

모바일 환경에서의 시각 음성인식을 위한 눈 정위 기반 입술 탐지에 대한 연구

A Study on Lip Detection based on Eye Localization for Visual Speech Recognition in Mobile Environment

송민규* · Thanh Trung Pham* · 김진영* · 황성택**

Song Min Gyu*, Thanh Trung Pham*, Jin Young Kim* and Hwang Sung Taek**

* 전남대학교 대학원 전자공학과

** 삼성전자 통신연구소 멀티미디어 Lab.

요 약

음성 인식 기술은 편리한 삶을 추구하는 요즘 추세에 HMI를 위해 매력적인 기술이다. 음성 인식기술에 대한 많은 연구가 진행되고 있으나 여전히 잡음 환경에서의 성능은 취약하다. 이를 해결하기 위해 요즘은 청각 정보 뿐 아니라 시각 정보를 이용하는 시각 음성인식에 대한 연구가 활발히 진행되고 있다. 본 논문에서는 모바일 환경에서의 시각 음성인식을 위한 입술의 탐지 방법을 제안한다. 시각 음성인식을 위해서는 정확한 입술의 탐지가 필요하다. 우리는 입력 영상에서 입술에 비해 보다 찾기 쉬운 눈을 이용하여 눈의 위치를 먼저 탐지한 후 이 정보를 이용하여 대략적인 입술 영상을 구한다. 구해진 입술 영상에 K-means 집산화 알고리즘을 이용하여 영역을 분할하고 분할된 영역들 중 가장 큰 영역을 선택하여 입술의 양 끝점과 중심을 얻는다. 마지막으로, 실험을 통하여 제안된 기법의 성능을 확인하였다.

키워드 : 립리딩, 눈 정위, GMM모델, K-means 집산화, 입술 검출

Abstract

Automatic speech recognition(ASR) is attractive technique in trend these day that seek convenient life. Although many approaches have been proposed for ASR but the performance is still not good in noisy environment. Now-a-days in the state of art in speech recognition, ASR uses not only the audio information but also the visual information. In this paper, We present a novel lip detection method for visual speech recognition in mobile environment. In order to apply visual information to speech recognition, we need to extract exact lip regions. Because eye-detection is more easy than lip-detection, we firstly detect positions of left and right eyes, then locate lip region roughly. After that we apply K-means clustering technique to devide that region into groups, than two lip corners and lip center are detected by choosing biggest one among clustered groups. Finally, we have shown the effectiveness of the proposed method through the experiments based on samsung AVSR database.

Key Words : Lip Reading, Eye localization. Gaussian Mixture Model, K-means clustering, Lip detection.

1. 서 론

IT기술의 획기적인 발전에 따라 HMI(Human Machine Interaction)를 이용한 제품들이 많이 등장하고 있다. 특히 과거에는 HMI 제품들을 다루기 위해 키패드를 이용하여 직접 조작하는 방식이 대부분이었으나 요즘은 터치패드나 음성인식과 같은 편리하고 간편한 방식들이 등장하고 있다. 그 중 음성인식기술은 손을 사용하지 못할 때나 멀리 떨어져있는 제품을 조작할 경우 발생하는 문제점을 해결케 한다. 과거 음성인식은 영화 속에서만 볼 수 있었던 것으로

현실에서는 낯선 용어였다. 하지만 최근 음성인식 기술이 발전함에 따라 음성인식을 이용한 ARS 서비스를 비롯하여 음성인식 게임, 음성인식 도어락, 음성인식 리모컨, 음성인식 기능이 탑재된 휴대폰 등 음성인식 기술을 응용한 상품들이 다양하게

출시되어 더 이상 음성인식이라는 말은 낯설지 않은 용어가 되었다. 하지만 실제 음성인식기술이 적용된 제품을 사용할 때 잘못된 음성인식으로 인해 오작동 되는 경우가 종종 발생하게 된다. 음성인식이 오작동 되는 가장 큰 이유는 실제 환경에서의 잡음문제이다. 잡음은 음성인식률 성능에 큰 영향을 준다[1]. 하지만 실생활에서 잡음은 항상 존재하기 때문에 이를 제거하는 것은 쉽지 않은 문제이다.

최근 이런 문제를 해결하기 위해 음성인식에 음성정보 이외 잡음에 영향을 받지 않는 영상정보를 접목시킨 시각 음성인식기술에 대한 연구가 활발히 진행되고 있다. 시각 음성인식기술은 사람의 입모양을 보고 음성을 인식하는 기

접수일자 : 2008년 11월 24일

완료일자 : 2009년 8월 6일

본 연구는 삼성전자 프로젝트와 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음(IIITA-2009-(c1090-0903-0008))

술로 'audio-visual speech recognition'이라는 분야로 연구되고 있다[2-4]. 이 시각 음성인식을 위해서는 정확한 입술의 탐지에 대한 연구가 선행되어야 한다.

본 논문에서는 모바일환경에서 시각 음성인식을 위한 입술탐지의 방법으로 눈의 위치를 탐지한 후 그 위치를 이용하여 입술의 후보영상을 찾은 후 그 후보영상에서 입술의 정위 및 추적 방법을 제안한다.

2. 관련 연구 및 제안 방법

눈과 입의 위치를 찾는 기존 연구로는 색 정보를 이용하여 얼굴, 눈 영역들을 추출하고 두 눈의 대칭성 및 위치를 반영하는 평가함수를 적용하여 눈을 인식하는 방법[5-7], 얼굴 영역에서 정해진 제약조건을 만족하는 눈동자 한 쌍이 찾아질 때까지 임계값을 변화하며 이진화를 반복 수행함으로써 눈을 찾는 방법[8], 고유얼굴 개념을 확장한 고유눈과 고유입을 이용하여 눈과 입을 찾는 방법들이 있다[9-10]. 본 논문에서는 그림1과 같이 모바일환경이라는 특수한 환경에서 입력 영상의 눈의 위치는 특정 위치에 존재할 것이라는 가정 하에 눈의 자유도에 제한을 두고 제한된 영역 안에서 눈의 위치를 탐지 후 탐지된 눈의 정위를 기반으로 얼굴의 기하학적 위치 관계를 이용하여 입술 후보영역을 구한 후 그 후보영역에서 정확한 입술 영역을 탐지한다. 눈 탐색은 적응 임계값을 이용하여 영상을 이진화 한 후 눈 후보영역으로 분할하고, GMM 기반으로 검증한다[11]. 그 후 탐지된 눈 정위를 기반으로 양 눈과 입의 기하학적 구조를 기반으로 입술 후보영상을 구한 후 K-means 집단화를 통해 입술의 영역을 분할하고 분할된 영역들에 필터를 적용하여 입술을 탐지하였다. 그림2는 제안된 방법의 흐름도를 보여준다.

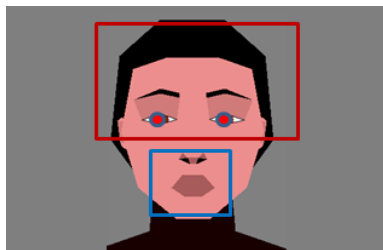


그림 1. 모바일환경에서 영상
Fig. 1. Image in mobile environment

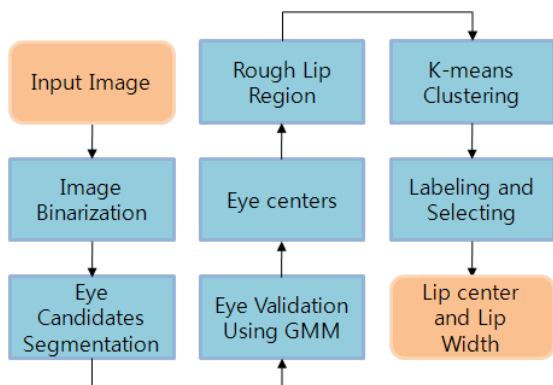


그림 2. 입술 탐색 흐름도
Fig. 2. Lip Detection Flowchart

3. 눈 탐색 및 검증

3.1 영상의 전처리

최초 입력 영상은 RGB좌표의 영상이다. RGB좌표 영상은 우선적으로 R, G, B 각각에 대한 정보를 다 가지고 있어야 하기 때문에 정보량이 많아지고 주변 조명의 변화에 상대적으로 민감하다. 그래서 이를 Y-휘도(밝기정보)와 Cb, Cr-색차신호(색상정보) 좌표로 변환하여 사용하였다. YCbCr좌표 영상은 CbCr이 상대적으로 Y신호에 비해 적은 정보를 가지고 있기 때문에 Y신호만 사용이 가능하며, RGB좌표 영상에 비해 조명의 변화에 보다 강인하다. 다음 식들은 RGB값을 Y값으로 바꾸는 변환식이다.

$$Y = 0.299R + 0.587G + 0.114B$$

3.2 눈 후보영역의 분할

눈의 위치를 찾는 시작점은 “눈은 항상 피부나 다른 부분에 비해 어둡다”는 점이다. 이 점을 이용하여 우리는 눈 입력 영상에서 눈으로 생각될 수 있는 영역을 분할하였다.

모바일 환경에서 입력 영상은 정면을 응시하고 있는 한 명의 얼굴 영상이라 가정할 수 있다. 가정된 사실에 기반을 두어 우리는 입력 영상에서 눈이 위치하는 곳을 대략적으로 유추할 수 있다. 따라서 우리는 눈이 존재할 곳으로 유추되는 영역에 마스크를 적용, 그 영역만 계산하여 계산 속도를 향상시키고 오류를 줄였다.

다음 그림3은 적응 분할 알고리즘이다. 마스크가 적용된 영상은 임계값 적용방법을 이용하여 이진화한다. 각 픽셀들은 픽셀의 Y값이 임계값보다 작으면 1, 크면 0을 가진다. 이렇게 이진화된 영상을 이용하여 우리는 눈 후보영역을 분할한다. 일반적으로 Y값은 조명의 변화에 민감하게 변화하기 때문에 임계값이 고정적일 경우 분할이 잘 되지 않을 수 있다. 따라서 우리는 분할이 제약조건을 만족할 때까지 임계값에 변화를 주는 적응 분할 방법을 도입하였다.

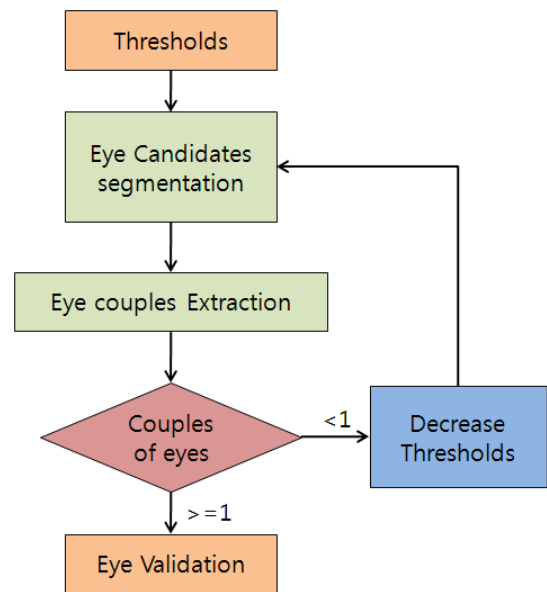


그림 3. 적응 분할 알고리즘
Fig. 3. Adaptive Segmentation Algorithm

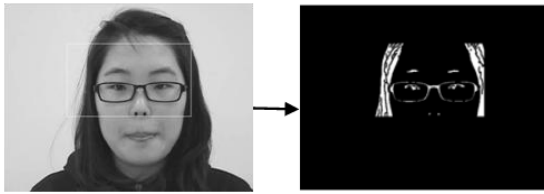


그림 4. (좌) 입력영상 Y좌표 이미지 (우) 분할된 눈 후보영역

Fig. 4. (L) Input Y Image (R) Eye Candidates Regions

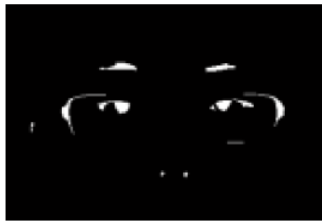


그림 5. 최종 분할된 눈 후보영역
Fig. 5. Eye Candidates Remaining

3.3 GMM을 이용한 눈 영역 검증

분할된 눈 후보 영역은 실제 눈 영역 이외 다른 영역들도 포함되어 있다. 실제 눈 영역을 찾아내기 위해 우리는 GMM을 이용하여 실제 눈 영역을 검증하였다.

먼저, 검증단계에서 시간을 줄이기 위해 실제 눈의 기하학적 특징을 이용하여 미리 오류영역을 제거하였다. 예를 들어 후보영역 중 크기가 너무 작거나 큰 경우, 높이가 너비보다 3배정도 큰 경우는 미리 제거 되었다. 다음 그림4는 입력영상에 분할 알고리즘을 적용하여 눈의 후보영역을 분할한 그림이고, 그림5는 분할된 영상에서 눈 영역이 아니라 간주되는 영역이 제거된 그림이다.

GMM(Gaussian Mixture Model)은 화자인식에 많이 사용되고 있는 방법으로 여러 개의 가우시안 확률밀도(Gaussian probability density) 함수들에 각각의 가중치를 준 다음, 이를 선형 결합함으로써 임의의 모양을 갖는 확률밀도 함수를 표현할 수 있는 확률분포 모델이다.

일반적으로 분할된 후보영역들 중 눈과 눈썹 영역은 매우 유사하여 종종 잘못된 결과가 유발한다. 따라서 우리는 눈 탐지의 정확도를 높이기 위해 학습 데이터와 입력데이터에 눈만 존재하는 데이터가 아닌 두 눈과 눈썹을 포함하는 데이터를 사용하였다. 그림6은 입력데이터들로서 최종 분할된 눈 후보영역들 중에서 거리적 특징을 이용하여 좌, 우 영역을 쌍으로 묶은 것을 보여준다. 각 영역은 중심점을 기준으로 높이를 너비의 1.5배 정도 넓게 구하여 눈썹 영역이 포함되게 하였다. 쌍으로 합쳐진 영역들은 36x8크기로 다운 샘플링(down-sampling)을 통하여 일정한 크기로 정규화하였다. 특징 값 추출 부분에서는 Harr wavelet 변환을 사용하였고 추출된 특징 값의 차수를 줄이기 위해 PCA를 사용하였다. 우리는 학습 데이터로 얻은 특징값을 이용하여 Gaussian Mixture 확률 밀도 함수를 얻은 후 각각의 눈 후보영역에 대한 확률을 계산하고 그 중에 가진 가장 큰 확률을 가지는 후보영역을 눈 영역으로 결정하였다. 그림7은 눈의 선택하는 방법의 흐름을 보여준다.

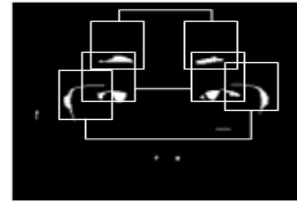


그림 6. 쌍을 이룬 입력 데이터
Fig. 6. Couple of Input data

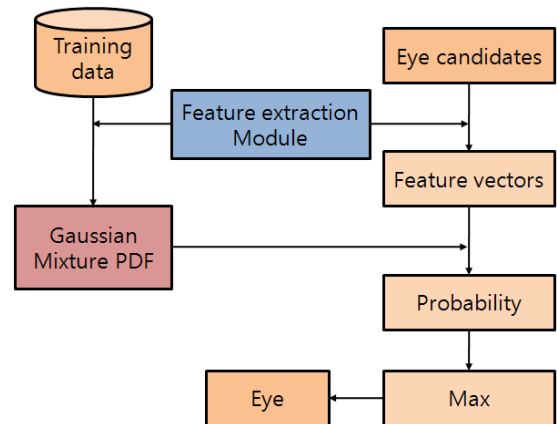


그림 7. GMM기반 눈 선택 흐름도
Fig. 7. Eye select flowchart based on GMM



그림 8. 단일 눈의 데이터
Fig. 9. Single eye samples



그림 9. 두 눈의 데이터
Fig. 9. Coupled eye samples



그림 10. 눈썹을 포함한 두 눈의 데이터
Fig. 10. Coupled eye with eyebrow samples

3.4 눈 영역 검출 실험

우리는 실험을 통해 제안한 눈 탐지 방법의 성능을 검증하였다. 실험은 삼성시청각 음성인식DB를 사용하여 진행하였다. 이DB는 모바일 환경에서는 얼굴이 항상 중앙에 존재한다는 가정 하에 105명의 사람이 총3가지 환경(표준 환경, 실내 환경, 실외 환경)에서 정면을 바라보는 640 x480크기로 제작된 영상이다. 제안된 방법의 성능 비교를 위해 그림8,9,10과 같은 3종류의 데이터들을 만들어 단일 눈만 검증, 두 눈을 쌍으로 검증, 두 눈과 눈썹을 같이 검증하는 경우로 각각의 학습데이터에 맞게 입력데이터를 만들어 실험을 하였다. 실험결과 제안된 방법의 성능이 가장 좋은 탐지 성공률을 보였다. 표1은 실험 결과를 비교한 것이다. 실험결과 제안된 방법은 표준 환경에서 98.1%, 실내 환경에서 96.2%, 실외 환경에서 94.3%로 총 96.2%의 탐지 성공률을

보였다. 그림10은 다양한 조건에서의 눈 검출 결과를 보여 준다.

표 1. 다양한 눈 검증 방법들의 실험결과 비교
Table 1. Comparison of different validation methods

	Standard DB	Indoor DB	Outdoor DB	Number of Samples
Single Eye	60%	--	--	105
Couple eye	90.3%	87.2%	85.6%	105
Couple eye with eyebrow	98.1%	96.6%	94%	105

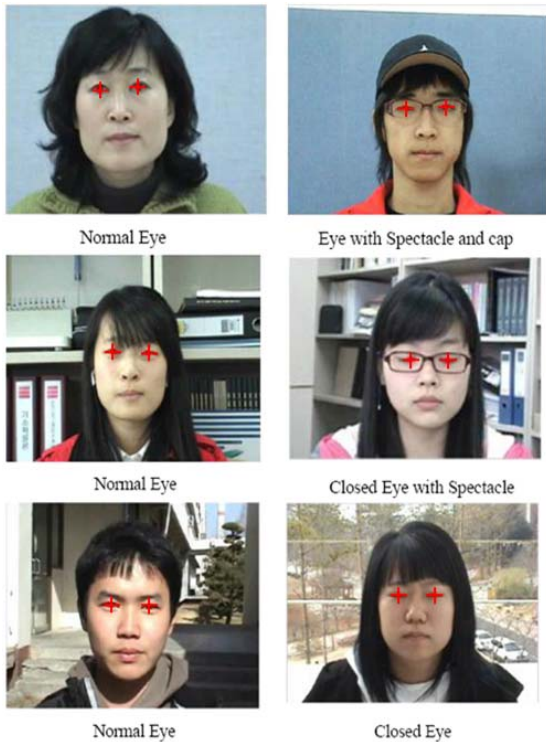


그림 11. 다양한 조건에서의 눈 검출 결과
Fig. 11. Eye detection results in variable environment

4. 입술의 정위

입술을 탐지하는 일반적인 방법으로는 색상 정보를 이용하는 방법이 있다. 피부색 정보를 이용하여 얼굴영역을 탐지하고 그 영역에서 입술의 색상정보로 입술 탐지가 이루어진다. 하지만 색상 정보만을 이용할 경우 주위 환경의 조명 상태에 색상정보의 변화가 심해 입술을 탐지하는데 많은 오류를 야기한다. 또한 피부색이나 입술색을 정의하는데 있어 임계값을 어떻게 정해야 할지에 대한 문제도 발생하게 된다.

본 논문에서는 입술 탐지를 위해 탐지된 눈 정위에 기반을 둔 입술 탐지 방법을 제안한다. 앞 장에서 언급한 눈의 탐지방법을 이용, 눈의 위치를 찾은 후 얼굴의 기하학적 정보를 이용하여 대략적인 입술 후보영역을 구한다. 탐지된 입술 후보영역에 K-means 알고리즘을 적용 몇 개의 그룹

으로 분할 한 후 가장 어두운 집단들을 선택하고 이를 레이블링 후 각 객체를 검증하여 입술의 위치를 탐지한다.

4.1 얼굴의 기하학적 특징에 근거한 입술 후보영상 탐지

일반적으로 사람의 얼굴에서 좌우 눈은 코와 입술을 연결하는 축을 기준으로 대칭으로 존재한다. 통계학적으로 좌측 눈의 안쪽의 끝부터 오른쪽 눈의 안쪽의 끝까지의 거리는 대략적으로 입술의 너비와 비슷하다. 또한 좌측 눈의 바깥쪽의 끝부터 오른쪽 눈의 바깥쪽의 끝까지의 거리는 두 눈의 중심에서부터 입의 중심까지의 거리와 비슷하다. 이점을 이용하여 입이 존재하는 영역을 찾는다. 그림12는 얼굴의 기하학적 특징을 이용하여 두 눈으로부터 찾아진 대략적인 입술 영상을 보여준다.

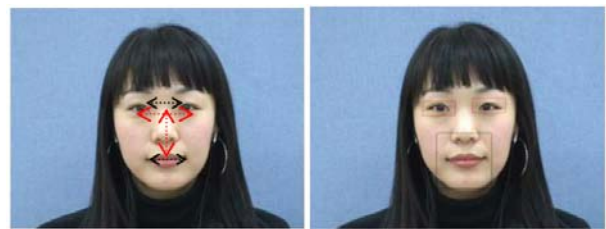


그림 12. 눈과 입의 거리 관계를 이용한 입술후보영상
Fig. 12. Rough cropped lip region using distance relation of eye and mouth

4.2 K-means 알고리즘을 이용한 분할

K-평균 알고리즘(K-means algorithm)은 주어진 데이터를 특정 성질에 기초해서 k 묶음으로 나누는 방법 중 하나이다[12-13]. 데이터가 벡터 공간을 이룬다고 가정하고, 이 알고리즘은 각 묶음의 분산도를 최소화한다. 이것을 식으로 나타내면 다음과 같다.

$$V = \sum_{i=1}^k \sum_{j \in S_i} |x_j - \mu_i|^2$$

여기서 k개의 묶음은 $S_i, i = 1, 2, \dots, k$ 으로 표시되고, μ_i 는 각 묶음에 속한 점들의 무게중심이다

맨 처음, 각 점들을 k개 집합으로 나눈다. 이 때 임의로 나누거나, 어떤 휴리스틱을 사용할 수도 있다. 그 다음 각 집합의 무게 중심을 구한다. 그 다음, 각각의 점들을 방금 구한 무게중심 가운데 제일 가까운 것에 연결 지음으로써 새로이 집합을 나눌 수 있다. 이 작업을 반복하면 점들이 소속된 집합을 바꾸지 않거나, 무게중심이 변하지 않는 상태로 수렴될 수 있다.

다음은 K-means 알고리즘의 간단한 요약 및 흐름도이다.(그림 13)

K-means 알고리즘 - 몇 개의 그룹인지를 결정하고 다음 3단계를 수렴 될 때까지 무한 반복한다. 수렴 조건은 그룹화 될 객체가 존재하지 않을시 까지 이다.

1. 각 그룹의 중심을 결정한다.
2. 각 객체들과 중심들 사이에 거리를 계산 한다.
3. 최소거리에 기반을 두어 객체를 그룹화 한다.

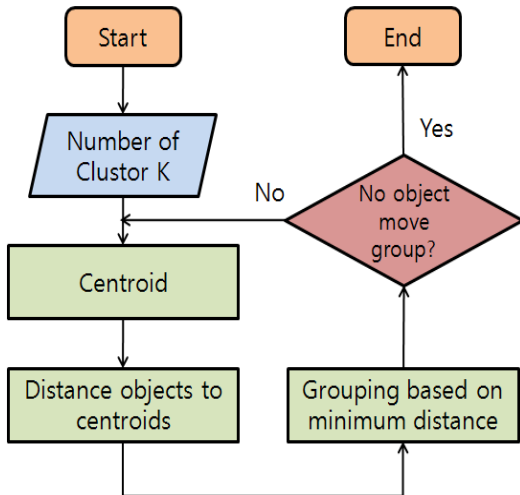


그림 13. K-means 알고리즘
Fig. 13. K-means Algorithm

K-means 알고리즘은 실제 이용할 때 빠르게 수렴하기 때문에 널리 쓰인다. 다만 이 알고리즘을 시작할 때 k값을 적절히 정해 주어야 하는 문제점이 있다. 만일, 데이터가 자연스럽게 분포되어 있지 않으면 이상한 결과를 얻을 수도 있다.

우리는 계산속도의 향상을 위해 입술 후보영상을 2:1로 다운샘플링(down-sampling) 한 후 K-means 알고리즘을 적용하여 영역을 분할하였다. K-means 집단화를 위해 RGB값을 이용한 3차원 벡터를 사용하였고 k값은 실험을 통해 가장 잘 분리가 되는 값으로 4를 선택하여 사용하였다. 분할된 영역들 중 가장 어두운 집단들을 선택하여 입술 후보를 얻을 수 있다. 그림 14는 입술 후보영역들을 구하는 모습을 보여준다.

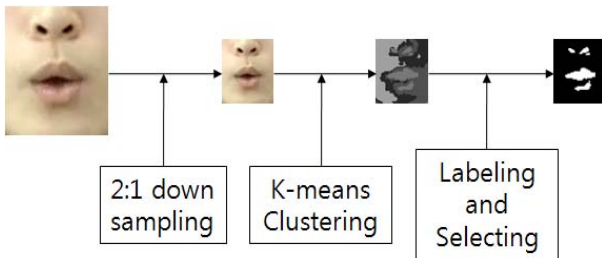


그림 14. 입술 후보영역 추출
Fig. 14. Lip candidates region extraction

4.3 Labeling 기법을 이용한 특징점 추출

마지막으로 입술의 위치 및 크기를 구하기 위해 우리는 분할된 입술 후보영역들을 레이블링 후 가장 너비가 넓은 객체를 선택하고 그 객체를 이용하여 입술의 너비와 중심을 구한다. 일반적으로는 구해진 객체의 무게중심을 구해 중심점을 얻어내지만 구해진 객체가 좌우 대칭이 아닐 경우 잘못된 중심점을 얻을 수 있다. 따라서 우리는 좌우 끝 점을 이용하여 입술의 너비를 구하고 끝점들의 중앙을 입술의 중심으로 구하였다. 다음 그림 15는 분할 후 선택된 객체들 중 가장 너비가 넓은 객체를 이용하여 입술의 양끝 점을 구하는 모습을 보여준다.

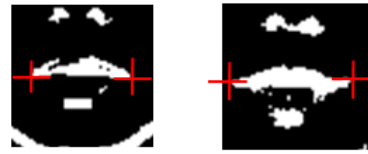


그림 15. 입술의 좌, 우 끝점 검출
Fig. 15 Lip corners detection

표 2. 입술 탐지 실험 결과

Table 2. Results of detect Lip

	Standard DB		Indoor DB	
	제안방법	색상방법	제안방법	색상방법
성공	103	99	100	92
실패	2	6	5	13
탐지율	98.1%	94.2%	95.2%	87.6%
	Outdoor DB		Total	
	제안방법	색상방법	제안방법	색상방법
성공	97	87	300	278
실패	8	18	15	27
탐지율	92.4%	82.9%	95.2%	88.3%

5. 실험 및 결과 고찰

우리는 실험을 통해 제안된 눈 탐지 알고리즘의 성능을 검증하였다. 실험은 2.4절에서 설명한 삼성 시청각 음성인식DB를 이용하여 진행하였다. 우리는 색상기반의 입술 탐지 방법을 이용하여 제안된 방법과 성능 비교를 하였다. 표 2는 제안된 방법을 이용한 입술 탐지 실험결과이다. 실험 결과 표준, 실내, 실외의 모든 환경에서 제안된 방법이 색상을 이용한 방법에 비해 높은 탐지율을 보였다. 특히 배경이 단순한 표준은 탐지율이 약 4%정도 높았지만 주위 배경이 복잡한 실내나 실외 환경에서는 탐지율이 약 10%정도 더 높은 성능을 보였다. 전체적으로 탐지율이 7%정도 높은 성능을 보였다.

제안된 방법에 의한 실험 중 오류가 발생한 경우의 대다수는 눈의 정확한 탐지가 이루어지지 않았을 경우 발생하였으며, 눈을 찾았을 경우, 입술의 탐지율은 99%의 높은 성능을 보였다. 따라서 보다 강인한 눈 탐지 알고리즘을 사용할 경우, 실험 결과 보다 더 높은 입술 탐지율을 나타낼 것으로 예상된다. 그림16, 17, 18는 각각의 환경에서 탐지된 눈과 입술을 보여준다.

6. 결론 및 향후 과제

본 논문에서는 모바일 환경에서 시각 음성인식을 위한 눈 정위 기반의 입술 탐지방법을 제안하였다. 시각 음성인식을 위해서는 정확한 입술 탐지방법이 요구되기 때문에 우리는 우선 눈의 위치를 탐지하고, 탐지된 눈의 위치 정보를 이용하여 입술의 위치를 탐지 하였고 삼성 시청각 음성인식DB를 이용한 실험을 통해 제안된 방법의 입술 탐지 성능을 검증하였다.



그림 16. 표준 환경 영상에서의 입술 탐지
Fig. 16. Lip detection in standard environment



그림 17. 실내 환경 영상에서의 입술 탐지
Fig. 17. Lip detection in indoor environment

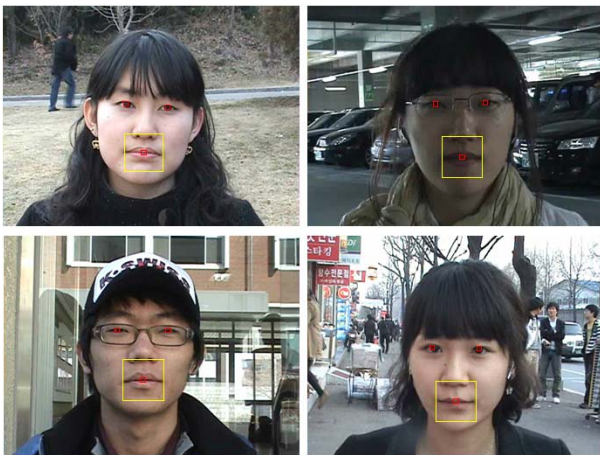


그림 18. 실외 환경 영상에서의 입술 탐지
Fig. 18. Lip detection in outdoor environment

제안된 방법의 가장 큰 문제점은 입술 탐지의 선행 조건으로 눈의 위치를 이용하는데, 눈의 탐지가 제대로 이루어지지 않을 경우 잘못된 입술 탐지 결과를 보이는 것이다.

특히 안경을 착용한 사람의 경우 안경테의 두께가 너무 두꺼울 경우나 안경테가 눈을 가렸을 경우, 혹은 선글라스를 착용하여 눈이 안보일 경우 등에는 탐지 성능이 크게 저하된다. 이런 경우에 발생하는 문제점들은 차후 선택적 방법, 즉 제한된 영역에서 눈을 탐지하지 못하였을 경우 다른 알고리즘을 이용하는 방법을 추가하여 수정할 계획이다.

또한 제안된 알고리즘을 실제 스마트폰에 적용할 경우, 일반 컴퓨터와 달리 제안된 사양에서 작업이 수행되기 때문에 부동 소수점(float point) 연산에 대한 연구나 수행속도 향상을 위한 연구들에 대해서도 병행하여 진행할 계획이다.

참 고 문 헌

- [1] Pedro J. Moreno, "Speech Recognition in Noisy Environment," *Ph.D. Thesis, ECE Department, CMU*, May 1996.
- [2] S. Dupont and J. Luetttin, "Audio-Visual Speech Modelling for Continuous Speech Recognition," *Proceedings of IEEE Transactions on Multimedia*, pp.141-151, 2000
- [3] J. N. Gowdy, A. Subramanya, C. Bartels, J. Bilmes, "DBN-based multi-stream models for audio-visual speech recognition." *proc. IEEE Int. conf. Acoustics, Speech, and Signal Processing*, pp.993-996, 2004
- [4] Jeff A. Bilmes and Chris Bartels, "Graphical Model Architectures for Speech Recognition," *IEEE Signal Processing Magazine*, vol.22, pp.89-100, 2005
- [5] E. Saber, A. M. Tekalp, "Frontal-view Face Detection and Facial Feature Extraction Using color, Shape and Symmetry Based Cost Function," *Pattern Recognition Letters* 19, pp.669-680, 1998
- [6] N. Eyoeno, A. Caplier, P. Y. Coulon, "A New Color transformation for Lips Segmentation," *Proceedings of IEEE International conference on Acoustic, Speech, Signal Processing*, pp.557-560, 1993
- [7] R. Stiefelhagen, U. Meier, J. Yang, "Real-Time Lip-Tracking for Lip Reading," *Proceedings of Eurospeech 97, 5th European Conference on Speech Communication and Technology*, 1997
- [8] J. Yang, R. Stiefelhagne, U. Meier, A. Waibel, "Real-Time Face and Facial Feature Tracking and Application," *Proc. Auditory-Visual Speech Processing*, pp.79-84, 1998
- [9] B. Moghaddam, A. Pentland, "Probabilistic Visual Learning for Object Detection," *IEEE ICCV'95*, pp.786-793, 1995
- [10] B. Moghaddam, W. Wahid, A. Pentland, "Beyond Eigen Faces : Probabolostoc Matching for Face Recognition," *IEEE Conf. Automatic Face and Gesture Recognition*, pp.30-35, 1998
- [11] T. T. pham, J. Y. Kim, S. Y. Na, S. T. Hwang, "Robust Eye Localization for Lip Reading in Mobile Environment," *Proceddings of SCIS&ISIS*

in Japan, pp.385-388, 2008

- [12] MacQueen, J. B. "Some Methods for Classification and Analysis of Multivariate Observations," *In Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281-297. 1967
- [13] Andrew W. Moore, "K-means and Hierarchical Clustering", *Tutorial Slides in School of Computer Science Carnegie Mellon University*, <http://www.csc.mcmu.edu/~awm>, <http://www.autonlab.org/tutorials/kmeans11.pdf>

저 자 소 개



송민규(Song Min Gyu)

2004년 : 전남대학교 전자정보통신공학과
공학사
2006년 : 동 대학원 전자공학과 석사
2006년~현재 : 동 대학원 전자공학 박사
과정

관심분야 : 시청각 개인추적, 시청각 신호처리, 임베디드 시스템
Phone : 062-530-0370
Fax : 062-530-1759
E-mail : smg686@lycos.co.kr



Thanh Trung Pham

2007년 : B.S degree in Mathematics and Informatics from HCMC University of Natural Sciences
2007년~현재 : M.S course in ECE from Chonnam National University.

관심분야 : DSP, Computer vision and Pattern Recognition
Phone : 062-530-0370
Fax : 062-530-1759
E-mail : trung.ptt@gmail.com



김진영(Jin Young Kim)

1986년 : 서울대학교 전자공학 공학사
1988년 : 동 대학원 전자공학 석사
1994년 : 동 대학원 전자공학 박사
1995년~현재 : 전남대학교 교수

관심분야 : 음성처리, 시청각 신호처리
Phone : 062-530-1757
Fax : 062-530-1759
E-mail : beyondi@jnu.ac.kr



황성택(Hwang Sung Taek)

1988년 : 서강대학교 물리학 학사
1990년 : 동 대학원 물리학 석사
2001년 : 동 대학원 물리학 박사
1993년~현재 : 삼성전자 통신연구소 수석
연구원

관심분야 : 광통신, FTTH시스템, OCR, 모션인식
Phone : 031-279-5157
Fax : 031-279-5255
E-mail : shwang@samsung.com