

# 다시점 영상 시스템을 위한 얼굴 추적

한 충 신\*, 장 세 훈\*, 배 진 우\*, 정희원 유 지 상\*

## Face Tracking for Multi-view Display System

Chung-shin Han\*, Se-hoon Jang\*, Jin-woo Bae\*, Ji-sang Yoo\* *Regular Members*

### 요 약

본 논문에서는 관찰자의 시점에 적응적인 다시점 영상 합성 시스템을 위한 얼굴 추적 기법을 제안한다. depth 카메라를 이용하여 텍스처 영상과 깊이 정보를 획득한 후, 회전(rotation)과 이동(translation) 등 기하학적인 변환을 이용하여 관찰자의 위치를 고려한 다시점 영상을 생성할 수 있다. 관찰 시점의 위치와 각도에 의한 입체감(motion parallax cue)을 제공하기 위하여 주 관찰자의 얼굴을 추적하는 기법을 제안한다. 제안된 기법에서는 모니터에 설치된 얼굴 획득용 카메라로부터 얻은 영상에서 얼굴 색상의 통계학적 특성과 변형적 형식(template)을 이용하여, 실시간으로 초기에 설정된 주요 관찰자(dominant face)의 얼굴영역을 추적하게 된다. 실험 결과 복잡한 배경 하에서도 얼굴 영역의 위치를 성공적으로 검출 및 추적하여 관찰자 시점에 해당하는 3차원 입체 영상을 디스플레이 할 수 있었다.

**Key Words** : Face Tracking, Multi-view synthesis, statistical characteristics of face colors

### ABSTRACT

In this paper, we proposed a face tracking algorithm for a viewpoint adaptive multi-view synthesis system. The original scene captured by a depth camera contains a texture image and 8 bit gray-scale depth map. From this original image, multi-view images can be synthesized which correspond to viewer's position by using geometrical transformation such as a rotation and a translation. The proposed face tracking technique gives a motion parallax cue by different viewpoints and view angles. In the proposed algorithm, tracking of viewer's dominant face initially established from camera by using statistical characteristics of face colors and deformable templates is done. As a result, we can provide motion parallax cue by detecting viewer's dominant face area and tracking it even under a heterogeneous background and can successfully display the synthesized sequences.

### I. 서 론

인간의 끊임없는 욕구는 보다 사실적이고 실감 있는 영상을 추구하게 되었다. 따라서 현실성 및 자연성이 최대로 반영된 3D 입체영상 콘텐츠 생성 기술은 지속적으로 연구 개발되어야 할 분야이다. 한 명의 관찰자에게만 입체감을 제공하는 스테레오 방

식의 경우 관찰자에게는 두 개의 평면 영상이 필요하고 따라서 이들을 처리하는데 있어서 기존의 동영상 압축 표준인 MPEG 또는 H.261, H.263 등을 적용할 수 있는 장점이 있다. 그러나 이러한 스테레오 방식의 경우 시점이 단 한 곳으로 고정되어 있기 때문에 관찰자가 이 시역을 벗어나게 되거나 인체 구조상의 차이로 인해 정확한 시역에 두 눈이

\* 광운대학교 전자공학과 디지털 미디어 연구실 whale11@image.gwu.ac.kr, justdash@image.gwu.ac.kr, bjw8751@image.gwu.ac.kr, jsyoo@daisy.kw.ac.kr,

논문번호 : KICS2004-11-256, 접수일자 : 2004년 11월 18일

※본 연구는 대학 IT 연구센터 육성지원 사업의 연구결과로 수행되었음

고정되지 못하게 되면 입체감을 느낄 수 없거나 심한 피로감을 느끼게 되는 단점이 있다.

많은 관찰자에게 자연스러운 입체감을 주기 위한 다시점 영상 디스플레이 기법은 현재 가장 효율적으로 입체감을 표현해 줄 수 있는 기법이다 그러나 다시점 영상 콘텐츠를 얻기 위해 다수의 카메라를 사용하는 경우에는 다수의 카메라 사용에 따른 부가적인 영상처리로 인한 오류가 존재하게 된다 뿐만 아니라 카메라 수와 카메라 간격의 제한으로 인해 시점의 이동시 불연속성이 발생할 수 있고 데이터 양이 많아지는 단점을 가지고 있다 본 논문에서는 기존의 여러 대의 카메라를 사용하는 다시점 영상처리 기술과는 다르게 한 대의 카메라에 깊이 센서를 장착하여 RGB 텍스처와 깊이 정보를 동시에 획득할 수 있는 depth 카메라를 사용하였다<sup>11)</sup>. 또한 획득된 깊이 정보를 이용하여 두 개 이상의 시점에 대해 영상을 생성할 수 있는 기법과 좌 우 영상으로부터 중간 시점에 해당하는 영상을 인위적으로 생성하는 기법을 사용하였다

본 연구에서는 관찰자 시점에 따라서 총11개 시점에 해당하는 스테레오 영상을 생성하고 생성된 영상 중 관찰자의 시점과 일치되는 영상을 디스플레이 하기 위하여 사용자 인터페이스로서 얼굴추적 기법을 새로이 제안하였다 제안된 기법에서는 현재 프레임 내에 관찰자의 위치를 검출한 후, 검출된 시점에 따라 3차원 물체 및 공간에 대한 다양한 시점을 화면상으로 정의하고, 시점 변화에 따라 해당되는 입체 영상을 디스플레이 장치(stereoscopic display)를 통하여 표현한다. 따라서 관찰자는 시점이 고정된 입체 디스플레이를 통해 임의 시점에 해당하는 스테레오 영상을 볼 수 있게 된다.

본 논문에서는 얼굴 추적을 위해서 색상 정보 기반의 방법을 사용한다 색상 정보를 이용한 얼굴 검출 방법은 얼굴의 특징을 빠르고 쉽게 얻을 수 있고, 다른 방법에 비해 계산량을 줄일 수 있어 실시간으로 구현될 수 있는 장점을 가지고 있다 그러나 색상 정보를 이용한 방법은 조명과 같은 외부 환경에 민감하여 영향을 받으며, 얼굴과 유사한 색상의 분포를 모두 검출해 낸다는 단점을 가지고 있다 따라서 본 연구에서는 실시간의 전체 조건을 만족하면서 강인성을 좀 더 보강할 수 있는 시스템과 기법의 구현에 연구의 목표를 두었다

본 논문의 2장에서는 전체적인 시스템 개요를 소개하고, 3장에서는 얼굴의 검출 과정 및 추적을 위한 기법을 기술하며, 4장에서는 제안한 시스템 및

모의실험 결과를 분석하였다. 5장에서는 본 논문의 결론과 앞으로의 연구방향을 언급하였다

## II. 시스템 개요

일반적으로 입체 영상 디스플레이에서 가장 이상적인 방식은 홀로그래프이다<sup>2)</sup>. 홀로그래프 방식은 동시에 다수의 관찰자에게 입체감을 줄 수 있을 뿐만 아니라 눈에 피로감을 주지 않는다. 그러나 현재의 기술로는 실시간 홀로그래프 시스템을 구현하는 것은 불가능하다. 현재 가장 많이 쓰이는 실시간 3차원 디스플레이 시스템은 스테레오 시스템과 다시점 시스템이다. 스테레오 시스템은 두 대의 카메라로 획득한 좌, 우 영상을 한 장의 입체영상으로 합성하여 디스플레이 한다. 이 때 좌, 우 눈의 양안시차에 의해 깊이감을 느끼게 된다. 그러나 관찰자의 위치가 변화한다면 그림 1과 같이 디스플레이 되는 영상은 변화된 시점에 해당하는 입체 영상이 아니고 변화되기 전 시점에 해당하는 입체영상이 디스플레이 된다. 즉 움직임 시차에 의한 입체감을 제공하지 못하는 것이다.

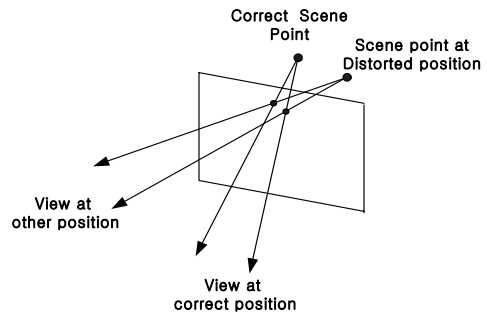


그림 1. 스테레오 시스템에서 기하학적인 오류

반면에 다시점 영상 디스플레이는 관찰자의 시점에 해당하는 입체 영상을 디스플레이 한다. 그래서 다시점 시스템은 스테레오 시스템의 깊이감 뿐만 아니라 관찰자 이동에 따른 움직임 시차까지 제공한다. 다시점 영상을 디스플레이 하는 방법에는 두 가지가 있다. 첫 번째는 다시점 영상을 동시에 디스플레이 하는 것이다<sup>3)</sup>. 그러나 이 방법은 움직임의 제약뿐만 아니라, 움직임 시차도 연속적이지 않다. 두 번째 방법은 관찰자 시점에 해당하는 한 쌍의 스테레오 영상을 디스플레이 하는 방법이다 우선 두 대의 카메라를 이용하여 스테레오 영상을 획득

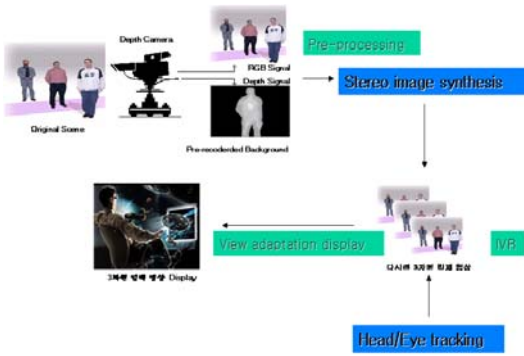
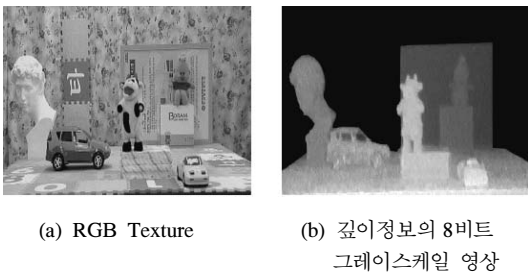


그림 2. 제안된 다시점 디스플레이 시스템



(a) RGB Texture (b) 깊이정보의 8비트 그레이스케일 영상

그림 3. Depth 카메라의 출력

한다. 획득 된 영상들을 전송한 후에 한 쌍의 스테레오 영상을 가지고 기하학적인 변환에 의해서 다시점 스테레오 영상을 생성하게 된다 다시점 디스플레이 시스템에서는 생성된 다시점 영상 중 관찰자 시점에 해당하는 입체 영상을 디스플레이 한다

그림 2는 본 논문에서 제안하는 다시점 영상 디스플레이 시스템의 개요도를 보여주고 있다 이와 유사한 연구들이 최근 유럽의 PANORAMA 프로젝트에서 활발히 연구되고 있다<sup>[4]</sup>. 이 시스템에서는 먼저, depth 카메라를 이용하여 그림 3에 보인 것과 같은 RGB 텍스처 영상과 깊이 정보를 가지고 있는 8비트 그레이스케일 영상을 획득한다. 획득된 영상으로부터 스테레오 영상을 먼저 생성한다 이때 획득된 RGB 텍스처 영상을 좌 영상으로 정의하고 대응되는 우 영상을 생성하기 위해 깊이 정보를 시차 정보로 변환하여 변환된 시차 정보를 가지고 우 영상을 생성하게 된다 또한 회전(rotation)과 이동(translation)등 기하학적인 변환을 통하여 수평 방향으로 11개 시점에 해당하는 다시점 스테레오 영상을 생성하게 된다 그리고 최종적으로 생성된 다시점 스테레오 영상 중에서 관찰자 시점에 해당하는 입체 영상을 디스플레이 하기 위해서 본 논문에서 제안하는 얼굴 추적 기법을 이용하였다

다시점 시스템에서 영상들은 실제 카메라로부터 획득된 영상을 기초한 가상의 카메라로부터 영상을 획득하게 된다. 이러한 가상의 카메라로부터 정확한 영상을 획득하기 위해서는 기하학적으로 정확한 영상의 생성 기법이 필요하다 RGB 텍스처와 깊이 정보를 이용하여 다시점 스테레오 영상을 생성하는 기법은 [5]에서 제안한 방법을 이용하였다 본 논문에서는 [5]의 방법으로 다시점 스테레오 영상을 생성한 후 관찰자의 시점에 해당하는 입체영상을 디스플레이 하기 위해 관찰자의 얼굴을 추적 하는 기법을 제안한다.

일반적인 얼굴 추적 기법은 주변 환경이나 조명과 같은 외부 환경에 많은 영향을 받는다 이러한 외부 환경에 의한 오류는 입체 영상 디스플레이 부분에 잘못된 시점 전달하게 되며 오류에 의해서 잘못된 입체 영상이 디스플레이 된다 따라서 본 연구에서는 이러한 오류를 최소화하기 위해서 환경에 강인하고 실시간으로 정확한 얼굴 추적 기법을 제안한다.

### III. 얼굴 추적 시스템

실시간 얼굴 추적 시스템은 HCI(human computer interaction)에 많은 응용 분야를 가지고 있다 여기서 얼굴 추적은 단순히 컴퓨터의 주의를 사용자의 얼굴에 맞추게 하는 것 외에 사용자의 얼굴을 인식하기 위한 전제 조건으로서 분할(segmentation)을 수행하기 위해서도 연구 되어지고 있다 이는 인식의 정확도를 높이기 위해서 필수적인 요소기술이다. 최근의 얼굴 추적 기법의 동향을 살펴보면 Wang과 Brandstein<sup>[6]</sup>은 머리의 경계선은 타원형이고 2계 도함수가 목 부분에서 가장 크다는 가정 하에 머리의 움직임에서 기인한 프레임 차(Frame Difference)에 기초한 얼굴 추적 기법을 제안하였다 그러나 카메라의 운동 등에 의한 배경의 움직임에 의해서 얼굴 영역이 정확히 분리되지 않을 수 있고 얼굴이 타원형이라는 가정 또한 부분적으로 가려지거나 모자를 착용하는 등의 상황에 의해서 깨질 수 있다. Hager와 Belbumeur<sup>[7]</sup>는 기저 영상(basis image)들을 이용해서 여러 가지 다른 조명 상황을 모델링 하여 상관 계수법에 의한 추적 기법을 제안하였다. 그러나 이 기법은 추적을 위해서 초기 위치를 설정해 주어야 한다는 단점을 가지고 있고 기저 영상이 깨끗한 정면 얼굴에 대해서만 만들어져 있으므로 부분적으로 가려지거나 얼굴 방향이 바뀌는

것에 대해 민감하다. McKenna 와 Gong<sup>[8]</sup>은 움직임이 있는 영역을 검출하여 정면 얼굴을 찾아내는 방법을 사용하였는데 이 또한 카메라의 움직임 등에 의한 배경 영역의 움직임과 얼굴이 부분적으로 가려지는 것의 영향을 많이 받는다. Quian<sup>[9]</sup>는 색상을 이용하여 속도 높이기 위하여 x, y 축으로 각각 투사된 색상의 함수 값을 이용하여 2차원의 문제를 1차원으로 낮추어서 확률적인 접근 방법으로 얼굴을 추적하는 기법을 제안하였다. 그러나 이는 문제를 너무 단순화하였기 때문에 제약 조건이 많고 환경 변화에 대한 색상 모델의 적응성에 관하여 언급이 없다. Yang 과 Waibel<sup>[10]</sup>은 피부색 여과 (skin color filtering) 방법에 기초한 얼굴 추적 기법을 제안하였다. 여기서는 얼굴 색상 모델의 적응 (adaptation)에 관한 방법을 소개하였는데 그 방법이 단순한 누적 통계 수치에 의한 것이었으므로 순간적인 조명 변화에 대한 해결책이 되지 못하였다.

위의 사례에서 알 수 있듯이 실시간 얼굴 추적은 실시간의 조건을 만족 시키다 보면 적용범위가 좁아진다거나 강인성이 보장되지 않는 등의 문제점이 있다. 따라서 본 연구에서는 실시간의 전제 조건을 만족시키면서도 강인성을 더 보장할 수 있는 시스템의 구현에 목표를 두었다. 영상으로부터 가장 빠르고 쉽게 얻을 수 있는 정보가 바로 색상 정보이며, 사물을 판단함에 있어서 가장 효율적이면서 컴퓨터의 계산량을 줄일 수 있다는 장점을 갖고 있기 때문에 얼굴영역 검출에 많이 이용된다. 그러나 색상 정보를 사용하는데 있어서 조명이나 주변 환경 변화에 의해서 야기되는 문제가 있다. 이러한 문제를 해결하기 위해서 얼굴 영역 추출 후에 얼굴의 특징점 추출 및 얼굴의 기하학적인 특성을 이용하여 정확한 얼굴 영역을 검증하게 된다. 검증된 얼굴 영역에 대하여 통계적 색상 모델 및 변형적 형판을 이용하여 얼굴을 추적한다.

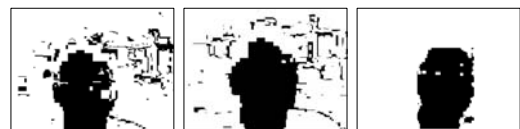
### 3.1 색상 정보에 의한 얼굴 영역 추출 기법

디스플레이 장치에 장착된 관찰자 얼굴 획득용 카메라로부터 얻는 입력영상은 색상정보를 포함한 320×240 크기의 컬러 영상이다. 입력 영상을 RGB 형태에서 YCbCr의 형태로 색 공간을 변환한 후 색상 정보를 이용한 얼굴 영역 추출 기법을 적용한다. 먼저 입력 영상의 Y, Cb, Cr 정보 중에서 Cb, Cr 성분에 대해 색상 영역 분할(color segmentation) 과정을 거쳐 살색 값을 갖는 영역만을 추출해 낸다. 살색 정보를 이용하기 위해서 영상에서 살색 부분

만을 뽑아내어 살색이 차지하는 Cb, Cr 값의 범위를 히스토그램으로 구한다. 이러한 히스토그램을 여러 개의 영상에 적용하여 살색이 차지하는 범위를 통계적으로 결정할 수 있다. 본 논문에서 구한 Cb, Cr의 선택 범위는 각각 다음 식과 같다.

$$f(x,y) = \begin{cases} 0 & \text{if } (137 < Cb < 152) \text{I } (123 < Cr < 137) \\ 255 & \text{otherwise} \end{cases} \quad (1)$$

색상 영역 분할과정을 거친 입력 영상의 결과 영상을 그림 4(a)에 나타내었다. 그림 4(a)의 영상을 보면 세밀한 잡음 형태가 존재하는 것을 볼 수 있다. 이런 잡음을 제거하기 위해 먼저 형태학적 필터 (morphological filter) 중에서 열림(opening) 연산을 수행하였다<sup>[11]</sup>. 그림 4(a)의 영상에 열림 연산자를 적용한 결과 영상이 그림 4(b) 영상이다. 다음 과정으로 수평 스캔을 하게 되는데 수평 방향으로 0의 값을 갖는 화소의 수를 세고, 그 결과가 임계값 미만의 값을 갖는 영역은 모두 255로 화소 값을 조정한다. 여기서 임계값은 얼굴 영역의 수평크기가 영상 전체의 절반정도 크기라는 가정 하에서 최대 값의 절반으로 하였다. 이러한 가정을 할 수 있는 이유는 디스플레이 장치 앞에 관찰자의 위치가 대부분 어느 정도는 정해지기 때문이다. 수평 방향으로의 스캔 작업이 끝나면 같은 방법으로 수직 방향으로 스캔 한다. 스캔 작업이 모두 끝나면 그림 4(c)의 영상을 얻을 수 있다.



(a)분할된 원영상 (b) 잡음제거 (c) 얼굴 영역  
그림 4. 얼굴 영역의 분리

### 3.2 얼굴 특징점의 검출

얼굴 영역 주변의 다른 후보 살색 영역을 제거하기 위해서, 검출된 후보 살색 영역이 얼굴 영역인지 아닌지에 대한 판단을 한다. 얼굴 영역의 확인 과정을 위해서 얼굴의 특징점 검출을 한다. 일반적인 얼굴영역 추출을 위한 환경과 본 연구에서 가정한 환경이 다른 점은 얼굴을 추출하려는 임의의 관찰자가 검은색 3D 편광 안경을 항상 착용한다는 것이다. 그러므로 눈을 특징점으로 하는 기존의 일반적인 얼굴 영역 추출 및 특징점 추출 기법으로는 정확한 얼굴 영역 판단이 어렵다. 이 경우 잘못된 시

점의 전달로 인해 재구성 된 영상이 잘못 표현될 수 있다. 본 연구에서는 이러한 제약 조건을 극복하기 위해, 얼굴의 기하학적인 특징 및 형판을 통해 특징점을 검출하였다. 먼저 정확한 얼굴 영역 추출 및 특징점 검출을 위해 추출하는 대상, 즉 관찰자의 얼굴 크기, 위치에 대한 정보를 저장한다 이러한 주요 관찰자에 대한 정보는 관찰자 이외의 다른 얼굴 영역이나, 손 같은 살색 영역 같은 후보 영역들이 존재할 경우, 주요 관찰자 얼굴 영역만을 추출하기 위한 정보로 이용된다 얼굴 특징점 검출을 위해서 형판(template)이 이용되었다 본 연구에서는 편광 안경을 착용하므로 눈을 검출할 수 없으므로 그림 5에 보인 입을 위한 형판을 설계하였다

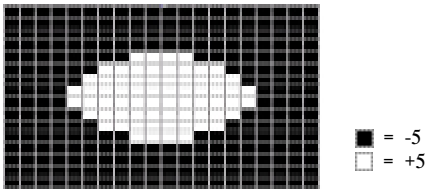


그림 5. 제안된 입 형판 50×16

입의 검출척도로 사용하기 위해 matching value (MV)를 제안하였다. MV는 0~255 사이의 흑백 영상과 -5와 5의 값을 갖는 입 형판과의 일종의 상관도(correlation) 값으로 볼 수 있다.

$$MV(m,n) = \frac{1}{X_i \times Y_i} \sum_{x=0}^{X_i-1} \sum_{y=0}^{Y_i-1} (E[x][y] \times I[x+(m-8)][y+(n-10)])$$

$$8 \leq m < X_i - 8, \quad 25 \leq n < Y_i - 25$$

(2)

여기서  $E[x][y]$ 는 사용된 입의 형판  $X_i$ 와  $Y_i$ 는 입 형판의 수직 및 수평크기를 각각 나타내며  $I[x][y]$ 는 입력 영상의 얼굴 영역을,  $X_i$ 와  $Y_i$ 는 입력 영상에서 검출된 얼굴 영역의 수직 및 수평 크기를 각각 의미한다. MV의 최대 값은 얼굴 영역에 형판과 일치하는 패턴이 있을 때 발생하게 된다 입이 검출되면 검출된 입의 위치를 기준으로 얼굴 구조에 맞게 편광 안경의 위치를 추정한다 그리고 추정된 편광 안경의 위치를 설정한 후 편광 안경 검출 판별 여부는 편광 안경 영역내의 분산 값과 화소 값으로 결정하는데, 그 이유는 편광 안경을 쓰게 되면 상대적으로 쓰지 않았을 때보다 화소 값들의 변화가 심하지 않기 때문에 분산 값이 작게 되기 때문이다. 편광 안경이 검은색이므로 화소 값이 다른

영역에 비해 작다. 분산의 임계 값을 실험적으로 최적화된 값을 선택하였다 본 시스템에서 얼굴의 특징적인 입과 편광안경을 추출한 후에는 추출된 입과 편광안경의 위치 및 거리를 측정한다 측정된 값이 얼굴의 기하학적인 특성과 비교하여 오차가 크면 얼굴 영역이 아닌 후보 살색 영역으로 검증한다. 이러한 과정을 반복적으로 수행함으로써 손이나 다른 사람 얼굴 같은 후보 살색 영역이 아닌 정확한 관찰자 얼굴을 추출할 수 있게 된다

### 3.3 얼굴 추적

얼굴 영역을 추출한 후, 추출된 얼굴 영역을 주요 관찰자로 정의한다 그 다음 주요관찰자에 대한 얼굴 추적이 이루어진다 본 논문에서는 얼굴 추적을 위해서 색상 기반의 통계적 모델 및 변형적 형판(template)을 이용하는 방법을 사용한다<sup>[2]</sup>. 얼굴 영역은 통계학적으로 얼굴 영역을 제외한 배경 영역과 다른 특징을 보인다. 얼굴 영역의 주요 색상은 살색과 머리카락 색이며, 배경영역은 기타 여러 가지 색상으로 이루어져 있다. 이러한 얼굴 영역에 대한 분산과 편차를 구한 후, 현재 프레임의 어느 특정한 위치의 화소에 대한 분산과 편차를 구하면서 얼굴 영역과 배경으로 분리한다. 얼굴의 두 가지 주요 색상인 살색과 얼굴의 머리 색상에 대한 통계적 모델을 나타내면 다음 식과 같다

$$p(x | w_i) = \alpha_{i1} N(u_{i1}, C_{i1}) + \alpha_{i2} N(u_{i2}, C_{i2}) \quad (3)$$

여기서  $p(x | w_i)$ 는  $x$ 가  $w_i$  영역에 속해 있을 확률 밀도 함수 값이다  $x$ 는 현재 프레임의 화소값이고,  $w_i$ 는  $i=1$  이면 전경인 얼굴 영역이고,  $i=2$  이면 배경인 비 얼굴 영역이다  $N(u, C)$ 는 가우시안 함수이고  $u$ 와  $C$ 는 각각 평균값과 분산 값을 나타내며,  $\alpha_{i1}$ ,  $\alpha_{i2}$ 는 가우시안 함수의 가중치이다. 이런  $p(x | w_i)$  함수 값은 특정한 화소 값이 어디에 존재하는지를 반영한다.

또한 변형적인 형판을 사용하여 최대한 얼굴 영역의 화소들을 많이 포함 시키는 영역을 찾아가게 된다. 즉 다음 식 (4)에서 에너지 값이 최소가 될 때 형판의 위치 값이 얼굴의 위치가 된다

$$f(\mathcal{R}) = \sum_{r \in \mathcal{R}} \log \left( \frac{p(x_r | w_2)}{p(x_r | w_1)} \right) \quad (4)$$

여기서  $r$ 은 형판 지역  $\mathcal{R}$ 의 화소이고,  $x_r$ 은 화소값이다. 얼굴영역을 찾기 위해서  $f(\mathcal{R})$ 의 에너지가 최소화 되는 곳의 변형적 형판 영역의 크기를 조절하면서 탐색한다. 변형적 형판은 프레임 기반 영상 분석 방법으로 좋은 성능을 가지고 있다

#### IV. 시스템의 구현 및 실험 결과

##### 4.1 시스템 구현

구현된 시스템은 주요 관찰자의 얼굴에 따른 위치를 인식한 후, 주요 관찰자 위치에 대응하는 스테레오 영상을 디스플레이 하는 시스템이다. 그림 6은 본 논문에서 구현한 디스플레이 시스템을 보여주고 있다. Depth 카메라로부터 획득된 텍스처 영상 한 장과 깊이 정보를 갖고 있는 8 비트 영상 한 장을 입력 받은 후, 디스플레이 장치에 장착된 카메라로부터 실시간으로 획득된 관찰자 영상에 대한 얼굴 추적 및 얼굴 추적 영상을 디스플레이 한다.

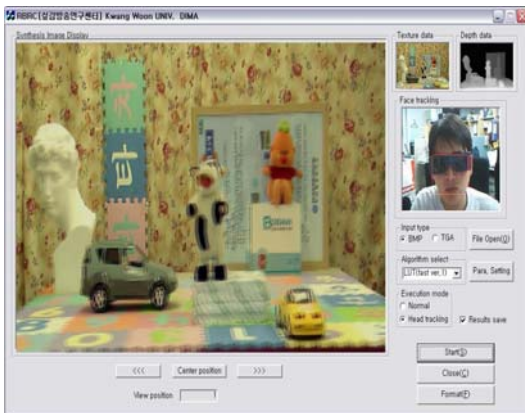


그림 6. 구현 프로그램

임의 시점에 해당하는 3차원 입체 영상을 디스플레이 하기 위해서 관찰자의 시점을 파악해야 하며 이를 위해서 얼굴 영역 추출 및 추적 두 단계를 거친다. 얼굴 영역 추출은 처음 입력되는 약 20 프레임의 영상에서 이루어진다. 정확한 얼굴 영역 추출을 위해서 얼굴의 특징점 검출 및 기하학적인 특징을 이용하여 후보 살색 영역을 제거한다. 정확한 얼굴 영역을 추출한 후, 추출된 얼굴 영역의 색상 모델 및 변형적 형판을 이용하여 얼굴 추적을 한다. 각 프레임별로 최적의 얼굴 영역을 찾은 후, 얼굴 영역의 현재 프레임에서의 위치값(view point)을 다음 식 (5)에 의해서 생성하게 된다

$$View\ Point = (VP_{cur} - VP_{first}) / (VP_{total} / p) \quad (5)$$

여기서  $VP_{first}$ 는 처음 얼굴 영역의 위치를 나타낸다. 이는 영상 입력 후 20 프레임 동안의 관찰자의 기준 위치가 된다.  $VP_{cur}$ 는 현재 얼굴의 위치를 나타내며,  $VP_{total}$ 은 입력 영상의 수평 크기를 나타낸다. 구현된 시스템에서 관찰자의 11개 시점(=  $p$ ) 변화에 따른 다시점 영상을 디스플레이 하기 위해서, 디스플레이 장치의 카메라로부터 획득된 입력 영상을 총 11개 영역으로 나누었다. 관찰자가 중앙에 위치하게 되면  $\theta=0$ 으로 정의하고 가장 왼쪽에 위치할 때가  $\theta=-5$ , 가장 오른쪽에 위치할 때가  $\theta=5$ 로 하며 그 사이에 모두 11개의 view가 있다고 정의하였다. 그림 7과 같이 특정 영역에 관찰자 시점(view point)이 있으면, 다시점 영상 생성부에  $\theta$  값을 전달하여 3D 스테레오 영상을 디스플레이 하게 된다. 3D 스테레오 영상을 보기 위해서는 그림 8에 보인 편광안경과 3D 안경식 모니터를 이용할 수 있다.

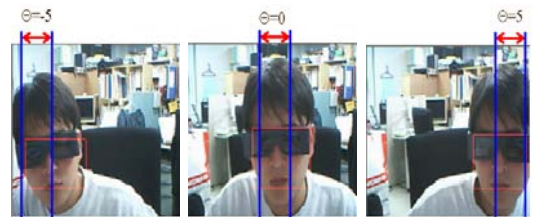


그림 7. 얼굴 위치에 따른 시점



그림 8. (a) 편광안경 (b) 3D 안경식 모니터

##### 4.2 실험 및 결과

본 연구에 사용된 실험 환경은 표 1과 같다. 디스플레이 장치에 장착된 카메라로부터 크기 320×240 영상을 입력받아 제안된 얼굴 추적 기법을 통해 실시간으로 관찰자의 얼굴 위치를 파악한다. 또한 관찰자의 시점에 해당하는 3차원 영상을 720×480 크기의 RGB 텍스처 영상과 깊이 정보를 이용하여 적절한 임의시점 영상을 재구성하게 된다

그림 9와 10은 본 논문에서 제안한 기법을 가지고 실험한 결과를 보였다. 그림 9(a)는 좌, 우 이동에 따른 얼굴의 정확한 추적결과를 나타내고, 그림 9(b)는 손이나 주변 살색 영역이 있어도 관찰자 얼굴을 제외한 다른 살색에는 영향을 받지 않는 것을 보여주는 결과영상이다. 그림 9(c)의 경우는 다른 얼굴이 영상에 나타났을 경우, 주요 관찰자만을 추적하는 결과를 보여준다.

표 1. 본 연구에 사용된 실험 환경

OS	Window XP
PC 사양	P-4 2.8 GHz
PC camera	센서 : 1/3 CMOS 화소수 : 640×480 전송속도 : 15 frame/sec
구현 수단	Visual C++ 6.0
실험영상	320×240 size, RGB 24bit , 10 frame/sec



그림 9. (a)~(c) 상황에 따른 얼굴 추적 결과(사각형 안)

그림 10은 관찰자의 얼굴을 좌, 우로 이동하였을 경우, 관찰자 각 시점(View Point)에 해당되는 입체 영상이 디스플레이 되는 것을 나타낸다. 그림 10의 결과는 그림 8에 보인 3D 모니터와 편광 안경을 이용하여 확인하였다. 실험 결과 그림 9와 그림 10의 결과와 같이 입체의 시점에 해당하는 입체영상을 디스플레이 하기 위한 인터페이스에서 실시간으로 정확하게 관찰자의 얼굴을 성공적으로 추적하였다

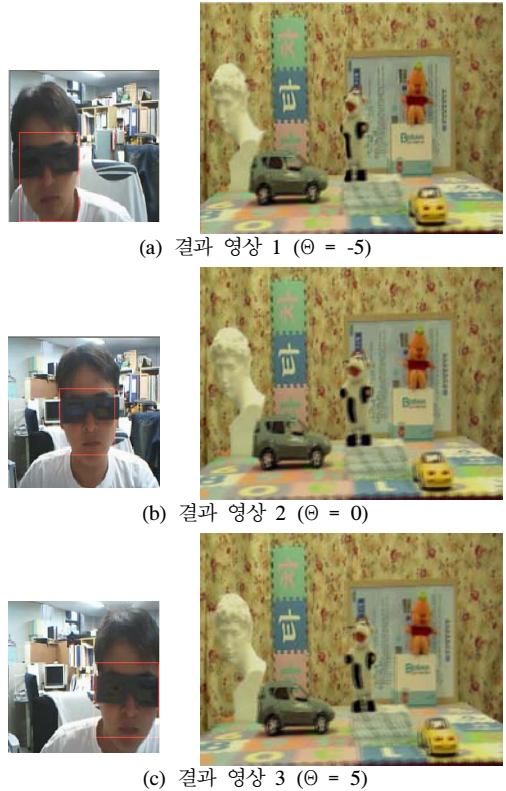


그림 10. (a)~(c) 시점 변화에 따른 다시점 영상 디스플레이

### V. 결론

본 논문에서는 다시점 영상을 디스플레이하기 위하여 여러 대의 카메라를 사용하지 않고 depth 카메라를 이용하여 한 장의 텍스처 영상과 그 텍스처 영상의 깊이 정보 영상을 가지고 총 11 시점의 스테레오 영상을 실시간으로 획득하였다. 획득된 영상을 관찰자 시점에 따라 디스플레이 하기 위하여 디스플레이에 장착된 카메라로부터 획득된 입력 영상으로부터 얼굴 영역을 검출한 후, 검출한 얼굴 영역을 실시간으로 정확하게 추적하는 기법을 새로이 제안하였다. 일반적인 사무실 환경에서는 사용자의 얼굴 뒤에 복잡한 배경과 다양한 외부 빛 등의 영향이 있다. 이로 인해 색상 모델 방법만으로 실시간으로 정확하고 완벽하게 얼굴을 추적하는 것은 불가능한 것으로 알려져 있다. 게다가 현재 얼굴 특징 추출 및 얼굴 인식에 있어서 성능이 우수한 것으로 알려진 Identix사의 FaceIt™ 역시, 인종 및 조명별로 얼굴 특징 추출 성능이 영향을 많이 받는 것으로 조사되고 있다. 이와 같이 색상 모델에 의한 한계를 모두 극복하는 것은 매우 어려운 일이지만 제

안한 기법은 다시점 영상 디스플레이 시스템에서의 임의 시점에 해당하는 영상을 실시간으로 디스플레이 하기 위한 인터페이스에서 좋은 성능을 보였다는 점에서 의미가 있다고 하겠다 향후 여러 가지 외부환경에 더 강인하며, 실시간으로 구현 할 수 있는 새로운 기법들이 연구 개발되어야 할 것이다

참 고 문 헌

[1] ATTEST, European IST Project <http://www.iti.gr/db.php/en/projects/ATTEST.html> 2002-2004

[2] I. Sexton and P. Surman, "Stereoscopic and autostereoscopic display systems," IEEE Signal Processing Mag., vol. 16, pp. 85-99, 1999.

[3] Heinrich-Hertz-Institut, Human Factors Dept., Berlin, Germany. (1999). 3D-displays. [Online]. Available: [http://atwww.hhi.de/3D\\_displays](http://atwww.hhi.de/3D_displays)

[4] J. R. Ohm, K. Grueneberg, E. Hendriks, M. E. Izquierdo, D. Kalivas, M. Karl, D. Papadimitos, and A. Redert, "A realtime hardware system for stereoscopic video-conferencing with viewpoint adaptation," Signal Processing: Image Commun., vol. 14, pp. 147-171, 1998.

[5] Jin Woo Bae, Hyok Song and Ji Sang Yoo, "An object based multi-viewpoint image synthesis using depth camera", ITC-CSCC 2004, Sendai, Japen, July. 5-8, 2004.

[6] C. Wang and M. S. Brandstein "A Hubrid Real-Time Face Tracking System", Proc. Acoustics, Speech, and Signal Processing Volume 6, 1998

[7] G. D. Hager and P. N. Belhumeur, "Real-time tracking of image regions with changes in geometry and illumination," Proc. Computer Vision and Pattern Recognition, pp.403-410, 1996.

[8] S. J. Mckenna and S. Gong, "Tracking Faces", Proc. Inernational conference on Face and Gesture Recongnition. pp.271-276 1996

[9] R. J. Qian, M. I. Sezan, K. E. Matthews, "A robust Real-time Face Tracking Algori-

thm", Proc International Conference on Image Processing, Vol. 1, page(S):131-135, 1998

[10] Jei Yang, Alex Waibel, " A real-time Face Tracker", Proceeding of WACV, pp142-147, 1996.

[11] J. Serra and P. Soille(eds.), Mathematical Morphology Its Applications to Image Processing, pp. 69-76, Kluwer Academic Publishers in the Netherlands, 1994.

[12] Rao, R.R.; Mersereau, R.M.," On merging hidden markov models with deformable templates" Image Processing, 1995. Proceedings., International Conference on , Volume: 3 , 23-26 Oct. 1995 Pages:556 - 559 vol.3

한 충 신(Chung-shin Han)

준회원



2004년 2월 광운대학교 전자공학과 졸업  
2004년 3월~현재 광운대학교 전자공학과 석사과정  
<관심분야> 영상인식, 3차원 영상처리, 동영상 코덱

장 세 훈(Se-hoon Jang)

준회원



2004년 2월 광운대학교 전자공학과 졸업  
2004년 3월~현재 광운대학교 전자공학과 석사과정  
<관심분야> 3차원 영상처리, 동영상 코덱

배 진 우(Jin-woo Bae)

준회원



1998년 2월 순천대학교 전자공학과 졸업  
2003년 2월 광운대학교 전자공학과 석사  
2003년 3월~현재 광운대학교 전자공학과 박사과정  
<관심분야> 웨이블릿 기반 영상처리, 동영상 코덱, 영상통신



유 지 상(Ji-sang Yoo)

정회원



1985년 2월 서울대학교 전자  
공학과 졸업(공학사)

1987년 2월 서울대학교 대학  
원 전자공학과 졸업(공학석  
사)

1993년 5월 Purdue 대학교  
전기공학과 졸업(Ph.D.)

1993년 9월~1994년 8월 현대전자산업(주) 산전연  
구소 선임연구원

1994년 9월~1997년 8월 한림대학교 전자공학과  
조교수

1997년 9월~2001년 8월 광운대학교 전자공학과  
조교수

2001년 9월~현재 광운대학교 전자공학과 부교수

<관심분야> 웨이블릿 기반 영상처리, 영상압축, 영  
상인식, 비선형 신호처리