

쾌 및 각성 차원 기반 표정 합성 시스템의 성능 검증 및 인터페이스의 효율성 비교*

Performance tests for the expression synthesis system
based on pleasure and arousal dimensions and efficiency
comparisons for its interfaces

한재현** 정찬섭**
(Jaehyun Han) (Chan-Sup Chung)

요약 내적상태 쾌 및 각성 차원에 근거한 표정 합성 시스템의 성능을 검증하고 이 시스템을 효과적으로 사용할 수 있는 인터페이스를 제안하였다. 우선 내적상태의 기저 구조에 대하여 차원모형을 가장한 것이 적절한가를 확인하고자 하였다. 이를 위해 평정자들로 하여금 차원 축 상의 17가지 표정들을 비교하도록 하였다. 참가자들의 비교 판단 결과로부터 시스템의 기본 가정에 대한 타당성이 확보되었다. 다음으로 시스템의 성능을 검증하기 위해 대표 표정 21가지를 선정하고 평정자들로 하여금 이들의 유사성을 판단하도록 하였다. 평균 유사성 평정값을 대상으로 다차원분석을 실시한 결과, 시스템의 성능이 신뢰로운 수준임을 확인하였다. 마지막으로 시스템 사용을 위한 최적의 조건을 찾기 위해 좌표평면 및 슬라이드바의 두 가지 방법으로 설계된 인터페이스들의 효율성을 비교하였다. 사용자들로 하여금 각각의 방식으로 25가지 특정 내적상태의 표정들을 직접 구현해보도록 하였다. 이들의 구현 결과들을 비교, 분석한 결과 내적상태 차원에 기반하여 표정을 합성하는 경우 좌표평면에서의 위치를 제시하는 입력 방법이 더 안정적인 것으로 확인되었다.

주제어 얼굴표정, 합성 시스템, 차원모형, 인터페이스, 효율성

Abstract We tested the capability of the pleasure and arousal dimension-based facial expression synthesis system and proposed the most effective interface for it. First, we tried to confirm the adequateness of the dimensional model as a basic structure of the internal states for the system. For it, subjects compared the 17 facial expressions on the two axes. The results validated the fundamental hypothesis of the system. Second, we chose 21 representative expressions from the system to test its performance and had subjects rate their similarities. We analyzed these data using multidimensional scaling methods and these results verified the system's reliability. Third, we compared the efficiencies of two interfaces -coordinate values and slide bars- to find the most suitable interface for the system. Subjects synthesize 25 facial expressions with each interface of it. The results showed that the visualization of two dimensional values into Cartesian coordinate is more stable as an input display of facial expression synthesis system based on dimensions.

Keywords facial expression, synthesis system, dimensional model, interface, efficiency

1. 서론

최근 영상 산업 분야에서 사이버캐릭터, 아바타 등의

사용이 급격히 증가함에 따라 얼굴표정 합성을 위한 다양한 측면에서의 접근이 시도되고 있다. 그러나 '사람 같은' 표정들을 표현하고자 하는 대부분의 표정 합성 방법들은 체계적인 DB나 시스템을 기반으로 하기보다는 실무자의 경험적 직관에만 의존한다는 근본적인 한계를 가지고 있다. 자연스러운 표정을 표현하기 위해서 수많은 시행착오를 반복하는 현재의 작업 과정은 시간적, 경

* 이 논문은 2002년 한국학술진흥재단의 지원에 의하여 연구되었음. (KRF-2002-074-HS1003)

** 연세대학교 인지과학 연구소

e-mail : jaehyun@psylab.yonsei.ac.kr, cschung@yonsei.ac.kr

제적 효율성을 떨어뜨리며 이는 경쟁력과 연결되기 때문에 문제의 해결 방안이 요구된다.

내적상태 정보로부터 자동으로 표정을 만들어내는 기술은 표정 합성 작업의 효율성 제고를 위한 대안이 될 수 있다. 이러한 이유로 자동 표정 합성 시스템을 개발하고자 하는 다양한 시도들이 계속되어 왔다. 지금까지 개발된 자동 표정 합성 기술들은 대부분 합성할 얼굴영상을 3차원 모형에 정합하고 모형을 변형시키는 방안을 채택하였는데 이들의 기저 알고리즘은 입력 영상을 정합한 모형과 그 변형 규칙에 의해 구분될 수 있다.

얼굴 움직임의 기술 체계인 FACS(Facial Action Coding System) [1]에 근거하여 표정을 합성하는 방법은 특정 정서의 얼굴표정에서 특징점들의 이동 거리를 계산하고 이를 바탕으로 임의의 얼굴에 표정을 표현해냄으로써 세밀한 표정 구현보다는 실시간 처리에 초점을 둔 응용시스템에 주로 사용되었다[2]. Yang, Kunihiro, Shimoda와 Yoshikawa [3]가 개발한 표정 합성 시스템도 FACS를 기반으로 설계되었다. 이 시스템은 복합적인 정서를 기본 얼굴 움직임들의 조합으로, 이 조합을 AU(Action Unit)들로, 그리고 AU들을 근육 수축을 통한 얼굴 표면의 변형으로 변환하는 세 가지의 하위 체계로 구성되었으며 이 과정을 거쳐 6가지의 기본 정서를 포함한 총 30개 이상의 내적상태를 표정으로 합성해내었다.

실제 얼굴표정은 피부 아래의 얼굴 근육들의 움직임에 의해서 발현된다는 사실에 기초하여 근육을 기반으로 표정을 합성하려는 접근들도 시도되어 왔다[4][5]. 근육 움직임에 의해 표정을 합성하는 방법들은 얼굴 피부와 근육을 용수철처럼 표현하고자 하였는데 이 방법은 지나치게 계산이 복잡하다는 단점이 있다. 그 중 한 예로 Ishikawa, Sera, Morishima와 Terzopoulos[6]는 사람의 피부와 근육 구조에 근거하여 얼굴 모형을 만들고 이마 부분과 입 주위에 각각 12개와 27개의 근육을 포함시켜 얼굴표정의 합성에 사용하였다. 이들은 표정을 짓고 있는 얼굴의 특징점에 부착된 마커(marker)로부터 2차원상의 움직임을 추출한 후 이 값을 입력값으로 하는 신경망에 의해 근육 움직임을 출력해내도록 시스템을 설계하였다. 이 시스템은 2차원 얼굴 움직임을 3차원으로 변환하여 자동적으로 표정을 합성하는데 성공적이었지만 마커를 붙인 얼굴 형태와 마커가 부착된 위치에 지나치게 의존적이라는 제한점이 있었다. 뿐만 아니라 근육에 기반한 다른 합성 방법들과 마찬가지로 합성에 요구되는 계산 시간이 길어 실시간 처리가 불가능하였다.

Aoki 등[5]은 근육에 기초한 대부분의 표정 합성 방법

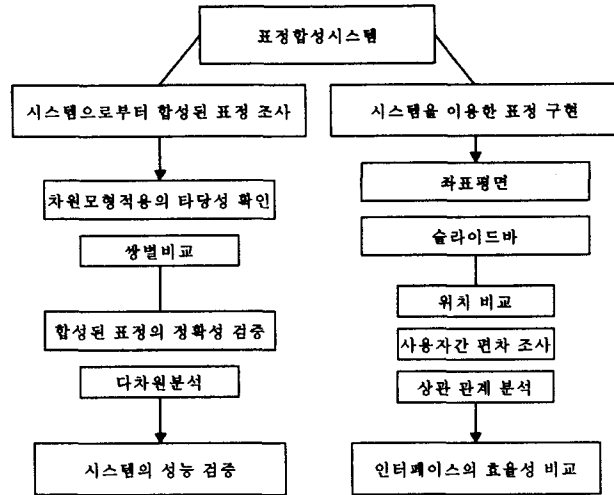
들이 사람 두상(skull)의 정확한 형태를 고려하지 않은 모형들을 사용하고 있기 때문에 자연스러운 얼굴표정을 나타내지 못한다고 지적하였다. 이와 같은 근본적인 문제점을 해결하기 위해서 이들은 컴퓨터 단층촬영으로 주사된(CT scan) 자료를 바탕으로 모형을 제작하고 두상과 근육 등의 내부 조직과 얼굴 표면의 변화와의 관계를 양적으로 분석한 후에 이를 근거로 '웃는', '화난', '슬픈', '놀란', '눈을 감는' 등의 얼굴표정을 합성하였다.

지금까지 연구된 표정 합성 시도들은 기저 모형 및 변형 알고리즘에 서로 차이점이 있지만 궁극적으로 자연스러운 표정을 만들어내야 한다는 공통적인 목적을 갖고 있었다. 그러나 개발된 대부분의 시스템은 인간의 정서 구조 및 정서와 얼굴표정의 연결 관계를 고려하지 않고 임의로 결정한 몇 가지 표정과 얼굴 움직임(예. 웃음, 눈을 감음)을 만들어내는데 그치고 있어 표정의 합성 결과라기보다는 단순한 근육 움직임의 표현 단계에 머물러 있다. 뿐만 아니라 사람들이 합성된 결과인 표정들로부터 어떠한 정서 상태를 추정하는가에 대한 확인 절차가 전혀 없어 시스템의 성능 평가 및 비교가 불가능한 상태이다.

얼굴표정은 서로 다른 문화권에 있는 사람들에게도 일반적으로 인식되는 횡문화적 보편성을 가진다는 연구 결과[7]나 다른 사람의 표정을 언뜻 보고도 얼굴표정이 의미하는 바를 쉽게 알 수 있다는 연구 결과[8]들은 특정 내적상태가 얼굴표정으로 표현되는 경우에 사람들이 의식하지 못하더라도 일정한 규칙에 근거한 움직임이 있다는 사실을 암시한다. 한재현과 정찬섭[9]의 표정 합성 시스템은 이와 같은 내적상태가 표정으로 표현될 때의 규칙과 표현된 표정으로부터 특정 내적상태를 추론하는 기준을 기초로 개발되었다. 이 시스템은 사람들의 내적상태는 쾌-불쾌 및 각성-수면의 두 기저 차원에 의해 표상될 수 있다는 가정 하에 다량의 경험적인 자료를 대상으로 내적상태와 얼굴표정의 관계 구조를 분석한 결과를 기반으로 했다는 점에서 이전의 시스템들과 차별화된다. 본 연구는 이미 개발된 자동 표정 합성 시스템의 이같은 특징적 기본 가정과 개발 과정의 타당성을 확인하고 최적의 인터페이스를 제안하기 위해서 설계되었다. (그림 1)에 제시한 연구 과정을 거쳐 시스템을 정교화함으로써 실제 작업에 활용될 수 있는 가능성을 타진해 볼 수 있다.

2. 표정 인식을 통한 합성 시스템의 성능 검증

자동 시스템의 정확성을 측정하기 위해서는 시스템의



(그림 1) 전반적 연구 설계

결과에 대하여 객관적인 평가가 수행되어야 한다. 표정 합성을 위한 자동 시스템의 경우 시스템의 사용자는 사람의 감정상태를 입력하게 되며 이로부터 도출되는 결과물은 표정이므로 합성된 표정이 사람들에게 어떻게 인식되는가를 조사함으로써 시스템의 성능을 검증해 볼 수 있다.

2.1 시스템 내 차원모형 적용의 타당성 확인

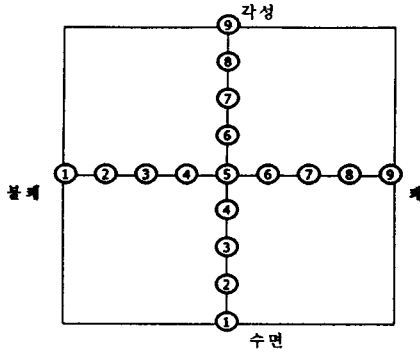
기존에 개발된 표정 합성 시스템들은 내적상태 범주를 선택하면 그에 대응되는 표정이 합성되는 구조를 갖고 있었다[3][5]. 그러나 검증 대상 표정 합성 시스템은 배우가 지은 표정에 대한 얼굴 특징변수의 측정값과 평정 자료를 기반으로 설계되었으므로 사용자의 입력값인 내적상태 정보가 쾌-불쾌 및 각성-수면 차원값으로 지정된다는 측면에서 이들과 구별된다. 시스템 내 내적상태 기저 차원 축 위에 구현된 표정들의 상대적인 쾌-불쾌 및 각성-수면 정도를 비교하여 평정하도록 하였다. 이 평정 결과에 기초하여 연속적으로 변화하는 얼굴의 각 특징변수들이 내적상태 차원상의 변화를 적절하게 반영하고 있는가를 확인하고 이로부터 사람의 내적 정서는 쾌-불쾌 및 각성-수면의 두 차원이 구성하는 공간에 의해 표상될 수 있다는 시스템 설계의 기본 가정이 타당한가를 확인하고자 하였다.

2.1.1 방법

참가자. 연세대학교 학부생, 대학원 심리학과 및 인지과학 협동과정에 재학중인 대학원생 30명이 평정자로 서 참가하였다. 이들은 개별적으로 평정을 수행하였다.

자료. 평정 대상 자료로서 검증 대상 표정 합성 시스템을 이용하여 쾌-불쾌 및 각성-수면 차원 위에 위치하는 표정들을 구현하였다. 쾌-불쾌 차원상의 표정은 각성-수면 차원값이 중간일 때 불쾌의 극단에서 쾌의 극단까지 균일한 간격으로 아홉 가지의 표정이 합성되었으며 각성-수면 차원상의 표정은 쾌-불쾌 차원값이 중간일 때 수면의 극단에서 각성의 극단까지 아홉 가지의 표정이 구현되었다. 중립 표정은 중복되었으므로 총 17개의 표정들이 시스템 성능 검증을 위한 평정 대상 자료로서 도출되었다. (그림 2)에 평정 대상 표정들의 좌표 평면에서의 위치를 도식적으로 표현하였다. 쾌-불쾌 차원 위에 위치한 아홉 가지의 표정으로부터 서로 다른 두 가지를 추출한 표정 쌍(pair) 36가지와 각성-수면 축 위의 표정 아홉 가지로부터 추출한 36쌍을 별도의 평정 단위(set)로 구성하였다.

절차. 평정자는 평정 요령에 대해 간단한 설명을 들은 뒤 의문 사항에 대해 질문하도록 하였다. 평정 방법을 숙지한 후에 시작 버튼을 눌러 평정이 시작되도록 하였다. 평정이 시작되면 평정자는 두 묶음의 평정 자료 중 먼저 쾌-불쾌 차원의 표정 쌍들에 대해 평정한 후에 각성-수면 차원의 표정 쌍에 대하여 평정하였다. 이 때 두 단위 내의 표정 쌍들은 무선회되어 제시되었다. 표정들은 200×333 화소의 크기로 좌우로 나란히 17인치 모니터에 제시되었으며 제시된 한 쌍의 표정을 비교하여(paired comparison) 상대적으로 더 쾌하거나 더 각성된 상태인 것을 선택하도록 하였다. 평정의 제한 시간은 없었으며 두 표정을 비교한 후 하나를 선택하면 다음 표정 쌍이 제시되도록 하였다.



(그림 2) 시스템 내 차원모형 적용의 타당성 검증을 위해 합성한 표정들의 위치 도식

2.1.2 결과 및 논의

쾌-불쾌 및 각성-수면 차원 축 위의 이홉 개 표정을 조합하여 만든 두 묶음의 표정 36쌍에 대하여 각각 평정자 30명이 상대적으로 더 쾌한 상태와 더 각성된 상태를 판단한 결과를 <표 1>에 제시하였다. <표 1>에서 볼 수 있듯이 쾌-불쾌 차원 상에서 불쾌의 극단(표정 1과 2) 또는 쾌의 극단(표정 7, 8과 9)에 위치하면서 인접한 두 표정간의 비교에 오판단(misjudgment)이 있었을 뿐 대부분 경우에 30명의 평정자 전원이 표정 합성

단계에서 의도된 대로 상대적인 쾌 정도를 평정하였다. 쾌-불쾌 차원의 양극단에서 오판단이 빈번하게 나타난 결과는 극단에 위치한 표정들이 쾌 또는 불쾌의 정서 상태를 극명하게 나타내고 있으므로 상대적인 쾌 정도를 판단하는 것이 더 어려웠기 때문인 것으로 해석할 수 있다. 전체적으로는 쾌한 표정들의 구분에서 불쾌한 표정들의 구분에서 보다 더 빈번하게 오판단이 나타나는 경향이 있었다. 이러한 결과는 사람이 불쾌한 감정의 인식에 더 민감하고 세분화되어 있다는 가설[10]을 지지하는 결과로 해석해볼 수 있다. 즉, 다른 사람의 표정으로부터 감정 상태를 인식하는 과정에서 불쾌한 감정을 추론해내는 것은 그로부터 원인을 찾고 이를 제거해야 할 필요가 있기 때문에 쾌한 감정을 추론해내는 것보다 생존과 직접 관련될 가능성이 높은 사실에서 비롯된 결과로 볼 수 있다.

각성-수면 차원상의 표정에 대한 쌍별 비교 판단 결과 쾌-불쾌 차원과는 달리 전체 차원에 걸쳐 인접한 표정들과의 비교에서 잘못된 판단을 하는 경우가 많았다. 오판단의 수도 총 40회로 기록되어 쾌-불쾌 차원에서의 판단에서보다 오판단의 빈도가 11회 많았다. 이같은 결과는 내적 감정상태의 기저 차원으로서 각성-수면 차원은 쾌-불쾌 차원에 비해 설명력이 약하며 덜 안정적인[11] 데에서 비롯되는 것으로 해석할 수 있다. 그러

<표 1> 쾌-불쾌 및 각성-수면 축 위의 표정쌍에 대한 평정자의 쌍별 비교 판단 결과

쾌-불쾌차원				각성-수면차원			
제시된표정쌍	오판단 빈도수	제시된표정쌍	오판단 빈도수	제시된표정쌍	오판단 빈도수	제시된표정쌍	오판단 빈도수
1	2	7	4	5	0	1	2
	3	0	6	0		3	3
	4	0	7	0		4	2
	5	0	8	2		5	0
	6	0	9	0		6	1
	7	0	5	6	1	7	0
	8	0	7	0		8	1
	9	0	8	1		9	0
2	3	0	9	0	2	3	5
	4	0	6	7	3	4	4
	5	0	8	0		5	2
	6	0	9	2		6	0
	7	0	7	8	5	7	0
	8	0	9	4		8	1
	9	0	8	9	3	9	0
3	4	0			3	4	3
	5	0	총오판단빈도	29		5	0
	6	0	오판단비율(%)	2.69		6	0
	7	0	평균	0.806		7	1
	8	0	(표준편차)	(1.658)		8	1
	9	1				9	0
			총오판단빈도	40			
			오판단비율(%)	3.70			
			평균	1.111			
			(표준편차)	(1.368)			

나 평정자 중 한 사람이 전체 36쌍 중 여덟 쌍에 대하여 잘못 판단하였을 뿐만 아니라 이 중 일곱 쌍에 대해서는 30명의 평정자 중 유일한 오판단을 하였으므로 내적 상태 각성-수면 차원의 변화를 민감하게 포착하지 못하는 특성을 가진 이 평정자에 의하여 전체 자료가 왜곡 되었을 가능성도 있다.

전체적으로 볼 때 표정 합성 시스템을 이용하여 도출된 표정들은 내적 감정상태의 쾌-불쾌, 각성-수면 차원 상의 변화를 민감하게 반영하고 있는 것으로 조사되었다. 이로부터 내적 감정상태를 쾌-불쾌 및 각성-수면 차원값으로 규정하고 이 차원값에 따라 연속적으로 얼굴 표정이 변화되도록 한 검증 대상 시스템의 접근 방식은 표정과 관련된 사람의 내적 감정상태를 적절하게 표상한 것으로 볼 수 있다.

2.2 시스템에서 구현된 표정의 정확성 검증

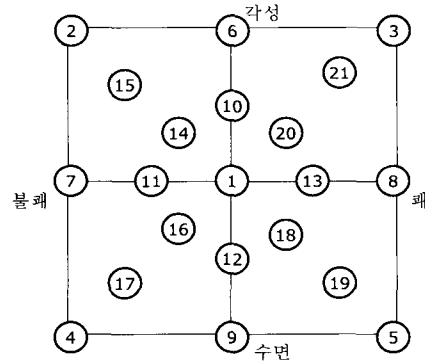
자동 표정 합성을 위한 검증 대상 시스템은 사람의 정서 구조에 대한 차원 모형을 가정하고 설계되었으며 이러한 접근은 비교적 타당한 것으로 확인되었다. 이에 근거한 시스템의 개발 과정과 성능의 타당성이 검증되기 위해서는 이와 같은 과정을 통해 도출된 표정 합성 알고리즘이 얼마나 정확하게 사람의 특징을 반영하고 있는가가 객관적으로 조사될 필요가 있다. 표정 합성의 궁극적인 목표는 사람같은 자연스러운 표정을 만들어내는 것이므로 표정 합성 알고리즘의 정확성은 사람들이 구현된 표정을 어떻게 인식하는가를 통해서 측정될 수 있다. 시스템을 통해 합성 가능한 모든 표정들에 대해 평가한다는 것은 불가능하므로 결과물들을 대표할 수 있는 표정들을 선정하여 이들을 검증 대상으로 결정하였다. 요구특성에 의한 효과를 배제하기 위하여 직접적인 내적상태 범주나 차원을 제시하지 않고 각 표정이 유사한 정도를 평정하도록 하였다.

2.2.1 방법

참가자. 연세대학교 학부생 및 대학원 인지과학 협동과정에 재학중인 대학생 30명이 평정자로서 참가하였다. 이들은 개별적으로 평정 과제를 수행하였다.

자료. 평정 대상 자료로서 검증 대상 표정 합성 시스템으로부터 구현되는 결과들을 대표할 수 있는 표정들을 선정하였다. 먼저 내적상태의 기저 차원인 쾌-불쾌 및 각성-수면 차원을 기준으로 중립과 양극단의 조합으로 아홉 개(3×3)의 표정을 선정하였다. 각 차원 축을 기준으로 이 얼굴들의 중간점에 해당하는 표정 네 개(2×2) 및 중립 표정과 양극단의 조합 표정인 네 가지의 표정 사

이에 등간으로 두 개의 중간 표정들을 선택하여 여덟 개 (4×2)의 표정을 선택하였다. 이 과정을 거쳐 평정 대상으로 선정된 표정은 (그림 3)과 같이 총 21개였으며 이들로부터 표정 쌍(pair) 210가지를 만들어내었다.



(그림 3) 성능 검증을 위해 선정한 시스템 내 대표 표정들의 위치 도식

절차. 평정자는 평정 요령에 대해 간단한 설명을 들은 뒤 의문 사항에 대해 질문할 수 있도록 하였다. 평정 방법을 숙지한 후 시작 버튼을 눌러 평정을 시작하도록 하였다. 참가자는 무선회되어 제시되는 표정 쌍들에 대하여 두 표정으로부터 추론되는 내적상태가 유사한 정도를 1점(매우 비슷함)에서 9점(매우 다름)까지의 척도에 평정하였다. 표정들은 319×512 화소의 크기로 좌우로 나란히 17인치 모니터에 제시되었으며 평정의 제한 시간은 없었고 두 표정을 비교하여 평정하면 다음 표정 쌍이 제시되도록 하였다.

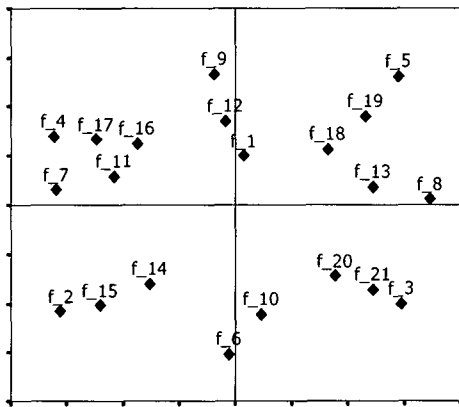
2.2.2 결과 및 논의

표정 합성 시스템으로부터 추출된 대표 표정쌍에 대한 유사성 평정 자료를 다차원 척도법(MDS: Multidimensional Scaling)으로 분석한 결과를 <표 2>에 요약하였으며 (그림 4)에 부적합도 수준(STRESS)과 설명력을 고려하여 2차원 공간상에 표현한 결과를 제시하였다.

<표 2> MDS 분석 결과 추출 차원 수에 따른 부적합도 (STRESS)의 크기 및 설명력과 그 변화량

추출차원수	STRESS	R ²	ΔSTRESS	ΔR ²
1	.372	.619		
2	.118	.925	.254	.306
3	.052	.978	.066	.053
4	.045	.985	.007	.008

유사성 평정 및 다차원 척도법을 이용하여 도출해 낸 표정들의 분포는 시스템 내의 분포와 매우 유사하였다. 이러한 결과로부터 합성 시스템으로부터 구현된 표정들은 시스템 개발자들이 표현해내고자 의도한 것과 유사한 방식으로 인식되고 있음을 확인할 수 있다. 다차원 척도법을 통한 분석 결과 추출된 두 차원은 시스템의 기저 모형인 쾌-불쾌 및 각성-수면 차원으로 명명될 수 있었다. 이들 두 차원은 전체 조사 대상 표정들의 변량 중 92.5%(STRESS : .118)를 설명하는 것으로 확인되었으며 이 중 첫 번째 차원인 쾌-불쾌 차원에 의한 설명량이 61.9%인 것으로 나타났다.



(그림 4) 유사성 평정에 기초하여 다차원 척도법을 이용, 두 차원에 표상한 결과

(그림 4)에 제시된 분석 결과를 (그림 3)에 나타난 합성 시스템의 도식과 비교하면 몇 가지 특징들을 확인할 수 있다. 다차원 척도법에 의한 분석 결과로 도출된 표정들의 분포는 전반적으로 쾌-불쾌 차원의 분포에 비하여 각성-수면 차원의 분포가 축소된 경향이 있었으며 이러한 경향은 특히 불쾌한 쪽에서 두드러졌다. 또한 시스템에서 중간 정도의 각성 수준으로 구현한 표정들은 조금씩 각성된 쪽으로 나타났다. 이러한 두 가지 특징은 모두 내적상태의 기저 차원으로서 각성-수면 차원이 쾌-불쾌 차원보다 덜 분명한 것에서 기인하는 것으로 해석해 볼 수 있으나 시스템 내 알고리즘의 편포 가능성도 검토되어야 할 것이다.

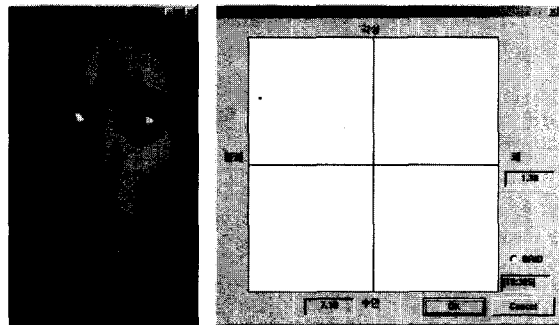
3. 표정 합성을 통한 인터페이스의 효율성 비교

표정 합성 시스템이 실제 표정을 만들어 내는 작업에 사용될 수 있기 위해서는 사용자에게 가장 효율적인 조

작 방법이 제공되어야 한다. 성능이 검증된 표정 합성 시스템의 사용 용이성 및 편의성을 극대화할 수 있는 인터페이스의 형태를 조사하고자 하였다. 이를 위해 사람들에게 서로 다른 입력 방식을 사용하여 시스템으로부터 표정을 구현해내도록 하고 그 결과들을 비교, 분석하였다.

3.1 좌표평면의 효율성 평가

쾌-불쾌 및 각성-수면의 두 차원에 의해 구성된 평면을 제시하고 그 평면상에서의 위치를 지정함으로써 차원모형으로 가정된 사람의 내적 정서 상태를 표현할 수 있다(그림 5). 이와 같은 시각화(visualization)가 표정 합성 시스템의 인터페이스 형태로서 적합한가를 확인하기 위해 사람들에게 내적상태 범주를 제시하고 좌표평면을 이용하여 그에 대응되는 표정을 합성해보도록 하였다.



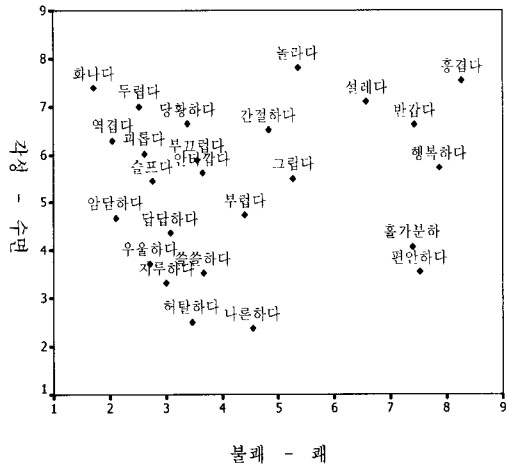
(그림 5) 좌표평면을 통한 인터페이스

3.1.1 방법

참가자. 연세대학교 심리학과 및 인지과학 협동과정의 대학원생 15명이 시스템의 사용자로 참가하였다. 이들은 개별적으로 표정 합성 시스템을 사용하였다.

자료. 내적 감정 상태를 기술하는 형용사 25개가 표정 합성 시스템 검증을 위한 범주로서 제시되었다. 이 형용사들은 김영아 등[11]이 제시한 83개의 내적 감정 상태 기술 형용사들로부터 선정되었다. 국어학 전문가가 이들 중 순수하게 감정 상태를 기술한다고 보기 어려운 형용사(예, 이상한)들을 제외하고 유사한 의미를 갖는 형용사들을 형용사군으로 묶은 후에 그 중 사용 빈도가 높아 각 묶음을 대표할 수 있는 범주들을 선택하여 결정하였다. 형용사 25개의 내적 감정 상태 쾌-불쾌, 각성-수면 차원값은 형용사군에 포함되었던 범주들의 차원값을 평균한 값이었으며 내적상태 사분면 상에

비교적 고르게 분포하고 있었다(그림 6).



(그림 6) 시스템을 이용하여 표정을 구현하도록 제시된 25개 내적상태 기술 어휘 및 분포

절차. 합성 시스템의 구조 및 사용법에 대한 간단한 설명을 들은 후 의문 사항에 대해서 질문하도록 하였다. 표정 구현 절차에 대해 숙지하고 난 후 시스템의 사용자는 무선적으로 25개의 형용사 범주들 중 하나를 제시 받고 그에 해당하는 표정을 구현해내도록 하였다. 이 때 주어지는 형용사 범주에 대한 쾌-불쾌 및 각성-수면 차원값이 참고 자료로서 시스템을 통해 함께 제공되었다. 사용자는 시간의 제한 없이 충분히 최적의 표정을 탐색할 수 있도록 하였다. 사용자가 쾌-불쾌와 각성-수면 차원이 구성하는 평면상의 임의의 위치를 마우스로 클릭하면 표정이 구현되어 즉시 표정의 변화를 확인할 수 있었다. 사용자가 주어진 범주에 대응되는 최적의 표정을 찾아내면 그 위치의 쾌-불쾌 및 각성-수면 차원 값이 기록되었으며 계속하여 다음 형용사 범주가 제시되었다.

3.1.2 결과 및 논의

좌표평면을 이용한 인터페이스의 효율성을 조사하기 위하여 세 가지 분석을 실시하였다. 먼저 제시된 형용사의 위치와 구현된 표정의 위치를 비교하였다. 다음으로 표정을 구현한 시스템 사용자들간의 평정 편차를 조사하였다. 마지막으로는 형용사의 차원값들과 구현된 표정의 차원값들 간의 상관 관계를 조사하였다.

표정 합성을 위해 제시된 형용사의 위치를 좌표평면 상에서 사분면별로 나누고 각 사분면별로 형용사의 위

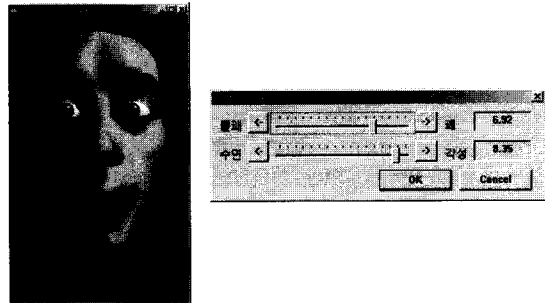
치와 구현된 표정의 위치를 비교하였다. 비교 결과 두 위치 사이의 관계는 제공된 형용사의 원위치에 따라 서로 차이가 있는 것으로 조사되었다. 시스템 사용자에게 주어진 형용사 범주가 불쾌한 내적 감정 상태를 나타내는 경우 표정은 더 불쾌한 쪽으로 치우쳐 합성되는 경향이 있었으며 전반적으로 사용자들은 제시된 형용사의 원 위치보다 조금 더 각성된 상태로 표정을 구현하였다.

시스템 사용자 15명의 표정 구현의 표준편차는 쾌-불쾌 차원에서 평균 0.74, 각성-수면 차원에서 평균 0.84로 각성-수면 차원에서의 구현의 편차가 더 큰 것으로 나타났다. 이러한 경향은 앞에서 논의된 것과 일관된 맥락에서 쾌-불쾌 차원에 비해 각성-수면 차원이 불안정한 것에서 비롯되는 것으로 해석할 수 있다.

스물 다섯 개의 내적상태 기술 형용사와 이들 형용사를 바탕으로 합성된 표정 25개에 대하여 이들의 쾌-불쾌 및 각성-수면 차원 상에서의 상관 관계를 조사하였다. 그 결과 쾌 차원은 .98, 각성 차원은 .93의 높은 상관성이 있는 것으로 조사되었다. 이러한 결과로부터 형용사와 이를 바탕으로 구현된 표정의 개별적인 위치에는 차이가 있었지만 형용사들과 표정간의 상대적인 분포는 상당히 유사함을 확인할 수 있다.

3.2 슬라이드바의 효율성 평가

쾌-불쾌 및 각성-수면의 두 차원에 의해 표상되는 내적상태는 각각의 값들을 수치화하여 조정함으로써 시스템에 입력될 수 있다(그림 7). 내적상태 기저 두 차원의 값들이 슬라이드바를 통해 조절되고 그에 따라 표정이 합성되는 방식의 인터페이스로서의 효율성을 조사하였다.



(그림 7) 슬라이드바를 이용한 인터페이스

3.2.1 방법

참가자. 연세대학교에서 개설된 심리학 과목 수강생

16명이 인터페이스의 사용자로서 참가하였다. 이들은 과목 이수를 위한 필수 요건으로 실험에 참가하였다.

자료 및 절차. 참가자는 쾌-불쾌와 각성-수면 차원을 움직이는 슬라이드바를 움직이거나 양쪽 끝의 화살표를 클릭함으로써 두 차원의 값을 조절하였다. 이를 제외한 모든 과정은 좌표평면의 효율성 평가 절차와 동일하였다.

3.2.2 결과 및 논의

슬라이드바를 이용한 인터페이스의 효율성을 조사하기 위해 좌표평면 인터페이스의 검증과 동일한 방법으로 사용자들의 사용경향을 분석하였다. 제시된 형용사의 위치를 기준으로 사분면별 형용사의 위치와 구현된 표정의 위치를 먼저 비교하였다. 그 결과 형용사의 위치에 관계없이 슬라이드바를 이용하여 표정을 합성하는 경우 표정은 더 불쾌한 쪽으로 치우쳐 합성되는 경향이 발견되었다. 제 1, 2사분면, 즉 각성된 상태의 표정의 경우 좌표평면을 이용한 인터페이스와 서로 다른 합성 경향을 보여 각성된 범주의 경우에는 덜 각성된 정도로 표정이 합성되는 경향이 있었다. 좌표평면과 슬라이드바를 이용하는 두 경우 모두 제시된 형용사의 쾌-불쾌 및 각성-수면 차원값과 구현된 표정의 쾌-불쾌 및 각성-수면 차원값 사이에는 통계적으로 유의미한 차이를 보이지 않았으며 두 인터페이스를 사용한 결과 사이에도 통계적으로 유의미한 차이가 발견되지 않았다.

시스템 사용자 16명의 표정 구현의 표준편차는 쾌-불쾌 차원에서 평균 0.82, 각성-수면 차원에서 평균 1.18로 좌표평면을 사용한 결과와 일관적으로 각성-수면 차원에서의 구현의 편차가 더 컸다. 슬라이드 바를 이용한 표정 구현 결과는 좌표평면을 이용한 합성 결과의 편차와 비교하여 통계적으로 유의미하게 더 큰 것으로 나타났다($t(98)=-2.591, p<.05$).

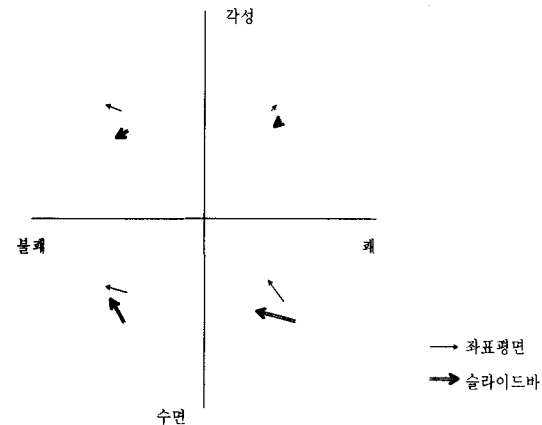
스물 다섯 개의 내적상태 기술 형용사와 이를 바탕으로 합성된 표정 25개에 대한 두 차원에서의 상관 분석 결과 각각 .96 및 .85의 상관이 있는 것으로 조사되었다. 이 결과는 좌표평면 인터페이스를 사용한 경우에 슬라이드바 인터페이스를 사용한 경우보다 형용사의 원차원값과 유사한 분포로 합성됨을 보여준다.

두 가지 인터페이스에 대한 효율성 조사 결과는 <표 3> 및 (그림 8)에 요약하였다. 이로부터 검증 대상 표정 합성 시스템을 이용하는 경우 사용자는 인터페이스의 종류와 무관하게 비교적 일관된 구현 결과를 얻을 수 있음을 확인할 수 있다. 그러나 독립적으로 두 차원을 제시하는 인터페이스보다는 좌표평면을 제시하고 평면

위의 위치를 입력하도록 하는 인터페이스가 더 안정적인 구현 결과를 산출해 낼 수 있었다.

<표 3> 좌표평면 및 슬라이드바를 이용한 인터페이스의 효율성 조사 결과 비교

	좌표평면	슬라이드바
사용자들의 평균 평정편차	0.79	1.00
쾌-불쾌차원	0.74	0.82
각성-수면차원	0.84	1.18
제시된-구현된 표정간 상관	0.96	0.91
쾌-불쾌차원	0.98	0.96
각성-수면차원	0.93	0.85



(그림 8) 제시된 형용사와 합성된 표정의 위치 차이 도식 (화살표의 각도와 길이는 차이의 방향과 거리를 나타냄)

4. 결론 및 종합논의

내적상태의 차원모형에 근거하여 개발된 표정 합성 시스템의 기본 가정 및 성능을 검증하고 인터페이스의 효율성을 비교하였다. 표정을 인식하고 합성해내는 경험적인 과정들을 통해서 표정에 의한 사람의 내적 감정 상태는 쾌-불쾌 및 각성-수면의 두 기저 차원에 의해서 효과적으로 표상될 수 있음이 확인되었다. 또한 이와 같은 두 차원에 기반하여 표정을 합성하는 경우에 사용자들은 두 차원의 값을 동시에 고려하여 좌표평면의 위치로서 입력하는 경우에 더 안정적으로 시스템을 사용할 수 있었다.

사람의 정서 범주들이 체계적으로 연관되어 있을 수 있다는 가능성이 제기되면서부터 쾌-불쾌 및 각성-수면 차원은 정서 범주 체계 내의 기저 차원으로서 밝혀

져 왔다[12][13][14]. 얼굴표정과 관계된 정서에 대해서도 이 두 차원 구조가 규명되어 왔는데[15] 이러한 경향은 범문화적으로 일관되었으며[16] 정서를 표현하는 단어의 특성에서 기인하는 것이 아니라는 사실도 확인되었다[17]. 검증 대상 시스템 내 표정들은 내적상태 캐 및 각성 차원 상의 변화에 따라 얼굴의 특징 변수들이 연속적으로 변형되어 구현되었는데 이들은 시스템의 입력값과 유사하게 사람들에게 인식된다는 사실이 확인됨으로써 표정과 관련된 내적상태의 기저 차원으로서 캐 및 각성 차원을 제안한 이전의 연구 결과들을 뒷받침하고 있다.

시스템으로부터 합성된 표정을 인식하는 경우나 시스템을 이용하여 표정을 구현하는 경우에 내적상태 각성-수면 차원 상의 변화는 캐-불쾌 차원만큼 표정을 통해 민감하게 전달되거나 표현되지 않아 상대적으로 불안정한 경향을 보였다. 한국어의 정서 어휘를 분석하여 내적 체계의 구조를 밝히고자 했던 국내의 연구들에서는 캐-불쾌 차원에 비해 각성-수면 차원의 설명력이 낮다는 것을 밝혀왔으며[11][18][19] 이들로부터 연구 결과가 인간의 정서 구조에 대한 일반적인 특징일 수 있음을 시사한다. 그러나 기저 두 차원을 밝혀낸 국외의 이전 연구들에서는 각 차원의 설명력을 구분하여 표시하지 않았으므로 이 특성이 한국어 또는 한국인에게 제한적인 것인지 추가로 조사될 필요가 있다.

표정 합성 시스템의 사용성은 정확성에 근거하여 평가하였다. 그러나 실제 시스템을 이용하는 경우 성능의 평가는 정확성과 함께 속도가 고려되어야 하고 이 두 특성은 대개의 경우 서로 상충적(trade-off) 관계에 있으므로 시스템을 사용하여 결과를 얻는 데까지 걸리는 시간의 문제도 검토되어야 한다. 뿐만 아니라 표정 합성 시스템을 전적으로 사용하지 않는 비전문가들에 의해서 효율성 평가가 수행되었다는 제한점이 있다. 표정을 만들어내는 작업을 전문적으로 하는 전문가들의 사용 방식과는 차이가 날 수 있으므로 그 차이가 어디에서 비롯되는 것인지, 두 집단이 서로 요구하는 것이 무엇인지 구분하여 조사할 필요가 있다. 표정 합성 시스템은 궁극적으로 모형에 의존적이지 않아 임의의 입력 영상에도 타당한 표정을 구현해낼 수 있는 시스템으로 발전될 필요가 있다. 이 단계에서야 비로소 글자를 대신하는 자극으로서, 미묘한 감정 변화의 지침 자료로서 유용하게 사용될 수 있을 것이다.

참고문헌

- [1] Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA : Consulting Psychologist Press.
- [2] Aizawa, K., Harashima, H., & Saito, T. (1989). Model-based analysis synthesis image coding(MBASIC) system for a person's face. *Signal Processing : Image Communication*, 1(2), 139-152.
- [3] Yang, D., Kunihiro, T., Shimoda, H., & Yoshikawa, H. (1999). A study of real-time image processing method for treating human emotion by facial expression. *IEEE Systems, Man, and Cybernetics '99 Conference Proceedings*, 360-364.
- [4] Terzopoulos, D., & Waters, K. (1993). Analysis and synthesis of facial image sequences using physical and anatomical models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(6), 569-579.
- [5] Aoki, Y., & Hashimoto, S. (1998). Physical facial model based on 3D-CT data for facial image analysis and synthesis. *Proceedings of the third IEEE international conference on Automatic Face and Gesture Recognition*, 448-453.
- [6] Ishikawa, T., Sera, H., Morishima, S., & Terzopoulos, D. (1998). Facial image reconstruction by estimated muscle parameter. *Proceedings of the third IEEE international conference on Automatic Face and Gesture Recognition*, 342-347.
- [7] Ekman, P., Friesen, W. V., O'Sullivan, M., Chan, A., Diacoyanni-Tarlatzis, I., Heider, K., Krause, R., LeCompte, W. A., Pitcairn, T., Ricci-Bitti, P. E., Scherer, K., Tomika, M., & Tzavaras, A. (1987). Universals and cultural differences in judgments of facial expressions of emotion. *Journal of Personality and Social Psychology*, 53(4), 712-717.
- [8] Russell, J. A., & Fehr, B. (1987). Relativity in the perception of emotion in facial expressions. *Journal of Personality and Social Psychology*, 116(3), 223-237.
- [9] 한재현, 정찬섭 (2002). 내적상태 차원모형에 근거한 얼굴표정 합성 시스템. *한국인지과학회 논문지*, 13(3), 11-21.
- [10] Mandler, G. (1975). *Mind and emotion*, New York :

John Wiley & Sons.

- [11] 김영아, 김진관, 박수경, 오경자, 정찬섭 (1998). 정서 관련 어휘 분석을 통한 내적 상태의 차원 연구. *한국감성과학회지*, 1(1), 145-152.
- [12] Schlosberg, H. (1952). The description of facial expressions in terms of two dimensions. *Journal of Experimental Psychology*, 44, 229-237.
- [13] Watson, D., Clark, L. A., & Tellegen, A. (1984). Cross-cultural convergence in the structure of mood: A Japanese replication and a comparison with U.S. findings. *Journal of Personality and Social Psychology*, 47, 127-144.
- [14] Zevon, M. A., & Tellegen, A. (1984). The structure of mood change: An idiographic/nomothetic analysis. *Journal of Personality and Social Psychology*, 43, 111-122.
- [15] Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, 39(6), 1161-1178.
- [16] Russell, J. A. (1983). Pancultural aspects of the human conceptual organization of emotions. *Journal of Personality and Social Psychology*, 45(6), 1281-1288.
- [17] Russell, J. A., & Bullock, M. (1985). Multidimensional scaling of emotional facial expressions: Similarity from preschoolers to adults. *Journal of Personality and Social Psychology*, 48(5), 1290-1298.
- [18] 이만영, 이홍철 (1990). 형용사 서술 의미의 구조에 관한 연구. *한국심리학회지: 실험 및 인지*, 2, 118-138.
- [19] 안신호, 이승혜, 권오식 (1993). 정서의 구조: 한국어 정서단어 분석. *한국심리학회지: 사회*, 7(1), 107-123.