

AAM 기반 얼굴 표정 인식을 위한 입술 특징점 검출 성능 향상 연구

한 은 정[†] · 강 병 준^{††} · 박 강 령^{†††}

요 약

AAM(Active Appearance Model)은 PCA(Principal Component Analysis)를 기반으로 객체의 형태(shape)와 질감(texture) 정보에 대한 통계적 모델을 통해 얼굴의 특징점을 검출하는 알고리즘으로 얼굴인식, 얼굴 모델링, 표정인식과 같은 응용에 널리 사용되고 있다. 하지만, AAM 알고리즘은 초기 값에 민감하고 입력영상이 학습 데이터 영상과의 차이가 클 경우에는 검출 에러가 증가되는 문제가 있다. 특히, 입을 다문 입력 영상 영상의 경우에는 비교적 높은 검출 정확도를 나타내지만, 사용자의 표정에 따라 입을 벌리거나 입의 모양이 변형된 얼굴 입력 영상의 경우에는 입술에 대한 검출 오류가 매우 증가되는 문제점이 있다. 이러한 문제점을 해결하기 위해 본 논문에서는 입술 특징점 검출을 통해 정확한 입술 영역을 검출한 후에 이 정보를 이용하여 AAM을 수행함으로써 얼굴 특징점 검출 정확성을 향상시키는 방법을 제안한다.

본 논문에서는 AAM으로 검출한 얼굴 특징점 정보를 기반으로 초기 입술 탐색 영역을 설정하고, 탐색 영역 내에서 Canny 경계 검출 및 히스토그램 프로젝트션 방법을 이용하여 입술의 양 끝점을 추출한 후, 입술의 양 끝점을 기반으로 재설정된 탐색영역 내에서 입술의 칼라 정보와 에지 정보를 함께 결합함으로써 입술 검출의 정확도 및 처리속도를 향상시켰다.

실험결과, AAM 알고리즘을 단독으로 사용할 때보다, 제안한 방법을 사용하였을 경우 입술 특징점 검출 RMS(Root Mean Square) 에러가 4.21픽셀만큼 감소하였다.

키워드 : AAM, 입술 검출, 캐니 경계 검출기, 히스토그램 투영, 얼굴 표정 인식

A Study on Enhancing the Performance of Detecting Lip Feature Points for Facial Expression Recognition Based on AAM

Eun Jung Han[†] · Byung Jun Kang^{††} · Kang Ryoung Park^{†††}

ABSTRACT

AAM(Active Appearance Model) is an algorithm to extract face feature points with statistical models of shape and texture information based on PCA(Principal Component Analysis). This method is widely used for face recognition, face modeling and expression recognition. However, the detection performance of AAM algorithm is sensitive to initial value and the AAM method has the problem that detection error is increased when an input image is quite different from training data. Especially, the algorithm shows high accuracy in case of closed lips but the detection error is increased in case of opened lips and deformed lips according to the facial expression of user. To solve these problems, we propose the improved AAM algorithm using lip feature points which is extracted based on a new lip detection algorithm.

In this paper, we select a searching region based on the face feature points which are detected by AAM algorithm. And lip corner points are extracted by using Canny edge detection and histogram projection method in the selected searching region. Then, lip region is accurately detected by combining color and edge information of lip in the searching region which is adjusted based on the position of the detected lip corners. Based on that, the accuracy and processing speed of lip detection are improved.

Experimental results showed that the RMS(Root Mean Square) error of the proposed method was reduced as much as 4.21 pixels compared to that only using AAM algorithm.

Keywords : Active Appearance Model, Lip Detection, Canny Edge Detector, Histogram Projection, Facial Expression Recognition

※ 본 연구는 2009년도 지식경제부 센싱 기반 감성서비스 모바일 단말 기술 개발 사업의 지원을 받아 수행하였습니다.

† 준 회 원 : 동국대학교 전자공학과 석사과정

†† 정 회 원 : 한국전자통신연구원 휴먼인식기술연구팀 선임연구원

††† 정 회 원 : 동국대학교 전자공학과 부교수(교신저자)

논문접수 : 2009년 3월 10일

수 정 일 : 1차 2009년 5월 7일, 2차 2009년 5월 27일

심사완료 : 2009년 5월 27일

1. 서 론

얼굴 인식[1], 홍채 인식, 지문 인식, 지정맥 인식 등과 같은 바이오 인식 기술은 개인만이 소유하고 있는 바이오 정보를 이용하여 개인의 신원을 확인하는 인식 기술이다. 이러한

바이오 인식 기술 중에 얼굴 인식은 사용자에게 친숙하고 영상 취득 시에 불편함이 없는 인식 기술로 알려져 있으며, 최근 공항에서의 출입국 관리 등에 많이 이용되고 있다. 얼굴 인식에서 얼굴의 특징점들을 정확하게 검출하는 것은 얼굴 인식 성능에 매우 큰 영향을 미치는 중요한 요소이며, 인간과 컴퓨터간의 상호작용에서의 감성 전달을 위해 최근 중요하게 연구되고 있는 표정 인식에서도 정확한 얼굴 특징점 검출은 표정 인식의 정확도를 결정하는 중요한 요소이다[2, 5]. 이러한 얼굴 특징점 중에서 특히 입술 특징점의 경우에는 음성 인식에서 음성 정보와의 결합을 통해 음성 인식 성능을 향상시키는데도 사용되고 있다[3,4]. 음성 인식에서 음성 정보만을 사용하는 경우에는 주변 잡음이 포함된 음성 정보에 의해 인식 성능이 저하되는 문제점이 발생되기 때문에 기존의 음성 인식 연구에서는 입술의 특징점 검출을 통한 입술의 형태 정보를 사용하여 이러한 문제들을 해결하였다[3,4].

최근 얼굴의 특징점을 검출하기 위해 주로 사용되고 있는 AAM 알고리즘은 사용자의 표정 변화에 의해 입의 형태가 변형될 경우에 검출 정확도가 감소되는 문제점이 있다. (그림 1)은 AAM 알고리즘에서의 입술 검출 오류의 예를 나타낸다. (그림 1)의 첫 번째 영상과 같이 입술을 다문 경우에 예외적으로 입술 검출 오류가 크게 발생한 경우가 있었지만, 대부분의 입술 검출 오류는 입의 형태가 변형된 경우에 발생되었다. 따라서 본 논문에서는 정확한 입술 특징점 검출을 위한 입술 윤곽선 검출 방법을 제안한다.

기존의 입술의 윤곽선 검출에 관한 연구는 Active contour를 이용한 방법[6, 7, 18], 컬러를 이용한 방법[8, 9], 컬러와 에지를 이용한 방법[10] 등이 있다. Active contour는 에너지 함수가 최소화되는 형태로 만들어 윤곽선을 검출하는 알고리즘으로 기존의 다른 방법들보다 높은 정확도를 보이나 반복적인 알고리즘으로 인한 많은 연산 처리량으로 인해 처리시간이 오래 걸리고 초기 윤곽선 모델이 부정확할 경우에 검출 성능이 저하되는 단점이 있다. RGB 컬러 정보를 이용하여 윤곽선을 검출하는 방법[8]은 피부와 입술의 컬러 특징 정보를 이용하여 빠른 처리시간 안에 입술 윤곽선을 검출하지만, 검출 성능이 조명에 민감한 문제점이 있다. 이러한 문제점을 해결하기 위하여 본 논문에서는 일정한 조명 아래에서 입력받은 컬러 얼굴 영상에서 컬러 정보와 에지 정보를 결합하여 빠른 처리시간 안에 입술의 윤곽선을 검출하는 방법을 제안한다.



(그림 1) AAM 알고리즘을 이용한 얼굴 윤곽선 검출 중 입 윤곽선을 잘못 검출한 예

본 논문의 구성은 다음과 같다. 2장에서는 AAM 기반으로 얼굴 특징 추출하는 방법에 대해 설명하고 3장에서는 본 연구에서 제안하는 방법의 전체 흐름과 제안하는 입술 검출 정보를 이용한 AAM 기반 얼굴 특징 추출 알고리즘의 과정에 대해 설명한다. 4장에서는 기존 AAM 알고리즘과 제안한 방법의 검출 정확도를 비교하고 분석하고, 최종적으로 5장에서는 본 연구의 결론과 향후 연구방향에 대해 언급하도록 한다.

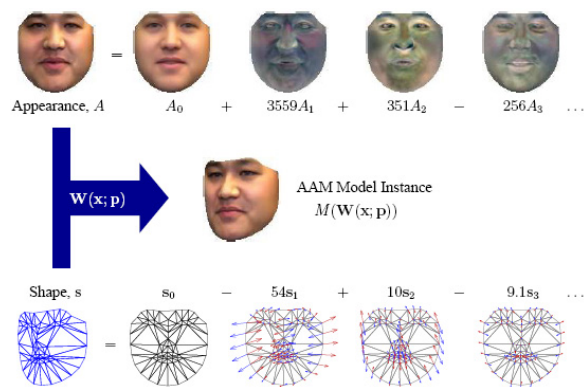
2. AAM 기반 얼굴 특징점 검출

AAM은 (그림 2)와 같이 PCA(Principal Component Analysis)를 기반으로 객체의 형태(shape)와 질감(texture) 정보에 대한 통계적 모델을 통해 얼굴의 특징점을 추출하는 알고리즘 [14]으로 (그림 2)와 같이 얼굴 형태를 나타내는 모델과 얼굴의 질감 정보를 나타내는 모델로 구성되어 있다. 얼굴 형태를 나타내는 모델은 식 (1)과 같이 얼굴 특징점의 평균(mean, base shape)과 얼굴 형태를 나타내는 기저(basis)들에 대한 계수(coefficient) 파라미터들로 구성되어있고, 얼굴의 질감 정보를 나타내는 모델은 식 (2)와 같이 얼굴의 픽셀 값을 나타내는 RGB값들의 평균과 얼굴의 질감 정보를 나타내는 기저들의 계수 파라미터들로 구성되어 있다[20].

$$s = s_0 + \sum_{i=1}^j c_i s_i \tag{1}$$

$$A(x) = A_0(x) + \sum_{i=1}^j c_i A_i(x) \tag{2}$$

여기서 s_0 와 $A_0(x)$ 는 각각의 모델의 평균을 나타내며, $c = (c_1, c_2, \dots, c_j)^T$ 는 각 기저들($s = (s_1, s_2, \dots, s_j)$, $A = (A_1, A_2, \dots, A_j)$)의 계수 파라미터를 나타낸다.



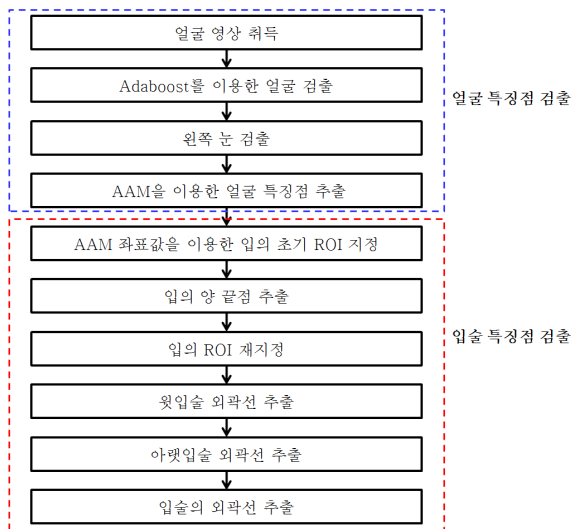
(그림 2) AAM의 shape 및 appearance model [20]

3. 제안하는 방법

3.1 제안하는 알고리즘의 전체 흐름도

(그림 3)은 본 연구에서 제안하는 AAM을 이용한 얼굴 특징 추출 및 입의 윤곽선 추출 방법의 흐름도이다. 먼저

일반적인 640 x 480 픽셀 급 카메라를 통해 가시광선 하에서 컬러 얼굴 영상을 취득한다. AAM 알고리즘의 초기 값을 결정하기 위해 입력 얼굴 영상에서 Adaboost 알고리즘을 이용하여 얼굴 영역을 검출한 후 검출한 얼굴 영역으로부터 사람 얼굴의 형태학적 사전 지식(prior knowledge)을 이용하여 (그림 4(b))와 같이 좌측 눈 후보 영역을 선택한다. 선택한 후보 영역 내에서 좌측 눈의 중심 좌표 값을 검출하여 AAM의 초기 값으로 결정하고, AAM을 이용하여 얼굴 전체 영역에 대한 특징점을 추출한다. 전술한 바와 같이 입술의 형태가 변형된 경우에는 AAM에 의한 입술 검출의 에러가 증가되는 문제점이 발생되기 때문에 본 논문에서는 (그림 3)과 같이 입술의 특징점을 검출하여 입술 검출 정확도를 향상시킨다. 입술 특징점 검출을 위해 먼저 AAM에 의해 비교적 높은 정확도로 검출되는 눈과 코 영역의 좌표를 이용하여 입의 초기 관심 영역(ROI, Region Of Interest)을 지정한 후 관심 영역 내에서 Canny 경계 검출 방법을 이용하여 입술의 양 끝점을 추출한다. 추출된 양 끝점을 기반으로 관심 영역을 재지정하고, 입술의 컬러 정보를 이용하여 윗입술 및 아랫입술의 윤곽선을 추출하고, 이를 기반으로 정확한 입술 경계를 검출한다.

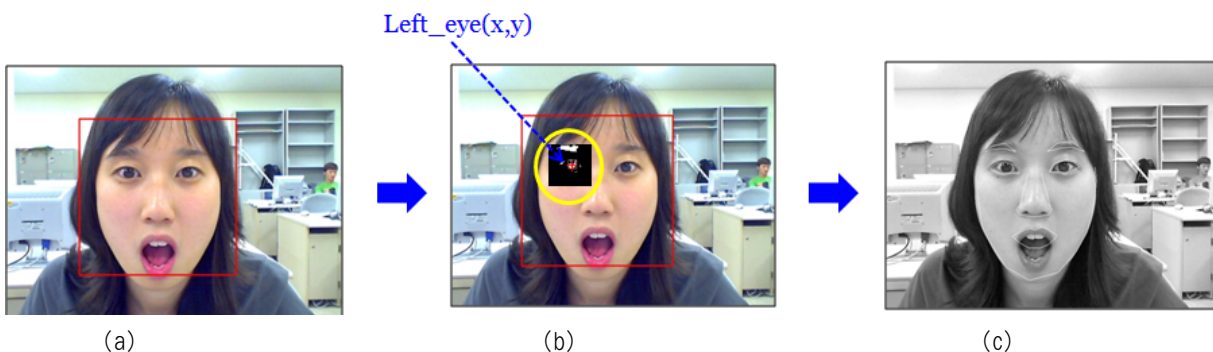


(그림 3) 제안하는 방법의 전체 흐름도

3.2 얼굴 특징점 검출

본 연구에서는 AAM 알고리즘을 기반으로 얼굴의 특징점을 추출하였다. 전술한 바와 같이 AAM은 초기 값에 매우 민감하기 때문에, 얼굴 특징점의 정확한 검출을 위해서 정확한 초기 값을 결정하는 것은 매우 중요한 문제이다[11]. 따라서 본 논문에서는 Adaboost 알고리즘을 기반으로 검출된 얼굴 영역의 좌측 눈 후보 영역으로부터 좌측 눈의 중심 좌표 값을 측정하여 AAM의 초기 값으로 사용하였다. 얼굴 영역을 검출하기 위해 사용되는 Adaboost 알고리즘은 여러 개의 약 분류기(weak classifier)를 조합하여 강 분류기(strong classifier)를 학습하는 알고리즘으로 학습시간이 오래 걸리는 단점이 있지만 처리 속도가 빠르고 단순하여 얼굴을 검출하는데 있어서 성능이 매우 우수한 것으로 알려져 있다. 본 연구에서는 OpenCV 영상처리 라이브러리에서 제공되는 정면 얼굴에 대해 학습된 Adaboost 분류기 모델을 사용하였다[12].

Adaboost 알고리즘을 이용하여 좌측 눈의 중심좌표를 추출하는 과정은 (그림 4)와 같다. 먼저 Adaboost 알고리즘을 이용하여 (그림 4(a))와 같이 사각 박스 형태의 얼굴 영역을 검출한 후, 박스의 좌측 상단 모서리로부터 사람 얼굴의 형태학적 사전 정보를 이용하여 왼쪽 눈 후보 영역을 선택한다. 선택된 영역 내에서 사람의 눈동자는 낮은 그레이 레벨 값을 나타내고, 그 외의 각막이나 피부 영역은 높은 그레이 값을 나타낸다는 특징을 이용하여 영상 이진화를 통해 (그림 4(b))와 같이 사람의 눈동자 영역과 그 외의 영역을 구분한다. 컴포넌트 라벨링(Component labeling) 방법을 이용하여 연결된 픽셀들을 하나의 영역으로 묶어 각 영역을 구분한 후, 영역의 크기가 미리 결정한 임계치 이하인 영역은 이진화에 의해 부분적으로 나타나는 영상잡음으로 판단하여 제거하였다. (그림 4(b))에서 보는 바와 같이 눈썹 영역도 역시 영상에서 낮은 그레이 레벨 값으로 나타나기 때문에 영상 이진화에 의해 검출된다. 하지만 눈동자는 원형의 형태를 지니며 눈썹은 길게 늘어난 형태를 지니기 때문에 컴포넌트 라벨링에 의해 구분된 각 영역의 원형 도를 측정하여 원형에 가장 근접한 영역을 사람의 눈동자 영역으로 추출하였다. 추출된 눈동자 영역의 무게 중심 좌표를 계산한 후 AAM의 초기 값으로 사용하여 (그림 4(c))와 같이 최종



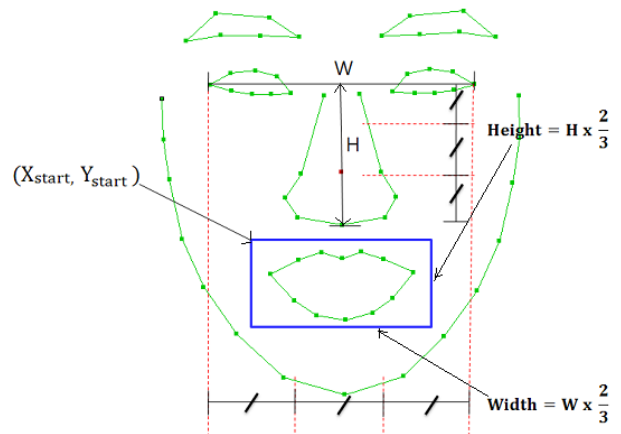
(그림 4) AAM 초기 값 설정과정: (a) Adaboost를 이용하여 얼굴 영역 검출 (b) 검출 영역 내에서 영상처리를 통한 이진화 및 컴포넌트 라벨링을 통한 왼쪽 눈 검출 (c)왼쪽 눈 좌표기반 AAM 특징점 검출

적으로 얼굴의 세부 영역 특징점들을 검출한다. AAM은 학습된 형태정보와 질감정보를 이용하여 얼굴의 특징점을 검출하기 때문에 얼굴 특징점을 검출하는데 있어서 우수하나, 입력된 영상이 학습 영상과의 차이가 큰 경우 검출 에러가 증가되는 단점이 있다. 특히, 사람의 입술의 경우는 표정에 따라 다양하게 변형이 발생되기 때문에 (그림 4(c))와 같이 입술에 대한 검출에러가 증가되는 문제점이 있다. 따라서 본 논문에서는 정확한 입술 특징점을 검출하기 위해 독립적으로 입술 영역을 검출한다.

3.3 입 영역에 대한 초기 관심 영역 지정

본 논문에서는 AAM 좌표를 이용하여 입에 대한 관심영역을 지정된 후 입술 영역을 검출하였다. 이렇게 입의 관심영역을 지정해 줌으로써 얼굴 전체 영역에서 입술의 윤곽선을 검출 하는 것보다 연산 수행 시간을 상당히 감소시킬 수 있고, 코와 턱과 같은 잡음으로 인해 잘못 검출될 확률도 감소시킬 수 있다. 기존 AAM 방법은 변형된 입에 대한 추출 정확도가 낮기 때문에, (그림 5)와 같이 정확도가 높은 양 눈과 코에 대한 특징점을 기준으로 관심영역을 지정하였다.

(그림 5)는 AAM 학습을 통해 측정된 평균 얼굴 형태식 (1)의 s_0 위에 얼굴비율에 따른 관심영역을 적용시킨 모습이다. 실제로는 사람마다 얼굴 형태가 다르기 때문에 그림과는 약간의 차이가 나타날 수 있지만 입술 영역이 왼쪽 눈의 바깥쪽 끝점과 오른쪽 눈의 바깥쪽 끝점 사이의 거리를 넘지 않고, 입의 높이는 양 눈과 코 끝 사이의 거리의 2/3 정도로 지정된다. 이러한 특징을 이용하여 관심영역의 기준 좌표를 아래 식과 같은 방법으로 계산한다. (X_{start}, Y_{start})는 관심영역의 시작점을 나타내며 X_{cl}, Y_{cl} 은 왼쪽 눈의 바깥쪽 끝점의 x좌표와 y좌표, X_n, Y_n 은 각각 코끝의 x좌표,



(그림 5) 얼굴의 영역별 비율에 따른 초기 관심 영역 지정

y좌표를 의미한다.

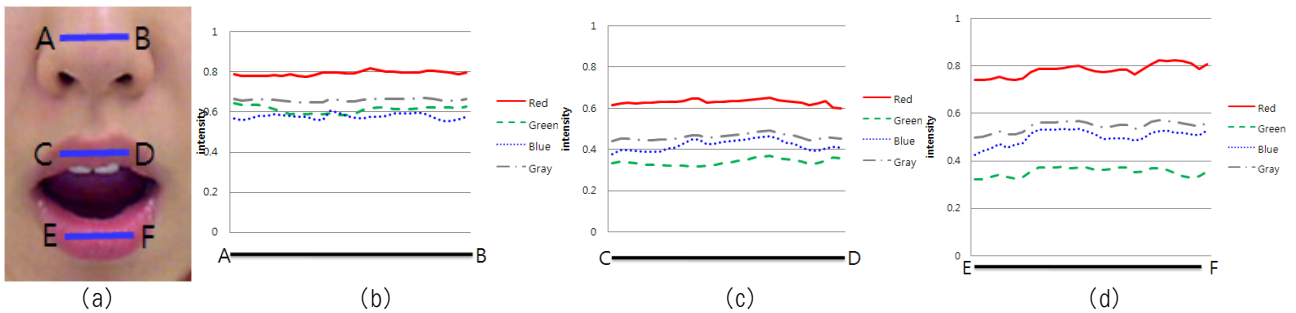
$$X_{start} = X_{cl} + (X_n - X_{cl})/3 \tag{3}$$

$$Y_{start} = Y_n + (Y_n - Y_{cl})/9 \tag{4}$$

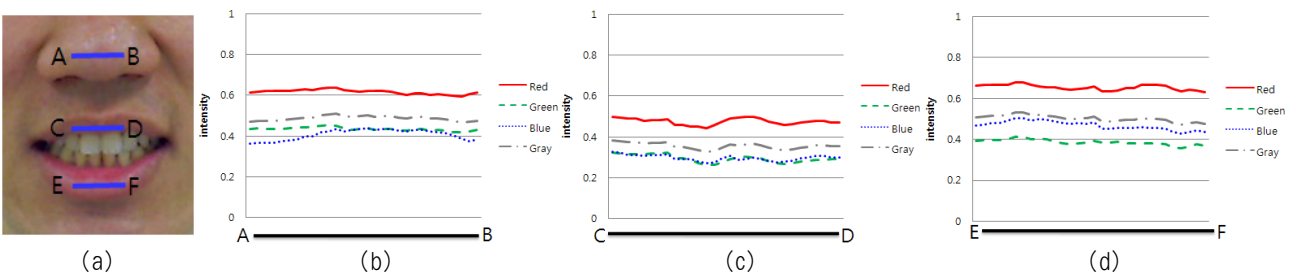
식 (3)과 식 (4)에 의해 입술에 대한 관심 영역의 시작점 좌표를 계산한 후, (그림 5)와 같이 좌·우 눈의 바깥쪽 끝점의 길이 비율과 눈과 코 사이의 길이 비율을 이용하여 초기 관심 영역의 크기를 결정한다.

3.4 입술의 컬러 특징 정보

본 논문에서는 RGB 컬러 정보를 이용하여 입술의 윤곽선을 검출하였다. (그림 6)과 (그림 7)은 여성과 남성의 각 위치별 정규화 시킨 RGB 성분을 나타내는 그래프이다. 남



(그림 6) 여성 영상의 컬러 프로파일 : (a) 원 영상 (b) 피부 (c) 윗입술 (d) 아랫입술



(그림 7) 남성 영상의 컬러 프로파일 : (a) 원 영상 (b) 피부 (c) 윗입술 (d) 아랫입술

성과 여성의 피부와 입술을 비교해볼 때 RGB 성분 값의 차이가 존재하지만 각 성분 간의 비율 차이는 거의 나타나지 않았으며, 동일한 사람의 경우라도 윗입술과 아랫입술의 RGB 색상 분포가 차이가 나는 것을 확인할 수 있었다. (그림 6)과 (그림 7)에서 보는 바와 같이 피부 영역에서는 G성분의 값이 B성분의 값보다 높게 나타나고 입술 영역에서는 두 성분의 크기가 동일하거나 B성분의 값이 G성분의 값보다 높게 나타났[9]. 입술 영역에서의 B성분의 값이 피부 영역에서의 G성분의 값과 유사하거나 R성분과 B성분의 값이 피부와 입술 영역에서 유사하게 나타나기 때문에 RGB 성분의 평균값에 의해 결정되는 그레이 영상의 경우에는 피부와 입술 영역간의 그레이 값 차이가 크게 나타나지 않는 반면에, G성분의 값은 (그림 6)과 (그림 7)에서 보는 바와 같이 피부 영역과 윗입술 영역간의 차이가 크게 나타난다. 따라서 본 논문에서는 입술 영역이 그레이 영상에서보다 G 영상에서 더 명확하게 나타나는 성질 [4]을 이용하여 입술의 양 끝점과 윗입술의 윤곽선을 검출 할 때는 그레이 영상 대신 G영상을 사용하였다. (그림 7)의 (b)와 (d)에서 보는 바와 같이 아랫입술의 경우에는 G 성분의 값이 피부 영역과 유사하게 나타나지만 피부나 윗입술 영역에 비해 R성분과 G성분의 차가 크기 때문에, 본 논문에서는 아랫입술을 검출 할 때는 R성분과 G성분의 차 영상을 사용하였다.

3.5 입의 양 끝점 추출

(그림 5)에서 지정한 입술영역에 대한 초기관심영역은 대략적인 입술의 존재 범위를 나타내므로, 이 영역 내에는 입술이외에 다른 특징점들이나 잡음들이 포함될 수 있다. 이러한 문제를 해결하고 정확한 입술 영역을 검출하기 위한 방법으로 본 연구에서는 지정된 초기관심영역에서 입의 양 끝점을 검출하여 관심영역을 재 지정하였다. 먼저, 실제 입의 너비를 구하기 위해 입의 양 끝점은 입을 벌렸을 때나 다물었을 때나 피부와 입술에 비해 어둡다는 특징을 이용하였다[4]. 초기관심영역 내에 포함된 잡음을 고려하여 Canny 경계 검출 알고리즘을 이용하였으며, (그림 8)과 같은 순서로 입술 양 끝점을 검출한다.

전술한 바와 같이 그레이 영상에서보다 G 영상에서 입술이 더 선명하게 나타나기 때문에 양 끝점 추출 시에 G 영상을 사용하였다[4]. 먼저 Canny 경계 검출 알고리즘을 사용하여 (그림 8(b))와 같이 입술의 경계면을 검출한 후, (그림 8(c))와 같이 닫힘 연산(closing) 연산을 통해 입술의 윤곽선이 중간에 끊어지지 않고 연결될 수 있도록 한다. 마지막으로

컴포넌트 라벨링을 수행하여 연결된 가장 큰 영역을 제외하고 나머지 영역을 제거함으로써 입을 제외한 나머지 영상 잡음을 제거한다. 본 논문에서는 (그림 8(d))와 같은 이진화 된 입술 영상에서 좌측 상단과 우측 상단에서부터 y축(상하) 방향으로의 탐색을 통해 입의 양 끝점을 검출한다. (그림 9)는 입의 양 끝점 추출 시에 나오는 대표적인 4가지 예와 양 끝점 추출 결과를 나타낸다. 좌측 상단과 하단의 그림은 입을 다물었을 때 영상이고 우측 상단과 우측 하단의 영상은 입을 벌리거나 이만 드러냈을 때의 모습이다.

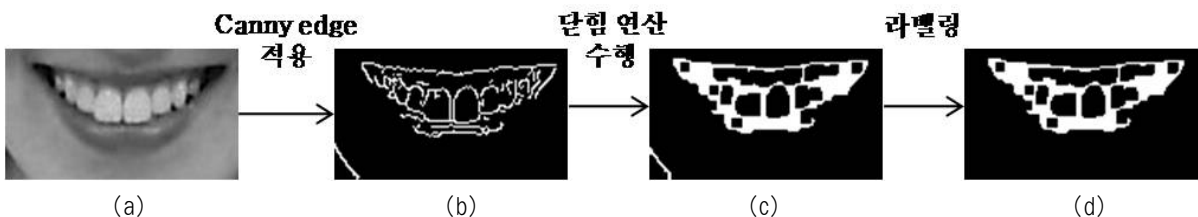
본 논문에서는 입술 영역의 높이와 너비를 측정하여 관심 영역의 높이와 너비를 결정함으로써 입술을 검출하기 위한 초기관심영역으로부터 관심영역을 재 지정하였다. 추출된 입술의 오른쪽 끝점의 x좌표 lip_r 과 왼쪽 끝점의 x좌표 lip_l 과의 차에서 제안하는 방법에 의해 검출된 양 끝점의 오차 (k)를 고려하여 식 (5)와 같이 관심영역의 너비를 측정한다.

$$lip_w = lip_r - lip_l + k \tag{5}$$

본 논문의 실험에서는 20명의 사람에 대한 총 200장의 실험 영상을 사용한 실험을 통해 측정된 최대 오차 값을 양 끝점의 오차를 고려하기 위한 k 값을 30으로 사용하였다. (그림 10)에서 보는 바와 같이 Canny 경계 검출 알고리즘을 적용한 영상을 수평축으로 투영(Projection)시킨 히스토그램을 통해 입술의 높이를 측정하였다. G 영상에서 아랫입술의 경계가 명확하지 않아 Canny 경계 검출 알고리즘을 통해 아랫입술의 경계가 검출되지 않는 예가 종종 발생되기 때문에, 본 논문에서는 수평축으로 투영시킨 히스토그램을 통해 측정된 입술의 높이를 이용하여 입을 벌린 영상인지 다문 영상인지 판단하고, 입을 벌린 영상인지 다문 영상인지에 따라 식 (5)로부터 측정된 입술의 너비에 서로 다른 가중치를 부여하여 식 (6)과 같이 관심영역의 높이(x)를 결정



(그림 9) 입술 양 끝점 추출 예



(그림 8) 입술 양 끝점 추출 과정 : (a) 원 영상 (b) Canny edge 적용한 영상 (c) 닫힘 연산 수행 후 영상 (d) 라벨링 후 영상

하였다.

$$x = \begin{cases} lip_w \times w_1 & , x < a \text{ 입을 다문 영상} \\ lip_w \times w_2 & , x > a \text{ 입을 벌린 영상} \end{cases} \quad (6)$$

본 논문의 실험에서는 20명의 사람에 대한 무표정 영상 총 40장의 실험 영상을 기반으로 측정한 평균 입술의 높이를 입을 다문 영상인지 입을 벌린 영상인지를 구분하기 위한 임계치로 사용하였다.(본 논문에서는 a 를 30으로 사용하였다.) 입을 다물었을 때는 상대적으로 입의 높이가 입의 너비에 비해 줄어들며 입을 벌리면 입의 너비에 비해 상대적으로 입의 높이가 커지는 점을 고려하여 20명의 사람에 대한 200장의 실험 영상을 가지고한 실험을 통해 w_1 을 0.6, w_2 를 1.2로 결정하였다.

3.6 입술의 윤곽선 추출

윗입술을 추출하는 과정은 (그림 11)과 같이 진행되며 다음과 같은 4단계로 나뉜다.

① 입술의 높이와 너비에 의해 결정된 관심영역의 G 영상으로부터 수평 방향의 Sobel 경계 검출 마스크를 이용하여 입술의 경계를 검출한다.

② Sobel 마스크를 적용한 영상에서 고정된 임계치를 사용하여 이진화 시킨다.

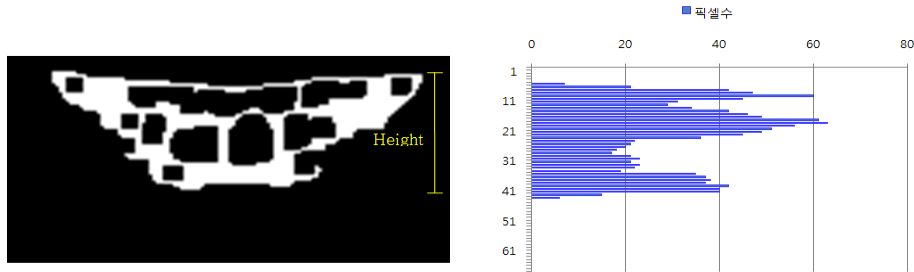
③ 컴포넌트 라벨링을 수행하여 가장 큰 영역을 제외한 나머지 영역을 제거함으로써 입을 제외한 나머지 영상 잡음을 제거한다.

④ 크기가 작은 잡음을 제거하기 위해 3×3 미디언 필터(Median Filter)를 적용한 후 영상의 위에서부터 아래로 탐색하여 윗입술의 좌표 값을 추출한다.

아랫입술은 윗입술 검출하는 방법과 같은 방법으로 검출할 경우 조명의 영향으로 인해 검출 정확도가 감소되는 문제점이 있다. 따라서 아랫입술을 검출 할 때는 전술한 바와 같이 R성분과 G성분의 차 영상을 이용하여 아랫입술을 검출한다. 기존 연구에서는 아랫입술 검출을 위해 변형된 형태의 R성분과 G성분의 차 영상을 사용하였으나[4], 본 연구에서는 이를 단순화 하여 연산의 처리량을 줄였다. (그림 12)에서 보는 바와 같이 R성분과 G성분의 차 영상으로부터 이진화를 통해 얻은 영상에서 영상의 아래로부터 위로 탐색하여 아랫입술의 좌표 값을 추출하였다.

4. 실험결과

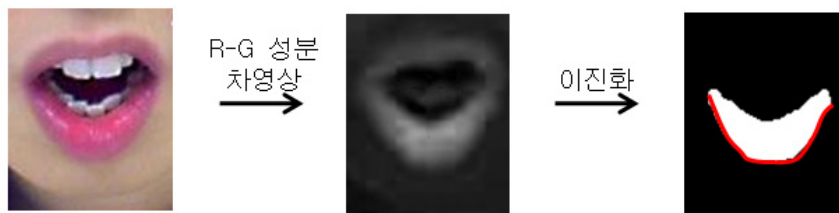
실험을 위해 Logitech Quickcam pro 5000 [19] 카메라를 사용하여 총 75명의 사람들로부터 각각 10장씩의 영상을 입력 받아서 총 750장의 얼굴 영상 데이터베이스를 구성하였다. 다양한 표정에 대한 얼굴 정보를 AAM 알고리즘 학습



(그림 10) 수평축으로 투영한 영상



(그림 11) 윗입술 추출 과정



(그림 12) 아랫입술 추출 과정

에 사용하기 위해 75명의 사람에 대해 각각 5가지 표정(무표정, 기쁨, 슬픔, 놀람, 화남)을 각각 2장씩 입력 받았으며, 각 영상은 640×480 픽셀 크기의 RGB 컬러 영상으로 구성되어 있다. 영상에는 안경 착용자 및 미착용자가 함께 포함되어 있으며, 표정이외의 요소를 없애기 위하여 실내에서 균일한 조명하여 영상 데이터베이스를 구축하였다. AAM 알고리즘은 DTU (Technical University of Denmark)에서 제공하는 AAM-API [13]을 이용하여 구현되었으며, (그림 13)과 같이 턱15개, 눈썹 12개, 눈 16개, 코 9개, 입 12개 등의 총 64개의 특징점을 사용한 얼굴 모델을 학습시켰다. 제안하는 입술영역 검출 방법은 학습과는 무관하기 때문에 학습의 유무에 영향을 받지 않는다. 따라서 학습에 쓰인 데이터베이스와 같은 얼굴 영상 데이터베이스로 테스트 실험을 하였다.

첫 번째 실험으로 기존의 AAM 알고리즘을 단독으로 사용하는 방법과 제안하는 “AAM + 입술 추출” 방법의 얼굴 특징점 추출 정확도를 비교 실험하였다. 실제 입술의 좌표 점과 제안하는 방법에서 검출된 입술 좌표 점들 간의 오차를 측정하는 것은 어려운 문제이다. 이는 눈이나 콧구멍 같은 얼굴 특징의 경우 ground-truth 점들과 검출된 점들이 각각 하나의 점들로 표시되므로, 이 두 점사이의 거리를 검출오차로 단순하게 계산하면 되지만, 입술경계나 턱선 같은 경우는 하나의 점들이 아닌 여러 점들이 포함된 선으로 검출되므로, ground-truth 점들에 대응되는 검출된 점들이 어떤 것인지 정하는 것이 어렵기 때문이다. 그래서 AAM으로 턱선 검출의 정확도를 측정하는 연구들에서는 턱 선상에서 등 간격으로 정해진 점들을 ground-truth 점들로 정하고, 검출된 점들 역시 이러한 규칙으로 정함으로써 상호간의 오차를 계산하였다. 하지만 본 연구에서와 같이 입술 경계의 경우, 입술 윤곽자체가 오목 및 볼록 형태가 혼합되어 있으므로, 단순히 등 간격으로 ground-truth 및 검출 점들을 정하

기는 문제가 있다. 따라서 본 연구에서는 ground-truth 점들로부터 입술 구간별 거리 비율을 측정한 후, 검출한 점들도 같은 비율을 적용하여 가장 근접한 비율일 때의 점들을 ground-truth에 대응되는 검출 점으로 인식하여 상호간의 RMS 오차를 측정하였다.

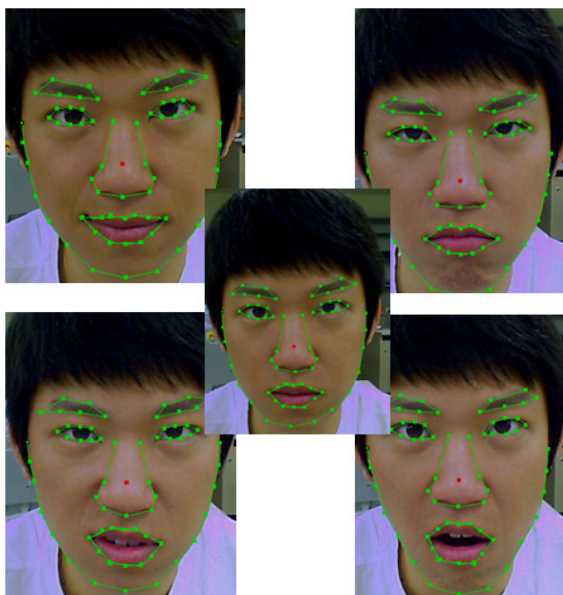
$$l_n = \sqrt{(x_{p_{n+1}} - x_{p_n})^2 + (y_{p_{n+1}} - y_{p_n})^2} \quad (7)$$

$$l = \sum_{k=a_1}^n l_k \quad (8)$$

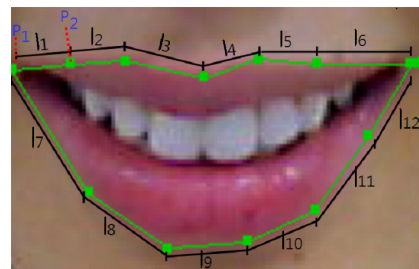
$$C_n = \frac{\sum_{k=a_1}^n l_k}{l}, \quad C'_n = \frac{\sum_{k=a_1}^{a_s} l_k}{l} \quad (9)$$

즉, 직접 손으로 그린 ground-truth 좌표 값은 AAM을 학습 시킬 때 생성되었던 좌표 값으로 사용하여 (그림 14)와 같이 실제 입술의 ground-truth 점들 간의 거리를 이용하여 구간별 비율을 구하였다. 식 (7) 및 (그림 14)에서 l_1 은 p_1 과 p_2 간의 직선길이를 나타낸다. 따라서 입술 외곽선의 전체 길이 l 은 식 (8)과 같이 구할 수 있고, 윗입술의 경우 a_1 이 1이고 n 은 6을 나타내고, 아랫입술의 경우 a_1 은 7이고 n 은 12를 나타낸다. 식 (9)를 통해 입술 외곽선 전체길이에 대한 위치별 입술 외곽선 길이 비율을 구할 수 있다. 이때 C_n 은 l_n 까지 합을 총길이로 나누어 구간까지의 비율을 나타낸다.

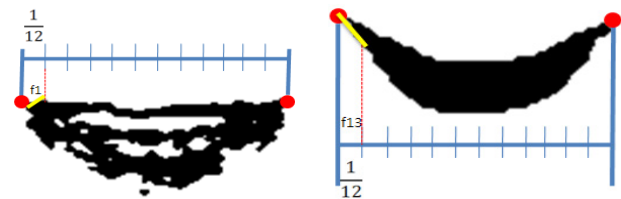
(그림 15)는 제안한 방법에서 추출된 입술 경계에서 ground-truth에 해당되는 좌표 값을 측정하는 방법을 그림으로 나타낸 것이다. (그림 15)와 같은 영상에서 x 좌표가 1씩 증가 할 때마다 y 좌표 값을 구하고 픽셀과 픽셀간의 거



(그림 13) 5가지 표정에 대한 영상 데이터의 예



(그림 14) 실제 입술의 ground-truth 좌표 값 (녹색 및 녹색 사각 점)



(그림 15) 제안한 방법에서 추출된 입술 경계에서 ground-truth에 해당되는 좌표 값을 측정하는 방법

리를 구하면 연산 량이 매우 많다. 따라서 12등분을 한 후 직선의 방정식을 적용하여 y 값을 구하였다. 직선의 방정식은 식 (10)과 같이 구성되어 있으며 입술 양 끝점을 포함하여 13개의 점을 이용하여 윗입술, 아랫입술 각각 12개의 직선의 방정식을 구성하였다. a_n 은 기울기를 나타내며 b_n 은 y 절편을 나타낸다.

$$f(x) = ax + b \tag{10}$$

$$a_n = \frac{y_n - y_{n-1}}{x_n - x_{n-1}}, b_n = y_{n-1} - (a_n \times x_{n-1}) \tag{11}$$

제안한 방법에서 추출된 입술 경계에서 ground-truth에 해당되는 좌표 값을 측정할 경우도 식(7), (8), (9)를 이용하여 l, l_n, C'_n 을 구한다. 이 때 C'_n 은 제안하는 방법에서 추출한 입술 경계점을 기반으로 계산된 구간별 길이 비율인데, 구하는 방식은 C_n 과 같지만, C_n 은 (그림 14)와 같이 총 7개의 특징점에 대해서만 구하는 데 비해 C'_n 은 입술의 왼쪽 끝점 좌표 값인 a_1 과 오른쪽 끝점의 좌표 값인 a_s 사이에 있는 모든 점들에 대해 구한다. 따라서 x 값은 왼쪽 끝점의 x 좌표에서부터 오른쪽 끝점의 x 좌표 사이에 존재하는 모든 점에 대해서 구하고, 그에 대응 되는 y 값은 식(10), (11)을 이용하여 구한다.

$$d = |C_n - C'_n| \tag{12}$$

식 (12)에서 d 의 값은 실제 손으로 찍은 ground-truth 좌표와 제안하는 방법으로 찾은 입술 좌표간의 구간별 길이 비율의 차를 나타내며, d 의 값이 0.01 이하 일 때의 (x, y) 의 좌표 값을 (그림 14)의 ground-truth 특징점 P_n 에 대응되는 제안한 방법을 검출된 입술 경계점으로 설정하였다. 이러한 방법으로 추출된 점과 ground-truth 좌표 간의 RMS(Root Mean Square) 오차를 측정하였다. <표 1>에서 보는 바와 같이 750장의 영상에 대해서 RMS 오차를 측정 한 결과, 기존의 AAM 알고리즘을 단독으로 사용했을 경우에는 입이 변형되거나 벌린 경우의 평균 RMS 오차가 10.09 픽셀로 입을 다문 경우의 평균 RMS 오차인 8.05 픽셀보다 2.04 픽셀만큼 큰 것을 알 수 있다. 하지만 제안하는 방법은 입술의 칼라 정보와 에지 정보를 결합하여 입술을 검출함으로써 입이 변형되거나 벌린 경우에 대해서도 정확한 검출이 가능하므로 입을 다문 경우와 입이 변형되거나 벌린 경우의 RMS 오차가 유사하게 나타났다. 전체 영상에 대해서 RMS 오차를 측정 한 결과, 제안하는 방법에 의해 평균 RMS 오차가 4.21 픽셀만큼 감소하는 것을 알 수 있었다.

(그림 16)은 얼굴의 특징점을 검출한 결과 영상을 나타내며, 입술의 형태가 다양한 형태를 나타내는 경우 제안한 방법을 통해 입술 검출 정확도가 보다 개선된 것을 확인할 수 있다.

<표 1> 기존 방법과 제안하는 방법의 입술 윤곽 점 검출의 평균 RMS 오차 비교 (단위 : 픽셀)

	기존 AAM	제안한 방법
입을 다문 경우	8.05	5.52
입이 변형되거나 벌린 경우	10.09	5.44
전체 영상	9.71	5.50



(a)



(b)

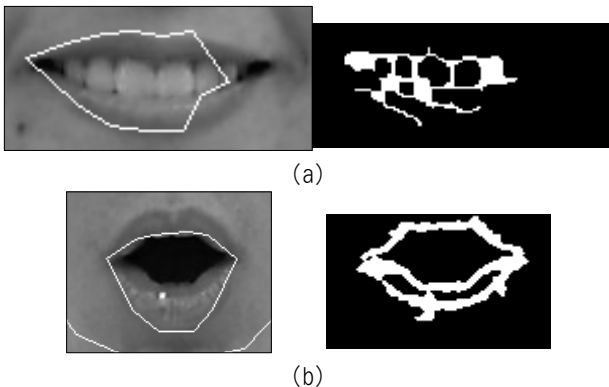
(그림 16) 결과 영상 : (a) 기존 AAM만 적용 결과 (b) 제안하는 방법 적용 결과

두 번째 실험으로 펜티엄-IV 3 GHz 프로세서와 3GB DRAM 환경에서 제안하는 방법의 평균 처리시간을 측정하였다. <표 2>에서 보는 바와 같이 Adaboost를 이용한 얼굴 검출과 AAM을 이용한 얼굴 특징점 검출을 위해 오랜 처리 시간이 요구되지만, 제안하는 입술 윤곽선 추출 방법은 9.5 ms로 매우 빠른 처리 속도로 처리되는 것을 알 수 있다.

(그림 17)은 제안하는 방법을 사용한 경우에도 입술의 검출 에러가 크게 발생하는 영상이다. (그림 17 (a))는 입술의 양 끝점을 추출하는 단계에서 닫힌 연산 수행 후에도 입술의 윤곽선들이 연결되지 않아 컴포넌트 라벨링 후 오른쪽 입술의 끝점이 있는 부분이 제거되는 문제가 발생하여 입술의 양 끝점 검출 에러가 발생된 경우이다. (그림 17(b))는 입술 양 끝점을 기반으로 관심영역 검출 시에 오류가 발생하는 문제점과 함께 윗입술 추출 하는 과정 중 윗입술 부분

<표 2> 단계별 처리시간

단계	처리시간(ms)
Adaboost를 이용한 얼굴 영역 검출	175.8
얼굴 영역 내에서 왼쪽 눈 좌표 검출	35.1
AAM을 이용한 얼굴 특징점 검출	84.2
제안된 입 윤곽선 추출	9.5
총 수행시간	304.6



(그림 17) 입술 검출 에러가 발생한 영상 : (a) 입술의 양 끝점 검출 에러 예 (b) 윗입술의 윤곽선 검출 에러 예

과 입의 안쪽 영역이 연결되지 않아 컴포넌트 라벨링 후에 윗입술의 윤곽선 부분이 제거되어 윗입술의 윤곽선 검출 에러가 발생한 경우이다.

5. 결 론

본 연구에서는 AAM을 기반으로 얼굴의 특징점을 추출하고 추가적으로 입술의 윤곽선 검출을 통해 입의 특징점을 추출함으로써 사람의 표정 변화에 의해 나타나는 변형된 입술에 대한 검출 정확도 감소의 문제를 해결하였다. 실험 결과, 기존의 AAM 알고리즘을 단독으로 사용하는 것과 비교했을 때, 제안하는 방법에 의해서 입술 영역에 대한 검출 RMS 에러가 4.21 픽셀만큼 감소하는 것을 확인할 수 있었다. 향후, 조명변화가 많거나 영상 해상도가 낮은 영상에 대해서 입술 윤곽선 검출 정확도 평가 및 성능 향상 연구를 수행할 계획이다.

참 고 문 헌

- [1] S. J. Lee, S. B. Jung, J. W. Kwon, S. H. Hong, "Face Detection and Recognition Using PCA," in Proceedings of IEEE Region 10th Conference (TENCON'99) pp.84-87, 1999.
- [2] Fangqi Tang, Benzai Deng, "Facial Expression Recognition using AAM and Local Facial Features," in Proceedings of International Conference on Natural Computation, pp.632-635, 2007.
- [3] Mehran Yazdi, Mehdi Seyfi Amirhossein Rafati, Meghdad Asadi, "Realtime Lip Contour Tracking For Audio-Visual Speech Recognition Applications," in Proceedings of World Academy of Science Engineering and Technology, Vol.30, pp.164-167, 2008.
- [4] 이종석, 심선희, 김소영, 박철훈, "제어되지 않은 조명 조건하에서 입술움직임의 강인한 특징추출을 이용한 바이모달 음성인식", Telecommunications Review, 14권, 1호, pp.123-134, 2004.
- [5] Ming-Hsuan Yang, D. J. Kriegman, and N. Ahuja, "Detecting Faces in Images: A Survey," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.24, pp.34-58, 2002.
- [6] Hadi Seyedarabi, Won Sook LEE, Ali Aghagolzadeh, "Automatic Lip Tracking and Action Units Classification using Two-Step Active Contours and Probabilistic Neural Networks," in Proceedings of Canadian Conference on Electrical and Computer Engineering, 2006.
- [7] M. LiDevin, P. Delmas, P.Y. Coulon, F. Luthon, V. Fristot, "Automatic Lip Tracking: Bayesian Segmentation and Active Contours in a Cooperative Scheme," in Proceedings of IEEE Conference on Multimedia, Computing and Systems, ICMCS'99, Fiorenza, Italy, June, 1999.
- [8] Md. Mahbubur Rahman, Faisal Mahbub Chowdhury, Rajib Kumar Rana, Mohammad Fazleh Elahi, Md. Aminul Islam, "A New Pseudo-automatic Outer Lip Contour Extraction Approach Based on RGB Components," International Journal of the Computer, the Internet and Management, Vol.15, pp.94-101, 2007.
- [9] N. Eveno, A. Caplier, and P. Y. Coulon, "A New Color Transformation for Lips Segmentation," in Proceedings of IEEE Fourth Workshop On Multimedia Signal Processing, pp.3-8, 2001.
- [10] R. M. Dansereau, C. Li and R. A. Goubran, "Lip Feature Extraction Using Motion, Color and Edge Information," in Proceedings of IEEE International Workshop on Haptic, Audio and Visual Environment and Applications, Sept., 2003.
- [11] Daehwan Kim, Jaemin Kim, Seongwon Cho, Yongsuk Jang, Sun-Tae Chung, Boo-Gyoun Kim, "Progressive AAM Based Robust Face Alignment," in Proceedings of World Academy of Science Engineering and Technology, Vol.21 pp.488-492, 2007.
- [12] OpenCV, <http://opencvlibrary.sourceforge.net/> (accessed on 2009. 5. 27)
- [13] M. B. Stegmann, B. K. Ersboll, R. Larsen, "FAME - A Flexible Appearance Modelling Environment," IEEE Transactions on Medical Imaging, Vol.22, Issue 10, pp.1319-1331, Oct. 2003.
- [14] T. Cootes, G. J. Edwards, and C. J. Taylor. "Active Appearance Models," in Proceedings of European Conference on Computer Vision, Vol.2, pp.484 - 498, 1998.
- [15] Rafael C.Gonzalez, Richard E.Woods, 'Digital Image Processing', the 2nd edition, Prentice Hall, 2003.
- [16] T.F. Cootes, G.J. Edwards and C.J. Taylor, "Active Appearance Models," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.681-685, June 2001.
- [17] 한은정, 이의철, 고유진, 박강령, 민경하, "AAM 모델에 기반한 다양한 입술 모양에 강인한 얼굴 특징 추출에 관한 연구", 대한전자공학회 2008년 추계학술대회, 31권, 2호,

pp.643-644, 2008.

- [18] 이경희, 변혜란, “얼굴 요소의 영역 추출 및 snakes를 이용한 윤곽선 추출”, 정보과학회논문지, 제27권, 제7호, 2000.
- [19] Logitech Camera, <http://www.logitech.com/> (accessed on 2009. 5. 27)
- [20] I. Matthews and S. Baker, “Active Appearance Models Revisited,” International Journal of Computer Vision, Vol.60, No.2, pp.135-164, 2004.



강 병 준

e-mail : kangbyj@gmail.com

2004년 상명대학교 소프트웨어학과(학사)

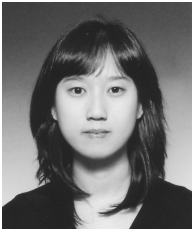
2006년 상명대학교 컴퓨터과학과(석사)

2009년 상명대학교 컴퓨터과학과(박사)

2009년 2월~2009년 5월 한양대학교 의공
학연구소 연구조교수

2009년 6월~현 재 한국전자통신연구원 휴먼인식기술연구팀 선임
연구원

관심분야 : 바이오인식, 영상신호처리, 컴퓨터비전 등



한 은 정

e-mail : hanej@dongguk.edu

2009년 상명대학교 디지털미디어학부(학사)

2009년 2월~현 재 동국대학교 전자공학과
석사과정

관심분야 : 바이오인식, 컴퓨터비전 등



박 강 령

e-mail : parkgr@dongguk.edu

1994년 연세대학교 전자공학과(학사)

1996년 연세대학교 전자공학과(석사)

2000년 연세대학교 전기/컴퓨터공학과(박사)

2000년 2월~2003년 2월 LG전자기술원
선임연구원

2003년 3월~2008년 2월 상명대학교 디지털미디어학부 조교수

2008년 3월~현 재 동국대학교 전자공학과 부교수

관심분야 : Biometric 영상 처리, 컴퓨터 비전 등