

# 실시간 영상으로부터 입술 검출에 관한 연구

\*김종수    \*\*한상일    \*\*\*서보국    \*\*\*\*차형태  
승실대학교, (주)씨앤비텍  
\*homat@mms.ssu.ac.kr

## Lip Detection from Real-time Image

\*Kim, Jong-Su    \*\*Hahn, Sang-Il    \*\*\*Seo, Bo Kug    \*\*\*\*Cha, Hyung-Tai  
Soongsil University, CNB Technology Inc.

### 요약

본 논문에서는 실시간 영상으로부터 입술 영역 검출 방법을 제안한다. 제안하는 방법은 영상으로부터 피부색 범위의 검출을 통하여 불필요한 잡음을 제거한 후 Harr-like 특징을 이용하여 얼굴을 검출한다. 다음 검출된 얼굴 영역으로부터 얼굴의 기하학적 정보를 이용하여 입술 후보 영역을 분리한 후 제안하는 Cb, Cr를 가지고 입술색 범위 검출해 낸다. 최종적으로 검출된 입술색 범위 영역에 Haar-like 특징을 다시 한번 적용하므로써 보다 정확한 입술 영역을 검출해낸다. 본 논문에서 제안한 알고리즘을 실험한 결과 기존의 알고리즘보다 검출률이 높았으며, 적용범위가 더 넓음을 실험을 통해 확인할 수 있었다.

### 1. 서론

유비쿼터스 시대에서 멀티미디어의 기술은 날이 갈수록 발전해 가고 있다. 또한 기술이 발전이 거듭됨에 따라 인간의 존엄성도 상당히 중요해졌고, 인간의 안전과 편의를 위해 로봇 기술은 엄청난 속도로 진화하고 있다.

과거의 로봇 기술이 일상생활에 밀접하게 연관이 있기 보다는 단지 산업용으로만 많이 사용되어지고 응용되었다면, 최근의 로봇 기술은 군사용, 산업용, 가정용 등으로 까지 발전 및 상용화가 된 상태이다.

가정용 지능형 로봇의 대표적인 예로서 1999년 일본 소니사에서 발표한 세계 최초의 감성 지능형 완구 로봇 애완견 "AIBO"를 들 수 있다.

AIBO는 애완동물을 대체하는 개념으로 시작되었으며, 현재 다양한 옵션과 지속적인 제품 개발에 의해 인간과 친구가 될 수 있는 하나의 새로운 개체로서 인정받고 있다. 기쁨, 슬픔, 성냄, 놀람, 공포, 혐오 등의 6가지 감성과 성애욕, 탐색욕, 운동욕, 충전욕 등의 4가지 본능이 구현되어 외부의 자극과 자신의 행동으로 인하여 감성과 본능 수치가 항상 변화한다.

지능형 로봇이란 이렇듯 인간과 비슷한 기본적인 욕구나 시각, 청각 등의 감각기관을 자율적으로 판단 및 행동하는 시스템을 말한다. 이러한 요구에 의해 얼굴 인식 및 제스처 인식을 비롯한 인간과 컴퓨터간의 상호작용에 관한 것으로서 HCI(Human-Computer Interacion) 기술들이 활발하게 연구 되고 있다[1].

HCI(Human-Computer Interacion) 기술 중에서 그 중 시각 정보는 인간이 정보를 인식하는데 있어 상당히 높은 비중을 차지하고 있으며, 이러한 인간의 시각을 대신하는 것은 컴퓨터의 카메라이다.

본 논문에서는 지능형 로봇에 사용되는 HCI(Human-Computer Interacion) 기술 중 하나로써, 캠을 가지고 인간의 시각정보를 이용한 얼굴 인식 기술을 사용하였으며, 인식된 얼굴 영역의 그 안의 요소로써 제안하는 알고리즘을 가지고 최종적으로 입술 영역 검출을 위한 전처리 작업을 제안하였다.

얼굴과 입술 검출을 하는데 있어서 여러 가지 검출 방법이 존재한다. Zang은 상, 하 입술에서의 밝기 성분 차이 및 분산 그리고 입 부분의 윤곽선을 이용하여 입술을 인식하였다[2]. Wark는 위, 아래 입술의 경계선을 추출하는데 있어서 입술의 윤곽선과 컬러 정보를 이용 하였다 [3]. Oliver는 2-D blob features와 HMM(Hidden Markov Model)를 이용하여 얼굴과 입술을 검출하였다[4]. 또한 Haar-like feature와 Adaboost 알고리즘을 이용하여 얼굴영역을 검출 하였다[5-7].

얼굴 검출은 캠으로 입력받은 영상 이미지를 가지고 조명에 강인한 YCbCr 컬러모델을 이용하여 피부색 범위를 검출하고, 검출한 피부색 영역에 Haar-like feature를 가지고 Adaboost 학습 알고리즘을 이용하여 최종적인 얼굴 영역을 검출 한다[9]. 그리고 검출된 얼굴 영역으로부터 얼굴의 기하학적 정보를 이용하여 입술 후보 영역을 분리한 후 제안하는 Cb, Cr를 가지고 입술색 범위 검출해 낸다. 마지막으로 얼굴 영역의 절반에 해당되는 입술 후보 영역에 한번 더 Haar-like feature와 Adaboost 학습 알고리즘을 시행함으로써 최종적으로 입술 영역을 검출하여 영상을 출력한다.

본론에서는 얼굴영역 검출과 YCbCr, Haar-like feature 그리고 Adaboost 학습 알고리즘에 대하여 전반적인 이론적인 설명과 제안한 알고리즘에 대해서 설명 하였으며, 본 논문에서 제안한 결과를 실험 결과 및 결론분석과 결론 부분을 최종 기술하였다.

"본 연구는 지식경제부 및 정보통신연구진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음" (ITA-2009-(C1090-0902-0007))

## 2. 본론

본 논문의 흐름은 얼굴 검출과 입술 검출로서 크게 2개의 부분으로 구성되어 있다. 그림 1은 본 논문의 제안하는 전체 시스템 흐름도를 보여준다. 배경과 얼굴 이미지가 캠으로부터 입력을 받고, 얼굴 검출을 하기 위해 RGB 컬러모델을 YCbCr로 변경해주어 피부색 범위를 검출한다. 검출된 피부색 영역에 Haar-like feature와 Adaboost 학습 알고리즘을 이용하여 최종 얼굴 영역을 검출 한다. 입술 후보 영역을 알기 위하여 검출된 얼굴 영역을 가지고 인간의 기하학정보를 이용하여 입술이 포함된 부분을 제외하고 나머지를 영상의 절반을 제거한다. 제거된 입술 후보 영역에 대해서 Cb, Cr값의 변경을 이용하여 입술색 범위를 검출하고 다시 한번 더 Haar-like feature와 Adaboost 학습 알고리즘을 시행함으로써 보다 정확한 입술 영역을 검출 한다.

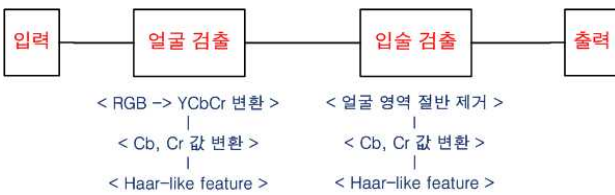


그림 1. 전체 시스템 흐름도

### 2.1 얼굴 영역 검출 방법

얼굴 영역 검출 방법에는 크게 4가지 분류 방식이 있다.[8] 다음의 표 1.는 얼굴 영역 검출 방법의 4가지 분류에 보여주고 있다.

표 1. 얼굴 영역 검출 방법

특징 기반 (Feature-Based) 방법	얼굴 요소, 색깔, 모양, 크기와 같은 얼굴 고유의 특징을 이용해서 얼굴 크기 및 위치를 추론하여 얼굴 영역을 검출하고, 얼굴 요소의 거리나 위치 등을 통해 얼굴인지 아닌지 판별하는 방법
지식 기반 (Knowledge-base) 방법	사람의 얼굴은 두 개의 눈, 한 개의 코와 입으로 구성되어 있고, 각 요소들은 기하학적 위치 관계로 구성됨을 전제로 하여 얼굴을 검출하는 방법
템플릿 매칭 (Template-matching) 방법	얼굴에 대한 표준 형판을 생성 한 다음, 입력 영상과의 차이를 이용하여 얼굴 영역을 검출하는 방법
외형 기반 (Appearance-base) 방법	학습 영상 집합에 의해 학습된 모델을 이용해서 얼굴을 검출하는 방법

### 2.2 RGB -> YCbCr 컬러모델로 변환

입력영상은 RGB 컬러 모델로서, 이는 피부색과 입술색을 검출함에 있어서 컬러공간 3채널을 모두 사용하기 때문에 많은 데이터 처리량과 넓은 피부색 분포로 인하여 적합하지 않다[9]. 따라서 본 논문에서는 RGB에서 YCbCr 컬러모델로 변환된 것을 이용했으며, 이는 외부 조명의 변화에 강하다는 장점을 지니고 있다.

RGB 컬러모델에서 YCbCr 컬러모델로 변환하는 방법은 다음 식(1)과 같다.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 0.299000 & 0.587000 & 0.114000 \\ -0.168736 & -0.331264 & 0.500000 \\ 0.500000 & -0.418688 & -0.081312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

다음 식 (2)은 피부색과 입술색 범위 검출을 위한 Cb, Cr 범위 값이다.

$$\text{SkinColor}(x,y) = \begin{cases} 1 & \text{if } (84 \leq C_b \leq 151) \cap (135 \leq C_r \leq 170) \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

$$\text{LipColor}(x,y) = \begin{cases} 1 & \text{if } (76 \leq C_b \leq 128) \cap (132 \leq C_r \leq 174) \\ 0 & \text{otherwise} \end{cases}$$

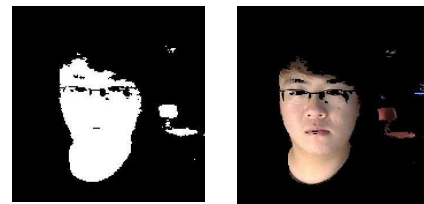


그림 2. 피부색 영역 검출 결과

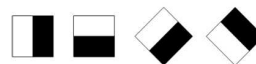


그림 3. 입술 후보 영역 검출 결과

### 2.3 Haar-like feature

Haar는 1910년에 간단한 직교 웨이블릿을 제안했다. 이를 확장한 Viola와 Jones가 처음 얼굴 검출에서 사용한 Haar-like feature는 가장 간단하면서도 효율적인 인자로 쓰인다. 이를 얼굴과 입술 검출에 사용하면 얼굴과 입술을 배경으로 판단할 확률이 1%보다 작아진다. 특징 값을 구하는데 요구되는 계산량은 단순히 합 연산밖에 없기 때문에 본 논문에서 사용하는 빠른 얼굴 및 입술 검출을 요구하는 응용분야에서 사용될 수 있다.

#### 1. Edge features



#### 2. Line features



#### 3. Center-surround feature



그림 4. Haar-lke Feature

## 2.4 Adaboost

Boost란 데이터들로부터 결과를 도출할 경우 약한 선택 기준들을 합쳐서 정확도가 높은 선택 기준을 만들어주는 알고리즘이다. Boost의 장점은 가정에 의해서 결과가 잘못 도출될 확률을 줄여주고 판단하기 어려운 문제를 정확하게 옳고 그름으로 판단할 수 있는 확률을 높여준다는 것이다. Adaboost 학습 알고리즘은 가장 잘 알려져 있는 boost 알고리즘이고 단순하면서도 효율적인 장점을 가지고 있다.

Adaboost 학습 알고리즘을 사용하려면 학습을 거듭할수록 얼굴 영역과 입술 영역을 잘 표현하는 특징값들을 많이 뽑아낼 수 있어 강력한 인식 알고리즘을 만들 수 있다. Adaboost 학습 알고리즘을 통해 뽑은 특징값들을 그림 5와 같이 단계별로 그룹화 한다.

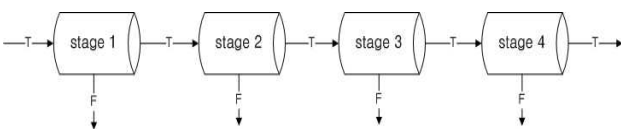


그림 5. N개의 stage를 갖는 분류기

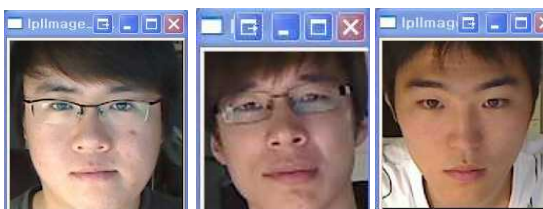
그룹화를 통하여 더 강력한 인식 알고리즘을 구현 할 수 있다. 첫 단계에서는 적은 수의 특징값들을 가지고 일정한 수준의 판별 능력을 가진 특징값들을 그룹화하고 다음 단계에서는 첫 단계에서보다 더 많은 수의 특징값들을 가지고 전 단계와 비슷한 판별력을 가지는 그룹을 만든다.

본 논문에서는 그림 4의 Haar-like feature와 그림 5의 Adaboost 학습 알고리즘을 이용하여 얼굴 영역 검출 및 최종 입술 영역을 검출하였다.

다음 그림 6과 그림 7은 각각 Haar-like와 Adaboost 학습 알고리즘을 사용하여 얻어진 얼굴 영역 검출과 입술 영역 검출의 결과이다.



(a) 검출 할 영역 표시



(b) 검출된 영역

그림 6. 얼굴 영역 검출



(a) 검출된 얼굴 영역 절반 제거



(b) 검출된 입술 영역 표시

그림 7. 입술 영역 검출

## 3. 실험 및 결과 분석

본 논문에서는 캠을 이용하여 실시간으로 촬영한 근거리 정면 영상을 사용하였으며, 제안한 입술 검출 알고리즘을 적용하여 성능을 평가하였다. 적용대상은 20-30대 성인 남녀를 기준으로 실험하였으며, 실험은 Logitech QuickCam Pro 9000, Intel(R) Core(TM)2 Duo CPU, MS Windows XP환경, Microsoft Visual C++ 6.0과 OpenCV Library 1.0을 이용하여 구현하였다.

본 논문에서 제안한 입술 검출의 실험 결과는 그림 8에 나타내었다. 실험 결과 입술 검출의 성공률은 95%로서 상당히 높은 정확성을 보였고, 5%의 실패율은 실외에서 촬영한 영상일 경우나 입력 영상의 포커싱 문제 혹은 피부색과 유사한 객체들로 인한 오인과 머리 염색 및 모양에 따른 비슷한 경우에 발생하였다. 그림 9는 검출 실패 영상을 보여준다.

입술 검출 결과

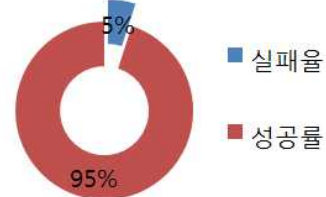
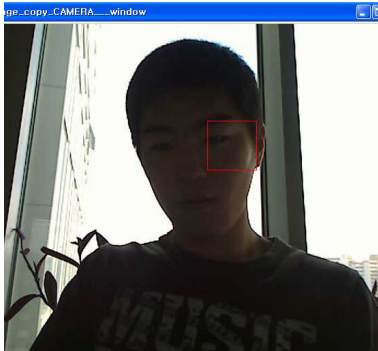


그림 8. 입술 검출 결과



(a)



(b)



(c)

그림 9. 실패 영상 예

#### 4. 결론

본 논문은 캠을 이용하여 실시간 입력받은 영상에서 인간의 얼굴 영역을 검출하여 입술 영역을 검출 하였다. 얼굴 영역 검출을 위해 우선 YCrCb의 컬러모델에서 밝기성분인 Y값을 무시하고 Cb, Cr값의 범위를 변경시켜줌에 따라 피부색의 범위를 검출하며, 검출된 피부색 범위에 Haar-like 특징을 가지고 Adaboost 학습 알고리즘을 적용시킴으로써 얼굴 영역을 검출한다. 입술 후보 영역은 검출된 얼굴 영역에 인간의 기하학 정보를 이용하여 입술 영역을 제외한 얼굴 영역의 이미지 절반을 제거 한 후 검출되어진다. 제거 된 입술 후보 영역에 대하여 Cb, Cr값의 범위를 변경시켜 입술색 영역을 검출하고, 최종적으로 입술색 영역에 다시 한번 더 Haar-like 특징과 Adaboost 학습 알고리즘을 적용 시켜줌으로써 최종 입술 영역을 검출하게 된다.

이는 기존의 얼굴 검출 영역만을 가지고 Haar-like 특징을 이용한 것과는 달리 입술 검출 영역에도 Haar-like 특징을 적용시켜줌으로써 본 논문에서 제안한 알고리즘을 실험한 결과 기존의 알고리즘보다 검출률이 높았으며, 적용범위가 더 넓음을 실험을 통해 확인할 수 있었다.

향후 연구는 입술을 검출함에 있어 입술 후보영역 뿐만 아니라 정확

한 입술모양의 검출을 할 예정이며, 조명이나 야외에서도 보다 강인한 검출률을 얻을 수 있도록 알고리즘을 향상 시킬 계획이다.

#### 참 고 문 헌

- [1] 홍석주, 이철우, “지능형 로봇을 위한 인간-컴퓨터 상호작용(HCI) 연구동향, 한국콘텐츠학회 2006년 제4권 제2호, 2006.
- [2] L. Zhang, “Estimation of the mouth features using deformable templates”, IEEE International Conference on Image Processing, Vol. III, pp. 328-331, 1997.
- [3] Wark T., Sridharan and V. Chandran, “An Approach to Statistical Lip Modelling for Speaker Identification via Chromatic Feature Extraction”, Proceedings of the 14th International Conf. on Pattern Recognition, Vol. 1, pp. 123-125, 1998.
- [4] Oliver N., A. Pentland, “LAFTER: Lips and Face Real Time Tracker”, Proceedings of the 1997 Conf. on Computer Vision and Pattern Recognition, pp. 123-129, 1997.
- [5] 한상일, 차형태, “Haar-like 특징과 템플릿을 이용한 귀 검출”, 방송공학회논문지 2008년 제13권 제6호, 2008.
- [6] 조성민, “피부색에 기반한 Cascade of Boosted Classifier working with Haar-like feature를 이용한 얼굴 검출”
- [7] 윤종일, “무선 로봇의 실시간 얼굴 검출 및 추적에 관한 연구”, 인제대 대학원, 2007.
- [8] 이병성, 전주철, “비전기반 실시간 입술 추적 및 입술색 변환”, 한국인터넷정보학회 2007 정기총회 및 추계학술대회 제8권 제 2호, 2007. 11.
- [9] M. Yang, D. J. Kriegman, N. Ahuha, “Detecting faces in Images: A Survey”, IEEE Trans. on Pattern Analysis and Machine Intelligence, pp. 34-58, VOL. 24, No. 1, January 2002.