

3차원 영상 객체 휴먼팩터 알고리즘 측정에 관한 연구

최 병 관*

A Research on the Measurement of Human Factor Algorithm 3D Object

Choi Byungkwan

〈Abstract〉

The 4th industrial revolution, digital image technology has developed beyond the limit of multimedia industry to advanced IT fusion and composite industry. Particularly, application technology related to HCI element algorithm in 3D image object recognition field is actively developed. 3D image object recognition technology evolved into intelligent image sensing and recognition technology through 3D modeling. In particular, image recognition technology has been actively studied in image processing using object recognition recognition processing, face recognition, object recognition, and 3D object recognition. In this paper, we propose a research method of human factor 3D image recognition technology applying human factor algorithm for 3D object recognition. 1. Methods of 3D object recognition using 3D modeling, image system analysis, design and human cognitive technology analysis 2. We propose a 3D object recognition parameter estimation method using FACS algorithm and optimal object recognition measurement method. In this paper, we propose a method to effectively evaluate psychological research techniques using 3D image objects. We studied the 3D 3D recognition and applied the result to the object recognition element to extract and study the characteristic points of the recognition technology.

Key Words : 3D Modeling, Human Factor, Image Processing, Object Recognition

I. 서론

4차 산업혁명을 기반으로 3차원 객체인식 기술은 정보통신, 방송, 의료, 영화, 게임, 애니메이션 등과 같은 기존의 모든 산업제품 개발에 광범위하게 응용되는 핵심기술로 광범위하게 자리잡고 있으며, 3D 응용 기술을 통한 정보기술이 기존 산업과 융합되면

서 새로운 패러다임을 창출하는 신 성장 동력산업으로 급부상 하고 있다. 특히, 객체인식을 통한 인지기 술은 사람의 얼굴 인식을 통한 연구가 급속히 개발 되고 있으며, 양안시차 원리를 통한 사람의 인지/정보 체계에서 3D영상인식 기법을 적용한 인식처리기술로 발전하고 있다. 또한, 휴먼팩터 알고리즘을 이용한 얼굴표정과 특징점 추출, 자료 분석과 Image Processing을 통한 외곽선추출 방법론등의 실험자들

* 가톨릭대학교 미디어기술콘텐츠학과 부교수

대상으로 설계방안 설계하고 분석하여 본 논문에서 각각의 상호관계에 있는 내용을 실험, 분석하여 제안하였다.

3D영상 객체인식 기술은 휴먼팩터 기반의 FACS 알고리즘을 적용하여 이미지 검출 및 모델링에 따른 실험과 분석을 하였으며, 객체인식에 대한 평가는 3D얼굴추적 및 양안시차 기술을 이용한 기술의 변화량을 분석하여 그 방법을 통한 얼굴검출의 핵심이 되는 이미지검출과 휴먼팩터를 통한 대칭점 추출과 객체인식기반의 이미지검출을 함께 적용하여 파라미터값을 측정하고, 3D 모델링을 통한 얼굴 인식 휴먼팩터 인지기술을 구현 적용하여 측정값에 대한 이론적, 수치적인 통계값을 연구하여 본 논문에서 이론과 연구동향을 고찰하였다.

마지막으로 3D영상 객체인식은 FACS알고리즘을 활용하여 휴먼팩터 측정을 위한 객체인식과 이론적, 기술적 3차원 인식기술 연구를 제안하며, 본 연구를 통하여 기존의 3D객체 인식을 기술의 차별성과 휴먼팩터 검출과 파라미터 측정기술에 대한 연구를 구현하고 증명하는 것으로 본 연구를 마무리하였다

II. 3차원 영상객체 이론

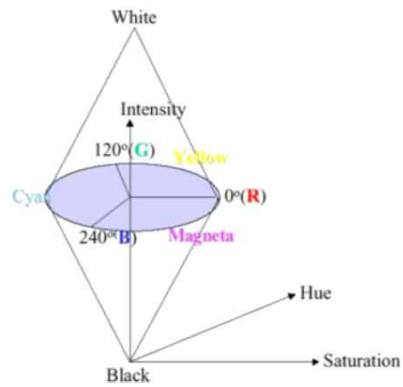
3D영상을 시청할 때에 시청자가 시각적 피로나 두통, 어지러움증, 멀미등을 경험한다면 이는 3D정보처리 매커니즘에 제시되는 3D자극이나 시청환경이 적절하지 못해서 발생하는 현상에서 나타난다. 시청자가 어지러움증과 구토등의 불편등을 해소하기 위해서는 3D자극의 특성이나 제시방식이 바뀌어야 한다. 이는 크게 3가지 측면에서 시청자의 3D시각적 피로에 영향을 미치는 물리적 파라미터로 볼 수 있다.

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$

$$\theta = \cos^{-1} \left[\frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right]$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$I = \frac{1}{3} (R+G+B)$$



<그림 1> HIS 색상계

첫째, 3차원 영상 이미지를 통한 객체인식에 대한 알고리즘 연구를 휴먼팩터 측정에 대입하여 Processing 할 수 있는 디스플레이와 연계하여 파라미터값을 얻는다. 이는 다른 변이들이 모두 적절하더라도 디스플레이 자체가 문제를 안고 있다면 시각적 불편함을 유발한다. Cross-talk, 좌안/우안 영상 제시 방식에 따른 좌안/우안 영상의 수직시차, 디스플레이의 resolution, 좌/우안 영상의 시간상에서의 제시 차이시간, 좌우안 영상의 밝기차이, 색상 차이 등이 디스플레이 요인과 관련된 물리적 파라미터들이다.

둘째, 3D 모델링을 통한 Image Processing과 디스플레이 시스템이 완벽하게 일치하도록 양안시차의 입체영상을 구현하여 적용할 수 있다. <그림 1>

3D입체 영상 HIS 처리 과정에서 보이는 색상계는 동영상까지 고려한 시간에 따른 화면시차의 변화량이 매우 중요할 수 있다. 디스플레이 자체가 좌/우 양안영상 차이를 유발할 수도 있겠지만 콘텐츠 자체가 다양한 좌,우양안 영상차이를 유발 할 수도 있다. 콘텐츠 요인과 밀접하게 관련된 요인으로는 영상 획득시의 카메라렌즈 효과 요인을 생각하면 된다.

<표 1> 휴먼팩터 연구 사례

구 분	내 용
영상정보처리시스템 HVS에 대한 연구	영상객체 정보를 인식하고 인지할 수 있는 깊이 특성 등을 이해함으로써, 시스템의 구조를 사람의 시각 시스템과 유사하게 설계 할 수 있다.
휴먼팩터 객체인식을 위한 단서 및 상호 관계에 대한 연구	휴먼팩터를 통한 주시각 조절과 현재 물체의 3D 위치 지점에 안구의 두께를 조절하여 눈의 초점을 주시하고 있는 물체에 맞추는 초점 조절이 연동되어 이뤄진다.
시각피로의 원인규명 및 해결책에 대한 연구	양안시차의 피로 유발 현상에 초점과 주시각 조절이 연동되어 피곤함을 느끼지 않고 3차원 깊이를 인식한다. 3D 영상을 보게 될 경우에는, 큰 시차에 의해 초점 및 주시 연동 파괴 현상이 발생한다.
양안시차 및 시점 변화에 대한 시청자의 민감도연구	영상의 해상도 차이와 공간적 해상도의 경우 50%의 해상도 차이에 대해 20%의 상위치 오차가 일어나며, 3D 영상의 해상도는 고해상도를 가지는 영상과 거의 같거나 약간 저하되는 정도로 인식된다.

셋째, 빅-데이터를 활용한 분산데이터를 수집하고 클라우드컴퓨팅 기술을 적용한 객체인식의 3D지각 인지에 대한 연구와 휴먼팩터 측정기술을 적용한 FACS 얼굴인식 범위의 벡터값을 얻어 경험적디자인(UX)을 통한 렌즈의 인지적용기술과 3차원 망막 조절(accommodation)을 통한 휴먼팩터 인식에 대한 객체분석을 명확히 찾아낼 수 있는 방법론 수치 해석을 통해 제시한다.

2.1 3차원 객체인식 휴먼팩터 연구

영상 객체를 통한 휴먼팩터 연구 기술을 제대로 반영하기 위해서는 휴먼팩터 관점에서 문제의 원인을 정확하게 진단하고, 해결방안을 제시해야 할 필요가 있다. 이를 위해서 다음과 같은 세부 연구들이 이루어져야 한다. 최적의 객체인식 영상처리 환경을 구현하기 위해서는 시청 환경에 대한 연구가 필요하다. 이를 위해 적절한 시청환경을 결정하는 파라미터를 정의하고, 이의 변화가 시청자에게 미치는 심리적 영향을 파악해야 한다. 일반적인 3D 툴을 이용하여 모델링을 하고 디지털 영상 소스의 카메라 모션 정보를 트래킹 기법을 통해 추적한 카메라 정보를 이용하여 렌더링한 영상과 디지털 영상 소스와 합성하게 된다. 이때 기존의 카메라 트래킹 데이터를 추출하는 기법을 사용한다.

2.2 객체인식 기반의 파라미터 연구

시청자가 자연스럽게 파라미터 측정을 통한 영상 인식 기술을 통해 영상을 시청할 수 있는 시청환경과 물리적 속성으로 적절한 시청거리, 영상의 왜곡을 줄여주는 관찰 시점, 편안한 시청을 제공할 수 있는 주변 광원도 등이 있다. 이러한 물리적 속성을 갖는 시청환경 하에서 3D 영상을 시청할 때 시청자들의 피로도, 깊이 및 모양의 왜곡, 심리적 편안함 등을 정신 물리학적으로 측정함으로써 3D 방송 시청 환경에 적합한 파라미터를 규명할 수 있다. <표 2>은 3D 방송 시스템의 시청환경 파라미터들과 권장되는 값 또는 범위를 정리한 것이다.

가장 기본적인 입체 변환 기법으로는 MTD(modified time difference)를 이용한 방법과 Garcia의 시간, 공간, 보간(spatial-temporal interpolation)

을 이용한 방법, 그리고 Matsumoto의 영상의 깊이 정보를 활용한 방법등이 있다. 하지만 기존 대부분의 방법들은 영상 변형을 통해 입체 영상의 화질이 떨어지고, 움직이는 물체의 속도와 방향을 결정해야 하며, 수직, 수평운동 방향을 구분하는 과정을 거쳐야 한다.

<표 2> 객체인식 파라미터 권장범위

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

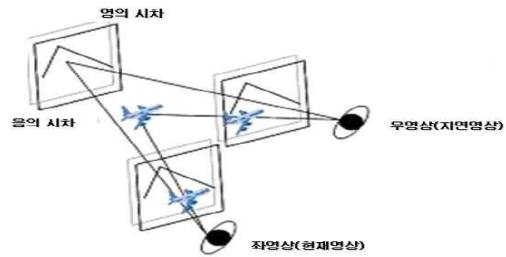
일반적으로 사람이 사물을 바라볼 때 움직임이 없는 배경보다는 움직임이 있는 객체에 대한 입체감을 더 느낀다. 따라서 객체의 움직임 방향에 영향을 받지 않고 객체를 구분하여 깊이감을 부여 함으로써 입체 변환시 부자연스러움을 최대한으로 줄여주는 변환 기법들이 연구되어 개발되고 있다.

2.3 시각적 피로 유발 원인 연구

현재까지의 기술로 개발된 입체영상 시스템은 영상을 시청할 경우, 시스템에 따라 약간씩 차이는 있으나, 대부분 20분을 전후하여 시청자가 불편할 정도의 시각적 피로나 두통 혹은 어지러움을 경험하게 된다. 이렇게 실제의 입체영상시스템은 3D 세상을 경험할 때에는 나타나지 않던 시각적 피로, 두통, 어지러움이 3D Image를 시청할 때 발생하는 정도 및 원인을 측정하고 규명하며, 이를 감소시키기 위한 여러 연구가 이에 속한다.

2.4 입장감 감성공학 연구

아이맥스영화를 감상해본 사람은 장면의 몰입감과 흥분을 기억할 것이다. 아이맥스 영화는 2D 영상임에도 불구하고 그러한 입장감을 제공함에 반해, 3D 이미지 영상은 기존의 2D 영상에 3D 양안 변이 정보를 첨가 하였음에도 불구하고 자연스럽게 앎고 입장감 또한 떨어진다. 따라서 주관적 감성에 영향을 미치는 파라미터와 이를 정량화 하는 기술을 감성공학적, 정신물리학적 관점에서 수행되는 연구이다. 이 연구에 대한 결과는 이상적인 시스템을 만들기 위한 기본 데이터로 활용할 수 있다.



<그림 2> Ross현상기반의 3D 변환원리

* 출처: Basic principle of 3D conversion based Ross phenomenon

Ross현상에 기반을 둔 입체 영상에서 시청자는 음의 시차와 양의 시차에 따라 영상 내 객체가 화면 바깥쪽으로 나와 보이거나 화면 안쪽으로 들어가 보이는 현상등으로 입체감을 느낄 수 있다. 일반적으로 디스플레이 장치를 통해서 재현하고 상영하는 영상들은 2D 영상이다. 이러한 2D 영상을 가지고 3D 효과를 내기 위하여 좌 영상과 우 영상으로 구성되는 3D 입체 영상을 생성하는 것이 3D 입체 변환의 기본 원리이다.

III. 휴먼팩터 측정 연구

3차원 객체 영상처리 기술은 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.

휴먼팩터 측정에 관한 시스템 연구의 평가라는 것은 공학적인 측면에서 평가할 수도 있지만, 개발된 시스템의 최종수용자인 사람, 즉 시청자에게 얼마나 친화적인가에 기반하여 평가를 하여야 하며, 이를 위한 것이 3D 영상 객체시스템의 평가 방법 연구이다. 이에 평가항목, 주관적 평가 방법, 객관적 평가 방법 등에 대한 연구가 이에 속한다.

3.1 평가항목 실험 연구

3차원 영상객체 시스템을 어떻게 측정하며, 평가해야 할 것인지는 그 평가항목에 따라 시스템의 우수성 기준이 달라지므로, 보다 객관성을 확보할 수 있도록 실험 환경에 대한 결정기준을 정하는 것이 중요하다. 휴먼팩터 알고리즘의 평가기준이 잘못 결정되면 시스템에 대한 객체 인식에 대한 평가는 불충분하게 되고 결국은 시스템이 문제를 지니고 있음에도 이를 발견해 내지 못할 가능성이 있다.

따라서, 이들 평가 방법에 대한 표준화가 필요하다. 현재까지의 연구에서 평가항목으로 많이 사용되는 것으로는 인식선명도, 인식깊이, 화질, 자연스러움(naturalness), 시각적 편안함(visual comfort) 등이 있다.

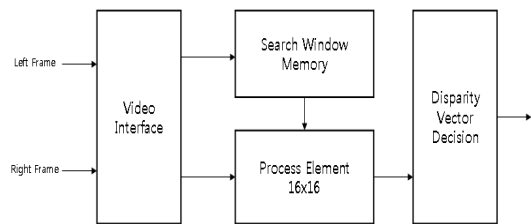
3.2 생리학적 평가방법 연구

휴먼팩터 측정을 위한 평가 방법의 기본적인 측정 방법은 정신물리학적 측정방법과 심리검사방법 모두를 기본적으로 시청자의 주관적인 자기보고서(self-report)에 기반을 두고 있다. 이 방법은 시청자가 시각적인 효과와 피로도를 느낄 수 있을 정도의 의도적이고, 시청자의 거짓 보고를 생리학적인 데이터값을 적용하였을 경우 이를 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 휴먼팩터 측정 및 연구에 필요한 시청자의 시청 위치와 양안의 거리 및 위치를 추적할 수 있는 휴먼인지 기술 개발에 대해 기술하였다.

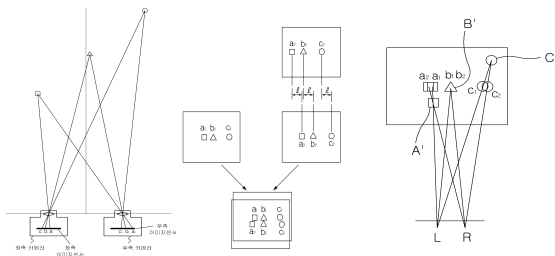
$$Error = \sum_{m=1}^M \sum_{n=1}^N |s_k(m, j) - s_{k-1}(m + j, n + j)|$$



<그림 3> 수평방향 변이

양안시차에 따른 다시점 디스플레이는 2D plus depth의 영상을 사용함으로써 높은 전송 효율을 가

질 수 있고 다양한 어플리케이션에 적용이 가능하다. 무안경식 다시점 디스플레이에서 깊이 영상의 공간적/시간적인 특성은 피로를 유발하는 주요 요소이다. 깊이 영상의 분산 및 평균 값을 사용하여 피로도 유발 요소인 공간적/시간적 복잡도, 깊이 생성 위치와 영상 전체의 움직임 크기를 측정하고 이를 이용하여 피로도도를 예측하는 방법이 제안되었다. 깊이 영상만을 사용함으로써 계산의 복잡도를 낮추고, 편안한 영상을 위한 깊이 값 조절도 용이하다.



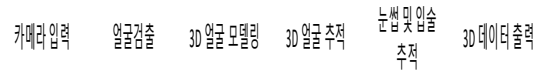
<그림 4> 3차원 객체 영상 입체 표현기법

3.3 휴먼팩터 객체인식 Landmark 측정

객체인식을 통해 얼굴검출 영역의 특징점 들을 추출하여 얼굴을 통한 Landmark를 설정하고, 이에 대한 3차원 이미지 모델링을 생성한다. 생성된 3차원 영상의 얼굴 검출 기반의 객체 영역을 기반으로 카메라 스트림 영역을 설정하고 이후 3D 얼굴 객체에 대한 위치를 추적한다. 얼굴 전체에 대한 추적 진행과 더불어 눈썹 및 입술에 대한 세부적인 영역에 대한 3D 객체 이미지를 추적한다.

본 논문에서 3차원 객체 인식의 입체영상 인식을 위해서 개발된 기술은 입력되는 한 개의 카메라 영상으로부터 얼굴 이미지를 검출한 후 3D 모델링을 통한 얼굴의 Landmark를 추출하여 시청자의 위치

와 응시 방향 및 각도를 추출하는 3D 얼굴추적 기반 휴먼인지 기술이다.



<그림 5> 3차원 객체인식 휴먼인지 흐름도

IV. 휴먼팩터 객체 모듈 검출방법

4.1 객체인식 모듈 실증연구

3차원 영상객체 기반의 휴먼팩터 알고리즘 측정에 관한 얼굴인식 시스템에서 가장 중요한 첫 번째 단계는 영상처리 기술에서 나타나는 얼굴을 표기하는 특징점(Landmark)의 위치를 정의하고, 설정된 값의 위치의 정확히 알아내는 것이다. 이러한 과정을 통틀어 객체인식 기반의 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.



<그림 6> 3차원 디지털 영상 합성 이미지

위의 <그림6>은 객체인식을 통한 3차원 입체영상의 합성 이미지를 나타내고 있다.

4.2 휴먼팩터 객체인식 실증 연구

PCA를 이용한 얼굴 객체의 인식 기술에서는 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.

객체인식 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 라 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 고유벡터

를 구한다.

$$XX^T XV_k = \lambda_k XV_k$$

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

또한 V_k 는 $k * k$ 행렬인데 m 개의 중요한 고유 벡터를 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.

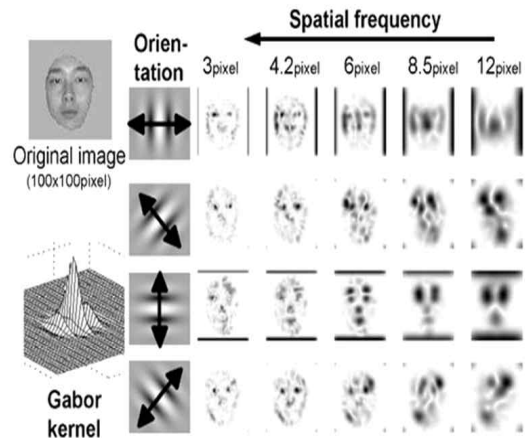
$$W_{mk} = \sum_{k=1}^K \sum_{m=1}^M E_{im}^T X_{ik}$$

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



<그림 7> 휴먼팩터 Landmark 측정 객체 검출

휴먼팩터 객체 모형을 통한 얼굴의 Landmark 측정 검출을 하여 얻어진 3차원 Image 데이터를 얼굴 모델링을 표현하도록 프로그래밍 처리를 하고 이에 대한 과정을 설정하여 렌더링하여 설계한다. 일반적으로 3D 얼굴 모델을 만드는 과정은 2개의 카메라를 이용하여 양안시차를 적용한 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.



<특징 추출 모듈(Gabor Wavelet Transformation)>

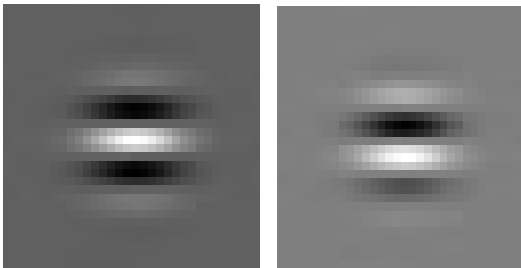
<그림 8> 3D 영상합성 이미지 추출

객체인식을 통한 얼굴의 특징점을 추출하기 위해서는 얼굴의 텍스처 정보를 이용하며, Gabor 웨이블릿 변환을 사용하여 얼굴 영상들을 처리한다. 각각의 영상은 한 쌍의 Gabor 커널(kernel)과 컨볼루션(convolution)된다. Gabor 웨이블릿은 다중해상도와 다중방향으로 이루어진 필터들로 하나의 영상에 대해 여러 개의 표상을 생성한다.

2차원 Gabor 웨이브렛 함수는 다음의 형식을 갖는다.

$$\psi_{\vec{k}}(\vec{x}) = \frac{k^2}{\sigma^2} \exp\left(-\frac{k^2 x^2}{2\sigma^2}\right) [\exp(i\vec{k}\vec{x}) - \exp(-\frac{\sigma^2}{2})]$$

2차원 웨이브벡터 $\vec{k} = (k, \theta)$ 는 필터의 공간 주파수 k 와 방향 θ 를 통제하며, 벡터 $\vec{x}(x, y)$ 는 영상을 이루는 각각의 화소를 나타낸다. σ 는 k 에 관련하여 Gaussian 윈도우의 폭을 조절한다. (그림1)은 영상에 적용된 Gabor필터 쌍이다. 이 필터들은 90도 위상관계를 이루고 있다. 필터쌍들이 위상 차이에 의하여 영상에 미치는 영향을 적게 하기 위하여 각각의 필터를 영상에 적용한 후, 생성된 2개의 컨벌루션된 영상 v_1, v_2 는 다음과 같이 처리된다.



<그림 9> 적용된 Gabor 필터 예

$$S = \sqrt{(v_1^2 + v_2^2)}$$

아래 그림은 Gabor 웨이브렛 계수 히스토그램의 평균값을 적용하여 얼굴의 주요 요소들의 경계선이 추출된 결과이다.



<그림 10> 객체인식을 통한 얼굴 경계선 검출

4.3 영상인식 분석 처리

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

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

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

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

$$t = 1, \dots, T$$

(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

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

여기서, 입력영상 x_i 가 h_j 의 값에 따라 올바르게 분류가 된 경우 $e_i = 0$ 올바르게 되지 않은 경우 $e_i = 1$,

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

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

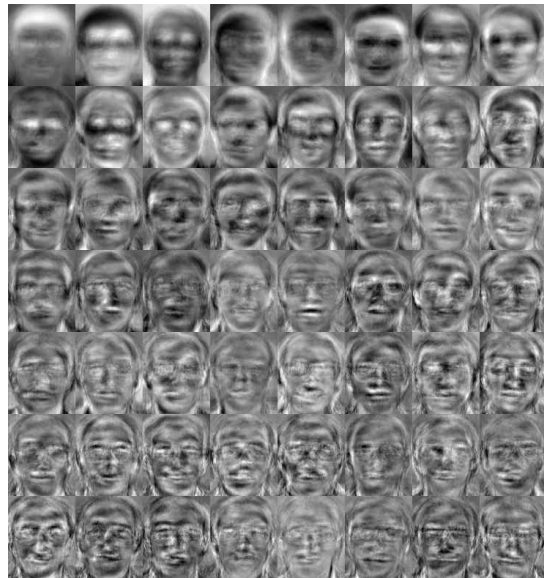
$$\text{여기서, } \alpha_t = \log \frac{1}{\beta_t}$$

은 다음과 같다.

- 개발언어 : Visual C++, JAVA
- 운영체제 : Windows10
- DataBase: MS_SQL, Oracle
- 입력장치: 웹 카메라(HD 720P/1500만화소 CMOS)
- 렌즈: F1 8f=45mm 5G렌즈 / 접사 = 30Cm

설 명:

Adaboost 학습 알고리즘을 통해 검출된 얼굴 영역의 이미지는 Gabor-kernel 기반의 PCA 특징 추출기 입력으로 사용된다. 객체인식 기반의 휴먼팩터 인식 시스템은 추출된 특징 데이터의 계수들과 등록되어 있는 인가자와 비인가자 간에 Euclidean 거리를 계산하고, 거리가 가장 가까운 얼굴로 매핑된다.



<그림 11> 객체 주성분 고유벡터(Eigen face DB)

4.4 휴먼팩터 객체인식 구현 결과

3차원 객체인식 기반의 휴먼인지 기술 개발환경

<4-6>의 그림은 자체 연구개발한 FACS 프로그램으로 다양한 사람들의 얼굴을 3차원 이미지로 인식하여 원 얼굴에 3D 와이어 프레임으로 오버레이하여 객체의 주성분 고유벡터 값이 제대로 인지되었는지 확인하기 위한 휴먼인지 테스트 결과를 보여주는 내용이다.



<그림 12> 3D얼굴특징 Landmark 추출 MAP

Landmark Shape를 이용하여 얼굴표정에 변화와 표시된 영상들에 대해서 FACS 데이터 추출을 통한 각 랜드마크에 해당하는 영상의 좌표 값을 포함하는 파일을 생성하였다. 이렇게 학습 할 영상과 랜드마크 좌표정보가 주어지면 능동 형태 모델 학습 알고리즘을 통해 학습을 수행하여 얼굴 특징에 대한 능동 형태모델을 생성하게 된다. 연구개발한 휴먼인지 프로그램의 성능은 다음과 같다.

성능은 실시간 인지를 기반으로 추적상태의 성능 및 인식 가능한 최소 얼굴 픽셀의 크기, 최대 인지 각도 및 각도에 따른 에러율과 CPU의 로드등을 구하였다.

V. 결론 및 향후 연구방향

본 논문에서는 영상객체 기반의 얼굴인식을 통한 휴먼인지 기술에 대한 검출 방법과 얼굴인식과 연계한 얼굴 검출을 위해서 Adaboost 알고리즘을 사용하였다. Adaboost 학습 알고리즘을 통해서 얻어진

특징점을 이용하여 다양한 입력 스트림에 대하여 얼굴 검출을 하는 과정을 연구하였고 이를 통해 다양한 검출방법과 검색기술을 통한 영상 처리기술의 가능성을 확인할 수 있었다.

3차원의 객체영상 인식을 통한 휴먼팩터 파라미터 측정은 FACS 휴먼인식을 통한 객체 검출과 인식시스템 기술을 통해 기존의 고정된 검출방식의 객체인식기술에 비해 실시간으로 움직이는 입체영상 값이 정보가 추가됨으로써 사람의 얼굴 객체인식을 통해 얼굴의 특징점을 찾은 다음 특징점(Landmark)에 필요한 특징 추출과 객체영역을 추출하는 방법을 통하여 얼굴 인식기술을 보다 정확성을 나타내어 적용한 방법으로 설계를 하였다. 이러한 영상객체 인식기술은 얼굴인식의 정도를 분석하기 위해 휴먼팩터 측정에 관한 객체인식 Adaboost 알고리즘을 통해서 객체인식의 검출에 대한 보다 정확도를 높이는 검색기술과 영상 압축기술이 보다 더 연구 되어져야 할 것으로 생각된다.

이를 위해서는 휴먼팩터에 대한 연구와 객체인식을 고려한 시스템 개발이 절대적으로 필요하다. 얼굴검출의 기술에 대한 연구는 그 동안 많은 분야에서 개발되어져 왔고 현재까지도 다양한 Adaboost알고리즘을 기초로 한 홀리스틱 분석방식을 이용한 연구가 다양한 분야에서 연구개발 되고 있다. 따라서, 본 논문에서 연구한 객체인식 기반의 휴먼팩터 측정 연구는 얼굴검출과 객체인식을 통해 2차 미분법을 적용하여 연구를 실행한 예가 없었기 때문에 객체인식을 통한 3차원 휴먼팩터 알고리즘의 적용 분석을 고화질 기술을 적용한 UHD, 4K급 실시간 객체인식 기술로 개발되어야 하며, 더불어 MPEG-21 기반의 압축영상 기술을 탑재한 연구가 지속적으로 이루어져야 될 것으로 판단된다.

참고문헌

- [1] Koni "Improved Face Detection Using Adaboost Algorithm tilted," Papers 2015.06 Article 12 No. 3.
- [2] Jeek "Using Adaboost for Iris Recognition in Mobile Environments A Study on Eye Detection," paper, 2008, 07, 2013-45CI-4-1.
- [3] "Jeek AdaBoost and circular multitarget classification technique based on the function Rox," thesis paper 2012.03, 2016-47CI-3-3.
- [4] Journal of Korea Multimedia Society "Adaboost algorithm, real-time face detection and tracking using," 2014.10, pp.1266-1275.
- [6] Stephen Milborrow, Fred nicolls, "Locating Facial Feature with an Extended Active shape Model," Lecture notes in computer science, 2013.
- [7] C.Padgett, G.Cottrell, "Representing face images for emotion classification," Advances in Neural Information Processing System, Vol.9, MIT Press. 2015.
- [8] Mark F. Bear, Barry Connors, and Michael Paradiso, Neuroscience: Exploring the Brain 3E, 2014.
- [9] V. Pavlovic, A.Garg, M. Rehg, Multimodel speaker detection using error feedback dynamic Bayesian network, in: Proc. IEEE Internet. Conf. Computer Vision and Pattern Recognition, 2016.
- [10] P. Ekman, W.V. Friesen. J.C. Hager, Facial Action Coding System(FACS): Manual, CD Rom, San Francisco, CA, 2017.
- [11] Y. Hu, D. Jiang. S. Yan, L. Zhang, H. zhang. "Automatic 3D reconstruction for face recognition," Proc. 6th IEEE Int'l Conf. on Automatic Face and Gesture Recognition, 2016, pp. 843-848.
- [12] Yongmian Zhang, Qiang Ji "Facial Expression Understanding in Image Sequences Using Dynamic and Active," Vision Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV 2013) 2-Volume Set 2003 IEEE.
- [13] Kipo "3-dimensional imaging technology," 2013 Technology Trends survey report, electrical/electronic field Article 1, 2013. 11, pp.80~82.
- [14] S.Pastoor and M. Wöpking, "3-D Displays: A review of current technologies," Displays, Vol.17, 2016, pp.100~110.
- [15] N.Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," SMPTE Journal, Dec. 2015, pp.1137~1144.
- [16] J.O.Merrit, "Stereoscopic display applications issues-Part 1: Human factor Issues," Course Notes, IS & TSPIE Symposium on Electronic Imaging Science and Technology, Feb. 2015.
- [17] 최병관, "다면기법 SPFACS 영상객체를 이용한 AAM 알고리즘 적용 미소검출 설계 분석," 디지털정보학회 논문지, 제11권, 제13호, 2015, pp.99-112.
- [18] 최병관, "영상객체 spFACS ASM 알고리즘을 적용한 얼굴인식에 관한 연구," 디지털정보학회 논문지, 제12권, 제4호, 2016, pp.1-12.

■ 저자소개 ■



최병관
(Choi Byungkwon)

현재 가톨릭대학교
미디어기술콘텐츠학과 부교수
2004년 10월 ㈜LG-NORTEL
상무/ICT 사업본부장
1991년 05월 ㈜한화/한화정보통신
기술중앙연구소
1981년 04월 한국과학기술원 시스템공학연구소
관심분야 : 인공지능, 빅데이터,
클라우드컴퓨팅, 사물인터넷(IOT)
AR/VR, 로봇공학, 3D Image SoC,
지능형자동차, 스마트공장, UI/UX
유비쿼터스/유씨티
E-mail : seri@catholic.ac.kr /
zcbk@naver.com

논문접수일 : 2018년 05월 23일
수정일 : 2018년 06월 04일
게재확정일 : 2018년 06월 05일