

# 입술 움직임 변화량을 이용한 실시간 화자의 클로즈업 및 트래킹 시스템 구현

양운모, 장언동, 윤태승, 곽내정, 안재형  
충북대학교 정보통신공학과

## Real Time Speaker Close-Up and Tracking System Using the Lip Varying Informations

Wun-Mo Yang, Un-Dong Chang, Tae-Sung Yun, Nae-Jung Kwak, Jae-Hyung Ahn  
Department of Computer and Communication Engineering  
Graduate school of Industry, Chungbuk National University

E-mail : ebaot@hanmai.net, udchang@hanbox.com, knj0125@hanmail.net, yuta@naver.com,  
jhahn@cbucc.chungbuk.ac.kr

### 요약

본 논문에서는 다수의 사람이 존재하는 입력영상에서 입술 움직임 정보를 이용한 실시간 화자의 클로즈업(close-up) 시스템을 구현한다. 칼라 CCD 카메라를 통해 입력되는 동영상에서 화자를 검출한 후 입술 움직임 정보를 이용하여 다른 한 대의 카메라로 화자를 클로즈업한다. 구현된 시스템은 얼굴색 정보와 형태 정보를 이용하여 각 사람의 얼굴 및 입술 영역을 검출한 후, 입술 영역 변화량을 이용하여 화자를 검출한다. 검출된 화자를 클로즈업하기 위하여 PTZ(Pan/Tilt/Zoom) 카메라를 사용하였으며, RS-232C 시리얼 포트를 이용하여 카메라를 제어한다. 실험결과 3인 이상의 입력 동영상에서 정확하게 화자를 검출할 수 있으며, 움직이는 화자의 얼굴 트래킹이 가능하다.

### 1. 서론

사람에 얼굴은 다른 사람들과 구분될 수 있는 다양한 특징을 가지며, 얼굴에 나타나는 표정으로 언어나 문자만으로 상세하게 전달 할 수 없는 여러 의미의 정보를 제공할 수 있다. 이러한 이유로 컴퓨터 비전 분야에서는 얼굴에 인식과 얼굴에 각 부분 검출을 매우 관심 있게 다루고 있으며, 최근 패턴 인식, 컴퓨터 비전, 신경망과 같은 다양한 분야에서 얼굴의 인식 분야가 활발히 연구되고 있다. 상업적, 법적으로 수많은 응용 분야를 가지고 있으므로, 얼굴 영상 처리 기술과 관련된 공학적인 측면의 연구가 선행되어야 한다[1].

인식과 관련된 연구들은 얼굴 영역 추출이 선행된 상태에서 수행되는 경우가 많았다. 많은 연구들이 단순한 배경이나 영상내의 얼굴의 크기를 머리에서 어

깨 사이의 크기로 고정시키는 방법을 쓰고 있다[2-4]. Delmas와 Coulon 등은 입술 모양 변화를 검출하여 립-리딩(lip-reading)에 적용하기 위한 적당한 입술 영상의 획득 문제를 해결하기 위해 각 사람의 머리에 카메라가 달린 헬멧을 써서 입술의 위치와 크기를 일정하게 유지하도록 조절하였다[5]. 그러나 실용화를 목표로 하는 얼굴 인식 시스템을 개발할 때, 영상 내의 얼굴 영역의 추출은 그리 간단한 문제가 아니다. 배경, 얼굴과 카메라의 거리, 카메라 시야 내에서의 얼굴의 위치 등을 시스템의 가정에 부합되도록 조정한다는 것은 매우 어렵기 때문이다[6]. 특히 화면 내에 여러 사람이 있을 경우 특정인에 대해 처리하는 것은 더욱 어려운 일이다.

본 논문에서는 두 대의 CCD 칼라 카메라를 이용 기준 카메라를 두어 전체적인 배경화면을 입력받고, PTZ카메라를 줌 카메라로 이용하여, 배경 화면에 나타나는 여러 사람들 중에서 현재 말하고 있는 사람의

얼굴을 클로즈업하는데 사용하는 방식으로, 각 사람의 얼굴 영역의 크기와 위치가 항상 적절한 크기로 유지되도록 하여, 얼굴 추출 및 인식 성능에 향상을 도모하였다.

본 논문에서 제안한 화자 검출 과정을 살펴보면, 우선 기준 카메라에 입력되는 화면 내에서 색상 정보를 이용하여 얼굴 색을 분류하고, 모양 정보를 이용하여 각 사람에 얼굴 영역을 검출[7]하였다. 검출된 얼굴 영역에서 입술 영역을 분리하여 입술 움직임이 가장 큰 사람을 화자로 판정한 후 좀 카메라를 화자로 이동시켜 확대된 얼굴 영역을 출력시키도록 하여, 확대된 얼굴 영역을 통해 얼굴 인식 및 립-리딩 시스템을 구현하는 것이 가능하도록 하였다.

## 2. 얼굴 영역 검출

본 논문에서는 특징점 추출이나 통계학적인 복잡한 알고리즘보다는 색상 정보를 이용한 신속한 알고리즘을 적용하여 얼굴 영역 검출을 하였다[7]. CCD 카메라의 조명에 민감한 특성을 최소화하기 위하여 영상의 RGB입력으로부터 YCbCr에 모델로 변환 후에 휘도 성분(Y)을 제외한 Cb와 Cr성분만을 이용 피부색 영역을 분리하였다. 전처리 과정으로서 잡음 제거를 위해 형태학적 필터링을 하였고, 수평투영을 하여 얼굴 영역에 함께 나타나는 목 부분을 줄였다. 그 후 레이블링을 통해 피부색 영역을 분리한 후 모양 제한을 하여 각 사람의 얼굴 영역만을 검출하였다. 그림 1은 얼굴 영역 검출 알고리즘의 과정이다. 변환되어진 Cb와 Cr성분을 가지고 피부색 분할을 수행한다. 식 (1)에 나타난 Cb와 Cr의 임계값은 Chai등이 제안한 모델이며, 임의의 표본 영상들로부터 얼굴 영역의 피부색 화소들만을 취하여 얼굴색 칼라 히스토그램을 이용하여 계산되었다[8].

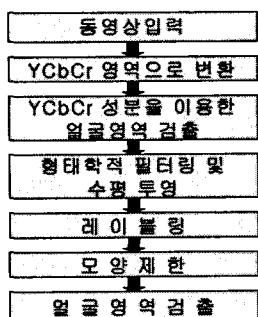


그림 1. 얼굴 영역 검출 알고리즘

$$B(x, y) = \begin{cases} 1 & \text{if } (77 \leq C_b \leq 132) \cap (133 \leq C_r \leq 171) \\ 0 & \text{Otherwise} \end{cases} \quad (1)$$

$B(x, y)$ 는 피부색으로 분할된 이진 영상이다.  $C_b$ 와  $C_r$ 성분이 임계값 내에 있으면 피부색으로 간주하여 1로 설정하고 다른 부분은 배경영상으로 간주하여 0으로 이진화 한다. 이렇게 이진화 되어진 영상에는 잡음 요소들이 포함되어 있는데, 이런 잡음들은 다음 과정인 레이블링의 연산량에 큰 영향을 준다. 그러므로 형태학적 필터링을 통하여 작은 잡음 요소를 제거하고 돌출 부분을 제거함으로써 영상을 단순화시킬 필요가 있다. 이러한 방법으로 확장 연산은 객체를 확장시키며 작은 구멍들을 채우고, 객체의 이음매 부분들을 연결시켜주는 연산을 한다. 수축 연산은 경계선을 수축 시킴으로써 아주 작은 물체를 제거하거나 감소시키는 연산을 한다. 본 논문에서는 제거 연산 후 채움 연산을 수행하였다. 이 후 수평 방향 투영을 수행함으로 임계치 보다 작은 값들을 제거함으로 얼굴 영역을 좀 더 정확하게 분리해 낼 수 있다. 그 후 연결되어지는 부분에 같은 값을 두는 레이블링을 통해 영상 내 다수의 얼굴 영역을 서로 분리하여 식별해낸다. 레이블링 과정을 통해 분리된 영역들은 얼굴뿐만 아니라 노출된 몸의 다른 부분이나 비슷한 색깔의 사물들을 포함하므로 형태학적인 구조 정보를 이용, 영역의 면적과 가로와 세로의 비를 제한하여 얼굴 영역을 검출하였다. 그림 2는 검출된 얼굴 영역을 나타낸다.



그림 2. 얼굴 영역 검출

## 3. 실시간 입술 움직임 정보 검출

일반적으로 실시간 영상 처리에 관련된 알고리즘들은 실시간으로 입력되는 동영상의 각 프레임에서 특정 값들을 추출하여 연속된 프레임 사이의 특징 값 변화 패턴을 분석하여 영상 신호를 처리하게 된다. 이런 기법으로 화소 값에 기반한 방법과 히스토그램에 기

반한 방법 등이 주로 사용되고 있다.

본 논문에서는 히스토그램 방법에 기반 하면서도 위치정보를 잊지 않는 방법을 제안한다. 검출된 얼굴 영역에서 다시 적절한 입술 영역을 설정한 후 얼굴 색으로 분리된 이진 영상에서 그 영역을 추적하면서 움직임에 따른 입술 영역의 화소면적 변화를 측정하여 움직임 정보를 추출하였다.

### 3.1 입술 영역 검출

대부분에 얼굴의 가로 세로 비가 약 1:1.48 정도의 비율을 가지는 타원형이며, 입술 영역은 전체 얼굴 면적의 중심점 아래 부분에 위치한다는 형태학적 구조 특징을 이용하여 얼굴 면적의 중심점 아래 부분을 입술 영역으로 추정한다. 각 사람의 얼굴 영역의 중심점을 익 (2)을 이용하여 구한다.  $C(x_c, y_c)$ 는 얼굴 영역의 중심 좌표이며  $x$ 와  $y$ 는 각각 X, Y의 좌표의 위치이며,  $A$ 는 얼굴 영역의 면적이다.

$$C(x_c, y_c) = \frac{1}{A} \left( \sum_{(x, y) \in R} x, \sum_{(x, y) \in R} y \right) \quad (2)$$

구해진  $C(x_c, y_c)$ , 중심점을 기준으로 아래 부분에 마스크를 설정하고, 마스크의 안쪽을 입술 영역으로 결정한다. 본 논문에서 설정한 마스크의 크기는 사람 얼굴의 형태학적인 분석 결과와 실험을 통하여 얼굴 면적의 크기를 700~1500화소로 정했을 때  $19 \times 19$ 가 가장 적합하였다. 이런 각 사람의 입술 영역 마스크에서 입술 안쪽 영역은 얼굴색 영역과 비교하여  $Cb$ 와  $Cr$  성분이 다르게 나타나므로 얼굴 색으로 인식되지 않는 부분이 발생한다. 그리고 화자가 말을 하게 되면, 설정된 마스크 내 영역에서 입술 안쪽의 음영과 치아에 의해 얼굴색이 아닌 화소의 수가 변화하게 된다. 즉, 피부색으로 분할된 이진 영상의 입술 영역 마스크 안에 0의 면적이 변화하게 된다.

본 논문에서는 입술 영역 마스크 내의 0의 면적 변화 특성을 이용하여 입술 움직임 정보를 획득하였다. 그림 3은 제안된 알고리즘을 적용한 결과를 나타낸다. 그림 3의 (a)와 (b)의 입술 영역 마스크 내에서 변화가 발생하면 그림 3의 (c)와 (d)의 이진 영상 내에서 입술 면적 변화가 감지된다. 그림 3의 (e)와 (f)는 두 이진영상의 입술 영역 마스크 내에서 수평 투영을 했을 때 얼굴색의 변화량을 나타낸 것이다. 각 프레임에서 얼굴색이 아닌 화소의 면적 즉 0의 면적을 계산한 후 두 프레임 간의 입술 영역 면적 차를 이용해 화자를 검출할 수가 있다.

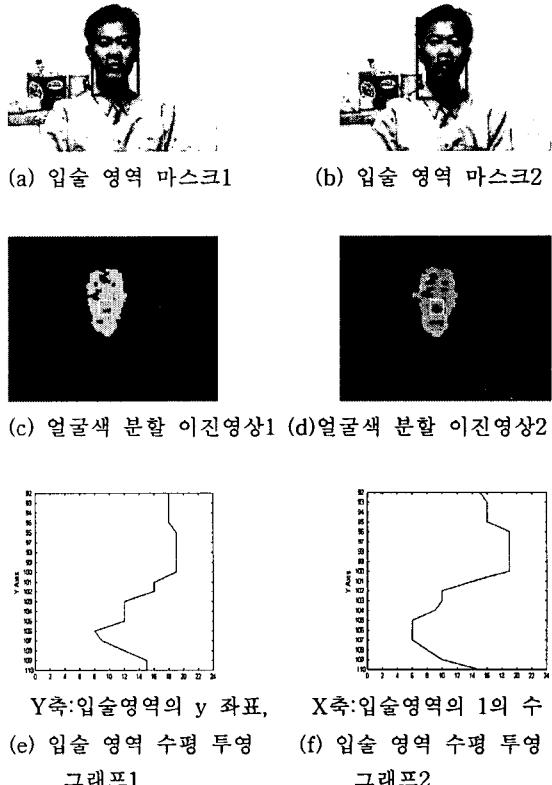


그림 3. 입술 영역 검출

### 3.2 입술 움직임 정보 검출

실시간 영상 분석 및 처리를 위해 추출된 마스크 내의 입술 움직임 정보 비교는 매 15 프레임 간격으로 수행하였다. 각 프레임마다 얼굴 영역의 좌표 및 크기가 변화할 수 있으므로 비교 프레임마다 얼굴 영역을 다시 검출하여 중심 좌표를 계산하고 입술 영역 마스크를 재 설정하였다. 그 후 각 사람별로 두 비교 프레임간의 입술 움직임 정보의 변화량을 검출하기 위해 식 (3)을 이용하였다.

$$\begin{aligned} \delta_n &= |S_n - S_{n-1}| \\ &= \left| \sum_{i=1}^{19} \sum_{j=1}^{19} f_n(x_i, y_j) \right. \\ &\quad \left. - \sum_{i=1}^{19} \sum_{j=1}^{19} f_{n-1}(x_i, y_j) \right| \end{aligned} \quad (3)$$

$S_n$ 과  $S_{n-1}$ 은 각 프레임에서 입술 영역 마스크의 0의 면적이며,  $\delta_n$ 은 입술 움직임 정보를 나타내는 두 프레임간의 마스크 내의 면적 변화량이다. 이때, 입술 움직임 정보  $\delta_n$ 의 값은 화자 후보의 입술이 움직이지

않더라도 조명 변화 등에 의한 화소 변화로 인하여 미세한 변화를 계속 일으켜 화자 결정에 오류를 일으킬 수 있다. 실제로 실험해 본 결과 영상 내의 사람이 말을 하지 않았는데도 2~4화소 정도의 변화량이 검출되었다. 이러한 오류를 방지하기 위하여 임계값  $T_{th}$  을 5로 정하여  $\delta_n$ 의 값이 임계값 미만의 변화를 일으켰을 때는 변화가 없는 것으로 처리하였다.

#### 4. 실시간 화자 클로즈업 시스템

##### 4.1 실시간 화자 클로즈업 시스템

본 논문에서 구현한 화자 클로즈업 시스템은 그림 4와 같다. 그림 4에서 결정된 화자의 얼굴 영역 중심점으로 줌 카메라의 초점을 변경하여 화자를 클로즈업하였으며, 이때 기준 카메라는 전체의 영상을 그대로 보존하여 계속적으로 입술 움직임 정보량의 변화를 감시한다. 만약 현재 클로즈업 된 화자의 입술 움직임 정보량보다 더 큰 정보량을 갖는 얼굴 영역이 검출된다면 그 사람으로 줌 카메라의 초점을 이동시켜 화면을 전환한다. 따라서 제안된 화자 클로즈업 시스템은 전체적인 영상과 클로즈업 된 두 동영상을 제공하게 한다.

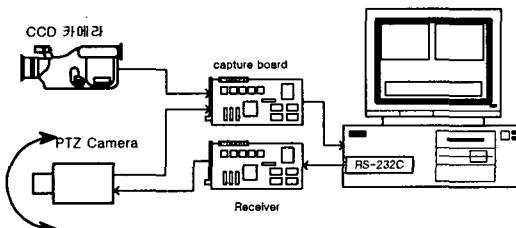


그림 4. 화자 클로즈업 시스템

##### 4.2 화자 결정

입술 움직임 정보를 이용하여 화자를 결정하는 순서는 그림 5와 같다. 화자 선정을 위하여 추출된 각 화자 후보들의 입술 움직임 변화량들을 연속적으로 비교하였으며, 실험을 통하여 임계치  $T$ 을 5로 설정하여 최대 변화량 횟수가 5가 되는 후보를 화자라고 결정하였다. 이렇게 하므로 순간적으로 변화량이 커지는 사람을 화자로 인식하는 것을 방지하였다. 일단 화자가 선정되면 선정된 사람을 제외한 모든 화자 후보들의 최대 변화량 발생 횟수를 0으로 리셋 시키고 선정된 화자에게는 1을 부여하여 화자 후보들의

일시적인 최대 입술 움직임 정보의 변화에도 지속적으로 현재의 화자가 선정되도록 하였다. 또한 화자를 교체하기 위해서는 입술 움직임 변화량을 지속적으로 관찰하면서 5회 이상 최대 변화량을 나타내는 화자에게로 줌 카메라가 이동하게 하였다.

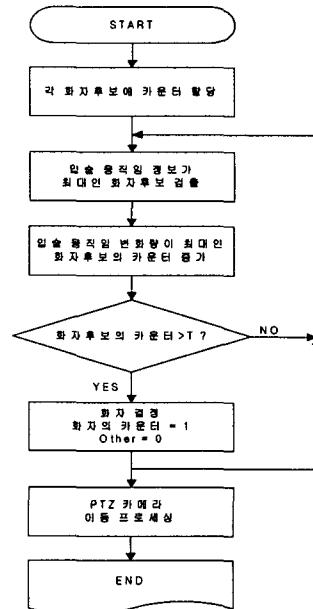


그림 5. 화자 검출 순서도

##### 4.3 PTZ 카메라 제어

본 논문에서 구현한 화자 클로즈업 시스템은 줌 카메라를 이동시키기 위하여  $320 \times 240$  크기의 기준 영상의 중심점을 0으로 하는 2차원 좌표계를 기준으로  $A(X_a, Y_a)$  지점에 위치하는 화자는  $X$  축으로  $X_a$ ,  $Y$  축으로  $Y_a$  화소 떨어진 지점에 위치한다고 가정한다. 따라서  $A$  지점으로 줌 카메라의 중심을 이동시키기 위하여  $X$ ,  $Y$  좌표의 화소 거리를 카메라 구동 시간 단위로 변환하여야 한다. 실험 결과 카메라의 이동 속도는 15ms/pixel로 계산되었다. 줌 카메라가 화자  $A$ 로 이동하였다면, 다음 이동의 기준점은 좌표계의 중심점  $O(0, 0)$ 에서 화자  $A$ 의 중심점  $A(X_a, Y_a)$ 로 변경된다. 따라서, 화자가  $B(X_b, Y_b)$  위치의 사람으로 바뀔 경우 카메라의 이동 거리는 식 (4)과 같이 계산된다.

$$X_M = X_a - X_b$$

$$Y_M = Y_a - Y_b \quad (4)$$

식 (4)에서 좌우축의 이동 방향은  $X_M$ 의 부호에 의해 결정된다.  $X_M$ 의 값이 양수이면 기준점의 좌측으로, 음수이면 우측으로 이동한다. 상하축의 이동 방향은  $Y_M$ 의 부호에 의해 결정된다.  $Y_M$ 의 값이 양수이면 상향으로, 음수이면 하향으로 이동한다. 이때, 고려해야 할 점은 현재 화자의 움직임이다. 화자는 말하면서 머리를 움직일 수도 있고 또 위치를 변경해 가면서 말을 할 수도 있을 것이다. 그러므로 현재 화자의 중심점 위치가 임계 화소값 이하만큼 움직인다면 위치 변화가 없는 것으로 간주하여 줌 카메라가 머리의 움직임에 민감하게 반응하여 움직이는 것을 방지했으며, 중심점 위치가 임계 화소값 이상 움직일 경우 줌 카메라는 화자를 추적하도록 설계하였다. 또 줌 카메라의 이동 중에 화자가 바뀌면 다 이동하지 못하고 다른 화자로 이동해 버리기 때문에 이동 거리 오차가 발생한다. 이를 방지하기 위해 플래그를 두어 이동 중에는 다른 이동 명령을 내리지 못하게 하였다.

## 5. 실험 결과 및 검토

연구를 위해 사용한 computer는 cpu가 Athlon(1 GHz)이며 Windows 98 환경 하에서 Visual C++ 6.0으로 프로그래밍 하였다. 사용된 두 대의 카메라는 SONY TRV900을 기준 카메라로 사용하였으며, 줌 카메라는 PTZ 기능이 있는 카메라를 사용하였다. 줌 카메라와 PC와의 통신은 RS-232C 포트를 이용하였으며 카메라 제어는 PTZ 카메라 리시버를 통하여 이루어진다. 또한 두 대의 카메라로부터 영상 신호를 입력받기 위하여 4채널 DVR 카드 Eyean-1000을 사용하여 초당 30 프레임의 동영상을  $320 \times 240$  크기로 획득하였다.

얼굴 영역의 검출 시간은 배경화면 내에 세 사람이 존재할 때 초당 6.25 프레임의 처리 성능을 나타냈다. 그러므로 화자를 검출해 낸 후 다시 실시간으로 화자를 인식하는 처리를 한다거나 립-리딩 처리를 할 여유 시간을 획득할 수 있었다.

두 입력 영상의 변화량을 검출하는 가장 간단한 방법 중의 하나는 두 영상의 화소 값과 화소 값을 비교하는 방법이다[10]. 두 영상의 화소값 차를 이용하여 만들어진 차영상을 통하여 물체의 움직임을 검출해 볼 수 있다. 그러나 이러한 방법은 입술의 움직임을 검출해 내는 데 단점을 보이고 있다. 본 논문에서 제안된 입술 움직임 검출 알고리즘의 특성 비교를 위하여 화소값 차 방법과 입술 움직임 결출에 대한 비교

실험을 하였다. 그림 6 (a)와 (b)의 두 영상에서 화자의 입술 영역의 정보 변화량 비교가 표 1에 나타나 있다. 편의상 화자 후보는 좌측에서부터 화자(A), 화자(B), 화자(C)로 간주한다.

화자(A)와 화자(B)는 입술 움직임이 있지만 화자(C)는 단지 고개만 돌리는 상태이다. 표 1에서 볼 수 있는 것과 같이 화자(A)와 화자(B)처럼 화자가 정면을 바라보면서 말을 할 때에는 화소값 차를 이용한 방법과 제안된 알고리즘 모두 입술 움직임을 검출할 수 있었다. 그러나 화자(C)처럼 말하지 않으면서 단지 얼굴을 좌우로 돌리거나 숙이는 경우와 같은 움직임이 발생할 때 화소값 차를 이용한 방법은 입술 움직임이 있는 것으로 판정하지만, 제안된 방법은 입술 움직임이 없는 것으로 판정하였다. 즉 제안된 알고리즘은 얼굴이 움직여도 입술 움직임 정보를 정확하게 검출할 수 있었다.

그림 7은 화자 결정 과정의 한 예이다. 동영상에서 15 프레임 간격으로 최대 입술 움직임 정보를 계산하여 화자가 결정되는 과정과 화자 교체 과정을 그래프로 나타냈다. 1201 프레임 동안의 최대 입술 움직임 정보를 계산하여 그라프로 나타냈다. 화자로 결정된

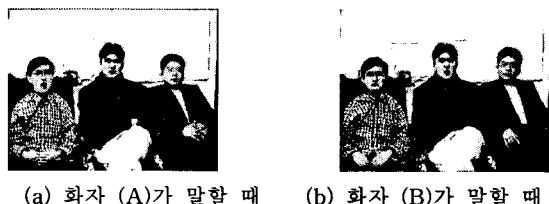


그림 6. 세 명의 화자 후보 영상

표 1. 입술 움직임 정보 변화량 비교 (단위:화소수)

화자 작용방법	화자 (A)	화자 (B)	화자 (C)
화소값 차	23	38	27
제안된 방법	9	6	1

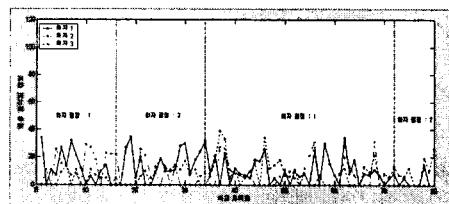


그림 7 화자 결정 과정

사람의 최대 입술 움직임 정보가 크게 변화하고 있음을 알 수 있다. 화자 교체의 경우 최대 입술 움직임 정보가 5회 이상 축적되는 화자 후보가 다음 화자로 결정됨을 알 수 있다. 첫 번째 화자가 결정된 프레임은 프로그램 실행 후 136번째 프레임부터이며, 약 4.5초가 소요되었다. 또한, 화자가 교체되어 클로즈업되는 시간은 4.5초 이내로 이루어졌으며, 이와 같이 화자 검출 및 카메라 위치 이동의 오류를 최소화하기 위한 시간을 실험을 통하여 확인하였다.

그림 8은 구현된 시스템의 시뮬레이터 화면이다. 왼쪽에 전체 화면이 있고 오른쪽에 화자가 검출되어 줌 카메라로 클로즈업된 상태이다. 아래 부분은 입술 움직임 정보 그리고 얼굴 영역과 입술 영역이 검출된 영상이다.

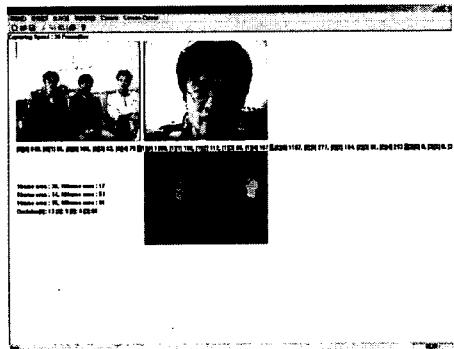


그림 8 화자 클로즈업 시뮬레이터 화면

## 6. 결 론

본 논문에서는 다수의 사람이 존재하는 입력영상에서 얼굴영역을 먼저 검출한 후 입술 움직임 정보를 이용한 실시간 화자 클로즈업 시스템을 구현하였다. 실시간 처리를 위하여 얼굴 영역 검출 기법으로 YCbCr 색상정보와 형태 정보를 이용하였고, 사람의 움직임으로 인한 화자 인식 오류를 방지하기 위하여 단순히 두 영상의 화소값 차를 이용하지 않고 히스토그램 기반 기법을 보완하여 입술의 움직임 정보를 검출하였다. 실험 결과 검출된 입술 움직임 정보를 이용하여 다수의 사람이 존재하는 입력영상에서 화자를 정확히 검출할 수 있었다. 또한 PTZ 카메라를 제어하여 영상처리에 적절한 이미지를 능동적으로 획득할 수 있었다.

본 논문에서 제안한 시스템은 영상회의, 얼굴인식,

립-리딩, 무인감시 시스템에 적용할 수 있다.

## 참 고 문 현

- [1] M. Kaneko and O. Hasegawa, "Processing of Face Images and Its Applications", IEICE Trans. on Information and System, Vol.E82-D, No.3, March 1999
- [2] Richard Harvey, Lain Matthews, J. Andrew Bangham and Stephen Cox, "Lip Reading from Scale-Space Measurements", IEEE Proc. Computer Vision and Pattern Recognition, pp.582-587, 1997
- [3] Juergen Luettin, Neil A. Thacker and Steve W. Beet, "Speaker Identification by Lipreading", IEEE Proc. Spoken Language Processing, ICSLP'96, pp.62-65, 1996
- [4] P.M. Antoszczyzyn, J.M. Hannah and P.M. Grant, "Local Motion Tracking in Semantic-Based Coding of Videophone Sequences", IEEE Proc. Image Processing and its application, pp.46-50, 1997
- [5] P. Delmas, P. Y Coulon and V. Fristot, "Automatic Snakes for Robust Lip Boundaries Extraction", 1999 IEEE International Conf, Vol.6, Acoustics, Speech and Signal Processing pp.3069-3072, 1999
- [6] 윤호섭, 왕민, 민병우, "눈 영역 추출에 의한 얼굴 기울기 교정", 전자공학회 논문지 제33권, 제12호, pp.1886-1898, 1996
- [7] 김영길, 한재혁, 안재형, "컬러 정지 영상에서 색상과 모양 정보를 이용한 얼굴 영역 검출", 한국멀티미디어 학회 논문지 제4권, 제1호, pp.67- 74, 2001
- [8] D. Chai and K. N. Ngan, "Locating Facial Region of a Head-and-shoulders Color Image", IEEE Proc. Automatic Face and Gesture Recognition, pp.124-129, 1998
- [9] 유태웅, 오일석, "색채 분포 정보에 기반한 얼굴 영역 검출", 한국정보과학회논문지, 제24권, 제2호, pp.180-192, 1997
- [10] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*, Addison-Wesley, pp.465-467, 1992