

컬러 정보를 이용한 실시간 표정 데이터 추적 시스템

Realtime Facial Expression Data Tracking System using Color Information

이윤정*, 김영봉**

부산대학교 U-port 정보기술사업단*, 부경대학교 전자컴퓨터정보통신공학부**

Yun-Jung Lee(leeyj01@pknu.ac.kr)*, Young-Bong Kim(ybkim@pknu.ac.kr)**

요약

온라인 기반의 3차원 얼굴 애니메이션을 위해서 실시간으로 얼굴을 캡처하고 표정 데이터를 추출하는 것은 매우 중요한 작업이다. 최근 동영상 입력을 통해 연기자의 표정을 캡처하고 그것을 그대로 3차원 얼굴 모델에 표현하는 비전 기반(vision-based) 방법들에 대한 연구가 활발히 이루어지고 있다. 본 논문에서는 실시간으로 입력되는 동영상으로부터 얼굴과 얼굴 특징점들을 자동으로 검출하고 이를 추적하는 시스템을 제안한다. 제안 시스템은 얼굴 검출과 얼굴 특징점 추출 및 추적과정으로 구성된다. 얼굴 검출은 3차원 YCbCr 피부 색상 모델을 이용하여 피부 영역을 분리하고, Harr 기반 검출기를 이용해 얼굴 여부를 판단한다. 얼굴 표정에 영향을 주는 눈과 입 영역의 검출은 밝기 정보와 특징 영역의 고유한 색상 정보를 이용한다. 검출된 눈과 입 영역에서 MPEG-4에서 정의한 FAP를 기준으로 10개의 특징점을 추출하고, 컬러 확률 분포의 추적을 통해 연속 프레임에서 특징점들의 변위를 구한다. 실험 결과 제안 시스템은 약 초당 8 프레임으로 표정 데이터를 추적하였다.

■ 중심어 : | 얼굴 애니메이션 | 컬러 | 얼굴 특징점 |

Abstract

It is very important to extract the expression data and capture a face image from a video for online-based 3D face animation. In recently, there are many researches on vision-based approach that captures the expression of an actor in a video and applies them to 3D face model. In this paper, we propose an automatic data extraction system, which extracts and traces a face and expression data from realtime video inputs. The procedures of our system consist of three steps: face detection, face feature extraction, and face tracing. In face detection, we detect skin pixels using YCbCr skin color model and verifies the face area using Haar-based classifier. We use the brightness and color information for extracting the eyes and lips data related facial expression. We extract 10 feature points from eyes and lips area considering FAP defined in MPEG-4. Then, we trace the displacement of the extracted features from continuous frames using color probabilistic distribution model. The experiments showed that our system could trace the expression data to about 8fps.

■ keyword : | Face Animation | Color | Facial Feature Point |

I. 서론

실시간 얼굴 애니메이션은 크게 애니메이션 데이터의 획득과 얼굴 모델 제어라는 두 가지 연구 분야로 나누어진다. 얼굴 애니메이션 데이터를 얻기 위한 방법으로 스크립트를 이용한 방법, 모션 캡처 등이 사용될 수 있다. 스크립트를 이용하는 경우 텍스트나 기타 문법적인 요소를 통해 애니메이션을 구동시키므로 직관적이지 못하고, 모션 캡처의 경우 실제 움직임을 캡처하므로 정확한 표정 데이터를 얻을 수 있는 반면 마커와 같은 특수 장치들이 필요할 뿐만 아니라 얼굴에 마커의 부착은 자연스러운 표정을 짓는데 방해물이 될 수 있으므로 온라인과 같은 다양한 응용에 적합하지 않다.

얼굴 애니메이션을 위한 표정 데이터를 얻는 가장 현실적인 방법은 실제 얼굴 표정을 그대로 사용하는 것이라고 할 수 있다. 따라서 최근에는 동영상 입력을 통해 연기자의 표정을 캡처하고 그것을 그대로 3차원 얼굴 모델에 표현하는 비전 기반(vision-based) 방법들에 대한 연구가 활발히 이루어지고 있다[1-10].

일반적으로 비전 기반 방법들은 동영상으로부터 얼굴을 검출하고 얼굴내의 특징점들을 추출한 후 이를 추적하는 과정으로 이루어진다. 이러한 방법들은 자연스러운 상태로 촬영된 동영상으로부터 기존의 이미지 분석 기법 등을 활용하여 표정 데이터의 추정이 가능하다. 그러나 얼굴이나 얼굴 특징을 검출하기 위한 기존의 알고리즘들은 주로 영상내의 얼굴에서 얼마나 정확한 데이터를 추출하는가에 주안점을 두고 있으므로 실시간 응용을 위해서 적합하지 않은 경우가 많다. 그러므로 각 단계에서 사용되는 이미지 분석 기법의 성능이 전체 성능을 좌우하는 문제가 있으며, 컬러, 예지 등 다양한 영상 정보를 활용하는 이미지 분석 기법들을 하나의 시스템으로 통합하는 문제도 고려해야만 한다. 따라서 추적 전 과정에서 영상에서 컬러 정보나 예지 정보 등과 같은 다양한 정보들을 사용하는 것도 좋지만 쉽게 얻을 수 있는 정보를 다양하게 활용하는 것도 처리 속도를 개선하는 좋은 방법이 될 것이다.

본 논문에서는 마커를 부착하지 않은 자연스러운 상태로 촬영된 동영상 입력으로부터 실시간으로 얼굴과

얼굴 표정 데이터를 추출하는 방법을 제안한다. 제안 방법은 얼굴 검출, 얼굴 특징 추출, 얼굴 추적의 과정으로 구성된다. 본 논문에서는 컬러 영상에서 가장 쉽게 얻을 수 있는 컬러와 밝기 정보만을 얼굴 및 얼굴 특징 검출 및 추적의 전 과정에서 활용한다. 그러나 컬러 영상에서 조명이나 주변 배경의 영향은 검출이나 추적 성능에 많은 영향을 주기 때문에 강건한 피부 영역 검출을 위해 3차원 YCbCr 색상모델을 사용한다. 또한 기존의 연구에서 얼굴 검출을 위해 많이 사용되는 Haar-like feature를 이용해 얼굴 검증을 한다. Haar-like feature는 P. Viola가 객체 검출을 위해 제안한 것으로 높은 검출율을 보여 현재 많이 사용되고 있다[11]. 그러나 처리 속도에 대한 부담이 높아 실시간 응용에는 적합하지 않으나 제안 방법에서는 피부색을 이용하여 얼굴 후보 영역을 검출한 후 얼굴 검증 단계에서 Haar-like featur를 적용함으로써 컬러 정보만을 이용했을 때 발생할 수 있는 오검출을 줄일 수 있다. 또한 다양한 인종과 아무런 제약도 가하지 않고 촬영된 얼굴 이미지 데이터베이스를 대상으로 한 실험을 통해 얼굴 및 특징 검출을 위한 기존의 알고리즘들을 실시간 응용에 적합하게 수정, 보완한 방법들의 수행 성능을 보인다.

본 논문의 구성은 다음과 같다. 2장에서는 비전 기반의 얼굴 애니메이션 시스템에 대한 기존 연구들을 살펴보고, 3장에서 제안시스템의 구성을 설명한다. 4장에서는 얼굴 검출 및 특징 추출 방법과 얼굴 추적방법에 대해 자세히 설명한다. 다음으로 5장에서 실험 결과를 보이고, 마지막으로 6장에서 결론과 향후 과제를 제시한다.

II. 관련 연구

이번 장에서는 비전 기반의 얼굴 애니메이션 시스템에 관한 기존 연구들을 좀 더 자세히 살펴본다.

Chai[6]는 얼굴 모션 추정과 얼굴 표정 제어가 통합된 얼굴 애니메이션 시스템을 제안하였다. 이 방법에서는 동영상 내의 얼굴 모션을 추정하기 위해 3차원 실린

더 모델과 광류를 사용하였고, 실제 사람 얼굴에 76개의 마커를 부착해 획득한 모션 캡처 데이터를 이용하여 각 얼굴에 대한 기저를 생성하고, 이를 데이터베이스에 저장하였다. 이와 같은 처리를 통해 만들어진 얼굴 모션 파라미터와 표정 파라미터를 3차원 아바타 얼굴 모델에 적용하여 얼굴 애니메이션을 생성하였다. 이 방법은 동영상에서 추적될 특징점의 초기 위치를 수작업으로 설정하여야 하며, 아바타 얼굴 모델 변형에 모션 캡처 데이터를 사용하므로 실제와 비슷한 얼굴 표정을 생성할 수 있지만 추정된 제어 파라미터와 모션 캡처 데이터 사이의 매핑이 필요하다.

한편 Chun[5]은 동영상으로부터 추출된 제어 파라미터만을 이용하여 3차원 얼굴 모델을 제어하는 방법을 제안하였다. 이 방법에서는 3차원 얼굴 애니메이션을 실시간으로 생성하기 위하여 동영상으로부터 얼굴 영역을 검출한 후, 검출된 얼굴에 3차원 실린더 모델을 투영하여 실린더 모델과 광류를 이용하여 얼굴의 포즈를 추적하였다. 이 방법은 HT 컬러로 표현된 피부색만을 이용해 얼굴을 검출함으로써 피부색과 비슷한 배경이 얼굴 영역으로 오검출 될 수 있는 문제가 있다.

마지막으로 제안 방법과 유사한 기존 연구로 배경완과 최권택의 연구를 들 수 있다[12][13]. 이 연구에서는 동영상내의 CAMShift 알고리즘을 이용하여 얼굴을 추적하고, 검출된 얼굴 영역 내에서 YCbCr 컬러 정보를 이용하여 입 영역을 검출한다. 검출된 입영역의 위치와 AAM을 이용하여 얼굴 모양을 추정하고 있다. 우선 동영상 초기 프레임에서 얼굴 영역 추정을 위해 Haar-like feature를 이용하고 있다. 그러나 콧수염이나 안경 등과 같은 변화 요소가 있을 경우 검출 오류가 높고, 전체 영상을 대상으로 Haar-like feature를 적용할 경우 오검출의 가능성이 높을 뿐만 아니라 처리 속도의 저하를 피할 수 없다. 또한 초기 프레임에서 얼굴 검출이 실패할 경우 AAM을 적용할 수 없다는 단점이 있다. 추적된 얼굴 영역에 AAM을 적용하여 얼굴 특징점들을 추출하고 있다.

본 논문에서는 배경완과 최권택의 연구에서와 비슷하게 얼굴 검출과 추적을 위해서 Haar-like feature와 CAMShift 알고리즘을 사용하고 있다. 그러나 본논문

에서는 Haar-like feature의 적용으로 인한 속도 저하를 방지하기 위해 전체 이미지를 대상으로 한 얼굴 검출 용도가 아닌 얼굴 후보 영역을 검증하기 위해 Haar-like feature를 사용하고 있으며 실험을 통해 처리 속도의 개선 및 오검출율을 낮출 수 있음을 보였다. 또한 배경완과 최권택의 연구에서 입 영역 검출 후 AAM을 사용하여 얼굴 윤곽을 추정하여 얼굴 특징을 추출하는 것과는 달리 제안 방법에서는 얼굴 영역 내의 밝기 정보를 이용해 눈 영역을 검출하고 이를 바탕으로 입 주변의 특징점들을 추정한다. 얼굴 윤곽에 대한 별도의 추정과정 없이 추출된 특징점들의 위치 관계를 이용하여 얼굴의 기울어진 정도를 추정한다.

본 논문에서는 사전 학습이나 동영상 내의 얼굴에 대한 사전 정보가 없는 상황에서 얼굴을 추적하고 얼굴 특징점들을 추출하는 것을 목적으로 한다.

III. 제안 시스템의 개요

본 논문에서 제안하는 비전 기반의 얼굴 검출 및 추적 시스템은 [그림 1]과 같이 동영상으로부터 얼굴 및 특징 검출, 그리고 추적 과정으로 구성된다.

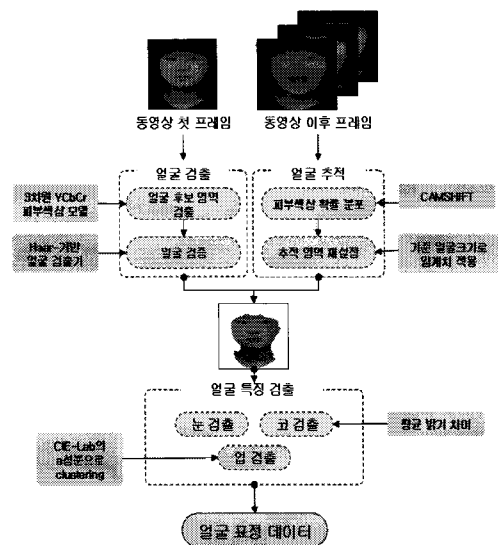


그림 1. 제안 시스템 개요

첫 번째 단계에서 얼굴 검출을 위해 머리에 부착되는 카메라나 얼굴에 부착된 마커가 없는 자연스러운 상태에서 촬영된 동영상 프레임에 입력받아 피부 픽셀을 검출한다. 피부색은 다른 물체와 구별되는 얼굴이 가진 고유한 특징 중의 하나로 색상 공간에서 일정 영역에 분포하는 특성을 지닌다. 여러 가지 색상 모델 중 YCbCr 컬러 모델이 동양인 피부색 검출에 우수한 성능을 보인다고 알려져 있다. D.Chai[14]는 컬러 영상에서 피부 픽셀 검출을 위해 YCbCr 칼라공간에서 밝기 값 Y를 제거한 CbCr에 대한 고정된 임계값을 제안하였다. Cb와 Cr 값에 대한 단일 임계치 적용은 구현하기 쉬우며, 비교적 높은 검출 성능으로 실시간 응용에서 얼굴 검출의 전처리 과정으로 널리 사용되고 있다. 그러나 컬러의 특성 상 조명의 영향을 배제할 수 없으므로 본 논문에서는 Y 성분을 함께 고려한 3차원 YCbCr 피부 색상 모델을 사용한다[7]. 3차원 YCbCr 피부 색상 모델은 피부 픽셀의 모든 밝기 Y에 대한 chrominance (Cr, Cb) 값을 별도로 가지므로 조명의 변화에 따른 피부색조의 변화에 대한 강건한 피부 픽셀 검출이 가능하다. 이렇게 검출된 피부 픽셀들의 불립을 형성하여 너무 크거나 너무 작은 불립들을 제외하고 남은 불립들을 얼굴 후보 영역으로 설정한다. 이 얼굴 후보 영역들을 대상으로 AdaBoost 알고리즘으로 학습된 Haar 기반 얼굴 검출기를 적용하여 검증함으로써 최종 얼굴 영역을 검출한다.

다음 단계로 영상 내에서 얼굴의 위치와 영역이 결정되면 이 영역 내에서 표정에 관계된 얼굴 특징점들을 검출한다. 제안 방법에서는 눈과 입 주변 그리고 코 끝의 위치를 특징점으로 사용한다. 각 특징 점들의 검출을 위해 탐색 영역을 설정하고, 탐색 영역 내에서만 각각의 특징점 검출을 시도한다. 눈 탐색 영역은 AdaBoost로 학습된 Haar 기반 양쪽 눈 검출기를 이용해 설정하였으며, 입과 코의 탐색 영역은 미리 검출된 양쪽 눈의 위치와 인체 측정학적 지식을 기반으로 설정하였다. 세 가지 특징 영역 모두 컬러 영상에서 손쉽게 얻을 수 있는 색상이나 밝기 정보를 이용하여 검출한다. 동영상의 첫 프레임에서 검출된 각 특징점들의 위치는 연기자의 얼굴 구조에 따른 표정 데이터의 정규화

를 위해서 사용될 것이다.

마지막으로 동영상의 두 번째 프레임부터는 첫 프레임에서처럼 얼굴 후보 영역과 검증 과정을 거쳐 얼굴 검출을 하지 않고 물체 추적 기법을 통하여 얼굴의 위치를 추적한다. 제안 방법에서는 기존에 제안된 여러 물체 추적 기법 중 비교적 속도가 빠르고 구현이 간단한 CAMShift 알고리즘을 기반으로 하며, 추적 시 얼굴과 비슷한 색상을 가지는 목이나 주변 배경이 포함되는 것을 막기 위해 프레임마다 추적 영역을 재설정하여 기준 크기 이상 추적 영역이 확장되는 것을 피한다. 추적된 얼굴 영역에서 특징점들을 검출하기 위해 추적 시 구해진 얼굴 회전 각도를 이용하여 얼굴 영역 이미지를 수평으로 보정하고 첫 프레임에서와 동일하게 각각의 특징점을 검출한 후 다시 회전시켜 특징점의 위치를 정한다.

IV. 얼굴 표정 추출

1. 얼굴 검출

영상내에서 피부 영역 검출을 위해 RGB 컬러 영상을 YCbCr 컬러 공간으로 변환 후, 3차원 YCbCr 피부 색상 모델을 적용하여 각 픽셀이 피부인지 아닌지를 판별한다. 3차원 YCbCr 피부 색상 모델을 이용하여 2진화된 이미지 내에는 피부색과 비슷한 컬러 특성을 가지는 많은 잡음 요소들이 포함되어 있다. 이러한 잡음 요소들을 제거하기 위해서 모폴로지 연산을 적용하고, 실제 얼굴이 될 가능성이 있는 얼굴 후보 영역을 검출하기 위해 피부 불립(blob)을 생성한다. 불립(blob)은 이미지에서 주변의 영역이나 픽셀보다 더 밝거나 어두운 시각적 모듈 영역을 가리키는 것으로 이미지 내에서 검출된 불립들은 주변과 연결되지 않는 독립적인 피부 영역으로 간주된다. 검출된 피부 불립들 중 불립 영역을 고려하여 너무 크거나 작은 불립들은 제외시키고 최종 얼굴 후보 영역을 검출한다. [그림 2]는 검출된 얼굴 후보 영역을 보여준다.

이렇게 검출된 얼굴 후보 영역 중에는 실제 얼굴뿐만 아니라 피부색과 비슷한 배경 영역 등이 포함될 가능성

이 높으므로 얼굴 후보 영역이 실제 얼굴인지 아닌지를 판별하기 위한 얼굴 검증과정이 필요하다. 본 논문에서는 얼굴 검증을 위해 얼굴 후보 영역에서 Haar 기반 얼굴 검출기를 적용하여 얼굴인지 아닌지 판단한다.



그림 2. 얼굴 후보 영역

얼굴 영역은 얼굴을 포함하는 바운딩 박스 형태로 검출되므로 이 영역 내에는 얼굴 주변의 머리카락과 같은 외부 잡음과 눈썹, 눈과 같은 내부 홀(hole)들이 존재한다. 이후의 얼굴 특징 영역 검출에서 얼굴 내부로 탐색 영역을 제한하기 위해 얼굴에 속하는 부분과 그렇지 않은 부분으로 구분되는 얼굴 마스크를 생성할 필요가 있다. 따라서 검출된 얼굴 후보 영역 검출에서 생성된 이진화 이미지에서 라벨링(labeling)을 통해 내부 홀들을 채워 얼굴 마스크를 생성한다. [그림 3]은 생성된 얼굴 마스크 이미지를 보여준다.



그림 3. 얼굴 마스크

2. 얼굴 특징 추출

2.1 눈 검출

얼굴 내에서 각 특징점들의 검출을 위해 탐색 영역을 설정함으로써 불필요한 탐색을 줄이고 오검출을 줄일 수 있다. 얼굴 내에서 눈은 얼굴 중심에서 상단 부분에 위치한다. 이러한 인체 측정학적 지식만을 이용해 눈

탐색 영역을 설정할 경우 머리카락이나 눈썹 등과 같이 눈과 비슷한 밝기 특성을 가지는 잡음으로 인해 강건한 눈 검출이 어렵다. 따라서 본 논문에서는 Bediz와 Akar[2]에 의해 제안된 Haar 기반의 양쪽 눈 검출기를 사용하여 눈 탐색 영역을 설정한다. [그림 4]는 Haar 기반 눈 검출기를 통해 설정된 눈 탐색 영역을 보여준다.

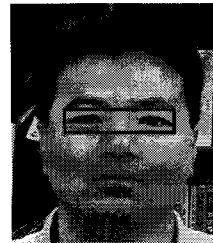
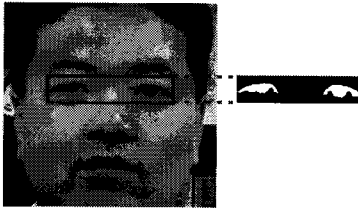


그림 4. 눈 탐색 영역

Haar 기반 검출기로는 정확한 눈의 위치를 알 수 없으므로 눈의 형태나 특성을 나타내는 다른 정보를 이용할 필요가 있다. 이미지 내에서 물체의 위치나 형태를 파악하기 위해 밝기 분포나 에지 등을 이용할 수 있다. 제안 방법에서는 눈은 주변의 다른 픽셀보다 현저히 어둡다는 점을 이용하여 눈 탐색 영역을 픽셀의 밝기 값을 기준으로 이진화 한다. 이진화를 위해 눈 탐색 영역을 그레이스케일 이미지로 변환하여 전체 영역의 평균 밝기와 각 픽셀의 밝기 값의 차이를 계산하고, 그 값이 임계치 이상이면 눈 영역으로 그렇지 않으면 피부 픽셀로 간주한다. 주변 픽셀과의 상대적인 밝기 차이에 임계치를 적용함으로써 전체적으로 밝거나 어두운 이미지에서도 강건하게 눈 영역을 검출할 수 있다. 눈 탐색 영역의 이진화는 식 1과 같다.

$$E(x, y) = \begin{cases} 255 & |V - V_{avg}| > Threshold \\ 0 & otherwise \end{cases} \quad (1)$$

여기서 V 는 (x, y) 픽셀의 밝기 값이고, V_{avg} 는 탐색 영역의 평균 밝기 값이며, $E(x, y)$ 는 이진화 된 눈 이미지를 나타낸다. 이진화 된 눈 영역 이미지가 [그림 5](a)에 나타나 있다. 눈 탐색 영역의 이진 이미지가 생성되면 흰색 픽셀의 히스토그램을 고려하여 왼쪽과 오른쪽 눈의 위치를 검출한다. [그림 5](b)는 최종 검출된 눈 위치를 보여준다.



(a) 이진화 된 눈 영역

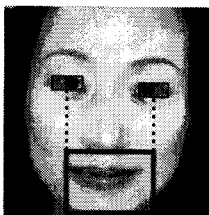


(b) 최종 검출된 눈 위치

그림 5. 눈 검출

2.2 입 검출

얼굴에서 입 주변은 비슷한 색상을 가지는 잡음 영역이 비교적 적고, 무표정한 상태에서의 입 영역은 얼굴이나 눈처럼 두드러진 텍스처 특성을 가지지 않기 때문에 Haar 기반 검출기의 검출 성능이 다른 영역에 비해 떨어지는 편이다. 따라서 검출된 양쪽 눈의 위치와 인체 측정학적 지식을 고려하여 얼굴 내에서 입 탐색 영역을 설정한다. 사람 얼굴에서 입은 대략적으로 수평 방향으로의 왼쪽 눈의 중심과 오른쪽 눈의 중심 사이에 위치하고, 수직 방향으로는 눈과 턱 사이 중 아래쪽 절반 부분에 위치한다고 할 수 있다. [그림 6]은 입 탐색 영역과 이진화된 입 마스크를 보여준다.



(a) 입 탐색 영역



(b) 이진화 된 입 탐색 영역

그림 6. 입 탐색 영역

[그림 6](a)은 제안 방법으로 설정된 입 탐색영역이다. 입 탐색 영역에서 정확한 입의 위치와 영역을 검출하기 위해 입 영역은 주변의 픽셀보다 붉은 색 강도가 강하다는 컬러 특성을 이용한다. 픽셀의 붉은 색 강도를 구하기 위해 여러 색상 모델 중 밝기 성분 분리 성능이 우수한 CIE-Lab 컬러를 이용한다. CIE-Lab 컬러 모델은 1976년 국제 조명 협회(Commission Internationale d' Eclairage : CIE)에서 색상 측정을 위해 국제 표준으로 제안한 uniform color space 중의 하나이다. Lab 색상은 장치 독립적이며 광도 즉 밝기 요소(L)와 두 색조 a와 b로 요소로 이루어진다. 두 색조 요소 중 a 성분은 Red와 Green 색의 정도, b 성분은 Yellow와 Blue 색의 정도를 나타낸다.

입술 영역 검출을 위해서 입 탐색 영역 이미지를 Lab 색상 공간으로 변환한 후 각 픽셀의 a 성분 값을 이용해 입술인 픽셀과 주변 픽셀로 분리한다. 사람마다 입술 색상에 많은 차이가 있으므로 단일 임계치로 입술 픽셀을 검출하기는 어려운 일이다. 그러나 입술 영역은 사람마다 색상의 차이가 있기는 하지만 피부 픽셀보다는 붉은 색상이 강하므로 단일 임계치 적용 대신 K-Means Clustering 알고리즘을 이용하여 입 탐색 영역을 두 개의 클러스터로 분리한다. K-means 클러스터링 알고리즘은 초기 클러스터의 중심 값을 어떻게 설정하느냐에 따라 클러스터링 결과가 민감하게 반응하므로 본 실험에서는 초기 클러스터의 중심 값을 탐색 영역의 각 픽셀의 a 성분의 평균값과 표준편차를 이용해 식(2)와 같이 설정한다.

$$\begin{aligned} Centroid[0] &= Avg_a - Std_a \\ Centroid[1] &= Avg_a + Std_a \end{aligned} \quad (2)$$

여기서 Avg_a와 Std_a는 탐색 영역의 a 성분 값의 평균과 표준 편차이다. 두 개의 클러스터로 분리된 것 중 붉은 색 강도가 높은 것을 입술 영역으로 간주한다. [그림 6](b)는 두 개의 클러스터로 분리된 입 탐색 영역의 이진 영상을 보여준다. 이진화 된 영상에서 흰색 픽셀들의 히스토그램을 고려하여 최종 입 영역을 설정한다. [그림 7]은 최종 검출된 입 영역을 보여준다.

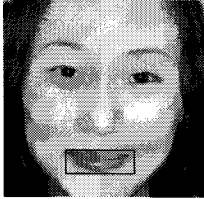


그림 7. 검출된 입 영역

검출된 눈과 입 영역에서 애니메이션에 필요한 얼굴 특징점은 MPEG-4에서 정의한 총 64개의 FDP 중에서 눈과 입의 움직임에 관련된 8개의 점으로 선택하였다. [그림 8]은 MPEG-4에서 정의한 FDP를 보여준다.

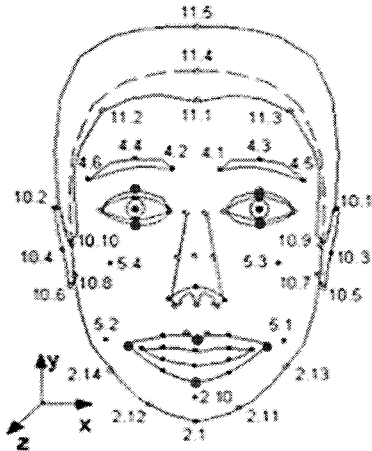


그림 8. 눈과 입 영역의 특징점

[그림 8]에서 붉은 점으로 나타낸 것이 제안 시스템에서 사용하는 눈과 입 영역의 8개의 특징점들이다.

3. 얼굴 추적

컴퓨터 비전 분야에서 얼굴 추적을 위해 예지나 피부색과 같은 다양한 특성들이 사용되고 있다. Snake나 Active Shape Model을 이용하는 방법 등은 얼굴의 예지 정보를 이용해 얼굴의 윤곽이나 형태 정보를 잘 추적할 수 있는 반면 복잡한 배경이나 경계가 명확하지 않은 물체의 추적에 적용하기 힘들고 계산 비용이 높아 실시간 처리 분야에는 부적합하다. 반면에 피부색은 가장 손쉽게 얻을 수 있는 얼굴 정보이며 얼굴이나 손 추적 등 여러 응용에서 효율적인 정보가 될 수 있다. 이

러한 컬러 정보를 이용한 물체 추적 알고리즘 중의 하나인 CAMShift 알고리즘은 상대적으로 적은 계산 복잡도로 실시간 응용에 적합하다고 할 수 있다. 그러나 CAMShift 알고리즘은 컬러 분포 확률을 이용하여 비슷한 색상 분포를 가지는 영역을 추적함으로써 추적과정이 진행될 때 비슷한 피부색을 가지는 목 부분이 얼굴 영역에 포함되어 추적되는 경향이 있다. 따라서 목 부분은 보통 옷으로 가리지 않는 한 얼굴과 연결되어 얼굴 영역에 포함되어 추적된다.

[그림 9]는 CAMShift 알고리즘을 사용하여 프레임별로 추적된 얼굴 영역을 보여준다.

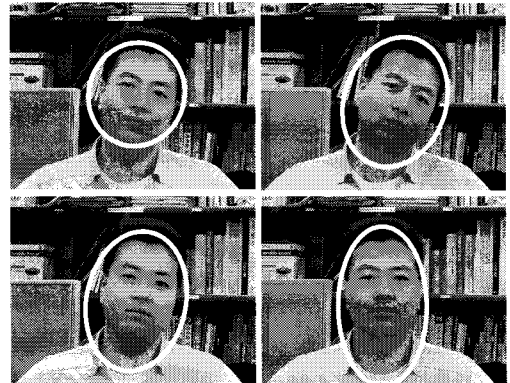


그림 9. CAMShift 알고리즘을 사용하여 추적된 얼굴 영역

그림에서 보듯이 추적이 진행 될수록 비슷한 색상 분포를 가지는 목 부분까지 얼굴 영역으로 포함되어 추적되는 것을 알 수 있다. 이뿐만 아니라 피부색과 비슷한 배경 영역이 얼굴과 연결되어 있을 경우 그 부분까지도 포함되어 추적되는 문제가 발생한다. 따라서 목 부분과 피부색과 비슷한 얼굴 영역이 얼굴 추적에 포함되지 않도록 하기 위해 첫 프레임에서 검출된 얼굴 영역의 크기를 고려하여 추적 영역을 재설정 한다. 현재 프레임의 얼굴 영역에서 특징점들을 검출한 후, 그 위치를 고려하여 추적 영역을 재설정한다. 이때 첫 프레임에서 설정된 얼굴 영역 크기를 기준 크기로 정하여 추적 영역이 확대된 채로 다음 프레임으로 진행하지 않도록 한다. [그림 10]은 얼굴 영역 재설정을 통해 추적된 얼굴 영역을 보여준다.

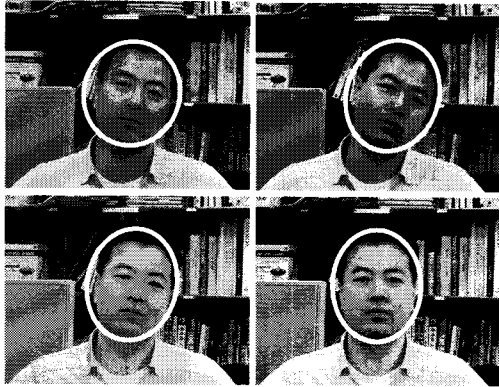
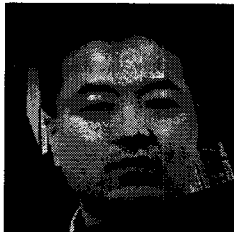


그림 10. 추적 영역을 재설정한 얼굴 추적 결과

위와 같이 추적된 얼굴 영역에서 눈, 코, 입 등의 얼굴 특징점을 추출하기 위해서는 얼굴의 회전 상태를 고려하여야 한다. 그림에서 보듯이 CAMShift 알고리즘으로 추적한 결과 얼굴의 회전 정보까지 추출할 수 있으므로 이 회전 각도를 얼굴 특징 검출에 이용한다. 추적된 얼굴 영역을 [그림 11]에서와 같이 회전 각도를 이용하여 역회전시킨 후 양쪽 눈의 수평 위치가 수평에 가깝게 보정된 얼굴 이미지를 생성한다.



(a) 기울어진 얼굴 영역



(a) 회전 각도를 이용해 양쪽 눈의 수평을 보정한 얼굴 영역

그림 11. 수평 보정된 얼굴 이미지

얼굴 영역내의 양쪽 눈 수평이 보정되면 얼굴 영역 내에서 눈, 코, 입 등의 특징 영역을 검출하고, 각 지점

들을 다시 얼굴 회전각도 만큼 재 회전시킴으로써 회전된 얼굴에서도 얼굴의 특징 영역들의 정확한 위치를 구해낼 수 있다. [그림 12]는 회전된 얼굴에서 검출된 얼굴 특징 영역들을 보여준다.

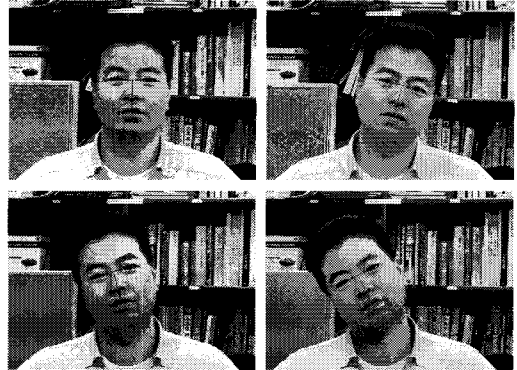


그림 12. 동영상 프레임에서 얼굴 및 얼굴 특징 영역 추적 결과

V. 실험 결과

본 실험은 세 가지로 수행되었다. 첫째는 피부색과 Haar 기반 분류기를 함께 사용한 얼굴 검출 성능, 둘째로 얼굴 특징 검출 성능, 마지막으로 얼굴 추적 성능으로 이루어진다. 실험에 사용된 이미지들은 Caltech의 정면 얼굴 DB와 인터넷에서 수집된 컬러 영상 그리고 직접 촬영된 영상으로 구성되었다.

1. 얼굴 검출

제안 방법의 얼굴 검출 성능 비교를 위해 Haar 기반 얼굴 검출기만을 사용해 얼굴 검출을 수행했을 때와 피부 색상 모델을 함께 사용했을 때의 얼굴 검출 속도를 비교하였다. 제안 방법에서 피부 색상 모델을 사용하여 얼굴 후보 검출 시 블립 형성을 위해 적용되는 임계치는 얼굴 영역 크기마다 달라져야 하지만 본 실험에서는 얼굴 최소 크기를 80×80로 하고, 최대 크기는 300×300으로 설정하여 블립에 포함되는 픽셀 수가 약 5,000개 이상에서 90,000개 이하가 되는 블립을 얼굴 후보로 간주하였다. 실험 영상은 다양한 배경에서 촬영된 50장의

컬러 이미지를 사용하였고, 이미지 내에는 얼굴이 하나만 있는 것으로 제한하였다. [그림 13]은 실험에 사용된 실험 영상을 보여준다.



그림 13. 얼굴 검출에 사용된 실험영상

Haar 기반 Classifier로는 Lienhart[15]가 제안한 정면 얼굴 분류기(FA1)를 사용하였고, 얼굴 후보 영역에 대해서만 얼굴 검증을 실시한 경우와 전체 이미지에서 얼굴을 검출했을 경우의 성능을 비교하였다. 실험은 Haar 기반 검출기를 이용해 얼굴을 검출한 경우와 제안 방법에서와 같이 얼굴 후보에 대해서만 Haar 기반 검출기를 적용했을 때의 두 가지 경우에 대한 처리 속도를 측정하였다. 이미지 크기와 피부색과 유사한 배경이 얼마나 포함되어 있느냐에 따라 차이가 있으나 50장의 실험 영상에서 측정된 검출 속도에서 평균적으로 20%정도 개선됨을 보였다. [표 1]은 실험 영상에 대한 평균 처리 속도이다.

표 1. 평균 얼굴 검출 속도

	제안방법	Haar	성능비
평균	0.334	0.424	79%

본 실험에서 피부색을 이용해 얼굴 후보를 검출하고, 얼굴 후보에 대해서만 Haar 기반 분류기로 얼굴 검증을 하는 것이 처리 속도면에서 더 효율적이라는 것을 알 수 있었다.

2. 얼굴 특징 검출

얼굴 내의 특징 영역 검출 성능 평가는 다양한 조명 조건에서 촬영된 450장의 컬러 얼굴 이미지를 사용하여 수행되었다.

2.1 눈 검출

눈 검출을 위해서 Bediz와 Akar에 의해 제안된 Haar 기반의 Classifier(EP3)를 사용하여 눈 탐색 영역을 설정하고, 탐색 영역 내의 밝기 값의 차이를 이용해 눈 위치와 눈 영역을 검출하였다. 이때 밝기 차이의 임계치는 50으로 설정하였다. [표 2]에 전체 실험 영상에 대한 검출율이 나타나 있다. 실험 결과 전체 실험 영상에서 약 78.5% 정도의 검출율을 보였다.

표 2. 눈 영역 검출 성능

	검출 성공	검출 실패	합계
이미지 수	353	97	450
검출율	78.5%	21.5%	100%

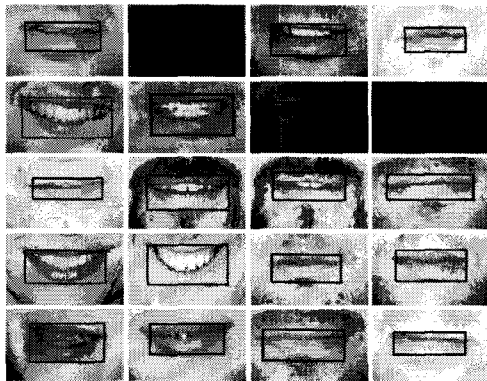
[그림 14]는 제안 방법으로 눈 검출 결과를 보여준다. 검출 실패한 요인으로는 안경을 착용한 경우, 부분 조명으로 인해 눈 부분에 진한 그림자가 생긴 경우, 허용치 이상 회전된 경우, 옅은 눈동자 색으로 밝기 값 차이가 크지 않은 경우 등이 있었다.

2.2 입 검출

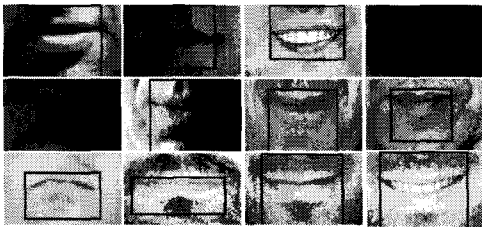
얼굴 영역 내에서 입 탐색 영역의 가로 범위는 왼쪽 눈의 중심에서부터 오른쪽 눈의 중심 영역까지로 설정하고, 세로 범위는 좌, 우 눈 영역의 세로 끝 지점부터 얼굴 끝 지점까지의 1/3 지점에서 시작하여 얼굴의 세로 끝 지점까지로 설정하였다. [그림 15]는 제안 방법의 입 검출 결과를 보여준다.

전체 실험 이미지 중 콧수염이 있는 경우에 검출율이

낮으나 전체적으로 조명이나 인종에 따라서는 거의 높은 검출률을 보임을 알 수 있다. 실험에서 검출에 실패한 경우는 주로 콧수염등의 부가적인 잡음이 부분적으로 있는 경우와 부분 조명으로 인해 입술의 특정 부분이 많이 가려지는 경우 오검출 됨을 알 수 있었다. [그림 15](a)는 입 검출에 성공한 샘플 이미지이며, [그림 15](b)는 검출에 실패한 샘플 이미지가 나타나 있다.



(a) 입 검출에 성공한 샘플 이미지



(b) 입 검출에 실패한 샘플 이미지

그림 15. 눈 검출 결과

제안 방법의 입 검출 성능이 [표 3]에 나타나 있다. 실험 이미지는 콧수염이 있는 경우의 이미지 133장과 보통의 이미지 317장으로 총 450장으로 구성되었다.

표 3. 입 검출 성능

	이미지 수	성공	실패	성공율
보통	317	275	42	87%
콧수염	133	88	45	66%
전체	450	363	87	81%



(a) 눈 검출에 성공한 예



(b) 눈 검출에 실패한 예

그림 14. 눈 검출 결과

450개의 이미지에서 입 검출의 성공한 것은 363개로 81%의 검출율을 보였고, 콧수염이 있는 경우는 88개가 검출에 성공하여 66%의 검출율을 보였다.

3. 얼굴 추적

얼굴 추적에 대한 실험 영상은 초당 15프레임으로 촬영된 컬러 동영상으로 각 프레임의 크기는 640×480 픽셀이며, 영상내의 얼굴은 하나로 제한하였다. 얼굴의 회전과 입 벌림 등의 모션으로 인해 프레임마다 얼굴의 크기가 조금씩 바뀔 수 있으므로 이를 고려하여 추적 영역의 임계치는 첫 프레임에서 검출된 얼굴 크기인 기준 크기의 가로, 세로 130%로 설정하였다. 본 실험에서는 얼굴의 움직임은 이동과 좌, 우 어깨 쪽으로 향하는 회전(rolling)으로 제한하였다. 얼굴 추적에 따른 처리 속도는 약 50fps 정도로 측정되었다. [그림 16]은 추적된 얼굴 영역에서 추출된 눈과 입 영역을 보여준다.



그림 16. 얼굴 추적과 얼굴 특징 추출

얼굴 추적 및 얼굴 특징 추출 성능은 초당 8 프레임 정도로 측정되었고, 프레임마다 추적 영역 재설정에 따른 추적속도 저하는 나타나지 않았다. 실험에서 얼굴의 이동이나 좌우회전(rolling) 외의 상하 회전(pitch) 등의 모션으로 인해 얼굴의 수평 보정이 이루어지지 않은 경우 얼굴 영역 추적은 가능하였으나 얼굴 특징 검출에 실패하였다.

VI. 결론 및 향후 과제

본 논문에서는 연기자의 얼굴에 아무런 제약도 가하지 않은 자연스러운 상태에서 촬영된 동영상 프레임에서 얼굴과 얼굴 특징점들을 추출하고, 추적하는 시스템을 제안하였다. 제안 시스템은 컬러 영상에서 3차원 YCbCr 피부 색상 모델을 이용하여 얼굴 후보를 검출하고, Haar 기반 얼굴 검출기를 통해 검증함으로써 인종에 관계없는 강건한 얼굴 검출이 가능하였다. 또한 얼굴 영역 내에서 각각의 특징 영역을 검출하기 위해 얼굴 내에서 탐색 영역을 설정하여 각 영역의 밝기와 색상 특성을 이용하여 특징 영역을 검출하였다. 눈 영역

검출을 위해 Haar 기반 양쪽 눈 검출기로 눈 탐색 영역을 설정한 후 주변 픽셀과의 밝기 차이를 이용해 좌, 우 눈 위치를 검출하였으며, 입 영역 검출을 위해서 인체 측정학적 지식을 기반으로 입 탐색 영역을 설정하고 CIE-Lab 컬러의 a 성분에 대한 K-means clustering 알고리즘 적용으로 입 영역을 검출하였다. 실험 결과 조명의 변화나 수염과 같은 잡음 요소가 포함된 경우에도 80% 이상의 검출율을 보였다. 얼굴 추적을 위해서는 피부 색상의 확률 분포를 기반으로 하는 CAMShift 알고리즘을 이용하였다. 또한 기존의 CAMShift 알고리즘에 첫 프레임에서 검출된 얼굴 영역 크기를 기준으로 추적 영역을 재설정함으로써 얼굴과 비슷한 피부 색상을 가지는 목이나 배경이 얼굴 영역 포함되는 것을 막을 수 있었다. 본 논문에서 제안한 얼굴 검출 및 특징 영역 추적은 전 과정이 사용자의 개입 없이 자동으로 처리되며 실시간 처리가 가능하므로 온라인 기반의 응용에 활용될 수 있을 것으로 기대된다.

향후 과제로는 제안 방법으로 추적된 특징점들의 범위 정보를 이용하여 3차원 아바타 얼굴 모델에 직접 적용하여 실시간 3차원 얼굴 애니메이션을 생성하고자 한다.

참고 문헌

- [1] J. Ostermann, "Animation of Synthetic Faces in MPEG-4," Computer Animation, pp.49-51, 1998.
- [2] Y. Bediz and G. B. Akar, "View point tracking for 3d display systems," 3th European Signal Processing Conference, 2005.
- [3] 이형수, "얼굴의 색상과 모양정보를 이용한 조명 변화에 강인한 얼굴 추적 시스템 구현", 정보처리학회 논문지B, 제10-B권, 제6호, pp.711-718, 2003.
- [4] 김정기, 민경필, 전준철, 최용길, "현실감 있는 3차원 얼굴 애니메이션을 위한 실시간 표정 제어", 인터넷정보학회 논문지, 제7권, 제2호, pp.23-35, 2006.
- [5] 권오륜, 전준철, "얼굴모션추정과 복제에 의한 3

차원 얼굴 애니메이션”, 정보처리학회 논문지, 제 14권, 제4호, pp.311-320, 2007.

[6] J. X. Chai, X. Jing, and H. Jessica, "Vision-based control of 3D facial animation," Proceedings of the 2003 ACM SIGGRAPH/Eurographics symposium on Computer animation, pp.193-206, 2003.

[7] G. M. Park and Y. B. Kim, "3D Skin Color Model For Asian Skin Color Segmentation," 2007 MITA Conference Proceedings, pp.165-168, 2007.

[8] H. P. Graf, T. Chen, E. Petajan, and E. Cosatto, "Locating Faces and Facial Parts," Proc. First Int'l Workshop Automatic Face and Gesture Recognition, pp.41-46, 1995.

[9] H. P. Graf, E. Cosatto, D. Gibbon, M. Kocheisen, and E. Petajan, "Multimodal System for Locating Heads and Faces," Proc. SecondInt'l Conf. Automatic Face and Gesture Recognition, pp.88-93, 1996.

[10] J. Yang and A. Waibel, "A Real-Time Face Tracker," Proc. Third Workshop Applications of Computer Vision, pp.142-147, 1996.

[11] P. Viola and M. J. Jones, "Rapid Object Detection using a Boosted Cascade of Simple Features," IEEE CVPR, Vol.1, No.2, pp.511-518, 2001.

[12] 배정완, 최권택, 변혜란, "Pan-Tilt-Zoom-Camera에서 AAM과 CAMShift를 이용한 얼굴 검출 및 추적", 한국정보과학회 2005 가을 학술발표 논문집, 제32권, 제2호, pp.931-933, 2005.

[13] 최권택, 변혜란, "얼굴 영역 추적과 입 검출을 이용한 AAM 얼굴 모양 파라미터 추정", 한국정보과학회 2005 가을 학술발표 논문집, 제32권, 제2호, pp.928-930, 2005.

[14] D. Chai and K. N. Ngan, "Face Segmentation Using Skin-Color Map in Videophone Applications," IEEE Transactions on Circuits

and Systems for Video Technology, Vol.9, No.4, 1999(6).

[15] J. C. Terrillon, M. Shirazi, H. Fukamachi, and S. Akamatsu, "Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images," in Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000.

[16] R. Lienhart and M. Jochen, "An extended set of haar-like features for rapid object detection," IEEE ICIP 2002, Vol.1, pp.900-903, 2002.

저자 소개

이 윤 정(Yun-Jung Lee)

정회원



- 1995년 2월 : 부경대학교 전자계산학과(이학사)
- 1999년 2월 : 부경대학교 전산정보학과(이학석사)
- 2008년 8월 : 부경대학교 전자계산학과 (공학박사)

• 2008년 9월 ~ 현재 : 부산대학교 U-Port 정보기술사업단 박사후연구원

<관심분야> : 얼굴 애니메이션, 웹 시각화

김 영 봉(Young-Bong Kim)

종신회원



- 1987년 : 서울대학교 계산통계학과(학사)
- 1989년 : 한국과학기술원 전산학과(석사)
- 1994년 : 한국과학기술원 전산학과(박사)

• 1994년 3월 ~ 1995년 2월 : 삼성전자 정보기술 연구소 선임연구원

• 1995년 3월 ~ 현재 : 부경대학교 컴퓨터멀티미디어 공학부 교수