

3D영상 객체인식을 통한 얼굴검출 파라미터 측정기술에 대한 연구

최병관*, 문남미*

Object Recognition Face Detection With 3D Imaging Parameters A Research on Measurement Technology

Choi, Byung-Kwan*, Moon, Nam-Mee*

요약

본 논문에서는 첨단 IT융,복합기술의 발달로 특수 기술로만 여겨졌던 영상객체인식 기술분야가 스마트-폰 기술의 발전과 더불어 개인 휴대용 단말기로 발전하고 있다. 3D기반의 얼굴인식 검출기술은 객체인식 기술을 통하여 지능형 영상검출 인식기술로 진화되고 있음에 따라 영상인식을 통한 얼굴검출기술과 더불어 개발속도가 급속히 발전하고 있다. 본 논문에서는 휴먼인식기술을 기반으로 한 얼굴객체인식 영상검출을 통한 얼굴인식처리 기술의 인지 적용기술을 IP카메라에 적용하여 인가자의 입,출입등의 식별능력을 적용한 휴먼인식을 적용한 얼굴측정 기술에 대한 연구방안을 제안한다. 연구방안은 1)얼굴모델 기반의 얼굴 추적기술을 개발 적용하였고 2)개발된 알고리즘을 통하여 PC기반의 휴먼인식 측정 연구를 통한 기본적인 파라미터 값을 CPU부하에도 얼굴 추적이 가능하며 3)양안의 거리 및 응시각도를 실시간으로 추적할 수 있는 효과를 입증하였다.

▶ Keyword : 3D영상객체인식, 영상검출, 휴먼인식

Abstract

In this paper, high-tech IT Convergence, to the development of complex technology, special technology, video object recognition technology was considered only as a smart - phone technology with the development of personal portable terminal has been developed crossroads. Technology-based detection of 3D face recognition technology that recognizes objects detected through the intelligent video recognition technology has been evolving technologies based on image

• 제1저자 : 최병관 • 교신저자 : 문남미
• 투고일 : 2011. 06. 07, 심사일 : 2011. 07. 25, 게재확정일 : 2011. 08. 16.
* 호서대학교 벤처전문대학원(Hoseo Graduate School of Venture)
* 호서대학교 벤처전문대학원(Hoseo Graduate School of Venture)

recognition, face detection technology with through the development speed is booming. In this paper, based on human face recognition technology to detect the object recognition image processing technology is applied through the face recognition technology applied to the IP camera is the party of the mouth, and allowed the ability to identify and apply the human face recognition, measurement techniques applied research is suggested. Study plan: 1) face model based face tracking technology was developed and applied 2) algorithm developed by PC-based measurement of human perception through the CPU load in the face value of their basic parameters can be tracked, and 3) bilateral distance and the angle of gaze can be tracked in real time, proved effective.

▶ Keyword : 3D image object recognition, image detection, and human perception

1. 서론

3D기반의 얼굴인식기술(3D Face recognition technology)이란? 사람의 얼굴 특징점을 추출하고 얼굴검출을 통한 인지 식별 추출데이터를 통하여 얼굴식별 인가자의 신상정보와 데이터 그리고 등록된 값에 대한 설정, 편집, 삭제등 각각의 얼굴 검출기술에 필요한 적용할 알고리즘을 통하여 영상 감시, 검색, 저장을 하고 검출된 시각차에 의해 얻어진 데이터 이미지로부터 생성한 시각차에 따른 차이를 두고 보여줌으로써 사람이 마치 영상이 제작되고 있는 장소에 있는 것과 같은 생동감과 현실감을 느낄 수 있게 만드는 시각적 기술을 3D 입체 영상 기술이라고 한다.

3D입체 영상기술은 3D인지를 비롯하여 정보통신, 방송, 의료, 영화, 게임, 애니메이션 등과 같은 기존의 모든 산업제품 개발에 광범위하게 응용되는 핵심기술로 자리잡을 것이며, 최근에는 3D 정보기술이 기존 산업과 융합되면서 새로운 패러다임을 창출하는 신(新)성장 동력산업으로 급부상하고 있다.

3D입체 영상 인식기술을 적용하여 양안시차 원리를 통한 사람의 인지/정보 체계에서 3D영상인식 기법을 적용한 인식 기술을 인지적용 기술이라 한다. 좌/우의 2D 영상은 각도 차이를 갖는 카메라 리그를 이용하여 획득 및 제작되며, 디스플레이에서 좌안/우안용 영상을 좌/우에 분리시키는 장치(특수 안경, 무안경식 모니터)를 이용하여 두 눈에 전달하게 된다. 각각의 눈에 입력된 영상은 뇌에 전달되어 사람의 인지/정보 체계에서 3D입체로 인식하게 된다. 좌/우 영상이 바뀌어 전달되거나 두 영상 사이의 깊이감이 매끄럽지 못할 경우 입체감 인식이 떨어지거나, 시각적 불편함이 발생할 수 있다.

여러가지 기술적인 문제가 있겠지만 가장 중요한 문제는 휴먼인식의 문제점이 분명하고 3D 휴먼인식의 핵심은 시각적

피로이다. 따라서 사람이 보기 쉽고, 사용하기 쉬운 3D 방송 시스템의 개발을 위해서는 휴먼인식기술에 대한 연구가 전제되어야 한다. 본 논문에서는 3DTV 방송 기술의 활성화에 필수적인 휴먼인식 기술의 기술 요소 및 동향에 대해서 살펴보고 휴먼인식측정을 위한 시청자 인식용 휴먼인지 기술 개발에 대해 기술하였다.

본 논문에서 기술하는 3D영상객체인식을 통한 휴먼인지 기술을 적용하여 기존의 고정된 얼굴인식 기반에서만 검출했던 문제점을 보완하고 휴먼인지 기술의 문제점을 보완하고 다양한 휴먼인식 파라미터 측정을 적용한 기술을 적용하여 얼굴 검출 기술에 대한 방법론을 제안한다.

2. 휴먼인식 기술요소

3D입체영상 검출시스템은 실제의 공간에서 깊이와 대상의 크기가 실제대로 지각되고 얼굴 감지를 통한 객체인식을 적용할 수 있어야 한다. 이러한 영상인식 기술을 적용, 실현하기 위해서 휴먼인식에 대한 연구가 시작되었으며, 다음과 같은 3가지 기술 요소들을 연구하고 있다.

첫째, 시스템 설계 단계에서부터 휴먼인식을 고려하는 것이다. 모든 경우에 그런 것은 아니지만 일단 시스템의 하드웨어 개발이 일정 수준의 궤도에 오르고 나면, 추가적으로 휴먼팩터를 고려하기 위해 시스템을 수정하기가 곤란하다.

특히, 고려해야 할 휴먼팩터가 기존 시스템의 구조만으로 해결하기 어려운 근본적인 문제일 때 더욱 그러하다.

둘째, 3D영상인식 기반의 파라미터 검출값을 규명하고 이에 대한 감성공학 기술로서의 움직이는 물체를 인식하고 인출한 데이터를 통하여 인지식별을 할 수 있는 환경에 대한 최적의 영상객체 인지에 대한 연구이다.

셋째, 3D얼굴 검출시스템에 대한 심리학적 평가 기술로서

이상적 시스템을 개발하기 위한 여러 단계를 거치면서 각 단계마다 개발된 시스템의 완성도를 시청자에게 제시하여 개발된 시스템의 심리학적 성능을 효과적으로 평가하기 위한 연구이다.

본 논문의 핵심내용인 휴먼인지 기술을 이해하기 위하여 이들 세 기술 요소에 대하여 좀더 상세하게 알아보기로 한다. 따라서 개발 도중에 수정을 하는 것보다는 설계 단계에서부터 얼굴인식을 통한 휴먼팩터를 측정하거나 고려하여 시스템 기능이나 사양을 정하는 것이 비용절감 효과를 얻을 수 있다.

2.1 3D영상시스템 설계에 필요한 휴먼팩터 연구

3D영상시스템 설계에 휴먼팩터를 제대로 반영하기 위해서는 휴먼팩터 관점에서 문제의 원인을 정확하게 진단하고, 해결방안을 제시해야 할 필요가 있다. 이를 위해서 다음과 같은 세부 연구들이 이루어져야 한다.

<표1>은 3D영상인식시스템의 파라미터들과 권장되는 값 또는 범위를 정리한 것이다

<표1> 휴먼팩터 연구 사례

구 분	내 용
사람의 3D시각 정보처리 시스템 HVS에 대한 연구 (HVS: Human Visual System)	3D 시각정보를 인식하고 인지할 수 있는 깊이 특성 등을 이해함으로써, 시스템의 구조를 사람의 시각 시스템과 유사하게 설계할 수 있다. 이를 위해 변이가 0으로 되는 호랍터(horoptor)에 관한 연구, 교차 변이(crossed disparity) 및 미교차 변이(uncrossed disparity)의 인식범위 등에 대한 연구를 말한다.
3D 깊이 인식을 위한 단서 및 상호 관계에 대한 연구	3D 깊이를 인식하는 과정은 주시하고 있는 물체에 양쪽 눈을 회전시켜 시선(line of sight)을 모으는 주시각 조절과 현재 물체의 3D 위치 지점에 안구의 두께를 조절하여 눈의 초점을 주시하고 있는 물체에 맞추는 초점 조절이 연동되어 이뤄진다. 이 때 초점 심도(depth-of-focus or field: DOF) 범위를 벗어난다는 진경이나 배경에 존재하는 물체들은 블러링(blurring)되어 보여 진다는 초점 및 주시각 조절연동 관계가 알려져 있다.
시각피로의 원인규명 및 해결책에 대한 연구	양안식 3D영상객체인식시스템에서 발생하는 대표적 영상인식검출기술등 초점과주시연동파괴(accommodation-convergence breakdown), 혼신(cross talk) 현상 등이 있다[1]. 초점과 주시연동 파괴 현상은 앞에서 기술한 바와 같이 실제 세계에서 우리가 물체를 볼 때에는 초점과 주시각 조절이 연동되어 피곤함을 느끼지 않고 3차원 깊이를 인식한다. 그러나 기존의 양안식 3D 영상 디스플레이를 통하여 우리가 3D 영상을 보게 될 경우에는, 큰 시차에 의해 초점 및 주시연동 파괴 현상이 발생한다.
양안시차 및 시점 변화에 대한 시청자의 민감도연구	사람이 차이를 감지할 수 있는 해상도(resolution), 양안 시차, 시점 변화(view adaptation)에 대한 연구를 수행함으로써, 사람의 감지 범위를 초과하지 않는 민감도(sensitivity) 범위 내에서 시스템이 개발되어야 한다. 현재까지의 연구 결과에 의하면 좌우 영상의 해상도 차이는 공간적 해상도의 경우 50%의 해상도 차이에 대해 20%의 상위치 오차가 일어나며, 특히 고해상도와 저해상도의 양안 영상을 이용하여 3D 영상을 디스플레이 하였을 때 인식되는 3D 영상의 해상도는 고해상도를 가지는 영상과 거의 같거나 약간 저하되는 정도로 인식된다.

3. D영상인식 환경 파라미터 규명 및 입체영상 인지감성 공학기술

최적의 3D영상인식 환경을 위해서는 검출기술에 대한 환경연구가 필요하다. 이를 위해 적절한 영상객체 환경을 결정하는 파라미터를 정의하고, 이의 변화가 인가지 또는 비인가자에게 미치는 심리적 영향을 파악해야 한다.

3.1.2 적절한 영상환경의 파라미터 연구

얼굴 검출을 통한 3D입체영상 기술은 검출환경에 따라 영향을 주는 적절한 물리적 속성에 따라 검출거리, 영상의 왜곡을 줄여주는 관찰시점, 여러명을 검출하는 10:1의 다인검출 등과 주변의 광도등이 있다. 이러한 물리적속성을 갖는 영상객체 인식기반의 3D영상인식기술은 입,출입 또는 검출자의 인지능력을 손쉽게 정확하게 검출하는 데 그 목적이 있었으나 검출 기술의 깊이 및 모양의 왜곡, 심리적 편안함 등을 정신적, 물리학적으로 측정함으로써 3D영상인식 환경에 적합한 파라미터를 규명할 수 있다.

3.2 영상인식 적용에 따른 다른 얼굴 검출정도의 정량화 연구

현재까지의 기술로 개발된 영상인식시스템은 약간씩의 편차와 고정된 입체각을 통한 검출등 불편할 정도의 시각적 검색기능을 적용하였다. 그러나 이렇게 실제의 3D영상기술을 적용한 검출기술을 적용하게 되면 검색인지에 대한 정확성과 3D얼굴인지 입체기술을 통한 검출을 하게되면 인지정도 및 원인을 측정하고 규명하며, 이의 파라미터 값을 통한 정확한 연구가 이에 속한다.

영상객체 파라미터	권장범위 또는 값
표시화각	60deg ~ 70deg
시청거리	3H
표시화면	시청거리에 의존, 34"50인치 적당
평판효과	양안 간격의 0.6 ~ 1.3배

<표2>는 3D입체인식시스템의 시청환경 파라미터와 권장 범위 및 값의 정리한 것이다.

3.2 3D영상 객체인식시스템 심리학적 평가방법 연구

3D영상시스템을 비롯한 일반적 시스템 개발에 있어 첫번째 시도에 의해 완벽한 시스템을 구현한다는 것은 거의 불가능하다. 따라서, 시스템을 개발하면서 그 때마다 발생하는 문제점을 파악하고, 그 단계에서의 시스템 완성도를 판단할 필요가 있다. 이를 위한 가장 간단한 방법은 시스템을 개발한 공학자의 주관적 판단에 따라 시스템을 평가하는 방법이다.

시스템의 평가라는 것은 공학적인 측면에서 평가할 수도 있지만, 개발된 시스템의 최종수요자인 사람시청자에게 얼마나 친화적인가에 기반하여 평가를 하여야 한다. 이를 위한 것이 3D영상 객체 인식시스템의 평가 방법 연구이며, 이에는 평가항목, 주관적 평가 방법, 객관적 평가 방법 등에 대한 연구가 이에 속한다. 이러한 평가항목과 방법을 사용함으로써 개발할 시스템의 우수성에 대한 객관성을 확보할 수 있을 것이다.

3.2.1 평가항목 개발 연구

시스템을 어떤 측면에서 평가해야 할 것인지는 그 평가항목에 따라 시스템의 우수성 기준이 달라지므로, 객관성을 확보할 수 있도록 결정하는 것이 중요하다. 평가항목이 잘못 결정되면 시스템에 대한 평가는 불충분하게 되고 결국은 시스템이 문제를 지니고 있음에도 이를 발견해 내지 못할 가능성이

있다. 따라서, 이들 평가 방법에 대한 표준화가 필요하다. 이때 주의할 것은 평가항목을 하드웨어적인 측면뿐 아니라 심리적 측면에서도 평가가 가능하도록 선택해야 한다는 것이다. 현재까지의 연구에서 평가항목으로 많이 사용되는 것으로는 인식선명도, 인식깊이, 화질, 자연스러움(naturalness), 시각적 편안함(visual comfort) 등이 있다.

3.2.2 생리학적 평가방법 개발 연구

정신물리학적 측정방법과 심리검사방법 모두 기본적으로 시청자의 주관적인 자기 보고법(self-report)에 기반을 두고 있다. 이 방법엔 시청자가 의도적으로 거짓 보고를 하는 경우를 잡아내기 위한 기법들이 사용되어 정신물리학적 측정방법 및 심리검사방법을 이용하더라도 상당히 신뢰성이 높고 타당한 평가방법을 개발할 수 있다.

생리학적 평가방법의 예로는 뇌파 및 안구운동의 변화를 3차원 영상을 시청할 때와 시청하지 않을 때에 대하여 각각 측정함으로써 시청자가 시각적 피로와 두통을 느낄 때 뇌파와 안구 운동 변화를 알아낼 수 있고, 이를 평가방법으로 활용할 수 있다.

본 논문에서는 얼굴인식을 통한 휴먼팩터 측정 및 연구에 필요한 객체인식기술을 적용한 위치와 거리 및 검출검출을 추적할 수 있는 휴먼인지 기술 개발에 대해 기술하였다.

4. 3D얼굴 추적 기반 휴먼인지 기술

본 논문에서 시청자 인식을 위하여 개발된 기술은 입력되는 한 개의 카메라 영상으로부터 얼굴의 검출한 후 3D 모델링을 통한 얼굴의 3D 모델을 추출하여 시청자의 위치와 응시 방향 및 각도를 추출하는 3D 얼굴추적 기반 휴먼인지 기술이다.



<그림 1> 3D 얼굴추적 기반 휴먼인지 흐름도

4.1 3D 얼굴,눈썹,입술 추적구현

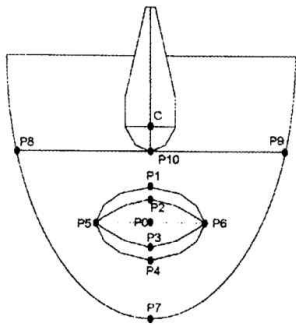
입력 스트림이 들어오면 이 스트림의 이미지에서 피부색과 타원 얼굴 모델로부터 얼굴 후보 영역을 구하게 된다.

검출된 얼굴 영역을 다음 그림과 같은 미리 만들어진 표준 3D 얼굴 모델(Wire Frame Model)의 정면 얼굴과 정합하여 특징점(눈, 코, 입등)별 Adaboost에 의한 위치를 확정한다.

본 논문에서는 특징점 검출을 위해 영상의 휘도치 분포를

다단계로 임계화하는 방법을 사용한다. 임의의 대상 영상들로부터 획득된 확률·통계적 분석에 의해 다단계 임계화된 구간을 수직·수평으로 투영하여 얼굴부위에 해당하는 휘도치 분포의 경계값을 결정한다.

이 결정된 임계값에 따라 얼굴이외의 성분으로부터 얼굴을 분할하고, 분할된 얼굴부위를 바탕으로 안면 요소 특징점들(눈, 코, 입등)을 추출한다.



(그림2) 표준 3D 입술 특징점

생성된 3D 얼굴 모델은 8개(양눈끝에 4개 코에 2개, 입양끝에 2개)의 기본적인 포인트를 포함하게 된다.

눈썹과 입술의 특징점까지 포함하는 경우는 16개의 포인트를 포함한 3D 모델로 구성하게 된다.

위의 그림에서 볼 수 있듯이 먼저 카메라로부터 입력되는 영상 스트림에서 얼굴 영역을 검출한다. 검출된 얼굴 영역의 특징점 들을 추출하여 얼굴에 대한 3D 모델을 생성한다. 생성된 3D 얼굴 모델을 기반으로 카메라 스트림으로부터 3D 얼굴 위치를 추적한다. 얼굴 전체에 대한 추적 진행과 더불어 눈썹 및 입술에 대한 세부적인 영역에 대한 3D 위치도 추적한다. 추적된 3D 얼굴 데이터를 화면에 디스플레이하거나 메타데이터 형식으로 저장한다.

4.2 얼굴 검출모듈(Face_Detect)

4.2.1 기능:

로드된 프레임 정보에서 얼굴을 검출하는 모듈이다.

4.2.2 설명:

얼굴 인식 시스템에서 중요한 첫 번째 단계는 영상에서 얼굴이 있는 위치를 알아내는 것이다. 이러한 과정을 통틀어 얼굴 검출이라고 하는데 최근 실시간 얼굴 인식을 목적으로 하는 다양한 시스템이 선을 보이고 있고, 이에 따라 실시간 응용에서의 인식을 위한 얼굴 영역을 검출하는 과정이 더욱 중요하게 된다. 사람의 얼굴은 응시하는 방향에 따른 정면 혹은

측면의 각도, 고개를 좌우로 기울이는 정도, 다양한 표정, 카메라와의 거리에 따른 얼굴 영상의 크기 등과 같은 형태적 변화와 조명에 따른 얼굴 내에서의 발기 정도의 차이, 복잡한 배경 혹은 얼굴과 구분이 어려운 색상의 다른 객체 등과 같은 외부적 변화에 따라 매우 다양하게 나타날 수 있기 때문에 영상으로부터의 얼굴 검출 연구는 많은 어려움과 한계성을 포함하고 있다. 이러한 한계 요소로 인해 얼굴 인식 분야에서 얼굴 영역 및 성분 검출연구가 상당히 중요한 요소로 다루어지고 있고, 초기 얼굴 인식 연구에서 전처리 단계로 취급되던 검출 부분이 하나의 독립적인 연구 분야로 다루어지고 있는 실정이다.

아래 <그림3>는 얼굴 인식 시스템에서 취급되는 한계 요소를 나타내고 있다.

pose		lighting	
expression		decoration	
occlusion		session	

<그림 3> 얼굴인식을 어렵게 만드는 한계 요소들

4.2.3 구현: (Algorithm: Adaboost face detection)

1). 입력

훈련 영상 집합 $S = (x_1, y_1), \dots, (x_n, y_n)$

여기서, $x_i \in R^k$, $y_i \in \begin{cases} 0 & \text{배경 영상} \\ 1 & \text{얼굴 영상} \end{cases}$

2). 비중 초기화

$$w_{(l,i)} = \begin{cases} \frac{1}{2m} & \text{for } y_i = 0 \\ \frac{1}{2l} & \text{for } y_i = 1 \end{cases}$$

여기서, m 은 훈련영상 집합 S 내의 배경영상의 개수이며, l 은 얼굴영상 개수

3). 반복 $t = 1, \dots, T$

(a) 비중의 표준화

$$w_{(t,i)} = \frac{w_{(t,i)}}{\sum_{j=1}^n w_{(t,j)}}$$

여기서, $w_{(t,i)}$ 는 t 번째 약한 분류기에 입력되는 i 번째 훈련영상의 비중을 의미

(b) 약한 분류기 (h_j)의 에러 ϵ_j

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i|$$

(c) 분류기의 선택

가장 낮은 에러율을 ϵ_t 라 하고, ϵ_t 를 가지는

약한 분류기 h_j 를 선택.

(d) 비중 업데이트

$$w_{(t+1,i)} = w_t \beta_t^{1-e}$$

여기서, 입력영상 x_i 가 h_j 의 값에 따라

올바르게 분류가 된 경우 $e_i = 0$

올바르게 되지 않은 경우 $e_i = 1$,

그리고 $\beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$

4. 최종 강한 분류기

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & otherwise \end{cases}$$

여기서, $\alpha_t = \log \frac{1}{\beta_t}$



<그림 4> 얼굴검출 모듈의 결과물

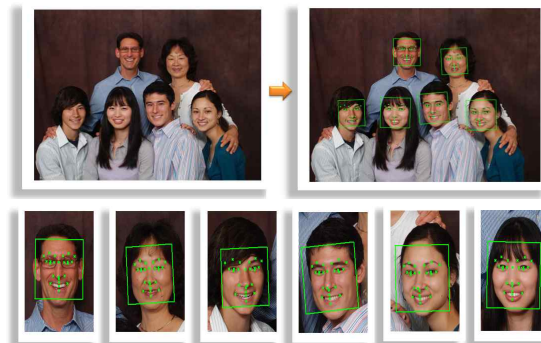
4.3 얼굴검출 및 3D 얼굴 모델링

얼굴을 검출하여 3D 얼굴 모델을 만드는 과정은 다음 그림과 같다. 일반적으로 3D 얼굴 모델을 만드는 과정은 2개의 카메라를 이용하여 위상차를 이용한 깊이 정보를 계산하는 과정을 통하여 생성하게 되는데, 제안된 논문에서는 단일 카메라 입력을 통해 얻어진 얼굴 영역을 정면 얼굴 모델을 이용한 3D 변환행렬을 구성하여 3D 얼굴 모델에 매핑하여 3D 좌표를 얻어내는 방법을 사용하고 있다.

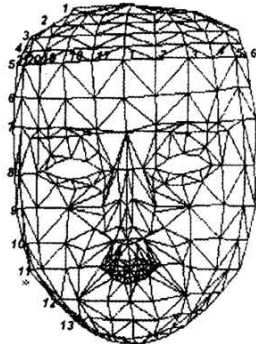


<표3>는 Adaboost 알고리즘을 이용한 얼굴검출 순서도

얼굴 검출을 위해서 Adaboost 알고리즘을 사용하였다. AdaBoost 학습 알고리즘을 통해서 얻어진 특징점을 이용하여 다양한 입력 스트림에 대하여 얼굴 검출을 하는 과정은 다음 그림과 같다.



<그림 5> Adaboost를 이용한 얼굴 검출방법



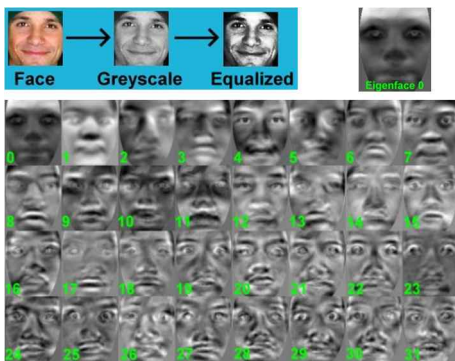
<그림6>표준3D와이어 프레임모델(Wire Frame Model)

생성된 3D 얼굴인식 모델은 8개(양눈끝에 4개 코에 2개, 입양끝에 2개)의 기본적인 포인트를 포함하게 된다. 눈썹과 입술의 특징점까지 포함하는 경우는 16개의 포인트를 포함한 3D 모델로 구성하게 된다.

4.3 분석 방법

PCA를 이용한 얼굴 인식에서는 얼굴을 가장 잘 표현할 수 있는 얼굴 공간(Face space)을 만드는 것으로 가능해진다. 이러한 얼굴 공간들의 벡터가 주성분(principal component)이 되고, 전체 영상공간에서 얼굴을 가장 잘 표현할 수 있는 벡터를 찾는 것이 목적이다. 얼굴영상에 일치하는 공분산 행렬의 고유벡터를 구하고, 새로운 영상을 구해진 고유벡터를 이용해 부수공간으로 투영시켜 학습 집합과 비교함으로써 인식을 수행한다.

각 개인의 얼굴은 고유벡터들의 선형 결합에 의해 표현되며, 고유벡터의 계산량을 고려하여 가장 큰 고유값과 대응하는 것으로부터 M 개의 고유벡터만을 사용함으로써 고차원의 영상 데이터를 저차원으로 차원 감소시킨다.



<얼굴 인식 모듈(Principal component analysis)>

다음은 주성분 분석을 이용한 얼굴인식 과정을 나타낸다. 학습 얼굴 영상을 1차원 벡터로 변환한다.

$$A = [a_{ik}]$$

여기서, i 는 학습 집합의 얼굴 영상의 전체 픽셀 수이고, a_k 는 학습 집합의 k 번째 얼굴이다.

학습 얼굴 영상들의 얼굴 평균 Ψ 을 구한다.

$$\Psi = \frac{1}{K} \sum_{i=0}^{n^2} \sum_{k=1}^K a_{ik}$$

각 얼굴 영상의 차영상 x_{ik} 을 평균 영상으로부터 구한다.

$$x_{ik} = a_{ik} - \Psi \quad (k = 1, \dots, K)$$

$$X = [x_{ik}]$$

X 는 학습 집합을 정의하는 i^*k 차원의 행렬이다. 행렬의 분산을 최대화하기 위한 공분산 행렬(covariance matrix)을 구한다.

$$C = XX^T$$

이 행렬은 차원이 크므로(영상의 픽셀 수와 동일) 계산하기가 어렵다. 따라서 XX^T 와 $X^T X$ 의 고유치는 같고, 고유 벡터는 X 를 곱하고 정규화 한 것과 동일하다는 사실을 이용해 계산을 쉽게 하고자 한다. 위 식의 고유값들을 λ_k 라고 하고, 고유벡터들을 E_k 라 가정한다.

변환된 방법으로 공분산 행렬을 구한다.

$$Q = X^T X$$

Q 의 고유벡터를 V_k 와 고유값 λ_k 을 구한다.

$$X^T X V_k = \lambda_k V_k$$

양변에 X 를 곱해 원래 데이터 집합의 고유벡터를 구한다.

$$XX^T X V_k = \lambda_k X V_k$$

여기서, $X V_k$ 는 C 의 고유벡터이므로, $E_k = X V_k$ 가 된다.

또한 V_k 는 k^*k 행렬인데 m 개의 중요한 고유벡터를 선정하면 V_{km} 으로 표현할 수 있다. 따라서 V_{km} 에 해당되는 고유벡터를 구하기 위한 식은 다음과 같다

$$E_{im} = \sum_{m=1}^m \sum_{k=1}^K x_{ik} V_{km}$$

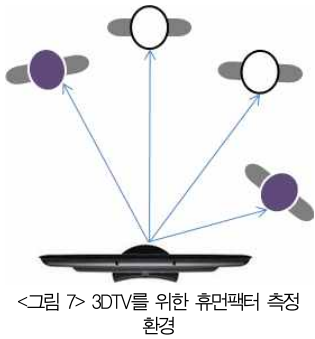
각 학습 얼굴 영상은 고유벡터 공간의 선형결합으로 표시한다.

$$W_{mk} = \sum_{k=1}^K \sum_{m=1}^M E_{im}^T X_{ik}$$

주성분 분석에 의해 구해진 가장 큰 고유치들과 상응하는 고유벡터들을 영상의 크기와 같은 형태로 변환한 것을 보여주며, 이러한 고유 벡터의 영상 형태를 고유 얼굴(Eigen face)라 부른다.

4.4 3D얼굴추적 기반 휴먼인지 기술구현 및 결과

제안된 연구는 다음 그림과 같은 휴먼 팩터 측정 환경을 고려하여 개발되었으며, 동일한 환경에서 실험하였다.

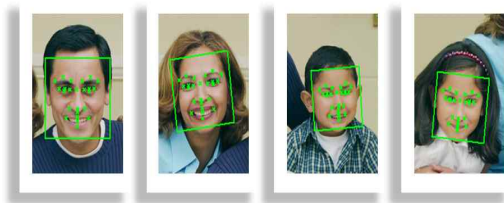


<그림 7> 3DTV를 위한 휴먼팩터 측정 환경

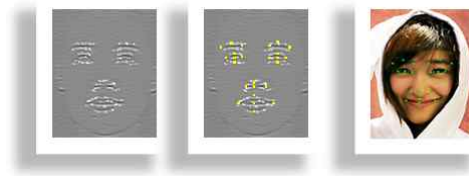
3D 얼굴인식 기반 휴먼인지 기술 개발환경은 다음과 같다.

- 개발언어 : Visual Studio 2005
- 운영체제 : Windows XP or Windows 7
- 입력장치 : 웹 카메라 (640x480 이상)

다음 그림은 개발된 프로그램으로 시청자의 얼굴을 3D로 인식하여 원 얼굴에 3D 와이어 프레임으로 오버레이하여 제대로 인지되었는지 확인하기 위한 휴먼인지 테스트 결과 화면이다.



<그림8> 3D 얼굴인식을 통한 3D얼굴 모델 추적화면



<그림 9> 3D얼굴특징 추출 MAP

현재 구현된 휴먼인지 프로그램의 성능은 다음과 같다. 성능은 실시간 인지를 기반으로 추적상태의 성능 및 인식가능한 최소 얼굴 픽셀의 크기, 최대 인지 각도 및 각도에 따른 에러율과, CPU의 로드등을 구하였다. 표를 보면 개발된 기술이 휴먼팩터 측정에서 요구하는 성능 기준을 넘어서고 있음을 알 수 있다.

5. 결론 및 향후 연구방향

3D영상 객체인식을 통한 휴먼인식 기술은 기존의 고정된 검출방식의 객체인식기술에 비해 실시간 움직이는 입체영상 깊이 정보가 추가됨으로써 얼굴객체인식을 통한 인식기술을 보다 현실감 있게 적용한 기술을 제공한다. 이러한 영상객체인식기술은 실용로봇제어분야 또는 상용화를 위해서 개발중인 CCTV를 이용한 보안분야, 자동차첨단융합기술 Auto-IT Automobile분야등에 적용할 수 있으며, 얼굴검출에 대한 보다 정확도를 높이는 검색기술이 보다 보완 되어져야 할 것으로 생각된다. 이를 위해서는 휴먼팩터에 대한 연구와 휴먼팩터를 고려한 시스템 개발이 절대적으로 필요하다. 휴먼팩터에 대한 연구를 위해서는 휴먼인지 기술이 전제되어야 하며, 본 논문에서는 영상객체인식을 위한 얼굴인식 기반의 휴먼인지 기술에 대해 기술하였다.

현재 개발된 휴먼팩터 측정을 위한 3D 얼굴인식 기반의 휴먼인지 기술은 고사양의 CPU를 기본으로 하고있으며, 한 사람의 얼굴만 추적하고 있다. 그러나, 3D기반의 얼굴 검출 기술을 위해서는 다수의 얼굴을 실시간 추적하는 기술로의 연구가 보다 보완되어져야 할 필요성을 갖고 있으며, 이러한 기술들을 하드웨어적으로 ASIC화하여 영상인식기반의 객체인식기술이 모든 산업분야에서 보다 첨단화되고 기술적으로 획기적인 기술로 자리잡기를 기대하며, 보다 완전한 기술로 개발되어져야 할 것으로 판단한다.

<참 고 문 헌>

- [1] Kics "IPTV over the face and eye tracking on-screen interface controlled study, " Papers 10-35-06-09 2010.06
- [2] Koni " "Improved Face Detection Using Adaboost Algorithm tilted " Papers 2008.06 Article 12 No. 3
- [3] Jeek " Using Adaboost for Iris Recognition in Mobile Environments A Study on Eye Detection "paper, 2008. 07 papers 2008-45CI-4-1
- [4] Jeek" AdaBoost and circular multi-target classification technique based on the function Rox "thesis paper 2010.03 2010-47CI-3-3
- [5] Journal of Korea Multimedia Society "Adaboost algorithm, real-time face detection and tracking using" 2006 . 10 (pp. 1266-1275)
- [6] Kics "IPTV over the face and eye tracking on-screen interface controlled study, " Papers 10-35-06-09 2010.06
- [7] Koni " "Improved Face Detection Using Adaboost Algorithm tilted " Papers 2008.06 Article 12 No. 3
- [8] Jeek " Using Adaboost for Iris Recognition in Mobile Environments A Study on Eye Detection "paper, 2008. 07 papers 2008-45CI-4-1
- [9] Jeek" AdaBoost and circular multi-target classification technique based on the function Rox "thesis paper 2010.03 2010-47CI-3-3
- [10] Journal of Korea Multimedia Society "Adaboost algorithm, real-time face detection and tracking using" 2006 . 10 (pp. 1266-1275)
- [11] Kipo "3-dimensional imaging technology, "2001 Technology Trends survey report, electrical / electronic field Article 1, 2001. 11, pp.80 -82
- [12] S. Pastoor and M. Wöpking, "3-D Displays: A review of current technologies," Displays, Vol.17, 1997, pp.100- 110.
- [13] N. Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," SMPTE Journal, Dec. 1993, pp.1137-1144.
- [14] J.O. Merrit, "Stereoscopic display applications issues-Part 1: Human factor Issues," Course Notes,IS&T/SPIESymposiumonElectronic Imaging ScienceandTechnology,Feb.1993.
- [15] S. Pastoor, "Human factors of 3D displays in advanced image communications," Displays, Vol. 14, No.3, 1993, pp.150-157.
- [16] W.A. Ijsselstein, P. J. H. Seutiëns and L. M. J. Meesters, "State-of-the -art in human factors and quality issues of stereoscopic broadcast television," Deliverable ATTEST/WP5/01, Aug. 2002, pp.43-57.
- [17] ETRI, CRC, "Binocular vision system, reduce fatigue Techniques, "Final Report of the International Joint, 2002. 7.
- [28] W. Blohm, I.P. Bledie, K. Schenke, K. Fazel, S. Pastoor, "Stereoscopic image representation with synthetic depth of field, " Journal of the SID, Vol.5, No.3, 1997, pp.307-313.
- [19] P.Engeldrum, Psychometric Scaling Imcotek Pre ss, Winchester, Massachusetts, USA.,2000.
- [20] S. Yano and M. Emoto, "Two factors in visual fatigue caused by stereoscopic HDTV images," Proceedings of SPIE, Vol.4864, Aug. 2002,

저 자 소 개**최 병 관**

1991 : 서울과학기술대학교 컴퓨터학과
공학사
1998 : 홍익대학교 컴퓨터학과 공학석사
1981-1986 : 한국과학기술원 연구원
1993-2004 : (주)한화/정보통신중앙
연구소 수석연구원
2004-2009 : (주)초이스티비 부사장
2010-현재 : (주)캐프 전무/ICT사업
본부장
관심분야: 디지털방송신호처리, 데이터
베이스, 3D 응용 기술,
Automotive-IT응용기술
Email : kaist62@paran.com



문 남 미

1987 : 이화여자대학교대학원 컴퓨터
학과 석사.

1980 : 이화여자대학교대학원 컴퓨터
학과 박사.

1999~2003 : 이화여자대학교 조교수

2003~2008 : 서울벤처정보대학원대
학교 디지털 미디어
학과 교수

2008~현재 : 호서대학교 벤처전문대
학원 IT응용 기술학과
교수

관심분야 : 양방향서비스, 필터링, HCI
메타데이터, e-러닝

Email : mnm01@korea.com