Journal of Korea Multimedia Society Vol. 9, No. 2, February 2006(pp. 234-243)

간단한 얼굴 방향성 검출방법

정민교[†], 박지숙[††], 엄성용[†††], 조현희[††††]

요 약

최근 급속한 HCI(Human-Computer Interaction) 및 감시 기술의 발달로, 얼굴영상을 처리하는 다양한 시스템들에 대한 관심이 높아지고 있다. 그러나 이런 얼굴영상을 처리하는 시스템들에 대한 연구는 주로 얼굴인식이나 얼굴 표정분석과 같은 분야에 집중되었고, 얼굴의 방향성 검출과 같은 분야에는 많은 연구가 수행되지 못하였다. 본 논문은 두 눈썹과 아래 입술로 구성된 얼굴삼각형(Facial Triangle)이라는 특징을 이용하여 얼굴의 방향성을 쉽게 측정하는 방법을 제안한다. 특히, 하나의 이미지만을 사용하여 얼굴의 수평 회전각과 수직 회전각을 구하는 간단한 공식을 소개한다. 수평회전각은 좌·우 얼굴삼각형간의 면적비율을 이용하여 계산하고, 수직회전각은 얼굴삼각형의 밑변과 높이 비율을 이용하여 계산한다. 실험을 통해, 제안하는 방법은 오차범위 ±1.68° 내에서 수평회전각을 구할 수 있었고, 수직회전각은 회전각이 증가할수록 오류가 줄어드는 경향을 보여주었다.

A Simple Way to Find Face Direction

Min Gyo Chung[†], Jisook Park[††], Seong Yong Ohm[†††], Hyun Hee Jo[††††]

ABSTRACT

The recent rapid development of HCI and surveillance technologies has brought great interests in application systems to process faces. Much of research efforts in these systems has been primarily focused on such areas as face recognition, facial expression analysis and facial feature extraction. However, not many approaches have been reported toward face direction detection. This paper proposes a method to detect the direction of a face using a facial feature called *facial triangle*, which is formed by two eyebrows and the lower lip. Specifically, based on the single monocular view of the face, the proposed method introduces very simple formulas to estimate the horizontal or vertical rotation angle of the face. The horizontal rotation angle can be calculated by using a ratio between the areas of left and right facial triangles, while the vertical angle can be obtained from a ratio between the base and height of facial triangle. Experimental results showed that our method makes it possible to obtain the horizontal angle within an error tolerance of ±1.68°, and that it performs better as the magnitude of the vertical rotation angle increases.

Key words: Face Direction Detection(얼굴 방향성 검출), Facial Triangle(얼굴 삼각형), HCI

1. 서 론

최근 급속한 HCI(Human-Computer Interaction) 기술과 감시(surveillance) 기술의 발달로, 얼굴을 처리하는 다양한 응용 시스템들에 대한 관심이 고조되고 있다. 이런 시스템에 대한 연구 개발은 주로 얼굴

※ 교신저자(Corresponding Author) : 정민교, 주소 : 서울시 노원구 공릉 2동 126번지(139-774), 전화 : 02)970-5753, FAX : 02)970-5981, E-mail : mchung@swu.ac.kr
접수일 : 2005년 2월 25일, 완료일 : 2005년 9월 7일
[†] 정회원, 서울여자대학교 정보미디어대학 조교수
[††] 종신회원, 서울여자대학교 정보미디어대학 조교수 (E-mail : jspark@swu.ac.kr)
[†††] 서울여자대학교 정보미디어대학 교수 (E-mail : osy@swu.ac.kr)
[††††] 서울여자대학교 컴퓨터학과 석사과정 (E-mail : lilly@swu.ac.kr)

의 특징 추출이나 자동 얼굴 인식, 얼굴 표정 분석 [1-5, 9-12] 등과 같은 분야에 집중되고 있다. 얼굴 특징 추출에 관한 연구는 다시 잡음이 많은 배경에서 얼굴 자체를 찾아내는 연구나, 얼굴에서 눈, 코, 귀, 입, 눈썹과 같은 특징들을 검출하는 연구로 세분할 수 있다. 얼굴을 자동으로 인식하는 연구에는 많은 다양한 개념이나 기술들이 발표되었는데, 대표적으로는 통계적 해석에 의한 방법, 얼굴 특징들 간의 구조적 관계에 의한 방법, 신경망 이론에 의한 방법들이 있다. 마지막으로, 얼굴 표정에 관한 연구는 각기 다른 얼굴에 나타나는 인간의 다양한 표정들에 관한 인식이나, 그러한 표정들을 두드러지게 표현하는 캐리커처(caricature) 알고리즘의 응용 및 개발과 같은 분야에 집중되고 있다.

위와 같은 얼굴에 대한 다양한 연구에도 불구하고, 얼굴의 방향성 측정에는 많은 연구가 진행되지 못하였다. 게다가, 국내외에서 발표된 기존의 얼굴 방향성 측정 방법들은 여러 가지 비현실적인 가정이나 단점을 내포하고 있었다. 예를 들어, 특정 모델 비율(model ratio)을 사용함으로써 얼굴에 대한 특수 모델을 가정한 경우[6], 항상 두장의 스테레오 이미지가 필요한 경우[7], 얼굴 특징점간의 거리를 미리 알고 있어야 하는 경우[8], 최초의 입력영상이 항상 정면 영상이어야 하는 경우[13], 얼굴의 방향성 값을 몇 가지 부류로만 추정하고 판별하는 경우(정면 포즈, 좌측 포즈, 우측 포즈 등)[14], 얼굴의 수평 회전각만을 검출하는 경우[15] 등 다양하다.

본 논문에서는 컴퓨터 비전 및 이미지 처리 기술에 기반하여, 하나의 이미지만을 사용하여 간단하게 얼굴 방향성을 측정하는 방법을 제안한다. 특히, 두 개의 눈썹과 아래 입술로 이루어진 얼굴 삼각형이라는 특징을 이용하여, 얼굴의 수평 회전각과 수직 회전각을 쉽게 검출할 수 있는 계산식을 소개한다. 본 논문에서 제안하는 방법은 사람의 얼굴 구조에 대한 특별한 가정이 없을 뿐만 아니라, 실험을 통해 매우 정확하게 얼굴 회전각을 측정할 수 있음을 알 수 있었다.

제안된 얼굴 방향성 측정 기술은 인물 사진 자동 정렬, 유사한 방향성을 가진 얼굴 찾기, 사람이 머리를 좌우 또는 상하로 의도적으로 움직이는 상황을 자동으로 인식하는 헤드 제스처(head gesture) 인식 기술 등과 같은 다양한 HCI 응용 분야에서 유용하게 사용될 수 있을 것으로 생각된다. 본 논문에서 새롭게 제안하는 얼굴 방향성 측정 방법과 기존 방법간의 비교는 (표 1)에 잘 나타나 있다. 그리고 기존 방법에 관한 자세한 설명은 다음 절에 잘 나타나 있다.

본 논문은 다음과 같이 구성되어 있다. 먼저, 제 2절에서는 얼굴의 방향성을 검출하는 새로운 방법을 단계별로 자세하게 설명한다. 즉, 전처리 과정, 얼굴 삼각형 구축 과정, 얼굴 삼각형을 이용한 얼굴의 방향성 계산 등과 같은 단계를 상세하게 설명한다. 제 3절에서는 제안 방법의 유용성과 정확성을 입증하는 여러 가지 실험 결과를 제시하고, 마지막으로 제 4절에서는 본 논문의 결론을 짧막하게 맺는다.

표 1. 논문에서 제안하는 얼굴 방향성 측정방법과 기존방법간의 비교

방법	주요 특징 및 제약점	방향성 값 유형
Gee[6]	얼굴에 대한 특수 모델을 가정함. 즉, 4 가지의 형태의 얼굴특징 관련 수치($L_f/L_n/L_e/L_m$)를 사용한 특정 얼굴 모델 비율(model ratio)을 사용함.	얼굴평면 법선벡터
Minagawa[7]	2개의 카메라를 사용한 2개의 스테레오 이미지가 필요함.	얼굴평면 법선벡터
Ballard[8]	3D 공간상의 3개의 얼굴 특징점 사이의 거리를 미리 알고 있어야 함. 또한 이 방법은 사람의 얼굴이 바뀔 때마다 위 3개의 특징점 간의 거리에 대한 정보가 필요함.	얼굴평면 법선벡터
김[13]	복수의 얼굴 이미지가 필요함. 그리고 최초의 입력영상은 반드시 정면영상이어야 함.	수평 회전값 수직 회전값
이[14]	얼굴 방향성 값을 몇 가지 경우(정면, 좌회전, 우회전)로만 추정하여 판별함. 또한 많은 수의 학습용 얼굴 포즈 데이터를 필요로 함.	3가지 포즈값
박[15]	통계적 방법에 의해 얼굴 회전 값을 추정함. 따라서 분석적 모델이 필요함.	수평 회전값
제안 방법	얼굴삼각형을 이용한 간단한 얼굴 방향성 계산식을 사용함. 이 방법의 주요 특징은 방향성 계산의 단순함에 있음.	수평 회전값 수직 회전값

2. 관련연구

이 절에서는 얼굴의 방향성 검출과 관련하여 국내외에서 이미 발표된 몇 가지 방법들에 대하여 그 내용과 문제점을 간략하게 기술한다. 먼저 국외에서 발표된 연구논문들 중, Gee 방법은 얼굴을 나타내는 3개의 모델 비율(model ratio)을 사용함으로써 얼굴에 대한 특수한 모델을 가정하였다[6]. Gee가 사용한 3개의 모델 비율은 구체적으로 $R_m=L_m/L_f$, $R_n=L_n/L_f$, $R_e=L_e/L_f$ 이었는데, 여기서 L_f는 두 눈을 연결하는 수평축과 입을 통과하는 수평축간의 거리, L_n는 코밑에서 코끝까지의 높이, L_e는 두 눈 사이의 수평거리, L_m는 입과 코밑까지의 거리를 각각 나타낸다. 참고로, 서양인의 얼굴을 기준으로 실험에서 사용한 각 얼굴 모델 비율 값은 $R_m=0.4$, $R_n=0.6$, $R_e=1.0$ 이었다.

Minagawa는 얼굴 방향성 측정을 위해 2대의 카메라에서 찍은 2개의 스테레오 이미지를 사용하였다[7]. 먼저, 왼쪽 카메라로부터 획득한 스테레오 이미지에서 3개의 특징점, 즉 두 눈과 입을 검출하고, 각 특징점들의 2D 이미지 평면상의 위치를 구하였다. 같은 방법으로 오른쪽 카메라의 스테레오 이미지에서도 두 눈과 입을 검출하고, 각각의 2D 평면상의 좌표 값을 구하였다. 이렇게 좌우 카메라의 서로 대응하는 특징점들의 2D 평면상의 좌표 값을 알고, 카메라의 초점거리와 좌·우 카메라간의 거리를 알게 되면, 해당 특징점의 3D 좌표 값을 계산할 수 있게 된다. 즉, 두 눈과 입의 3D 공간상의 위치를 알 수 있게 된다. 끝으로, 이처럼 3D 위치를 알게 된 3개의 특징점을 동시에 통과하는 하나의 얼굴평면을 구성하여, 이 평면의 법선 벡터를 구하면, 이 법선 벡터가 바로 얼굴의 방향성을 나타내는 값이 된다. 그러나 Minagawa 방법은 얼굴의 방향성을 검출하기 위하여 항상 2개의 이미지가 필요하다는 단점을 가지고 있다.

Ballard는 3D 공간상에서 임의로 3개의 특징점을 정하고, 이 3개의 특징점 사이의 거리가 상수로 미리 주어졌을 경우, 얼굴 평면의 법선 벡터를 구하는 방법을 제시하였다[8]. 3개의 얼굴 특징점으로는 주로 두 눈과 입 또는 두 눈과 코를 사용하였다. 이 방법은 먼저 3D 공간의 원점(카메라의 투영점)과 2D 얼굴 이미지 상의 특징점 f_i(i=1,2,3)를 연결하는 3차원 벡터를 만들면서 시작된다. 그리고 이 벡터를 연장하여 2D 이미지상의 특징점 f_i 에 대응하는 3D 공간상의 얼굴 특징점 F_i(i=1,2,3)을 통과하도록 한 후, 상수로서 주어진 얼굴 특징점 F_i간의 거리를 이용하여 F_i의 3D 공간 좌표를 구하게 된다. 그러면 F_i를 포함하는 얼굴 평면이 만들어지게 되며, 따라서 얼굴 평면에 수직인 법선 벡터도 쉽게 구할 수 있게 된다. 그러나 Ballard 가 제시한 방법은 얼굴에 존재하는 3 가지 특징점간의 거리를 미리 알고 있어야 한다는 제약조건과, 사람의 얼굴이 달라질 때마다 3개의 특징점 사이의 거리에 대한 정보가 필요하다는 단점을 가지고 있다.

최근 국내에서도 얼굴 방향성 연구와 관련하여 몇 편의 논문들이 발표되었다[13-15]. 방법 [13]의 경우, 한 대의 카메라로 연속 촬영된 얼굴 영상의 회전 각도를 추론하는 방법을 제안하였으나, 최초의 입력 영상이 항상 정면 영상이 되어야 한다는 단점을 가지고 있었다. 따라서 임의의 좌·우 방향성을 가지는 얼굴 영상을 단독으로 처리하는 곳에는 사용이 불가능하다. 한편, 방법 [14]에서는 신경망을 사용하여 입력영상에 나타나는 얼굴의 위치와 포즈를 추정하였다. 그러나 이 방법은 포즈를 정면 포즈, 좌회전 포즈, 우회전 포즈의 3가지 경우로만 추정하여 판별하였고, 수직방향 얼굴회전과 관련한 포즈는 판별하지 못하였다. 게다가, 이 방법은 포즈 추정을 위해서 3000개(각 포즈별로 1000개씩) 이상의 포즈 학습 데이터를 필요로 하였다. 마지막으로, 논문 [15]에서는 몇 명의 실험자 얼굴을 통계적 방법으로 처리하여, 수평방향의 얼굴 회전각을 검출하는 식을 유도하고 있다(수직방향의 얼굴 회전각은 검출하지 못하고 있다). 그러나 이렇게 통계적 방법으로 유도된 식은 정확한 수학적 모델을 사용하여 분석적(analytical)으로 검증할 필요성이 제기되고 있다.

3. 새로운 얼굴 방향성 검출 방법

이 절에서는 본 논문에서 제안하는 얼굴 방향성을 측정하는 새로운 방법에 대하여 상세하게 설명하기로 한다. 먼저, (그림 1)은 제안 방법의 전체적인 구조를

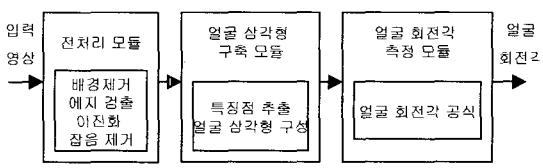


그림 1. 본 논문에서 제안한 얼굴 방향성 측정 방법의 논리적 구조

보여주고 있는데, 크게 3가지 단계 즉, 전처리 단계, 얼굴 삼각형 구축 단계, 얼굴 삼각형을 이용한 얼굴의 회전각 측정 단계로 구성되어 있다.

전처리 단계에서는 무엇보다도 먼저 입력영상에서 배경 영역과 얼굴 영역을 분할하여, 배경을 제거하고 얼굴 영역만을 추출하는 작업을 수행하게 된다. 이런 작업을 수행하는 다양한 영역 분할 알고리즘들이 이미 많이 개발되어 있지만, 본 논문에서는 Fleck[16,17]이 사용했던 스킨필터를 적용하여 입력 이미지로부터 얼굴영상을 추출한다. 이후 추출된 얼굴영상에 Sobel 연산자와 같은 미분 연산자(gradient operator)를 적용하여 에지(edge)를 검출하고, 검출된 에지를 담고 있는 에지맵(edge map)을 이진영상(binary image)으로 만든 후, 최종적으로 잡음(noise) 제거 방법을 사용하여 이진영상에 존재하는 잡음을 제거한다.

다음 단계에서는 위 전처리 단계에서 완성된 잡음 제거 영상으로부터 3개의 특징점 (즉, 두 눈썹과 아래 입술)의 위치를 검출하고, 이를 이용하여 얼굴 삼각형을 구성하게 된다. 여기서, 얼굴 삼각형(facial triangle)이란 위 3개의 특징점으로 형성된 얼굴 평면(facial plane) 위에 그려 놓은 가상의 다각형(virtual polygon) 이라고 생각할 수 있다. 실제로, 얼굴 삼각형을 만들기 위해서는 위 특징점 외에도. 3개의 얼굴 축(facial axe)도 같이 생각하게 되는데, 3개의 얼굴 축은 두 눈썹을 연결하는 수평축, 아래 입술의 중앙을 직각으로 관통하는 두개의 수평축과 수직축으로 이루어진다 ((그림 3) 참조). 그러면, 얼굴 삼각형은 두 눈썹을 연결하는 수평축이 얼굴 윤곽선과 만나는 두개의 교차점과 아래 입술의 중앙 점을 서로 연결하여 만들 수 있다 ((그림 4) 참조). 여기서, 수직축은 하나의 큰 얼굴 삼각형을 두개의 작은 얼굴 삼각형(즉, 좌우측 볼에 해당하는 삼각형)으로 분할하는 역할을 수행하게 된다.

마지막으로 얼굴 회전각 측정 단계에서는, 전 단계에서 구축한 얼굴 삼각형을 실측하여 필요한 데이터를 구하고, 실측 데이터를 본 논문에서 제안하는 얼굴 회전각 측정 공식의 입력으로 사용하여, 최종적으로 얼굴의 회전각을 계산하게 된다.

3.1 전처리 과정

전처리 단계의 첫 과정은 입력영상으로부터 얼굴이 존재하는 영역과 기타 배경 영역을 분리하는 작업이다. 본 논문에서는 Fleck의 스킨필터[16, 17]를 사용하여 입력영상으로부터 얼굴 영역을 검출한다. 이 방법은 먼저 영상의 R, G, B 값을 로그-보색 칼라 표현(log-opponent color representation) 값인 I, Rg, By으로 변환시킨다.

$$L(x) = 105 * \log_{10}(x + 1 + n)$$

$$I = L(G),\ Rg = L(R) - L(G),\ By = L(B) - \frac{L(R) + L(G)}{2}$$

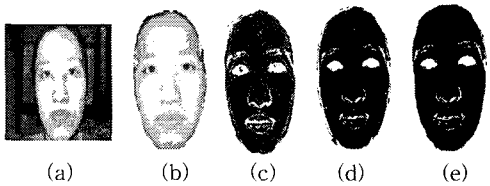
위의 $L(x)$ 식에서 n값은 [0,1] 사이에 존재하는 난잡음(random noise)을 나타낸다. 이렇게 I, Rg, By 값을 얻게 되면, 아래 식을 사용하여 색상(hue)과 채도(saturation) 값도 얻을 수 있게 된다.

$$Hue = \tan^{-1}(Rg/By),\ Satuartion = \sqrt{Rg^2 + By^2}$$

이제, 사람의 피부 색깔에 해당하는 색상과 채도 값이, a) 색상=[110, 150], 채도=[20, 60] 이거나, b) 색상=[130, 170], 채도=[30, 130]이라는 사실을 이용하여, 위 조건을 만족하는 모든 화소값을 마킹(marking)함으로써 입력영상으로부터 얼굴에 해당하는 영역만을 검출할 수 있게 된다. 위 스킨필터를 적용하여 입력영상으로부터 얼굴영상을 추출한 예가 [그림 2(a)]와 [그림 2(b)]에 잘 나타나 있다.

이렇게 영역 분할된 이미지는 에지를 검출하기 위해 미분 연산자를 적용하게 되는데, 미분 연산자는 연산자 마스크에 있는 가중치와 이웃 화소들의 화소값을 이용하여 현재 화소의 화소값을 구하는 연산자이다. [그림 2 (c)]는 Sobel 미분 연산자를 사용하여 에지를 검출한 영상의 예를 보여주고 있다. 여기서 이미지 명도 값은 미분값(gradient)의 크기를 나타낸다.

다음으로, 검출된 에지를 포함하고 있는 에지맵 영상은 이진화(binarization) 과정과 잡음을 제거하는 두 과정을 거치게 된다. 이진화 과정은 화소의



(a) (b) (c) (d) (e)

그림 2. 전처리 단계에서 생성되는 중간 이미지들 : (a) 원 이미지, (b) 배경을 제거한 이미지, (c) 에지맵(edge map), (d) 이진화된 에지맵, (e) 잡음을 제거한 이미지

미분값에 임계값(threshold)을 적용하여 수행한다. 즉, 미분값의 크기가 특정 임계값보다 크면 백색 화소로 처리하고, 그러지 않으면 검정색 화소로 처리한다. 그러나 영상의 히스토그램 분석을 통해서는 정확한 임계값 설정이 어렵기 때문에, 본 논문에서는 경험론적 임계치인 150 전후 값을 사용하여 이진화를 수행한다.

이진화 과정 후 얼굴에 존재할 수 있는 다양한 형태의 잡음을 제거하기 위해, 잡음 제거 방법 중의 하나인 모폴로지(morphology) 닫힘(closing) 연산을 사용한다. 닫힘 연산의 구조요소(structuring element)로서 반지름 1인 디스크(disk)를 사용함으로써, 영상 내에 존재하는 미세한 구멍(hole), 좁고 긴 갭(gap)과 같은 잡음을 효과적으로 제거하고, 얼굴의 윤곽선을 한층 부드럽게 한다. 이진화 결과 영상은 [그림 2(d)]에, 모폴로지 연산을 적용한 결과 영상은 [그림 2(e)]에 각각 잘 나타나 있다.

3.2 얼굴 삼각형 구축

본 절에서는 [그림 2(e)]에 있는 것과 같은 최종 이미지를 사용하여 얼굴 삼각형을 구축하는 과정을 좀더 자세하게 설명한다. 기본적으로 얼굴 삼각형은 3개의 특징점 즉, 두 눈썹과 아래 입술을 사용하여 구축하게 된다.

[그림 3]에 있는 얼굴 모델을 생각해보자. 여기서, 얼굴 이미지는 아래쪽에서 시작하여 위쪽으로 스캐닝 된다고 가정한다. 그러면 L_1은 각 라인을 왼쪽에서 오른쪽으로 스캐닝 할 때, 처음으로 만나는 백색 점의 위치를 나타내고, L_2는 라인의 스캐닝 방향이 오른쪽에서 왼쪽으로 반대로 바뀌었을 때, 만나게 되는 첫 번째 백색 점의 위치를 나타낸다. 개념적으로, L_1과 L_2는 아래 입술에 있는 두개의 최하위 점을 나타낸다. 비슷한 방법으로, 두 눈썹의 최상위 점을 구하고 각각 E_1 과 E_2로 표시한다. 이렇게 검출된 4개의 점 L_1, L_2, E_1, E_2을 사용하여 L축(Lip axis), E축(Eyebrow axis), S축(Symmetry axis), 3개의 축을 [그림 3]에 있는 것처럼 구한다. L축과 S축은 각각 아래 입술의 중앙인 $(L_1+L_2)/2$ 점을 직각으로 통과하는 수평축과 수직축을 나타내고, E축은 두 눈썹의 중간점 $(E_1+E_2)/2$을 통과하면서 L축과 평행한 축을 나타낸다.

얼굴 삼각형을 성공적으로 구축하기 위해서는 위 3개의 얼굴 축을 정확히 그릴 수 있어야 한다. 그러나 [그림 3(b)]와 같은 측면 얼굴 모습에서는, 왼쪽 또는 오른쪽 눈썹이 전혀 보이지 않거나 또는 부분적으로 모호하게 보일 수 있다. 이런 경우에도, 본 논문은 아래와 같은 절차를 사용하여 3개의 얼굴 축을 정확히 구축할 수 있었다.

1. 위에서 설명한대로 먼저 L_1, L_2를 구하고, 그 중점을 $L_m=(x_m, y_m)$이라고 한다. 즉, $L_m=(L_1+L_2)/2$.

2. 같은 방법으로, 왼쪽 눈썹 위치 $E_1=(x_1, y_1)$과 오른쪽 눈썹 위치 $E_2=(x_2, y_2)$를 각각 구한다.

3. 이제 x_m을 기준으로, x_1과 x_2의 상대적인 값에 따라 얼굴 축을 아래와 같이 결정한다.

(1) $x_1 \leqq x_m$이고 $x_m \leqq x_2$인 경우는, 두개의 눈썹이 분명하게 보이는 경우이다. 이 경우에는, 위에서 설명한 방법대로 L_1, L_2, E_1, E_2을 사용하여 3개의 얼굴 축을 쉽게 구할 수 있다.

(2) $x_1 \leqq x_m$이고 $x_2 \leqq x_m$인 경우는, 왼쪽 눈썹은 보이지만 오른쪽 눈썹은 보이지 않는 경우이다. 따라서 L축과 S축은 $L_m=(x_m, y_m)$에서 서로 직교하도록 구성하고, E축은 E_1을 통과하면서 L축과 평행하도록 그리면 된다.

(3) $x_m \leqq x_1$이고 $x_m \leqq x_2$인 경우는, 왼쪽 눈썹은 보이지 않지만 오른쪽 눈썹이 보이는 경우이다. 이 경우 얼굴 축을 구축하는 방법은 위 단계 (2)와 비슷하다. 다만, E 축이 E_1대신에 E_2를 통과하도록 그리면 된다.

(4) 마지막으로 $x_2 \leqq x_m$이고 $x_m \leqq x_1$인 경우는 현실생활에서 일어날 수 없는 경우이기 때문에 무시한다.

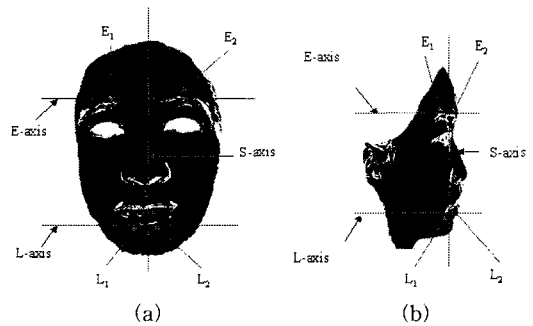


그림 3. 얼굴 모델 : (a) 정면 얼굴 모델, (b) 측면 얼굴 모델

[그림 4]는 위에서 설명한 세 개의 얼굴 축을 이용하여 얼굴 삼각형을 만드는 방법을 보여 준다. 여기서 E축과 얼굴 윤곽선이 만나는 교차점은 P_1과 P_2이고, S축은 L 축과는 P_3에서 만나고 E축과는 P_4에서 만나게 된다. 이때 P_1, P_2, P_3을 연결하는 삼각형을 얼굴 삼각형이라고 정의한다. 더욱이, 전체 얼굴 삼각형은 좌우 볼에 해당하는 두개의 작은 얼굴 삼각형으로 나뉠 수 있다. 왼쪽 볼에 해당하는 얼굴 삼각형은 기호로는 LFT(Left Facial Triangle)로 표시하고, 그 면적은 A_{LFT}로 나타낸다. LFT는 P_1, P_3, P_4 세 개의 점으로 이루어진다. 비슷하게, 오른쪽 볼에 해당하는 얼굴 삼각형은 P_2, P_3, P_4 세 개의 점으로 이루어져 있으며, 기호로는 RFT(Right Facial Triangle)를 사용하여 나타내고, 그 면적은 A_{RFT}로 나타낸다.

3.3 얼굴 회전각 측정

본 절에서는 위에서 구한 얼굴 삼각형에 기반하여 얼굴 회전각을 측정하는 방법을 기술한다. [그림 5(a)]는 얼굴의 수평 회전 각을 검출하기 위한 기하학적 모델이고, [그림 5(b)]는 수직 회전각을 검출하기 위한 기하학적 모델이다.

3.3.1 수평 회전각 측정

[그림 5(a)]는 사람머리를 위에서 내려다본 모양으로써, A_0, B_0는 사람의 두 귀를 나타내고, N_0는 코의 위치를 나타낸다. 또한 위 세 점이 카메라 이미지 평면에 각각 A_0', B_0', N_0' 로 매핑(mapping) 된다고 가정하자. 그러면, 머리가 수평 방향으로 임의의 각 h 만큼 회전 하였을 때, A_0, B_0, N_0는 각각 새로운

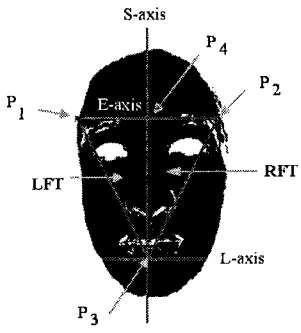


그림 4. 얼굴 삼각형 : 전체 얼굴 삼각형은 두개의 작은 얼굴 삼각형인 LFT와 RFT을 포함한다

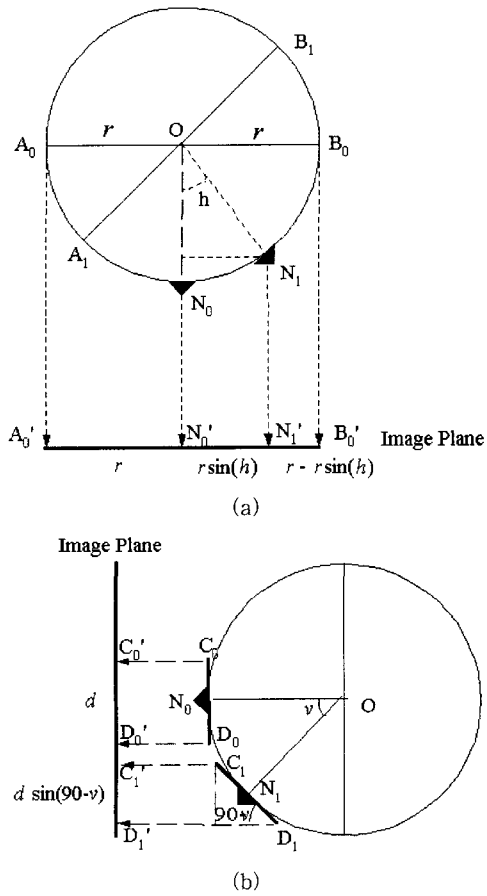


그림 5. 얼굴 방향성 검출을 위한 기하학적 모델 : (a) 수평 회전 모델, (b) 수직 회전 모델

위치인 A_1, B_1, N_1로 이동하게 된다. 특별히, 3D 공간상에서 코의 위치가 N_0에서 N_1로 움직일 때, 카메라 이미지 평면상에서 코의 위치는 N_0'에서 N_1'로 $r\sin h$ 만큼 이동하게 된다. 그런데, 코는 S축 상의 점이기 때문에, 카메라 평면상에서의 코의 이동은 실제로는 S축의 평행이동에 해당하게 된다. 이로 인해, 얼굴 삼각형을 구성하는 두 점인 P_3, P_4가 E축과 L축을 따라 이동하게 되는 결과를 가져오게 되고, 궁극적으로는 두 점의 평행 이동이 LFT와 RFT의 면적에 변화를 초래하게 된다.

P_3과 P_4사이의 거리를 s 라고 하면, 회전각과 LFT/RFT 면적 간의 관계는 다음과 같이 표현 될 수 있다.

$$A_{LFT} = \frac{1}{2} sr(1 + \sin h)$$

$$A_{RFT} = \frac{1}{2} sr(1 - \sin h)$$

$$A_{LFT} - A_{RFT} = sr \sin h$$

$$A_{LFT} + A_{RFT} = sr$$

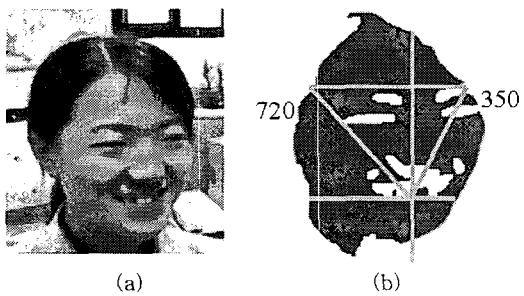
LFT와 RFT 면적간의 비율인 R_{area}을 아래와 같이 정의하면, 수평 회전 각 h는 아래 식 (1)과 같이 구할 수 있다.

$$R_{area} = \frac{A_{LFT} - A_{RFT}}{A_{LFT} + A_{RFT}} = \sin h$$
$$h = \sin^{-1}(R_{area}) \quad (1)$$

[그림 6]은 식 (1)을 이용하여 얼굴의 수평 회전각을 계산하는 예를 보여주고 있다. [그림 6(b)]는 [그림 6(a)]에 있는 입력영상에 대한 얼굴 삼각형과 그 LFT, RFT를 보여주고 있다. 여기서, $A_{LFT} = 720$이고 $A_{RFT} = 350$이다. 그러므로 $R_{area} = (720 - 350)/(720 + 350) = 0.346$이기 때문에, 수평회전각은 $h = \sin^{-1}(0.346) \simeq 20.25°$이 된다.

3.3.2 수직 회전각 측정

얼굴의 수직 회전각을 계산하는 방법을 이해하기 위해서, [그림 4]에 있는 얼굴 삼각형을 다시 살펴보기로 하자. 기본적인 아이디어는 머리를 수직으로 회전할 때, 수직 선분인 $\overline{P_3P_4}$는 크게 변화하는 대신, 수평 선분인 $\overline{P_1P_2}$는 변화가 거의 없다는 사실이다. [그림 5(b)]는 사람 머리를 옆에서 본 측면도이다. 여기에서, 선분 $\overline{C_0D_0}$는 얼굴 삼각형의 수직 선분인 $\overline{P_3P_4}$에 대응하며, 선분 $\overline{C_1D_1}$는 선분 $\overline{C_0D_0}$가 각도 v 만큼 회전했을 때의 새로운 위치를 표시한 것이다. 그러면,



(a) (b)

그림 6. 수평 얼굴 회전각 계산 예 : (a) 입력 이미지, (b) 좌우 얼굴 삼각형 LFT와 RFT

선분 $\overline{C_1D_1}$이 카메라 평면에 투사되어 나타나는 길이는 $d\sin(90 - v) = d\cos v$ 가 된다. 여기서, d는 선분 $\overline{C_1D_1}$의 실제 길이를 나타낸다. 이런 사실에 기초하여, 얼굴의 수직 회전각 v 을 계산하는 식을 아래와 같이 유도할 수 있다.

$$\hat{d} = d\cos v$$
$$v = cos^{-1}\left(\frac{\hat{d}}{d}\right) = \cos^{-1}\left(\frac{\hat{d}}{\hat{c}k}\right) \quad (2)$$

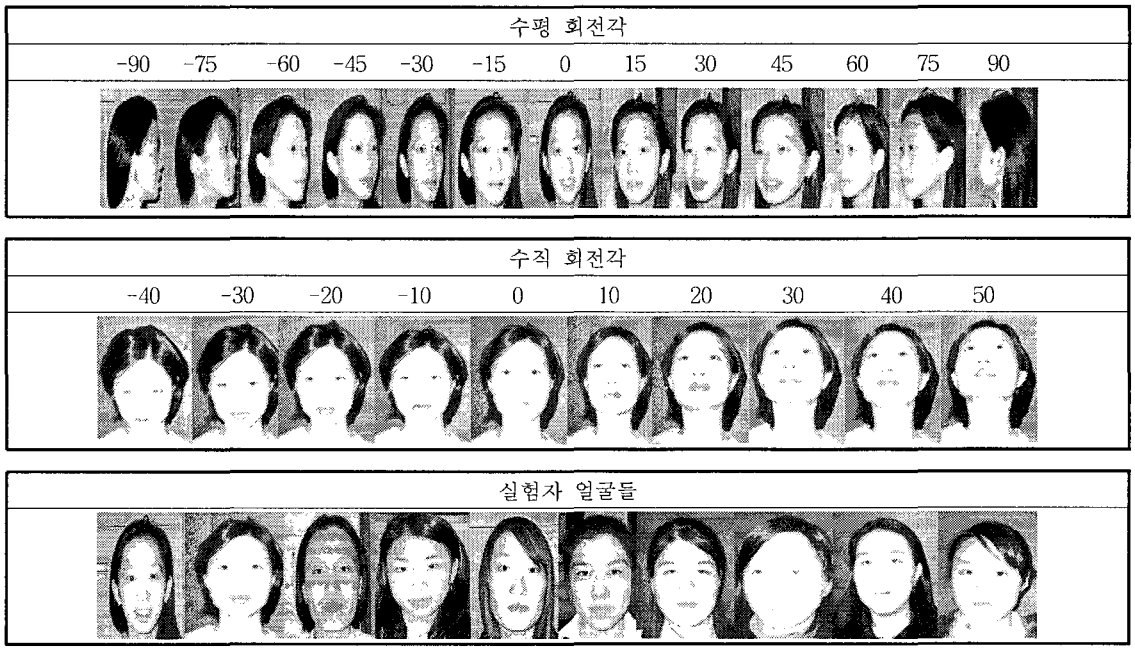
여기에서, 선분 $\overline{C_1D_1}$이 카메라에 투사되어 나타나는 길이인 $\hat{d}$는 얼굴 삼각형의 P_3과 P_4 사이의 거리를 측정하여 얻고, $\hat{c}$는 P_1과 P_2사이의 거리를 측정하여 얻는다. k는 상수로서 2/3을 사용하였다. 위식에서 우리는 $d = ck \simeq \hat{c}k$ 관계식을 사용하였다. 왜냐하면, 수평 선분인 $\overline{P_1P_2}$는 수직 회전각에 상관없이 일정한 길이를 유지하기 때문이다 (즉, $c \simeq \hat{c}$).

4. 실험 및 결과 분석

실험을 위한 얼굴 영상은 캐논 IXUS 400 디지털 카메라를 사용하여 획득하였다. 10 명의 실험자 얼굴을 사용하였고, 각 실험자마다 13개의 수평 회전(0, ±15, ±30, ±45, ±60, ±75, ±90) 이미지와 10개의 수직 회전(0, ±10, ±20, ±30, ±40, 50) 이미지를 획득하였다. [표 2]는 실험에 사용된 10명의 실험자 얼굴과 수평 및 수직회전 측정을 위한 샘플 입력영상을 보여주고 있다.

식 (1)에 제시된 수평 회전각 계산식의 정확성을 입증하기 위하여, 65개 이미지(5 사람마다 13개 수평 방향 이미지) 각각에 대하여 A_{LFT}과 A_{RFT}을 먼저 구하고, 이를 이용하여 R_{area}을 계산하였다. 그래서, 회전각에 대한 R_{area}값을 그래프로 표시하면, [그림 7]에 나타난 것과 같은 사인(sine) 곡선 관계가 나타나게 된다. 이런 사인 곡선 관계는 식 (1)에서 제시한 식이 얼굴의 수평 회전 각도를 측정하는데, 충분히 신뢰할 수 있는 수식임을 뒷받침하고 있다. 좀더 자세한 오류 분석과 실험 결과에 따르면, 본 논문에서 제안한 방법을 사용하여 R_{area}는 오차 범위 ±0.086% 범위 내에서 구할 수 있었으며, 수평 회전각은 오차 범위 ±1.68° 범위 내에서 구할 수 있었다.

표 2. 수평/수직회전 측정을 위한 샘플 영상과 실험에 참여한 10명의 실험자 얼굴들



눈썹이 가리지 않고 보이는 것이 얼굴의 회전각 측정에 매우 중요하다. 그런데, 눈썹은 여러 가지 요인들 때문에 가려져 보이지 않은 경우가 생길 수 있다. 예를 들어, 안경이나 이마에 늘어뜨린 긴 머리 때문에 눈썹이 부분적으로 가려지는 경우이다. 눈썹이 분명하게 보이는 경우에 비교하여, 눈썹이 부분적으로 가려지는 경우는 실제로 얼굴의 회전각 측정에 미소한 오류가 발생하였다. 실험결과에 의하면, 안경을 낀 경우에는 R_{area}을 ±0.098% 오차범위 내에서 구할 수 있었으며, 긴 머리를 가진 경우에는 R_{area}을 ±0.19% 오차범위 내에서 구할 수 있었다 (이전 문단에서 언급한 것처럼, 정상적인 실험자의 경우에는 R_{area}을 ±0.086% 오차범위 내에서 구할 수 있었다). 하지만, 이런 정도의 오류는 얼굴의 회전각을 측정하는데 큰 영향을 주지는 못하였다.

식 (2)에서 제시한 수직 회전각 측정 공식을 사용했을 때 발생한 오류는 [그림 8]에 나타나 있다. y축은 실제 각도와 식 (2)로 계산된 각도의 차이에 대한 절대값을 나타낸다. 식 (1)과 비교하여, 식 (2)는 상당히 큰 오류를 발생시키는데, 여기에는 부정확한 특징점 위치, 단순한 카메라 모델 등과 같은 불확실성 때문인 것으로 생각된다. 그러나 수직 회전 각이 커질수록 오류가 줄어드는 바람직한 경향을 보여주고 있다.

사람이 머리를 좌우 또는 상하로 의도적으로 움직이는 상황을 자동으로 인식하는 헤드 제스처(head

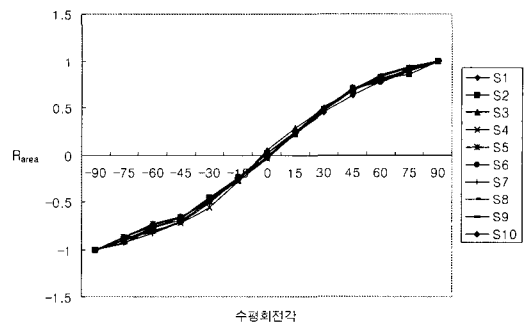


그림 7. 수평 회전각에 대한 R_{area} 값 그래프

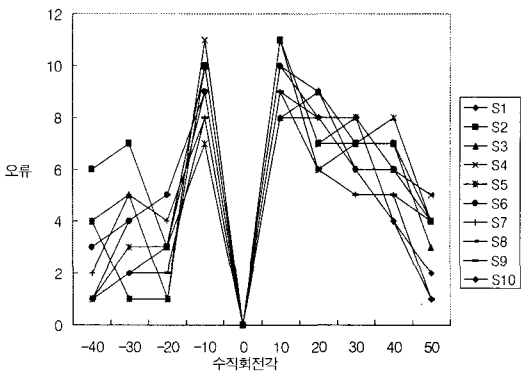


그림 8. 수직 회전각에 대한 오류 그래프

gesture) 인식기술은, HCI 응용 시스템을 구성하는 주요 요소 기술 중의 하나이다. 머리를 상하로 움직이는 제스처를 *head nod* 라고 부르며 주로 찬성이나 이해를 표시할 때 사용된다. 그러나 머리를 좌우로 움직이는 제스처는 *head shake* 라고 부르는데 반대나 부정을 표시할 때 많이 사용된다. 본 논문에서 제안하는 얼굴의 수직/수평 방향성을 측정하는 방법은 이와 같은 head nod 또는 head shake를 인식하는데도 유용하게 사용될 수 있다. 또한, 제안된 방법은 인물 사진 자동 정렬이나 유사한 방향성을 가진 얼굴 찾기 등과 같은 다양한 응용 분야에서도 편리하게 사용될 수 있다. 여기서, 인물 사진 자동 정렬이란 사진에 찍힌 얼굴의 방향성에 따라 사진의 좌우/상하 마진을 자동으로 설정하는 기능을 말한다.

5. 결 론

본 논문에서는 얼굴 삼각형이라는 특징에 기반하여, 하나의 입력 영상만을 사용하여 얼굴의 방향성을 측정하는 간단한 방법을 제안하였다. 얼굴 삼각형 구축을 위해 두개의 눈썹과 아래 입술을 주요 특징점으로 사용하였고, 위 특징점 들을 사용하여 가상의 3개의 얼굴 축을 또한 구축하였다. 여기서, 3개의 얼굴 축은 두 눈썹을 연결하는 수평축, 아래 입술의 중앙을 직각으로 관통하는 두개의 수평축과 수직축으로 이루어진다. 그러면, 얼굴 삼각형은 두 눈썹을 연결하는 수평축이 얼굴 윤곽선과 만나는 두개의 교차점과 아래 입술의 중앙 점을 서로 연결하여 만들어지는 삼각형으로 정의하였다.

구축된 얼굴 삼각형에 기반하여, 얼굴의 수평 회전각과 수직 회전각을 계산하는 간단한 공식을 유도하였다. 논문에서 제안하는 방법은 사람의 얼굴 구조에 대한 특별한 가정이 없을 뿐만 아니라, 실험을 통해 상당히 정확하게 얼굴 회전각을 측정할 수 있음을 알게 되었다. 제안된 얼굴 방향성 측정 기술은 인물 사진 자동 정렬, 유사한 방향성을 가진 얼굴 찾기, 헤드 제스처 인식 등과 같은 다양한 HCI 응용 분야에서 널리 유용하게 사용될 수 있으리라고 생각한다.

참 고 문 헌

[1] B. Moghaddam, "Principal manifolds and bayesian subspaces for visual recognition," *Proc. of the 7th International Conference on Computer Vision*," pp. 1131-1136, 1999.

[2] M. A. Turk and A. P. Pentland, "Eigenfaces for recognition," *J. Cognitive Neurosci.*, Vol. 3, No. 1, pp. 71-86, 1991.

[3] C. Liu and H. Wechsler, "Probabilistic reasoning models for face recognition," *Proc. of Computer Vision and Pattern Recognition*, pp. 827-832, 1998.

[4] G. Guo, S. Li, and K. Chan, "Face recognition by support vector machines," *Proc. of the International Conferences on Automatic Face and Gesture Recognitions*, pp. 196-201, 2000.

[5] H. Sahbi and N. Boujemaa, "From coarse to fine skin and face detection," *The 8th ACM International Conference on Multimedia*, pp. 432-434, 2000.

[6] Andrew Gee and Robert Cipolla, "Determining the gaze of faces in images," *Image and Vision Computing*, Vol. 12. No. 10, pp. 639-647, 1994.

[7] T. Minagawa, H. Saito, and S. Ozawa, "Face-direction estimating system using stereo vision," *Proc. of the 23rd Annual International Conference of the IEEE Industrial Electronics Society, New Orleans*, Vol. 3, pp. 1454-1459, 1997.

[8] Philippe Ballard and G. Stockman, "Controlling a computer via facial aspect," *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. 25. No. 4, pp. 669-677, 1995.

[9] 민경필, 전준철, 박구락, "스킨컬러와 변형모델에 기반한 컬러영상으로부터의 얼굴 및 얼굴 특성영역 추출," 한국인터넷정보학회 논문지, 제3권, 제6호, pp. 13-23, 2002.

[10] Yongsheng Gao and M. Leung, "Face recognition using line edge map," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 24, No. 6, pp. 764-779, 2002.

[11] 이찬호, 박주철, 최형일, "정규화된 웨이블렛 특징에 의한 얼굴 인식," 한국정보과학회 논문지,

제27권, 제10호, pp. 1046-1053, 2000.

[12] 김용균, 이옥경, 이창수, 오해석, "얼굴 캐리커쳐 생성시스템," 한국정보처리학회 추계학술대회, 제8권, 제1호, pp. 887-890, 2001.

[13] 김호영, "얼굴 회전 각도의 정량적 추론," 숭실대학교 석사학위 논문, 1998.

[14] 이재훈, 전인자, 이정훈, 이필규, "오류-역전파 신경망 기반의 얼굴 검출 및 포즈 추정," 한국정보처리학회 논문지 B, 제9-B권, 제6호, pp. 853-862, 2002.

[15] 박지숙, 동지연, "얼굴특징 정보를 이용한 얼굴 방향성 검출," 한국인터넷정보학회 논문지, 제4권, 제6호, pp. 57-67, 2003.

[16] Margaret Fleck, David Forsyth, and Chris Bregler, "Finding naked people," *European Conference on Computer Vision*, Vol. 2, pp. 593-602, 1996.

[17] David Forsyth and Margaret Fleck, "Automatic detection of human nudes," *International Journal of Computer Vision*, Vol. 32, No. 1, pp. 63-77, 1999.

정 민 교

1985년 서울대학교 컴퓨터공학과 졸업(공학사)
1987년 KAIST 컴퓨터학과 졸업(공학석사)
1996년 University of Iowa 컴퓨터학과 졸업(공학박사)
1987년~2001년 Korea Telecom 멀티미디어 연구소 (선임연구원)
2001년~2003년 Vivcom Inc. (Founder & Engineering Director)
2003년~현재 서울여자대학교 정보미디어대학 조교수
관심분야 : 컴퓨터 비전, 영상처리/코딩, 멀티미디어 검색, 컨텐츠 보호, 생체 인식

박 지 숙

1990년 한국과학기술원 전산학과 졸업(학사)
1992년 서울대학교 대학원 컴퓨터공학과 졸업(석사)
1998년 서울대학교 대학원 컴퓨터공학과 졸업(박사)
1999.5~2000.4 한국전자거래진흥원 EC진흥부 팀장
2000.5~2002.2 삼성SDS Biz. Modeling팀 과장
2002.3~현재 서울여자대학교 정보미디어대학 조교수
관심분야 : 멀티미디어 시스템, 전자상거래

엄 성 용

1985년 서울대학교 컴퓨터공학과 졸업(학사)
1987년 서울대학교 대학원 컴퓨터공학과 졸업(석사)
1992년 서울대학교 대학원 컴퓨터공학과 졸업(박사)
1992년~1993년 컴퓨터신기술공동연구소 특별연구원
1993년~1995년 University of California, Irvine에서 Post-Doc
1996년~현재 서울여자대학교 정보미디어대학 교수
관심분야 : 컴퓨터 그래픽스, 플래시메모리 시스템, 임베디드 시스템

조 현 희

2004년 서울여자대학교 멀티미디어통신공학과 졸업(학사)
2004. 3~현재 서울여자대학교 컴퓨터학과 석사과정

관심분야 : 멀티미디어 시스템, 영상처리