L. & Parker, D. M. (1990). Reaction time measures of feature saliency in schematic faces. *Perception*, 19, 661-673.
- [18] Snodgrass, J. (1992). *Judgment of feeling states from facial behavior: A bottom-up approach*. Unpublished doctoral dissertation, University of British Columbia.