

얼굴의 포즈 상태와 오토마타 기법을 이용한 헤드 제스처 인식

(Head Gesture Recognition using Facial Pose States and Automata Technique)

오 승 태 * 전 병 환 **

(Seung Taek Oh) (Byung Hwan Jun)

요 약 본 논문에서는 인식된 얼굴의 포즈 상태 옆에 오토마타 기법을 적용하여 다양한 헤드 제스처를 인식하는 방법을 제안한다. 얼굴 영역의 추출에는 YIQ의 I성분인 최적의 얼굴색 정보와 적응적인 차영상 정보를 이용하며, 눈 영역 추출에는 소벨 연산자와 투영 기법, 그리고 눈의 기하학적 위치 정보를 이용한다. 얼굴의 상태 인식에는 계층적인 특징분석 방법을 사용하며, 인식된 얼굴 상태 옆에 오토마타 기법을 적용하여 13가지 제스처; 준비, 상측, 하측, 좌측, 우측, 전진, 후퇴, 좌 윙크, 우 윙크, 좌 더블 윙크, 우 더블 윙크, 긍정, 부정제스처를 인식한다. 총 8명으로부터 1,488 프레임의 영상을 취득하여 실험한 결과, 99.3%의 얼굴 영역 추출률, 95.3%의 눈 영역 추출률, 94.1%의 얼굴 상태 인식률, 그리고 93.3%의 헤드제스처 인식률을 얻었다.

Abstract In this paper, we propose a method for the recognition of various head gestures with automata technique applied to the sequence of facial pose states. Facial regions are detected by using the optimum facial color of I-component in YIQ model and the difference of images adaptively selected. And eye regions are extracted by using Sobel operator, projection, and the geometric location of eyes. Hierarchical feature analysis is used to classify facial states, and automata technique is applied to the sequence of facial pose states to recognize 13 gestures: Gaze, Upward, Downward, Leftward, Rightward, Forward, Backward, Left Wink, Right Wink, Left Double Wink, Right Double Wink, Yes, and No. As an experimental result with total 1,488 frames acquired from 8 persons, it shows 99.3% extraction rate for facial regions, 95.3% extraction rate for eye regions, 94.1% recognition rate for facial states, and finally 93.3% recognition rate for head gestures.

1. 서 론

컴퓨터 관련 학문 및 기술의 발전으로 인하여 기존의 한정된 입력장치에 의존하는 사용자 인터페이스에 비해 보다 자연스럽고 편리한 사용자 인터페이스에 대한 요구가 증가하고 있다. 이러한 요구를 충족시키기 위해 많은 연구자들이 필기 인식, 음성 인식, 그리고 영상 인식 분야에 대해 연구를 진행하고 있다. 특히, 헤드 제스처 인식은 사용자와 컴퓨터간의 의사 소통을 기계적인 방

식(키보드, 마우스 등)에 의존하지 않고, 얼굴의 방향이나 고개의 움직임 또는 눈의 깜빡임 등과 같은 정보를 이용하여 보다 자연스러운 컴퓨터 사용 환경을 제공해주는 것이 가능하다.

얼굴 및 얼굴 요소를 추출하기 위한 초기 방법에는 추출된 에지(edge)에 적용하는 고정 또는 가변 형판 정합(template matching) 방법[1-3]이나 타원 적합(ellipse fitting) 방법[4] 등이 있으나, 주로 단순한 배경의 정면 얼굴에 한정되고 잡음에 민감하며 처리 시간이 오래 걸리는 단점으로 인해 주로 정지 영상에만 적용된다. 동영상에서의 추출을 위해서는 차영상(difference image)이나 색상 정보를 이용하는 방법[5-8], 광류(optical flow)에 기반한 방법[9], 스네이크(snakes; active contour models)를 이용한 방법[10-11] 등이 있는데, 광류 접근

* BK21 대전·충남 정보통신인력양성사업단 연구조교(RA) 수혜

† 학생회원 : 공주대학교 컴퓨터공학과
ost5156@hanmail.net

** 종신회원 : 공주대학교 정보통신공학부 교수
bhjun@kongju.ac.kr

논문접수 : 2001년 1월 5일
심사완료 : 2001년 9월 4일

방법은 처리 시간이 매우 오래 걸리고 스네이크 접근 방법은 스네이크의 초기 설정에 따라 성능이 크게 좌우되는 단점을 갖는다.

한편, 영상에서 3차원적인 헤드 포즈(head pose)를 분석하는 방법에 관련하여, 두 대의 카메라에서 입력한 스테레오 영상에서 눈동자와 입 가장자리의 점들을 추출하여 얼굴 방향을 분석하는 연구[12], 머리 영역과 영굴 영역을 추출한 후 코, 입, 눈, 눈썹의 특징점을 추출, 이에 색상정보와 퍼지이론을 이용하여 얼굴의 방향을 파악하는 방법[13], 각 얼굴의 방향별로 학습에 의해 STCM(State Transition Covariance Matrix)을 저장하고, 입력된 영상과 비교하여 얼굴의 포즈를 파악하는 방법[14], 다양한 조명이나 포즈의 얼굴을 추적하기 위해 얼굴영상의 텍스처가 매핑된 실린더를 모델링하고, 입력 얼굴 영상과의 매핑 에러가 최소가 되도록 하는 워핑 템플릿과 조명 템플릿의 조합을 분석하는 방법[15] 등이 있다. 그러나, 이러한 얼굴 방향 분석 기술은 기본적으로 가능한 모든 방향의 얼굴 포즈를 대상으로 하고 있기 때문에, 일련의 특정 얼굴 상태로 구성되는 헤드제스처를 실시간에 인식하는 연구에 그대로 적용하기는 어렵다. 이와 더불어 일련의 약속된 포즈 상태로 규정되는 헤드 제스처(head gesture)를 인식하기 위한 연구가 최근 활발하게 진행되고 있다. 문병선[16]은 색상정보에 의해 추출된 얼굴 영역에 대해 연속된 영상에서 단순히 위치 변화의 방향을 구별하여 긍정, 부정의 2 제스처를 인식한다. 그러나 이 방법은 다양한 제스처를 사용하는 경우에는 적용하기 어렵다. 최현일[17]은 모자이크 영상에 BP(Back Propagation) 신경망을 적용하여 얼굴 영역을 추출하고 세그멘테이션(segmentation), 라벨링(labeling), 이진화(binazation)를 이용하여 눈 위치를 추출한 후, 이 값을 은닉 마르코프 모델(HMM; Hidden Markov Models)의 입력 값으로 이용하여 긍정, 부정, 중립의 3 제스처를 인식한다. Morimoto[18]은 Yes, No, Maybe, Hello의 4 제스처를 정의하고 벡터 양자화 과정을 통해 생성된 매개 변수(Image-yaw, Image-curl, Image-pitch)를 HMM에 적용하여 인식한다. 그러나 HMM 접근 방법을 사용하려면 입력되는 일련의 제스처들이 미리 정확하게 분할되어야 하는데, 특히 다양한 제스처를 사용하는 경우에는 제스처간의 구분이 매우 어려워진다. 이상에서 살펴본 바와 같이, 기존에 연구된 헤드제스처들은 매우 한정적이다. 이러한 한정된 헤드 제스처들로는 자연스러운 사용자 인터페이스 구현에 많은 제약이 따르게 되므로 보다 다양한 종류의 헤드 제스처를 정의하고 인식할 필요가 있다.

본 논문에서는, 그림 1과 같이, 동영상에서 얼굴 및

얼굴 요소를 신속하고 정확하게 검출하고 다양한 헤드 제스처를 실시간에 처리할 수 있는 방법을 제안한다. 먼저, 최적의 얼굴색 정보와 적응적 움직임 정보를 이용하는 방법[19]을 적용하여 다양한 포즈에 관계없이 얼굴 영역을 추출한다. 추출된 얼굴 영역에 대한 에지 영상의 투영 기법으로 눈의 후보 영역을 추출하고, 눈의 기하학적 정보와 눈의 개폐 정도를 이용한 검증 과정을 통해 최종적인 눈 영역을 결정한다. 그런 다음, 얼굴의 포즈 및 링크 정보를 계층적으로 분석하여 얼굴의 9가지 상태를 구분하고, 이렇게 생성된 일련의 얼굴 상태 열(state sequence)에 오토마타 기법(automata technique)을 적용하여 총 13 종류의 헤드 제스처를 인식한다. 이러한 방법에서는 연속되는 제스처의 구분이 신속하고 정확하게 이루어지며, 얼굴 상태나 제스처의 추가가 비교적 용이하게 이루어진다.

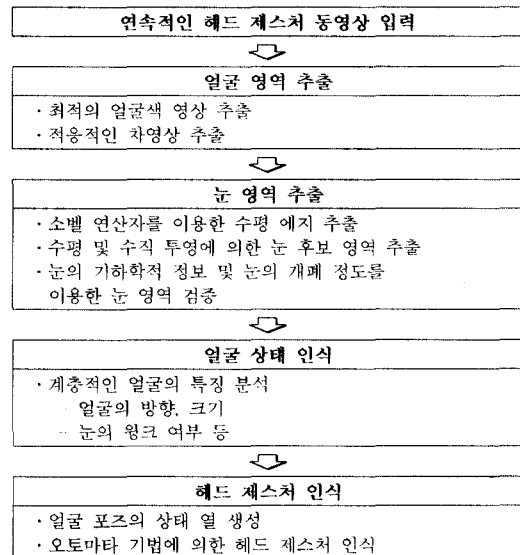


그림 1 헤드 제스처 인식의 전체 흐름도

2. 얼굴 및 눈 영역 추출

2.1 얼굴영역 추출

본 논문에서는 얼굴 영역을 추출하기 위해 최적의 얼굴색 정보와 적응적 움직임 정보를 이용한다[19]. 최적의 얼굴색 정보를 측정하는데는 YIQ 컬러 모델의 I 성분을 이용하였고, 움직임 정보의 경우에는 현재 영상 f_i 와 이전 영상 f_{i-1} 간의 차이가 일정 수준 이상으로 발생하지 않으면 영상 f_{i-1} 대신 그 다음 영상 f_{i-2} 와의 차이를 이용하는 방식을 사용했다. 그림 2는 얼굴 영역

추출의 과정을 보여준다.

2.2 눈 영역 추출

기존의 눈 추출 방법들[12, 20-22]은 대부분 배경이 단순한 정면 영상을 사용하여 복잡한 배경에서는 눈 추출이 어렵다는 점과 눈을 추출하기 위해 영상전체에 대해 눈 추출 알고리즘을 적용하여 처리시간이 오래 걸린다는 단점이 있다.

따라서 본 논문에서는 이미 추출된 얼굴 박스 영역에 대해서만 눈 영역 추출 알고리즘을 수행한다. 먼저, 수평 소벨 연산자를 이용하여 얼굴 영역내의 수평 에지 성분을 추출하여 이진화한다. 이진화된 얼굴 영역에 대해 수평방향으로 프로젝션을 수행한 후, 3개의 최고 마루를 기준으로 가장 근접한 최소 골을 찾아내어 영역 분리를 시도한다. 이렇게 분리된 영역들을 대상으로 수직방향으로 프로젝션을 수행하여 수평 영역 분리시 사용된 방법과 동일한 방법으로 눈 후보 영역들을 추출한다. 추출된 후보 영역들에 대해서 실제 눈 영역을 검증하기 위해 눈의 기하학적 위치 정보와 눈의 감은 정도를 분석한다. 눈의 감은 정도는 눈 영역에서 흰자위와 검은자위의 평균 명암도 차이를 이용한다.

그림 3은 전체적인 눈 영역 추출 과정을 나타낸다.

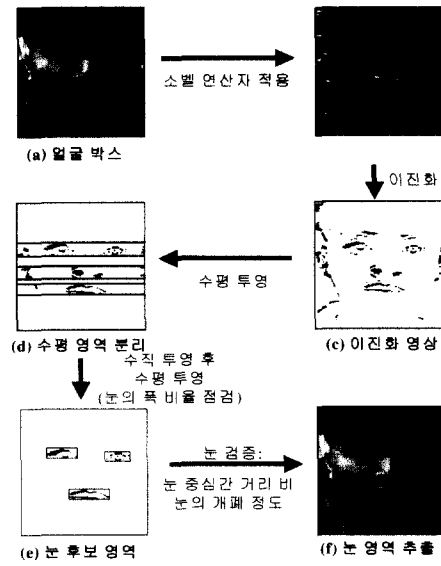


그림 3 눈 영역 추출 과정 및 결과

3. 얼굴의 상태 인식

3.1 얼굴의 9가지 상태 정의

본 논문에서 정의하는 얼굴의 9가지 상태는 정면, 상측, 하측, 좌측, 우측, 전진, 후퇴, 좌 윙크, 우 윙크 상태이며 그림 4와 같다.

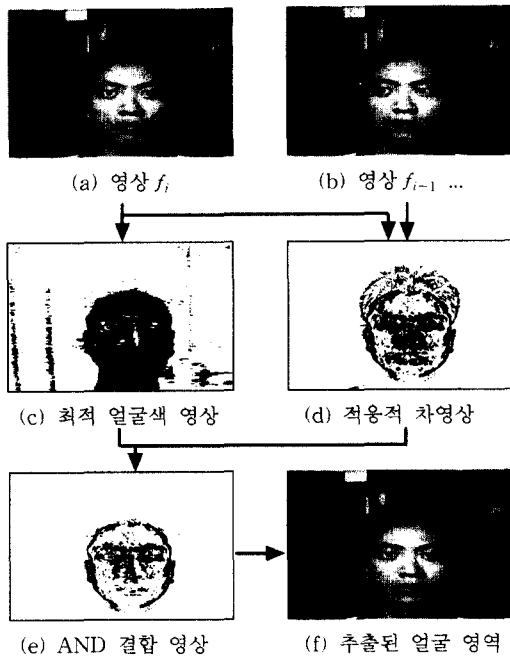


그림 2 최적의 얼굴색과 적응적 차영상 정보를 이용한 얼굴 영역 추출 과정

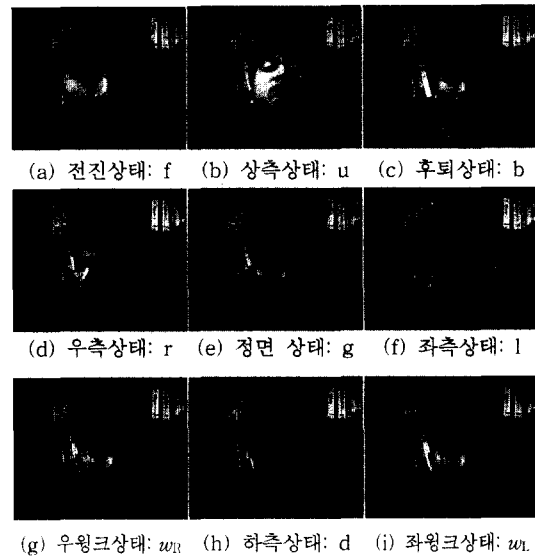


그림 4 얼굴의 9가지 상태에 대한 예

3.2 계층적인 특징을 이용한 얼굴 상태 인식

그림 5는 계층적인 특징을 사용하여 얼굴 상태를 인식하는 과정을 나타낸 것이며, 앞 절에서 정의한 얼굴의 각 상태를 구분하기 위해서는 표 1과 같은 특징 정보를 사용한다. 이때, 보조 특징(F_B)은 좌우 구분 특징(F_{LR}) 혹은 상하 구분 특징(F_{UD})과 함께 사용된다.

표 1 얼굴의 9상태를 구분하기 위한 특징

구분	특징	정의
주	전후 구분 특징 (F_{FB})	S: 얼굴 박스의 크기(면적) S_C : 최근 정면 상태의 얼굴 박스 크기 S_G : 현재 상태의 얼굴 박스 크기 최근 정면 상태의 얼굴 박스 크기에 대한 현재 상태의 얼굴 박스 크기의 비율 $R_{FB} = \frac{S_C}{S_G}$
	좌우 구분 특징 (F_{LR})	L: 두 눈 사이의 중심에서 얼굴 박스의 가장자리까지의 거리 L_L : 얼굴 박스의 좌측 가장자리까지의 거리 L_R : 얼굴 박스의 우측 가장자리까지의 거리 양 눈 사이의 중심에서 얼굴 박스의 좌측 가장자리까지의 거리에 대한 우측 가장자리까지의 거리 비율 $R_{LR} = \frac{L_R}{L_L}$
	상하 구분 특징 (F_{UD})	H: 두 눈 사이의 중심에서 차영상 얼굴 박스의 상단 끝까지의 거리 H_G : 최근 정면 상태에서의 거리 H_C : 현재 상태에서의 거리 양 눈 사이의 중심에서 차영상 박스의 상단 끝까지의 거리 비율 $R_{UD} = \frac{H_C}{H_G}$
특	링크 구분 특징 (F_W)	I: 눈 영역에서 화소값 상위 10%와 하위 10%의 평균 명암도 차이 I_L : 왼쪽 눈의 밝기 차이 I_R : 오른쪽 눈의 밝기 차이 눈 영역에서 화소값 상위 10%와 하위 10%의 차이를 구하고, 좌측 눈과 우측 눈의 차이값 $D_W = I_L - I_R$
	보조 특징 (F_B)	$C(x, y)$: 두 눈 사이의 중심 좌표 $C_C(x, y)$: 최근 정면 상태의 좌표 $C_G(x, y)$: 현재 정면 상태의 좌표 최근에 정의된 정면 상태의 눈 위치와 현재 상태의 눈 위치 차이 $D_F(x, y) = C_G(x, y) - C_C(x, y)$

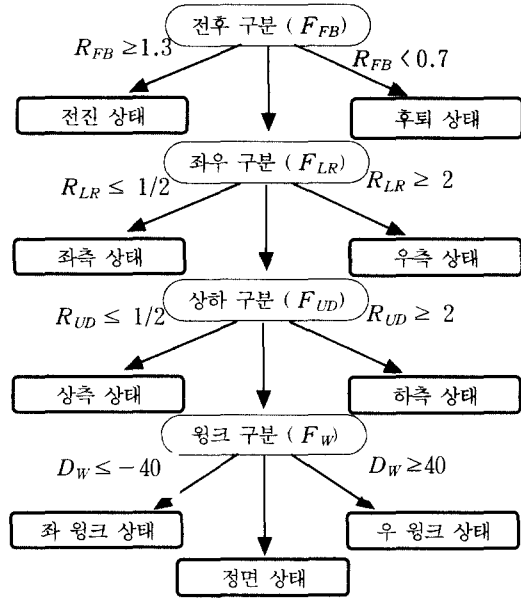


그림 5 계층적인 상태 인식 특징을 이용한 상태 인식도

4. 헤드 제스처 인식

4.1 헤드 제스처의 정의

본 논문에서 사용하는 헤드 제스처에 대한 정의 및 활용 효과는 표 2와 같다.

표 2 본 논문에서 정의한 13가지 헤드 제스처

헤드 제스처	코드	정의	효과
준비	G	정면 상태를 유지하는 채로 준비하는 제스처	시스템 초기화, 타 제스처 준비
상측이동	U	고개를 상측으로 들어올리는 제스처	화면의 상하좌우 이동
하측이동	D	고개를 하측으로 내리는 제스처	
좌측이동	L	고개를 좌측으로 돌리는 제스처	
우측이동	R	고개를 우측으로 돌리는 제스처	
전진	F	고개를 앞으로 이동하는 제스처	화면 확대
후퇴	B	고개를 뒤로 이동하는 제스처	화면 축소
좌 링크	WL	좌측 눈을 한 번 깜빡이는 링크 제스처	마우스 좌측 버튼의 클릭
우 링크	WR	우측 눈을 한 번 깜빡이는 링크 제스처	마우스 우측 버튼의 클릭
좌 더블 링크	WLD	좌측 눈을 두 번 연속으로 깜빡이는 링크 제스처	더블 클릭
우 더블 링크	WRD	우측 눈을 두 번 연속으로 깜빡이는 링크 제스처	더블 클릭
긍정	Y	고개를 상하로 끄덕이는 제스처	실행 및 Yes 응답
부정	N	고개를 좌우로 가로젓는 제스처	취소 및 NO 응답

4.2 오토마타 기법에 의한 헤드 제스처 인식

앞에서 정의한 헤드 제스처를 인식하기 위하여 본 논문에서는 오토마타 기법에 의한 헤드 제스처 인식 방법을 제안한다. 3장에서 정의한 얼굴의 상태를 헤드 제스처 인식을 위한 상태 열로 생성하고, 생성된 시퀀스들을 헤드 제스처로 인식하기 위해 표 3과 같은 오토마타 기법을 적용한다. 이때, 각 제스처는 준비 상태에서 시작한다. 그리고 상태의 반복 횟수를 나타내기 위한 첨자로 e, *, +를 사용하는데, e는 3회 생성, *는 0회 이상 생성(최대 5회까지 생성), +는 1회 이상 생성(최대 5회까지 생성)을 의미한다.

5. 실험

5.1 실험 환경

제안한 방법의 평가를 위해, PentiumIII 550MHz PC에 CCD 카메라를 이용하여 320×240 크기의 동영상 데이터를 획득하였고, Visual C++ 6.0으로 알고리즘을 구현하였다. 실험에 사용된 동영상 데이터는 뒷배경이 복잡하고 컴퓨터 화면을 주시하는 얼굴 영상이며, 총 8명으로부터 취득한 1,488 프레임의 영상 데이터로 구성되어 있다. 동영상 데이터는 정면, 좌측, 우측, 상측, 하측의 주시와 전진과 후퇴의 이동, 그리고 좌 윙크와 우 윙크의 제스처를 연속적으로 취하여 획득하였다.

표 3 헤드 제스처 인식에 사용되는 오토마타 기법

헤드 제스처	오토마타 기법
준비	$G \leftarrow g^e \mid Gg$
상측 이동	$U \leftarrow Gu^e \mid Uu$
하측 이동	$D \leftarrow Gd^e \mid Dd$
좌측 이동	$L \leftarrow Gl^e \mid Ll$
우측 이동	$R \leftarrow Gr^e \mid Rr$
전진	$F \leftarrow Gf^e \mid Ff$
후퇴	$B \leftarrow Gb^e \mid Bb$
좌 윙크	$W_L \leftarrow Gw_L^+ g$
우 윙크	$W_R \leftarrow Gw_R^+ g$
좌 더블 윙크	$W_{LD} \leftarrow W_L g^* w_L^+ g$
우 더블 윙크	$W_{RD} \leftarrow W_R g^* w_R^+ g$
긍정	$Y \leftarrow Gd^+ g^+ u^+ g^+ d^+ g^+ u^+ g$
부정	$N \leftarrow Gl^+ g^+ r^+ g^+ l^+ g^+ r^+ g$

5.2 실험 결과 및 분석

헤드 제스처 인식 이전 단계에서 행해지는 얼굴 영역 추출, 눈 영역 추출, 그리고 얼굴 상태 인식에 대한 결

과가 표 4에 나타나 있다.

얼굴 영역 추출의 경우, 99.3%의 추출 성공률을 보였는데, 얼굴 영역 추출에 실패한 총 11장(0.7%)의 얼굴 상태 영상은 모두 차영상 정보의 부족으로 발생하였다.

눈 영역 추출의 경우, 95.3%의 추출 성공률을 나타냈다. 실패한 영상은 총 70장(0.5%)인데, 이 중 얼굴 영역 추출에 실패한 영상을 제외하면 59장의 영상에서 눈 영역 추출에 실패하였다. 각각의 유형으로는 상측 상태에서 눈썹과 눈 영역이 결합되거나 눈썹을 눈으로 오추출한 경우가 18장(30.5%)이고, 눈을 감은 경우 눈에서 수평 에지 성분이 적게 추출되어 실패한 경우가 17장(28.8%), 윙크 상태에서 윙크 동작의 미숙으로 실패한 경우가 10장(17%), 눈을 머리카락이 가린 경우가 4장(6.7%), 그리고 기타의 경우가 10장(17%)으로 나타났다. 따라서, 보다 정교한 눈 영역 추출에 대한 알고리즘 개발이 필요하다.

표 4 얼굴 영역 추출, 눈 영역 추출, 상태 인식 결과 (추출 또는 인식된 프레임 수/총 프레임 수)

실험결과 상태	얼굴 영역 추출률	눈 영역 추출률	얼굴 상태 인식률
정면	99.9% (749/750)	98.3% (731/750)	97.5% (731/750)
상측	98.0% (98/100)	83.0% (83/100)	78.0% (78/100)
하측	99.0% (98/99)	95.0% (94/99)	94.0% (93/99)
좌측	97.9% (94/96)	92.7% (89/96)	87.5% (84/96)
우측	99.2% (117/118)	94.1% (111/118)	93.2% (110/118)
전진	98.9% (90/91)	94.5% (86/91)	94.5% (86/91)
후퇴	99.0% (94/95)	93.7% (89/95)	93.7% (89/95)
좌 윙크	100% (69/69)	92.8% (64/69)	92.8% (64/69)
우 윙크	98.6% (69/70)	92.9% (65/70)	92.9% (65/70)
상태별 총계	99.3% (1477/1488)	95.3% (1418/1488)	94.1% (1400/1488)

얼굴의 상태 인식의 경우, 얼굴의 9가지 상태를 인식한 결과 총 1,488장 중 1,400(94.1%)장의 영상에 대해 상태 인식에 성공하였다. 상태인식에 실패한 영상은 88장(5.9%)인데, 실험 결과에 의해 나타난 문제점은 크게 세 가지 원인에 의해서 발생하였다. 첫째, 얼굴 및 눈 추출에 실패하여 상태 인식에 필요한 정보를 얻지 못한 경우가 88장 중 70장(79.6%)이었다. 둘째, 얼굴의 상태를 구분하기 위해 사용한 계층적인 특징에서 주 특징과 보조 특징이 불일치하여 상태 구분이 모호한 경우로 14장(15.9%)이었다. 셋째, 윙크 동작이 미숙하여 눈의 열린 정도가 부정확한 경우가 3장(3.4%)이었고, 그리고 기타 1장(1.1%)으로 나타났다.

그림 6은 얼굴 상태 인식에 성공한 예이고, 그림 7은 실패한 예이다.

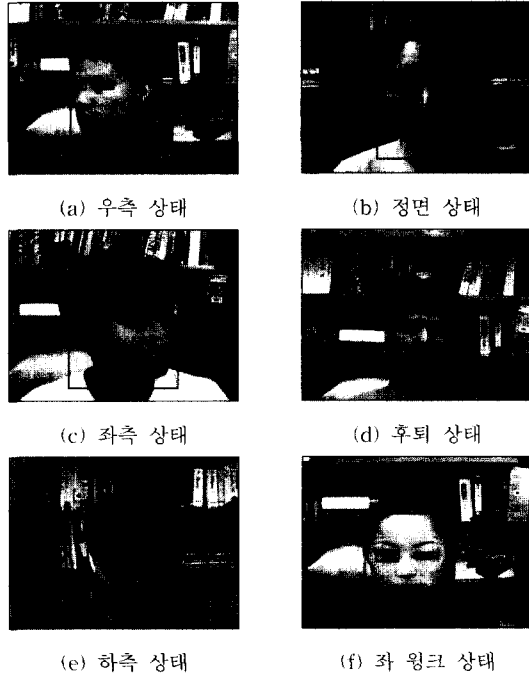


그림 6 얼굴 상태 인식에 성공한 영상의 예

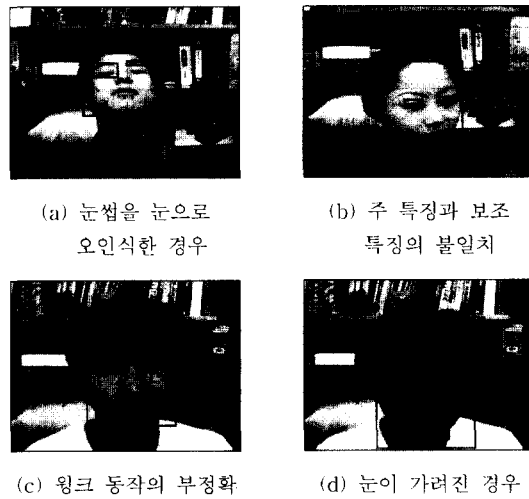


그림 7 얼굴 상태 인식에 실패한 영상의 예

본 논문에서 제안한 방법에 의해 헤드 제스처를 인식

한 결과가 표 5에 나타나 있다.

표 5 헤드 제스처 인식 결과

(인식된 제스처 수/총 제스처 수)

구 분		헤드 제스처 인식률		
헤드 제스처	포즈 제스처	준비	100% (8/8)	100% (56/56)
		상측	100% (8/8)	
		하측	100% (8/8)	
		좌측	100% (8/8)	
		우측	100% (8/8)	
		전진	100% (8/8)	
	행위 제스처	좌 윈크	87.5% (7/8)	85.4% (41/48)
		우 윈크	87.5% (7/8)	
		좌 더블 윈크	75.0% (6/8)	
		우 더블 윈크	87.5% (7/8)	
		긍정	87.5% (7/8)	
		부정	87.5% (7/8)	
		93.3% (97/104)		

실험 결과, 포즈 제스처(준비, 상측, 하측, 좌측, 우측, 전진, 후퇴)에서는 56개를 모두 인식하여 100%의 인식률을 얻었다. 이는 포즈 제스처의 경우, 일부 인식에 실패한 상태가 발생하더라도 주로 오토마타기법에 의해 제스처를 인식하는데 문제가 되지 않기 때문이다. 한편 행위 제스처(좌우 윈크, 더블 윈크, 긍정, 부정)의 경우에는 동일한 상태의 지속이 짧기 때문에 일부 상태의 미인식에 영향을 받기 쉬운 특성을 갖는다. 또한 실험용 동영상 데이터를 제공한 사람들 중 일부가 윈크 제스처의 동작에 미숙하여 제스처 미인식이 발생하는 경우가 많았다. 결과적으로, 행위 제스처는 총 48개중 41개를 인식하여 85.4%의 제스처 인식률을 나타냈다. 실패한 제스처 7개는 윈크 상태 미인식에 의한 윈크 제스처나 더블 윈크 제스처의 실패가 총 5개 발생했고, 하측 상태 혹은 좌우측 상태의 미인식에 의한 긍정·부정 제스처의 실패가 2개 발생했다. 결과적으로 93.3%의 헤드 제스처 인식 성공률을 나타냈다.

처리 시간을 PentiumIII(CPU clock = 550MHz)에서 측정된 결과, 얼굴 영역 추출의 경우 약 0.02초, 눈 영역 추출의 경우 약 0.01초, 상태 인식의 경우 약 0.005초로 나타났으며, 따라서 일련의 상태로 구성되는 제스처의 인식은 실시간에 처리될 수 있다.

6. 결론

본 논문에서는 헤드 제스처 인터페이스를 구현하기 위해, 헤드 포즈 및 윈크 정보에 의한 얼굴의 9가지 상

태를 정의하고, 인식된 얼굴 상태 열에 대하여 오토마타 기법을 적용하여 총 13가지의 헤드 제스처를 인식하는 방법을 제안하였다.

제안된 시스템에서는 다양한 제스처를 신속하게 인식하기 위해, 얼굴 영역 추출에는 최적의 얼굴색 정보와 적응적 움직임 정보를 이용하였고, 눈 영역 추출에는 기 추출된 얼굴 영역에 투영기법을 적용하여 눈의 후보 영역을 추출하고 이후, 눈의 기하학적 정보와 개폐정보를 이용하여 눈 영역을 결정하였다. 얼굴의 상태 인식에서는 얼굴의 방향, 크기 정보와 눈의 윙크 여부에 따라 얼굴의 여러 상태를 구분하였다. 마지막으로 헤드 제스처 인식에는 생성된 일련의 얼굴 상태 열에 오토마타 기법을 적용하였다.

향후에는 다양한 제스처 인식을 위해 정교한 얼굴 요소의 추출 방법에 대한 연구가 필요하다. 또한, 시선 추출 연구와의 상호 보완을 통해 보다 정교한 헤드 제스처 인터페이스를 구현하는 방법에 대한 연구가 요구된다.

참고 문헌

[1] R. Brunelli and T. Poggio, "Face Recognition: Features versus Templates," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 15, No. 10, pp.1042-1052, 1993.

[2] Y. H. Kwon and N. V. Lobo, "Face Detection Using Templates," *Proc. of Int'l Conf. on Pattern Recognition*, Vol. 1, pp.764-767, 1994.

[3] L. C. De Silva, K. Aizawa, and M. Hatori, "Detection and Tracking of Facial Feature by Using Edge Pixel Counting and Deformable Circular Template Matching," *Trans. on IEICE Information and System*, Vol. E78-D, No. 9, pp.1195-1207, 1995.

[4] S. A. Sirohey, "Human Face Segmentation and Identification," Tech. Report, CAR-TR_695, Center for Automation Research, Univ. of Maryland, 1993.

[5] H. H. Nagel, "Analysis Techniques for Image Sequences," *Proc. of Int'l Conf. on Pattern Recognition*, pp.186-211, 1978.

[6] T. C. Chang, T. S. Huang, and C. Novak, "Facial Feature Extraction from Color Images," *Proc. of Int'l Conf. on Pattern Recognition*, Vol. 2, pp.39-43, 1994.

[7] 유태웅, 오일석, "색채 분포 정보에 기반한 얼굴 영역 추출", *한국정보과학회논문지(B)*, 제24권, 제2호, pp. 180-192, 1997.

[8] C. H. Lee, J. S. Kim, and K. H. Park, "Automatic Human Face Location in a Complex Background using Motion and Color Information," *Pattern*

Recognition, Vol. 29, No. 11, pp.1877-1889, 1996.

[9] B. K. P. Horn and B. G. Schunck, "Determining Optical Flow," *Artificial Intelligence*, Vol.17, pp.185-203, 1981.

[10] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active Contour Models," *Int'l Journal of Computer Vision*, pp.321-331, 1988.

[11] 이경희, 변해란, "얼굴 요소의 영역 추출 및 Snakes를 이용한 윤곽선 추출", *한국정보과학회논문지*, 제27권, 제7호, pp.731-741, 2000.

[12] M. Xu and T. Akatsuka, "Detecting Head Pose from Stereo Image Sequence for Active Face Recognition," *Proc. of Int'l Workshop on Automatic Face- and Gesture-Recognition*, pp.82-87, 1998.

[13] Q. Chen, H. Wu, T. Shioyama, and T. Shimada, "Head Pose Estimation using Both Color and Feature Information," *Proc. of Int'l Conf. on Pattern Recognition*, pp.1051-1055, 2000.

[14] J. Sherrah and S. Gong, "Fusion of 2D Face Alignment and 3D Head Pose Estimation for Robust and Real-Time Performance," *Proc. of Int'l Workshop on Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, Corfu, Greece, pp.24-30, Sep. 26-27, 1999.

[15] M. L. Cascia, S. Sclaroff, and V. Athitsos, "Fast, Reliable Head Tracking under Varying Illumination: An Approach Based on Registration of Texture-Mapped 3D Models," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, No. 4, pp.322-336, Apr. 2000.

[16] 문병선, 오경환, "머리 움직임을 이용한 긍정/부정 의사 인식", *한국정보과학회 추계학술발표논문집*, 제25권, 제2호, pp.538-540, 1998.

[17] 최현일, 김수환, 이필규, "시각사용자 인터페이스를 위한 Multi-classifier 기반의 헤드 제스처 인식", *한국정보처리학회 추계학술발표논문집*, 제5권, 제2호, pp. 1295-1298, 1998.

[18] C. Morimoto, Y. Yacoob, and L. Davis, "Recognition of Head Gestures Using Hidden Markov Models," *Proc. of Int'l Conf. on Pattern Recognition*, pp.461-465, 1996.

[19] 박형철, 전병환, "동영상에서 최적의 얼굴색 정보와 움직임 정보에 기반한 얼굴 영역 추출", *한국정보과학회 논문지: 소프트웨어 및 응용*, 제27권 제2호, pp. 193-200, 2000. 2.

[20] 이상영, 함영국, 박래홍, "지식에 기초한 특징추출과 역전파 알고리즘에 의한 얼굴인식", *대한전자공학회 논문지*, 제31권, B편, 제7호, pp.928-937, 1994.

[21] 강영미, 정성환, "클라이언트/서버 환경에서의 내용 기반 얼굴 영상 검색 시스템의 구현", *한국정보과학회*

춘계학술발표논문집, 제25권, 제1호, pp.663-665, 1998.

- [22] A. L. Yuille, P. W. Hallinan, and D. S. Cohen, "Feature Extraction from Faces Using Deformable Templates," Int'l Journal of Computer Vision, pp.99-111, 1992.



오 승 택

2000년 2월 공주대학교 전자계산학과 졸업(이학사). 2000년 3월 ~ 현재 공주대학교 대학원 컴퓨터공학과 석사과정. 관심분야는 컴퓨터비전, 패턴인식, 영상처리, HCI 등



전 병 환

1989년 2월 연세대학교 전자공학과 졸업(공학사). 1991년 8월 연세대학교 대학원 전자공학과(공학석사). 1996년 8월 연세대학교 대학원 전자공학과(공학박사). 1996년 8월 ~ 1997년 2월 연세대학교 신호처리연구센터 전문연구원. 2000년 8월 ~ 2001년 6월 (주)모리아테크놀로지 연구소장. 1997년 3월 ~ 현재 공주대학교 정보통신공학부 컴퓨터전공 조교수. 관심분야는 컴퓨터비전, 패턴인식, 가상현실