

# 임베디드 환경에서의 실시간 립리딩 시스템 구현

김 영 운<sup>†</sup> · 강 선 경<sup>††</sup> · 정 성 태<sup>†††</sup>

## 요 약

본 논문은 임베디드 환경에서의 실시간 립리딩 방법을 제안한다. 기존 PC 환경에 비하여 임베디드 환경은 사용할 수 있는 자원이 제한적이므로, 기존 PC 환경의 립리딩 시스템을 임베디드 환경에서 실시간으로 구동하기는 어렵다. 이러한 문제를 해결하기 위해 본 논문은 임베디드 환경에 적합한 입술영역 검출 방법과 입술 특징추출 방법, 그리고 발성 단어 인식 방법을 제안한다. 먼저 정확한 입술영역을 찾기 위해 얼굴 색상정보를 이용해 얼굴영역을 검출하고 검출된 얼굴 영역에서 양쪽 두 눈의 위치를 찾아 기하학적 관계를 이용해 정확한 입술영역을 검출한다. 검출된 입술영역에서 주위 환경 변화에 따른 조명 변화에 강인한 특징을 추출하기 위해 히스토그램 매칭과 입술 폴딩, RASTA 필터를 적용하고 주성분 분석(PCA)을 이용한 특징계수를 추출해 인식에 사용하였다. 실험결과 CPU 806Mhz, RAM 128MB 사양의 임베디드 환경에서 발성 단어에 따라 1.15초에서 2.35초까지의 처리 속도를 보였으며, 180개의 단어 중 139개의 단어를 인식해 77%의 인식률을 얻을 수 있었다.

키워드 : 립리딩, 얼굴검출, 입술검출, 임베디드 시스템, 사용자 인터페이스

## Real Time Lip Reading System Implementation in Embedded Environment

Young-Un Kim<sup>†</sup> · Sun-Kyung Kang<sup>††</sup> · Sung-Tae Jung<sup>†††</sup>

## ABSTRACT

This paper proposes the real time lip reading method in the embedded environment. The embedded environment has the limited sources to use compared to existing PC environment, so it is hard to drive the lip reading system with existing PC environment in the embedded environment in real time. To solve the problem, this paper suggests detection methods of lip region, feature extraction of lips, and awareness methods of phonetic words suitable to the embedded environment. First, it detects the face region by using face color information to find out the accurate lip region and then detects the exact lip region by finding the position of both eyes from the detected face region and using the geometric relations. To detect strong features of lighting variables by the changing surroundings, histogram matching, lip folding, and RASTA filter were applied, and the properties extracted by using the principal component analysis(PCA) were used for recognition. The result of the test has shown the processing speed between 1.15 and 2.35 sec. according to vocalizations in the embedded environment of CPU 806Mhz, RAM 128MB specifications and obtained 77% of recognition as 139 among 180 words were recognized.

Keywords : Lip Reading, Face Detection, Lip Detection, Embedded System, Human Interface

## 1. 서 론

최근 휴대가 간편한 개인 모바일 장치들이 많이 사용되면서 사용자 친화적인 새로운 인터페이스 기술들이 요구되고 있다. 그중 음성인식 기술이 실용화 되었지만 소음으로 인한 인식률 저하가 큰 문제로 지적되고 있다. 이 문제를 개선하기 위한 방법으로 소음을 제거하는 방법 등 여러 가지

가 제안되었는데, 그 중의 한 방법은 화자의 입술 움직임을 분석하여 발성 단어를 인식하는 립리딩(lip reading)을 활용하는 것이다. 소음과 같은 잡음환경에서 립리딩과 음성인식 기술을 함께 사용함으로써 음성 인식률이 향상 되는 것을 알 수 있다[1-4].

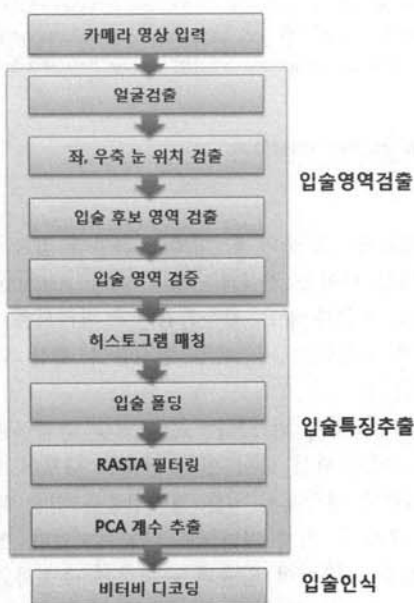
립리딩을 수행하기 위해서는 크게 입술 영역검출과 입술 특징추출, 그리고 발성 단어 인식 단계로 나누어 볼 수 있다. 먼저 입력된 영상에서 입술 영역검출과 입술 특징을 추출 방법은 크게 두 가지 기법으로 나눌 수 있다. 첫째는 기하학적 모델을 이용하여 입술의 윤곽선을 추출하고 추출된 입술 모델의 계수를 특징 값으로 사용하는 모델 기반 방법이다[5-7]. 둘째는 입력 영상에서 입술 영역을 추출한 다음

※ "이 논문은 2009년 교육과학기술부로부터 지원받아 수행된 연구임" (지역 거점연구단육성사업/헬스케어기술개발사업단)  
† 준 회 원 : 원광대학교 컴퓨터공학과 박사과정  
†† 준 회 원 : 원광대학교 컴퓨터공학과 공학박사  
††† 총신회원 : 원광대학교 컴퓨터공학과 교수(교신저자)  
논문접수 : 2009년 12월 4일  
수정일 : 2010년 3월 11일  
심사완료 : 2010년 3월 11일

에 픽셀 값에 주성분 분석, 선형 판별 분석, 이산 코사인 변환 등의 변환을 적용하여 특징의 크기를 줄여서 사용하는 픽셀 기반 방법이다[8-10]. 일반적으로 픽셀 기반 방법에 비하여 모델 기반 방법이 더 많은 계산을 요구하므로 본 논문에서는 임베디드 시스템 환경에 적합한 픽셀 기반 방법을 이용한 입술 영역검출과 입술 특징추출 방법을 사용하였다.

본 논문에서는 픽셀 기반 방법 중 주성분 분석(PCA: Principal Component Analysis)을 사용 하여 입술 영역의 특징을 추출 하였다. PCA는 다른 기법들에 비해 기저벡터를 계산하는데 있어 많은 계산량이 요구되지만 특징 정보의 차원 축소를 통해 정보를 압축해주는 알고리즘으로 가장 좋은 결과를 얻을 수 있다[11]. PCA는 부동 소수점 연산과 같은 많은 연산을 필요로 한다. PC 환경에서는 부동 소수점 연산에 제한이 없으나 임베디드 시스템에서는 부동 소수점 연산에 제한이 많이 있다. 이를 해결하기 위해서 부동 소수점 연산을 고정 소수점 연산으로 변환하여 처리 속도를 최적화 하였다.

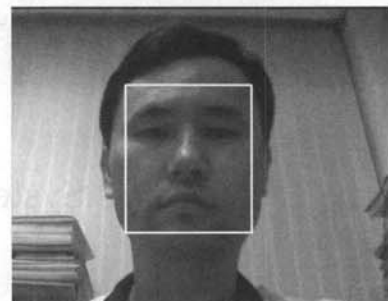
(그림 1)은 임베디드 환경에서 실시간으로 립리딩을 수행하기 위한 전체적인 알고리즘이 나타나 있다. 먼저 임베디드 시스템에 내장된 카메라를 이용해 영상을 획득하고 피부 색상정보를 이용해 얼굴영역을 검출한다. 검출된 얼굴 영역에서 좌, 우측 양쪽 눈의 위치를 검출하고, 기하학적 특징을 이용해 입술영역을 검출한다. 검출된 입술영역에서 히스토그램 매칭을 통해 조명변화에 대한 보상을 수행하고 입술 폴딩을 이용해 입술 영상을 절반으로 줄여 처리 속도를 향상시킨다[12]. 또한 RASTA(Relative SpecTrAl) 필터를 통해 불필요한 특징을 제거해 준다[12, 13]. 마지막으로 주성분 분석(PCA)을 이용해 특징계수를 추출하고 비터비 디코딩 알고리즘을 이용해 발성된 단어를 인식한다.



(그림 1) 임베디드 환경에서의 립리딩 알고리즘  
(Fig. 1) The Lip Reading Algorithm in Embedded Environment

## 2. 얼굴검출

본 논문에서는 임베디드 시스템의 전면부에 부착되어 있는 카메라를 이용해 발성자의 얼굴 영상을 입력받아 사용하였으며, 카메라와 얼굴 사이의 거리는 얼굴 전체가 화면에 나오도록 하였다. 카메라로부터 입력되는 영상은 조명과 주위 환경에 따라 얼굴 영역의 색상이 다양하게 변한다. 따라서 본 논문에서는 입력 영상의 첫 프레임에서 영상의 중앙 부분에 영역을 설정하여 영역 내에 포함된 픽셀들의 색상을 분석하여 이와 유사한 색상을 가지는 영역을 얼굴 영역으로 검출하였다[14]. (그림 2)에는 입력된 영상에서 얼굴 영역을 검출한 결과가 나타나 있다.

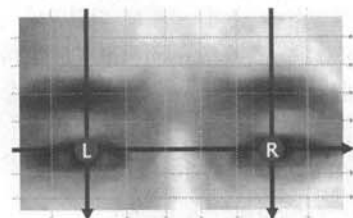


(그림 2) 얼굴 검출 결과  
(Fig. 2) Face detection result

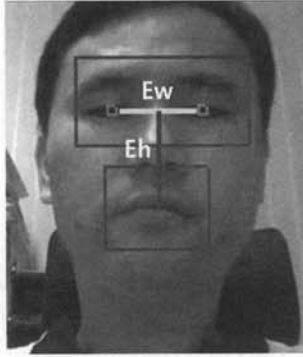
## 3. 입술검출

본 논문에서는 입술검출 방법으로 얼굴영역에서 양쪽 두 눈을 먼저 찾고 두 눈으로부터 기하학적 특징을 이용해 입술영역을 찾는 방법을 사용하였다[14]. 먼저 눈과 눈썹은 피부에 비하여 어두운 특성을 가지고 있으므로 명암도를 이용하여 눈을 검출한다. (그림 3)에는 명암도가 가장 작은 값을 탐색한 행이 굵은 선으로 표시되어 있고, 열 중에서 제일 작은 값과 그 다음으로 작은 값을 가지는 열이 굵은 선으로 표시되어 있다. 이와 같이 행과 열이 만나는 교차점 2개를 찾아 좌측 눈과 우측 눈의 위치로 결정한다.

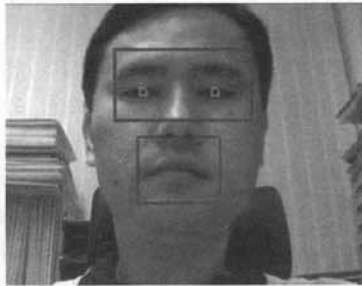
위와 같이 정확한 눈 위치를 찾은 후 기하학적 특징을 이용해 입술 영역을 찾는다. (그림 4)는 좌측 눈과 우측 눈 간의 거리( $E_w$ )에 1.4배를 계산해 입술 영역( $E_h$ )을 찾은 그림이다. 또한 (그림 5)는 눈동자와 입술 영역 검출 결과 그림



(그림 3) 좌, 우측 눈 위치 검출 결과  
(Fig. 3) Left and right eye detection result

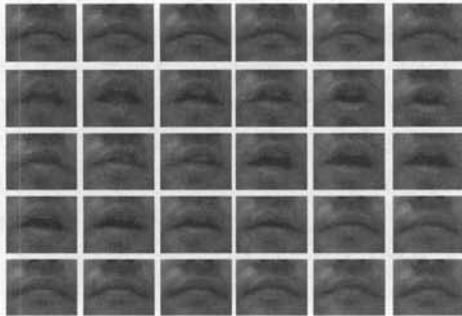


(그림 4) 입술 후보영역을 찾기 위한 기하학적 정보  
(Fig. 4) Geometrical constraint to search lip candidate area



(그림 5) 눈동자와 입술 영역 검출 결과  
(Fig. 5) Eye and lip area detection result

이다. (그림 6)에는 일련의 동영상으로부터 검출된 입술 영역의 예가 나타나 있다.



(그림 6) 검출된 입술 영역 예  
(Fig. 6) Samples of detected lip

#### 4. 특징추출

본 논문에서는 히스토그램 매칭(Histogram matching)과 입술영상 접기(Lip folding), RASTA(Relative SpecTrAl) 필터링을 적용하고 주성분 분석(PCA)을 이용해 입술 특징 정보를 추출하였다. 입술 영상의 색상정보는 조명에 따라서 그 분포가 쉽게 변한다. 본 논문에서는 이러한 문제를 해결하기 위한 방안으로 히스토그램 매칭을 사용하였다. 히스토그램 매칭은 서로 다른 두 영상을 비교하여 상대적으로 비슷한 히스토그램을 갖게 하는 방법이다. 학습 데이터 환경

의 히스토그램과 인식 환경의 히스토그램을 비교 했을 때 서로 다른 히스토그램 분포는 인식을 저하의 원인이 된다. 따라서 본 논문에서는 매핑 함수를 만들어 실제 인식 환경의 히스토그램을 학습 데이터 환경의 히스토그램과 비슷한 분포를 가지도록 매핑 하여 조명의 변화를 보상하였다. (그림 7)은 원본 입술 영상과 히스토그램 매핑을 적용한 입술 영상을 보여 주고 있다.

입술 영상은 화자의 머리 회전이나 상하좌우 기울어짐이 없다는 전제 하에 입술의 기하학적 대칭성에 의거하여 입술 영상을 절반으로 접을 수 있다. 입술의 모양이 좌우가 같은 대칭이라면 입술을 수직 축 기준으로 절반으로 접은 영상도 원래의 입술 영상이 같은 주요한 정보를 대부분 포함할 수 있을 것이므로 정보의 왜곡도 거의 없을 것이다. 이와 같은 근거로 입술 영상을 반으로 접게 되면 원래 36x28 크기의 영상이 18x28 크기의 영상으로 줄어들게 된다. 이는 데이터 처리량이 절반으로 줄어들고 PCA를 통해 생성되는 특징 파라미터들의 수를 감소시킬 수 있다. 또한 접어진 18x28 크기의 픽셀 값들은 대칭되는 픽셀들의 평균값들이므로 영상 잡음 요소 및 좌우 측면 조명의 불균형에 대한 강인함을 갖게 된다. (그림 8)은 입술 폴딩을 적용한 영상이다.

발성자는 실제 환경에서 말을 할 때 고정된 자세로 말을 하지 않는다. 이는 발성시 머리 움직임이나 발음 습관 등에 따라 다양한 입술 움직임이 발생하고 이로 인해 인식을 저하의 원인이 될 수 있다. 또한 주위 환경 변화에 따른 조명의 변화도 인식을 저하의 원인이 된다. 조명은 실생활에서 보면 시간대에 따라 조사강도와 방향이 다르며, 그에 따른 영상의 왜곡 또한 다르다. 조명은 카메라로부터 입력되는 영상에서 실제 색 정보를 왜곡시키며, 영상 분석을 통해 파라미터를 추출하여 인식하는 방식을 사용하는 경우 인식을 성능에 큰 영향을 미친다. 따라서 본 논문에서는 이러한 성



(그림 7) 히스토그램 매칭 적용 전과 후의 입술영상 (위: 적용 전, 아래: 적용 후)  
(Fig. 7) The Lip Images before and after Adopting the Histogram Matching(upper: before, below: after)



(그림 8) 입술 폴딩을 적용한 영상  
(Fig. 8) Examples of folded lip image

능저하를 보상하기 위해서 명암이나 조도변화에 강한한 RASTA 필터를 사용하였다.

RASTA 필터는 원래 음성과 잡음이 서로 다른 점을 이용하여 인식 향상에 강인하게 작용하는 필터로 연구되었다. 본 논문에서는 영상을 이용한 립리딩에 RASTA 필터를 적용하였으며, 다음 식은 고역통과 필터와 저역통과 필터 식이다.

고역통과 필터식 :

$$Y_t[n,m] = 0.9859 \times (X_t[n,m] - X_{t-1}[n,m]) - 0.9716 \times Y_{t-1}[n,m]$$

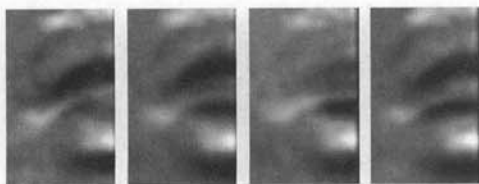
저역통과 필터식 :

$$Y_t[n,m] = 0.8638 \times (X_t[n,m] - X_{t-1}[n,m]) - 0.7257 \times Y_{t-1}[n,m]$$

여기서  $Y_t[n,m]$ 는 시간  $t$ 에서  $(n,m)$ 픽셀 좌표의 필터링된 이미지 출력값이다.  $X_t[n,m]$ 는 입력 이미지의 픽셀 값,  $X_{t-1}[n,m]$ 는 시간  $t$ 의 과거 값이 현재 입력에 영향을 주는 IIR필터이다. 저역 통과 필터 식은 대역 통과 필터링 수행을 위해 고역 통과 필터링의 출력 값을 입력으로 하여 실행되며, 실제 출력 값은 대역 통과 필터링을 수행한 결과 값과 동일하다.

하나의 영상 프레임은 주파수 영역으로 변환하면 변하지 않는 부분은 저주파 영역, 급변하는 부분은 고주파 영역으로 나타나게 된다. 이와 마찬가지로 입술이 변하는 동영상은 시간의 영역이 아닌 주파수 영역으로 살펴본다면, 변하지 않는 부분은 저주파 영역에 변화가 심한 부분은 고주파 영역에 나타날 것이다. 또한 시간의 흐름에 따라서 픽셀 값이 이전 픽셀 값과 차이가 나면 고주파 영역에 도시되고 그렇지 않으면 저주파 영역에 도시된다. 이를 이용해 적절한 필터를 사용하여 중요한 정보만을 추출하는 것이 필터링의 목적이다.

립리딩에서 발성자가 말을 할 때 시간의 흐름에 따라 입술 영상이 계속적으로 변화하는 부분과 변하지 않는 부분으로 나누어진다. 즉 단어를 발음하는 동안 입술 영역은 계속적으로 변하고 상대적으로 입술 주변 영역들은 변화가 적다. 이때 발음을 하면서 계속적으로 변하는 부분은 고주파 영역에서 나타나고 변화가 적은 부분은 저주파 영역에서 나타나게 된다. 이를 바탕으로 고역 통과 필터와 저역 통과 필터를 적용하면 보다 강인한 특징 파라미터들을 구할 수 있다. (그림 9)에는 RASTA 필터링을 적용한 입술 영상의 예가 나타나 있다.



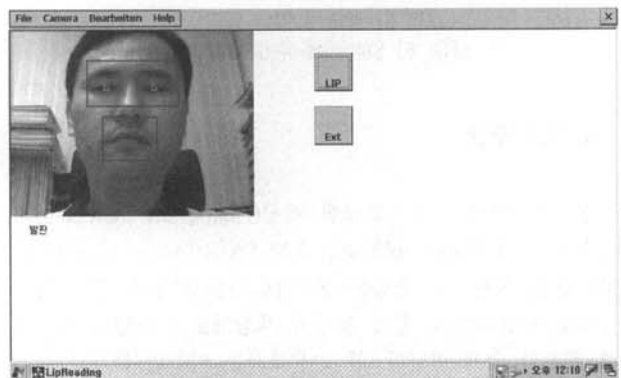
(그림 9) RASTA 필터링 적용한 입술 영상  
(Fig. 9) RASTA filtered lip images

마지막으로 RASTA 필터링을 적용한 입술 영상을 통계적 알고리즘인 주성분 분석(PCA)을 사용하여 중요한 몇 개의 특징 파라미터만을 추출해 처리함으로써 인식속도를 현저히 단축하고 인식률을 증가시킬 수 있도록 하였다.

### 5. 실험결과

구현 환경은 Microsoft Embedded Visual C++를 사용하였으며, HTK 3.0의 HMM 알고리즘을 사용해 학습 및 인식하였다. 실험에 사용된 마이크로비전 MV-320은 806Mhz CPU와 128MB Memory를 내장하고 있다. 본 논문에서는 전면에 부착 되어 있는 카메라를 사용해 실험하였으며, 카메라로부터 입력된 영상은 240x320 크기의 10 Frame/Sec 이내의 영상을 제공한다.

임베디드 환경에서는 실시간으로 입력되는 영상에서 특징 벡터를 추출하고 인식 하려면 시스템 자원의 부족으로 실시간 처리에 많은 문제점이 발생한다. 따라서 본 논문에서는 사용자가 발성을 시작하면 입술 영상을 발성이 모두 끝날 때 까지 36x28크기로 메모리에 저장한다. 메모리에 저장된 입술 영상을 본 논문에서 제안한 방법으로 특징 벡터를 추출하고 비터비 디코딩 알고리즘을 이용해 인식을 수행한다. 발성 단어의 인식 수행 속도는 발성 단어에 따라 차이가 있으나 3.62초~7.46초가 걸렸으며, 단어 발성 시간을 제외한 특징 추출과정에서 인식까지의 수행 속도는 1.15초~2.35초가 걸렸다. 본 실험의 수행 속도는 사용자가 발성을 시작할 때 버튼을 클릭하여 시작하고 발성을 다 마친 후 다시 버튼을 클릭하여 발성 단어를 인식 하였으므로 정확한 수행 시간과는 오차가 발생할 수 있다. 본 논문에서는 처리 속도 향상을 위해서 부동 소수점 연산을 고정 소수점연산으로 변환 하였으며, 빠른 입술 검출과 특징 추출 알고리즘을 사용하였다. 또한 묵음-단어-묵음의 3상태 HMM 모델과 비터비 디코딩 알고리즘을 임베디드 시스템에 탑재해 사용하였다. (그림 10)은 임베디드 시스템에서 실시간으로 립리딩을 수행해 인식 결과를 보여주고 있다.



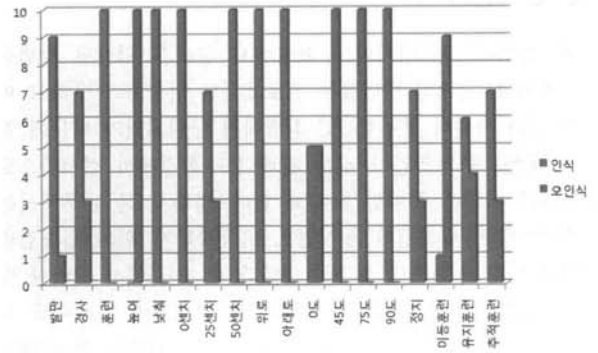
(그림 10) 임베디드 시스템에서 실시간 립리딩 인식 결과  
(Fig. 10) The Result of Real time Lip Reading Recognition by using a Embedded system

임베디드 시스템에서 립리딩 인식 테스트를 위해 18개 단어를 10번 반복해 총 180개의 단어를 학습 시켜 데이터베이스를 제작하였다. 18개의 단어는 헬스케어 기기인 경사침대를 제어하기 위한 명령어들로 이루어져 있다. <표 1>은 경사침대 조작에 필요한 18개의 발성단어 목록이다.

인식 테스트는 학습 환경과 비슷한 조명에서 18개 단어를 10번 반복해 총 180개의 단어를 사용해 인식 테스트를 수행하였으며, 테스트 결과는 180개의 단어 중 139개의 단어를

<표 1> 발성단어  
<Table 1> Spoken Words

순번	단어	순번	단어
1	발판	10	아래로
2	경사	11	0도
3	훈련	12	45도
4	높여	13	75도
5	낮춰	14	90도
6	0센치	15	정지
7	25센치	16	이동훈련
8	50센치	17	유지훈련
9	위로	18	추적훈련



(그림 11) 립리딩 인식 결과 그래프  
(Fig. 11) Results Graph of Lip Reading Recognition

인식해 77% 인식률을 얻었다. <표 2>는 180개의 단어를 인식 테스트하였을 때의 결과를 보여 주고 있으며, (그림 11)은 인식 결과를 그래프로 나타내고 있다. 실험 결과에서 보면 높여, 이동훈련, 0도와 같은 단어가 잘못 인식하는 비율이 높았으며, 이 중에서 높여 단어는 이동훈련과 50센치로 잘못 인식하는 결과를 보였다. 또한 이동훈련은 훈련으로 잘못 인식 하였으며, 0도는 위로와 0센치로 잘못 인식 하였다. 이와 같이 잘못 인식 하는 이유는 발음이 비슷하거나 입모양이 비슷한 경우에 인식률이 낮아짐을 알 수가 있다.

<표 2> 립리딩 인식 결과  
<Table 2> Results of Lip Reading Recognition

단어	발판	경사	훈련	높여	낮춰	0센치	25센치	50센치	위로
		아래로	0도	45도	75도	90도	정지	이동훈련	유지훈련
S01	발판	경사	훈련	이동훈련	낮춰	0센치	25센치	50센치	위로
	아래로	0도	45도	75도	90도	0센치	훈련	유지훈련	훈련
S02	발판	경사	훈련	50센치	낮춰	0센치	25센치	50센치	위로
	아래로	위로	45도	75도	90도	정지	훈련	추적훈련	추적훈련
S03	발판	0센치	훈련	50센치	낮춰	0센치	45도	50센치	위로
	아래로	0도	45도	75도	90도	정지	훈련	추적훈련	추적훈련
S04	발판	경사	훈련	50센치	낮춰	0센치	25센치	50센치	위로
	아래로	0도	45도	75도	90도	정지	훈련	유지훈련	추적훈련
S05	발판	발판	훈련	이동훈련	낮춰	0센치	25센치	50센치	위로
	아래로	위로	45도	75도	90도	정지	훈련	유지훈련	추적훈련
S06	발판	경사	훈련	50센치	낮춰	0센치	45도	50센치	위로
	아래로	0도	45도	75도	90도	0센치	훈련	추적훈련	추적훈련
S07	경사	경사	훈련	이동훈련	낮춰	0센치	25센치	50센치	위로
	아래로	0센치	45도	75도	90도	0센치	이동훈련	유지훈련	훈련
S08	발판	경사	훈련	50센치	낮춰	0센치	25센치	50센치	위로
	아래로	0센치	45도	75도	90도	정지	훈련	유지훈련	훈련
S09	발판	경사	훈련	50센치	낮춰	0센치	25센치	50센치	위로
	아래로	0도	45도	75도	90도	정지	훈련	추적훈련	추적훈련
S10	발판	0센치	훈련	이동훈련	낮춰	0센치	45도	50센치	위로
	아래로	위로	45도	75도	90도	정지	훈련	유지훈련	추적훈련



## 6. 결 론

본 논문에서는 임베디드 환경에서 실시간 립리딩 방법을 제안하였다. 기존의 방법들은 프로세서의 성능이 우수하고 메모리 등의 자원이 풍부한 PC 환경에서 구현되었기 때문에 자원에 대한 제약 조건이 많은 임베디드 환경에서 실시간으로 동작하기 어려운 문제를 가지고 있다. 따라서 본 논문에서는 임베디드에서 립리딩이 실시간으로 동작하기 위해 입술검출 방법과 특징 추출 방법을 제안하였다. 본 논문에서 제안한 방법은 계산량이 작고 자원을 적게 사용함으로써 임베디드 환경에서 실시간으로 동작할 수 있었다. 향후 과제는 현재 제안된 방식은 화자중속으로 제한된 사용자만 인식이 가능하나 향후 화자독립으로 립리딩 연구를 진행 하고자 한다.

## 참 고 문 헌

- [1] E. D. Petajan, "Automatic lipreading to enhance speech recognition," in Proc. of Global Telecommunications Conf., pp.265-272, Nov. 1984.
- [2] C. C. Chibelushi, F. Deravi, and J. S. D. Mason, "A review of speech based bimodal recognition," IEEE Trans. Multimedia, Vol.4, No.1, pp.23-37, Mar. 2002.
- [3] G. Potamianos, C. Neti, G. Gravier, A. Garg, A.W. Senior, "Recent advances in the automatic recognition of audiovisual speech," Proceedings of the IEEE, Vol.91, No.9, pp.1306-1326, Sept. 2003.
- [4] J. S. Lee and C. H. Park, "Robust Audio-Visual Speech Recognition Based on Late Integration," IEEE Trans. on Multimedia, Vol.10, No.5, pp.767-779, Aug. 2008.
- [5] T. Wark, S. Sridharan, V. Chandran, "An approach to statistical lip modelling for speaker identification via chromatic feature extraction," Proc. of Fourteenth International Conference on Pattern Recognition, pp.123-125, Aug. 1998.
- [6] P. Delmas, N. Eveno, M. Lievin, "Towards robust lip tracking," Proc. of 16th International Conference on Pattern Recognition, pp.528-531, Aug. 2002.
- [7] S. Dupont and J. Luetin, "Audio-visual speech modeling for continuous speech recognition," IEEE Trans. on Multimedia, Vol.2, No.3, pp.141-151, Sep. 2000.
- [8] C. Bregler and Y. Konig, "Eigenlips for robust speech recognition," in Proc. IEEE ICASSP, Vol.2, pp.669-672, 1994.
- [9] P. Duchnowski, U. Meier, and A. Waibel, "See me, hear me: Integrating automatic speech recognition and lip-reading," in Proc. ICSLP, Vol.2, pp.547-550, 1994.
- [10] G. Potamianos, A. Verma, C. Neti, G. Iyengar, and S. Basu, "A cascade image transform for speaker independent automatic speechreading," in Proc. 1st IEEE Int. Conf. Multimedia and Expo., Vol.2, pp.1097-1100, 2000.
- [11] S. Lucey, "An Evaluation of Visual Speech Features for the Tasks of Speech and Speaker Recognition," Proc. Int. Conf. Audio-Video-base Biometric Person Authentication, Guildford, UK, pp.260-267, 2003.
- [12] Jinyoung Kim, Joo-hun Lee, Katsuhiko Shirai, "An Efficient Lip-Reading Method Robust to Illumination Variations," IEICE TRANSACTIONS on Fundamentals of Electronics, Communications and Computer Sciences, Vol.E85-A, No.9, pp.2164-2168, September 2002.
- [13] Hynek Hermansky, Nelson Morgan, "RASTA processing of speech," IEEE Transactions on Speech and Audio Processing, Vol.2, No.4, October 1994.
- [14] 김영운, 강선경, 정성태, "모바일 장치에서의 립리딩을 위한 실시간 입술 영역 검출," 한국컴퓨터정보학회, 제14권, 제4호, 2009년 4월.
- [15] 한은정, 강병준, 박강령, "AAM 기반 얼굴 표정 인식을 위한 입술 특징점 검출 성능 향상 연구," 한국정보처리학회논문지B, 제16권, 제4호, pp.299-308, 2009년 8월.
- [16] 이종석, 박철훈, "컬러 입술영상과 주성분분석을 이용한 자동 독순," 한국정보처리학회논문지B, 제15권, 제3호, pp.229-236, 2008년 6월.
- [17] 소인미, 김명훈, 김영운, 이지근, 정성태, "조명 변화에 견고한 얼굴 특징 추출," 한국정보처리학회 학술대회논문집, pp.697-700, 2005년 11월.



### 김 영 운

e-mail : kyw1007@wku.ac.kr

2003년 2월 원광대학교 컴퓨터·정보통신 공학부(공학사)

2005년 8월 원광대학교 컴퓨터 공학과(공학석사)

2006년 3월~현 재 원광대학교 컴퓨터 공학과 박사과정

관심분야: 영상처리, 컴퓨터비전, 패턴인식, HCI, 컴퓨터그래픽스



### 강 선 경

e-mail : doctor10@wku.ac.kr

2000년 2월 원광대학교 전기·전자공학부

2004년 2월 원광대학교 정보·컴퓨터교육 학과 석사학위 취득

2010년 2월 원광대학교 컴퓨터 공학과 공학박사 취득

관심분야: 휴먼 컴퓨터 인터페이스, 영상처리, 영상인식, 임베디드시스템



### 정 성 태

e-mail : stjung@wku.ac.kr

1987년 2월 서울대학교 컴퓨터공학과(공학사)

1989년 2월 서울대학교 컴퓨터공학과(공학석사)

1994년 8월 서울대학교 컴퓨터공학과(공학박사)

1995년 3월~현 재 원광대학교 컴퓨터공학과 교수

관심분야: 영상인식, 영상처리, 컴퓨터 그래픽스