

표준형상모형 정합을 통한 얼굴표정 구조 분석¹⁾

한재현, 심연숙, 변혜란[†], 오경자*, 정찬섭*
연세대학교 대학원 인지과학 협동과정, 컴퓨터과학과[†], 심리학과*

Structural Analysis of Facial Expressions Measured by a Standard Mesh Frame

Jaehyun Han, Youn-Sook Shim, Hye-Ran Byun[†],
Kyung Ja Oh*, Chan-Sup Chung*
Cognitive Science Program, Department of Computer Science[†],
Department of Psychology*, Yonsei University

자동 표정인식 및 합성 기술과 내적상태별 얼굴표정 프로토타입 작성의 기초 작업으로서 특정 내적상태를 표현하는 얼굴표정의 특징적 구조를 분석하였다. 내적상태의 평정 절차를 거쳐 열 다섯 가지의 내적상태로 명명된 배우 여섯 명에 대한 영상자료 90장을 사용하여 각 표정의 특징적 구조를 발견하고자 하였다. 서로 다른 얼굴들의 표준화 작업과 서로 다른 표정들의 직접 비교 작업에 정확성을 기하기 위하여 각 표정 표본들을 한국인 표준형상모형에 정합하였다. 정합 결과로 얻어진 각 얼굴표정의 특징점에 대해 모형이 규정하고 있는 좌표값들만으로는 표정해석이 불가능하며 중립얼굴로부터의 변화값이 표정해석에 유효하다는 결론을 얻었다. 표정의 특징적 구조는 그 표정이 표현하는 내적상태가 무엇인가에 따라 발견되지 않는 경우도 있었으며 내적상태가 기본정서에 가까울수록 비교적 일관된 형태를 갖는 것으로 나타났다. 내적상태별 특징적 표정을 결정할 수 있는 경우에 표정의 구조는 얼굴표정 요소들 중 일부에 의해서 특징지어짐을 확인하였다.

최근 감성 인터페이스, 애니메이션, 게임과 관련된 영상 산업 분야에서 사이버캐릭터와 아바타를 사용하는 추세가 증가하면서 임의의 얼굴 표정을 자동으로 인식하거나 반대로 특정 내적상태로 인식되도록 얼굴표정을 변형, 합성하는 방법에 대한 연구 결과들이 광범위하게 축적되어 왔다. 표정의 자동인

식에서는 얼굴을 구성하고 있는 특징요소들을 구분하는 작업과 각 특징요소들의 생김새를 파악하는 작업이 필요한데 이를 위하여 얼굴표정을 구성하는 파라미터의 기하학적 거리 및 근육 움직임을 측정하는 방법(예. Lien, Kanade, Cohn 과 Li, 1998 ; Zhang, Lyons, Schuster 와 Akamatsu, 1998), 여러 크기 및 방향의 필터를 적용하여 특징 성분을 추출해내는 방법(예. Lyons, Akamatsu, Kamachi 와 Gyoba, 1998), 또는 윤곽선을 추출해내는 방법(예. 이경희, 고재필,

1) 본 연구는 과학기술부 감성공학기술개발사업 17-02-A-10의 연구비 지원을 받아 수행되었습니다.
e-mail : jaehyun@psylab.yonsei.ac.kr

변해란, 이일병과 정찬섭, 1998 ; Yokoyama, Yagi 와 Yachida, 1998) 등 다양한 방법들이 사용되어져 왔다. 표정의 자동합성에서도 근육에 기반한 방법(예. Ishikawa, Sera, Morishima 와 Terzopoulos, 1998), CT 자료를 이용한 방법(예. Aoki 와 Hashimoto, 1998), 표정의 2차원 자료에 근거한 방법(예. Mukaida, 1998 ; Mukaida, 1999) 등 다양한 접근이 시도되어 왔다. 그러나 어떠한 구체적인 기술을 이용하는지에 관계없이 자동적으로 표정을 인식하고 합성하는 시스템의 기반에는 사람들의 반응을 기초로 한 내적상태와 표정을 연결하는 모형이 설정되어 있어야 한다. 시스템 설계의 궁극적 목적이 사람과 동일하게 표정을 인식하고 사람처럼 자연스러운 표정을 합성하는 것이기 때문이다.

사람들이 표정에 반응하는 양상을 조사하고 그로부터 얼굴표정과 내적상태의 연결 관계를 형식화할 수 있기 위해서는 사람들의 표정에 대한 반응뿐만 아니라 관찰 대상 표정 자체에 대해서도 세밀한 분석이 필요하다. 그러나, Fernández-Dols 와 Ruiz-Belda(1997)가 지금까지 수행되어 온 다수의 표정에 대한 정서 연구들의 검토 결과에서 밝히고 있듯이 대부분의 연구들에서 표정 자체는 간과되고 그 표정에 대한 반응에만 초점을 두어 왔다. “표정을 짓고 있는 사람을 관찰하고 그 사람이 두 종류의 영화 중 어떤 것을 보았는지 추론해보라” 등과 같이 관찰자의 판단에만 의존한 실험 연구 결과들은 실제로 관찰자의 반응이 무엇에 근거한 것인가를 분명히 설명해주지 못하므로 그러한 연구에서 도출된 결과만으로는 표정과 내적상태의 관계 모형을 설정하기 어렵다는 한계를 가진다.

이와 같은 한계를 극복하고 표정의 자동 인식 및 합성시스템과 얼굴표정별 프로토타입 작성 작업에 기초 자료를 제공하기 위하여 이 연구에서는 얼굴표정자료에 대한 반응 및 표정 자체에 대한 분석을 통하여 내적상태별 얼굴표정의 특징적 구조를 발견하고자 하였다.

이를 위하여 여기서는 정찬섭, 오경자, 이일병과 변해란(1998)이 개발한 DB의 표정영상자료들을 분석 대상으로 사용하였다. 이 DB의 자료들은 연구자의

주문에 의해 포즈(posed)된 것이지만 가능하면 자연스러운 표정을 얻기 위하여 전문 연극배우 또는 연기 전공 학생들에게 해당 내적상태를 기술하는 스크립트를 주고 일방향거울에 비치는 자신의 표정을 관찰하면서 표정을 짓도록 하여 수집된 것이다. 뿐만 아니라 정찬섭 등(1998)은 촬영 당시에 배우들에게 지시했던 내적상태를 그 표정을 기술하는 것으로 단정하지 않고 각 표정들에 대해 다수의 사람들이 내적상태를 평정한 값을 사용함으로써 각 표정에 대한 내적상태 기술의 타당성을 높일 수 있도록 하였다.

표정에 대한 분석기준으로 Ekman과 Friesen(1978)의 FACS(Facial Action Coding System)를 사용해 온 기존의 연구들은 서로 크기 및 모양이 다른 사람들의 얼굴을 표준화하지 못하여 직접적인 비교가 어려우며 모든 표정을 AU단위로만 나누었기 때문에 분류 및 비교가 조악하다는(coarse) 약점이 있었다. 이러한 점들을 보완하기 위해서 선정된 분석대상 표정자료들을 MPEG4 - SNHC 표준기반 한국인 표준 형상모형(정찬섭 등, 1998)에 정합하여 얻은 FAP (Facial Animation Parameter)의 좌표값들을 얼굴표정의 분석 기준으로 사용하기로 하였다.

표정의 분석 기준 결정 및 내적상태 특정

표정의 상관 구조 분석

내적상태별 얼굴표정에 따른 특징적 구조를 발견해내기 위한 사전 작업으로서 수집된 얼굴표정 자료들을 어떠한 기준으로 분석할 것인가를 결정하였다. 수집된 분석대상 자료들을 표준형상모형에 정합하여 얻은 형상모형 꼭지점의 좌표값들을 기준으로 먼저 Pilowsky, Thornton 과 Stokes(1986)가 고안한 차원값들을 측정, 비교하였다. 이 결과를 바탕으로 중립얼굴에서부터 표정얼굴로의 변화값을 기준으로 표정의 특징적 구조를 분석하기로 결정하였다. 표정이 표현하고 있는 내적상태에 따라 특징적 구조가 발견되는 것과 그렇지 않은 것이 있음을 확인하였다.

방법

자료. 이전의 표정연구들에서 주로 사용되었던 여섯 가지의 기본정서의 표정만으로는 내적상태와 표정간의 관계에 대한 일반적인 모형을 설정하기에

는 무리가 있다는 판단 하에 조사 대상이 되는 표정 기술단어의 수를 늘렸다. 표정을 통해서 내적상태를 기술하는데 일상적으로 사용되고 있는 한국어 단어들을 수집하고 단어들의 쾌-불쾌, 각성-수면 차원 상에서의 분포를 분석한 김영아, 김진관, 박수경, 오경자와 정찬섭(1998)의 연구 결과를 참조하여 각 사분면 상에 고루 분포하면서 비슷한 위치의 다른 단어들보다 사용 빈도가 높고 대표성이 있는 열 다섯 가지의 내적상태 기술 단어를 선정하였다. 이들 열 다섯 가지의 내적상태 기술 단어는 근심, 기대, 기분 좋은, 꺼림직한, 놀라운, 만족, 미안한, 미워하는, 불쾌한, 뿌듯한, 속상한, 실망스러운, 안타까운, 울적한, 졸리운이었으며 이들의 내적상태 두 차원이 구성하는 공간상에서의 분포는 그림 1에 제시하였다

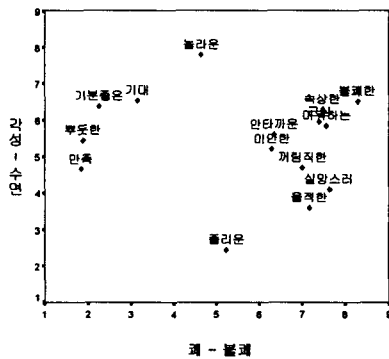


그림 1. 분석대상 열 다섯 가지의 내적상태 기술 단어 분포

분석에 사용된 표정 자료는 연세대학교 표정DB에서 표집된 6명 배우의 열 다섯 가지 내적상태와 중립 표정에 해당되는 96장의 사진이었다.

형상모형 기반 좌표값 측정. 각 표정자료들을 MPEG4-SNHC 표준기반 한국인 표준형상모형에 정합하여 FAP 점들의 좌표값을 구하였다. 표준형상모형은 3차원을 가정하고 있으나 정합된 표정자료는 2차원 정보만을 포함하고 있었으므로 얻어진 좌표값들 중 얼굴 표고에 해당하는 z축의 좌표값은 전적으로 모형에 의존한 추정값이었기 때문에 분석에서 제외시켰다.

분석. 각 표정자료의 FAP 측정값들에 대하여 두 가지 분석이 수행되었다. 첫 번째 분석은

MPEG4 - SNHC 기반의 FAP 좌표값들로 규정된 Pilowsky 등(1986)의 측정값들이 각 내적상태에 해당되는 표정 구조의 특성을 제대로 포착할 수 있는지를 알아보기 위한 것이었다. 이를 위하여 각 얼굴표정의 2차원 좌표값들을 근거로 Pilowsky 등(1986)이 고안한 열 두 차원 중 위, 아래 눈동자의 폭을 제외한 열 차원의 값을 측정하였다. 만일 이와 같은 차원값으로 표정 구조를 파악할 수 있다면 각 내적상태별 배우간 상관성이 높아야 하며 서로 구분되는 내적상태의 표정간에는 상관성이 낮아야 한다는 가정 하에 이들 열 차원 값들의 내적상태 표정간 및 배우 간 상관관계 분석을 하였다. 두 번째 분석에서는 중립 표정에서 특정 내적상태 표정으로의 변화를 토대로 표정구조의 특성을 분석하였다. FAP 기준 39개 점들 중 코의 가장 높은 부분에 해당하는 점은 표준 모형에 정합될 때 항상 원점으로 가정되었으므로 분석에서 제외하고 변화량으로서 형상모형 38개 꼭지점들의 변화거리와 각도를 계산하였다. 내적상태 특정 표정 구조가 있다면 이들 꼭지점들의 배우간 변화 각도 및 거리가 유사할 것이라는 가정 아래 배우간 상관분석을 하였다.

결과 및 논의

분석 대상 표정자료들을 표준형상모형에 정합하여 얻은 FAP의 좌표값들을 기준으로 한 Pilowsky 등(1986)의 차원값은 얼굴표정을 분석하는 기준으로는 변별력이 전혀 없음이 발견되었다. 열 여섯 가지의 측정값들의 상관성은 표정들 간, 배우들 간에 .95 이상의 매우 높은 상관을 보였으며 이 결과들 중 일부를 표 1에 제시하였다.

표 1. FAP로 측정된 Pilowsky 등(1986) 차원값들의 얼굴표정 및 배우 간 평균 상관계수

(a) 각 배우별 표정 간 차원측정값들의 상관

표정	배우	김선심	김승현	유연수	이종민	이지현	황인성
기분좋은-만족		0.98	1.00	0.98	0.99	0.96	0.99
기분좋은-불쾌한		0.99	0.98	0.98	0.99	0.98	0.98
놀라운-졸리운		0.99	0.99	0.99	1.00	0.97	1.00
만족-실망스러운		0.99	1.00	0.99	1.00	0.99	1.00
불쾌한-속상한		1.00	1.00	0.99	0.99	0.99	1.00
실망스러운-울적한		0.99	0.98	0.99	0.98	0.99	0.99

(b) 각 표정별 배우 간 차원축정값들의 상관

배우	표정	놀라운	만족	불쾌한	속상한	울적한	졸리운
김선심-김승현		0.99	1.00	1.00	0.99	1.00	1.00
김승현-유연수		1.00	1.00	0.98	0.99	0.99	0.99
유연수-이종민		0.99	1.00	0.98	0.99	0.99	0.99
이종민-이지현		0.99	1.00	1.00	1.00	0.99	0.99
이지현-황인성		0.99	0.99	0.99	0.99	1.00	1.00
김선심-황인성		0.99	1.00	1.00	0.99	0.99	1.00

표 1에서의 결과는 표정이 얼굴이라는 한정된 공간 내에서 특징요소들의 미세한 움직임에 의해 표현되는 것이며 그러므로 표정의 정점에서 각 차원값들만으로는 표정별 특징구조를 설명할 수 없음을 나타낸다. 표정이 가지는 이와 같은 특성을 고려하여 표정별 구조의 특징을 밝히기 위해서는 동일한 인물의 중립얼굴과 정점 표정 얼굴간의 차이를 중심으로 표정을 분석하는 것이 필요하다는 결론을 얻고 이에 맞추어 분석 자료를 FAP의 38개 각 점들이 중립얼굴에서 표정얼굴로 변화할 때의 변화 각도 및 변화거리로 재설정하였다.

표 2에서 알 수 있듯이 중립얼굴로부터 표정얼굴로 변화할 때에 FAP 38개 점들이 변화하는 각도 및 거리에 대한 배우들 별 상관은 표정얼굴이 어떠한 내적상태를 표현하는 것이냐에 따라 큰 차이를 보였다. 즉, 모든 배우들간에 각 점들이 유사한 각도와 정도로 변화하여 비교적 동일한 하나의 패턴으로 표정을 나타낼 수 있는 내적상태가 있는 반면에 배우들마다 변화정도가 상이하여 공통점을 찾을 수 없는 표정도 있었다. 이와 같은 결과로부터 서로 다른 사람들에서 하나의 내적상태가 표현될 때 공통적이고 특징적인 패턴이 발견되는지의 여부는 그 내적상태에 의존적임을 알 수 있다. 표 2는 표정으로 표현되는 내적상태가 Ekman(1972)이 가정하고 있는 기본정서(기쁨, 놀람, 공포, 분노, 혐오, 슬픔)와 유사할수록 배우들간에 평균 .60이상의 높은 상관을 보임을 알려준다.

표 2. 배우 간 중립얼굴에서 표정얼굴로의 변화각도 및 변화거리의 평균 상관

(a) FAP 각 점들에 대한 중립얼굴에서 표정얼굴로의 변화 각도의 평균 상관 및 상관의 표준편차

표정	변화 각도의 평균상관	상관의 표준편차	표정	변화 각도의 평균상관	상관의 표준편차
놀라운	0.89	0.06	졸리운	0.52	0.28
울적한	0.74	0.21	꺼림직한	0.47	0.31
속상한	0.72	0.07	뿌듯한	0.43	0.29
기대	0.66	0.21	만족	0.42	0.23
불쾌한	0.65	0.13	미워하는	0.41	0.20
실망스러운	0.64	0.11	근심	0.41	0.34
기분좋은	0.63	0.18	미안한	0.20	0.35
안타까운	0.55	0.16	전체평균	0.56	0.21

(b) FAP 각 점들에 대한 중립얼굴에서 표정얼굴로의 변화 거리의 평균 상관 및 상관의 표준편차

표정	변화 거리의 평균상관	상관의 표준편차	표정	변화 거리의 평균상관	상관의 표준편차
기분좋은	0.99	0.01	안타까운	0.35	0.54
기대	0.74	0.36	뿌듯한	0.34	0.23
불쾌한	0.64	0.49	꺼림직한	0.32	0.36
속상한	0.61	0.56	미워하는	0.29	0.43
놀라운	0.56	0.64	졸리운	0.28	0.41
미안한	0.48	0.42	울적한	0.26	0.24
만족	0.41	0.31	실망스러운	0.24	0.40
근심	0.38	0.24	전체평균	0.46	0.38

내적상태 특정 표정구조의 세부 분석

내적상태 특정 표정의 상관 분석 결과로부터 기본정서와 유사한 내적상태가 표정으로 표현되는 경우에 표현하는 사람이 누구인가에 관계없이 비교적 유사하게 얼굴 특징요소들이 변화함을 알 수 있었다. 이러한 결과를 바탕으로 비교적 일관된 특징적 변화 양상을 가지는 것으로 조사된 표정들이 구체적으로 어떠한 구조적 속성을 가지는가를 조사하고자 하였다.

방법

자료. 표정의 상관구조 분석 결과 변화각도 또는 변화거리의 배우들 간 평균상관이 .60을 초과하는 내적상태에 대한 표정자료와 중립얼굴 자료를 사용하였다. 기대, 기분좋은, 놀라운, 불쾌한, 속상한, 실망스러운, 울적한의 7가지 내적상태에 대한 6명의 표정자료와 각각의 중립얼굴자료를 포함한 총 48장의 자료를 분석대상으로 선정하였다.

분석. 표정의 상관 구조 분석을 위해서 조사된 표정별 FAP 38개 점들에 대한 변화 정보 중 개인적 특성 또는 내적상태의 정도에 의해 유연성이 있는 변화거리보다 그렇지 않은 변화각도에 초점을 두었다. 각 표정별, 배우별로 38가지의 점들에 대한 변화각도를 FAP에서 규정한 그룹으로 나누어 공통적 변화 경향을 발견하고자 하였다. 얼굴의 38개 점은 눈썹, 눈, 코, 귀, 입, 입술, 뺨, 턱의 8개 그룹으로 분류되었으며 분류된 각 그룹은 각각 6개, 4개, 2개, 4개, 8개, 8개, 4개, 2개의 점들을 포함하고 있었다.

결과 및 논의

일곱 가지의 내적상태를 표현하는 표정들에 대하여 얼굴의 각 그룹별 중립얼굴로부터의 변화 각도의 분포를 조사한 결과 기대, 기분좋은, 놀라운, 불쾌한, 속상한, 울적한의 여섯 가지 내적상태에 대한 표정들로부터 비교적 공통된 변화패턴을 발견할 수 있었다. 이들에서는 적어도 하나 이상의 얼굴표정 요소들이 유사한 패턴으로 변화하였으며 각 요소들의 특징점을 그림 2에, 내적상태별 특징적 변화 구조를 그림 3에 제시하였다.

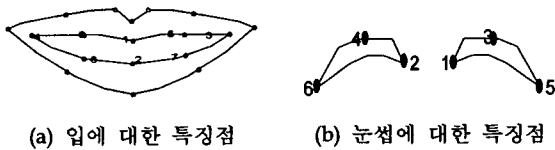


그림 2. 얼굴요소들에 정의된 특징점 분포

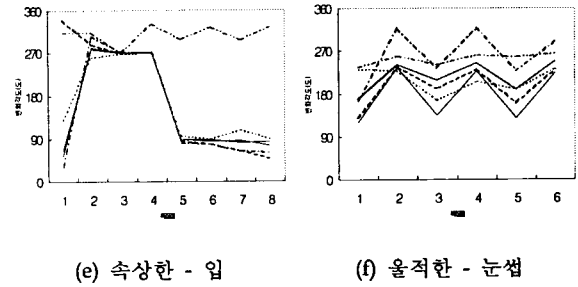
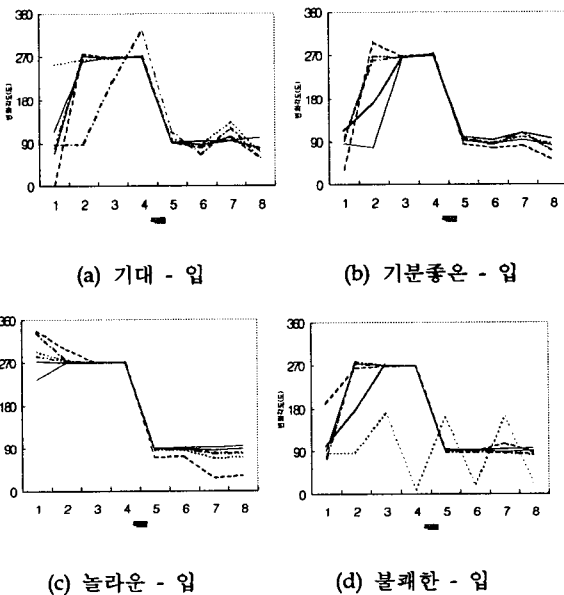


그림 3. 내적상태별 표정 특징적인 특징요소의 변화경향

그림 3의 결과로부터 하나의 내적상태에 대한 특징적인 표정은 얼굴 전체보다는 요소들의 특징적인 변화 패턴에 의해서 결정된다고 볼 수 있다. 또한 그 특징적인 변화는 대개 그 얼굴의 입의 모양에 의한 것임을 알 수 있다.

종합논의

내적상태를 표현하는 표정들 중 하나의 대표적인 패턴으로 수렴되지 않는 표정들도 있었다는 것은 하나의 내적상태에 대하여 하나이상의 얼굴표정이 존재하는 경우도 있음을 암시한다. 이러한 내적상태와 얼굴표정의 관계에 대한 모형은 정서와 얼굴움직임의 관계를 단순한 일대일대응 관계가 아니라 “or 관계” (or-relationship)로 결론지은 Ellgring, Scherer와 Wallbott(1997)의 연구 결과와 일관된다. 또한 분석을 통해 밝혀진 특징적 구조는 특정한 내적상태에 대한 전형적인 표정을 결정할 수 있는 경우라도 실생활에서 표현되는 표정은 얼굴 전체의 움직임보다는 일부분의 움직임에 의해 특징지어질 수 있다는 Carroll과 Russell (1997)의 연구 결과와도 일치한다.

기본정서와 유사한 내적상태들에서 비교적 배우간 표정의 상관이 높은 것으로 나타났는데 이는 이들 기본 정서와 관련된 표정인식과 합성은 비교적 신뢰할 수 있지만 그 이외의 내적상태들에 대해서는 그렇지 못할 수도 있음을 암시한다.

참고문헌

- 김영아, 김진관, 박수경, 오경자, 정찬섭 (1998). 정서 관련어휘분석을 통한 내적상태의 차원연구. 한국감성과학회지, 1(1), 145-152.
- 이경희, 고재필, 변혜란, 이일병, 정찬섭 (1998). 얼굴

- 표정인식을 위한 얼굴요소 추출. *한국감성과학회지*, 1(1), 33-40.
- 정찬섭, 오경자, 이일병, 변혜란 (1998). 표정/제스처에 의한 감성측정기술 및 DB개발. 과학기술부 감성공학기술개발사업 최종보고서, G - 17 - 01 - 09.
- Aoki, Y., & Hashimoto, S. (1998). Physical facial model based on 3D-CT data for facial image analysis and synthesis. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 448-453.
- Carroll, J. M., & Russell, J. A. (1997). Facial expressions in Hollywood's portrayal of emotion. *Journal of Personality and Social Psychology*, 72(1), 164-176.
- Ekman, P. (1972). Universal and cultural differences in facial expressions of emotions. In J. K. Cole (Ed.), *Nebraska symposium on motivation; 1971*. Lincoln : University of Nebraska Press.
- Ekman, P., & Friesen, W. V. (1978). *Facial action coding system*. Palo Alto, CA : Consulting Psychologists Press.
- Ellgring, H., Scherer, K. R., & Wallbott, H. G. (1997). The manifold facial expressions of emotions. *Electronic Proceedings of the 7th European Conference facial expression-measurement and meaning*.
- Fernández-Dolz, J. M., & Ruiz-Belda, M. A. (1997). Spontaneous facial behavior during intense episodes : Artistic truth and optical truth. In J. A. Russell & J. M. Fernández-Dols (Eds.), *The psychology of facial expression*. New York : Cambridge University Press.
- Ishikawa, T., Sera, H., Morishima, S., & Terzopoulos, D. (1998). Facial image reconstruction by estimated muscle parameter. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 342-347.
- Lien, J. J., Kanade, T., Cohn, J. F. & Li, C. C. (1998). Automated facial expression recognition based on FACS action units. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 390-395.
- Lyons, M., Akamatsu, S., Kamachi, M. & Gyoba, J. (1998). Coding facial expressions with Gabor wavelets. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 200-205.
- Mukaida, S. (1998). Face image generation system -FUTON system-. *Proceedings of ATR Symposium on Face and Object Recognition '98*, 39-40.
- Mukaida, S. (1999). Investigation of human sampling accuracy of facial image generator (FUTON system). *Proceedings of ATR Symposium on Face and Object Recognition '99*, 67-68.
- 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*. Dordrecht: Martinus Nijhoff Publishers.
- Yokoyama, T., Yagi, Y. & Yachida, M. (1998). Facial contour extraction model. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 254-259.
- Zhang, Z., Lyons, M., Schuster, M. & Akamatsu, S. (1998). Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron. *Proceedings of the third international conference on automatic Face and Gesture Recognition*, 454-459.