

영상객체 spFACS ASM 알고리즘을 적용한 얼굴인식에 관한 연구

최 병 관*

ASM Algorithm Applid to Image Object spFACS Study on Face Recognition

Choi Byungkwan

〈Abstract〉

Digital imaging technology has developed into a state-of-the-art IT convergence, composite industry beyond the limits of the multimedia industry, especially in the field of smart object recognition, face - Application developed various techniques have been actively studied in conjunction with the phone. Recently, face recognition technology through the object recognition technology and evolved into intelligent video detection recognition technology, image recognition technology object detection recognition process applies to skills through is applied to the IP camera, the image object recognition technology with face recognition and active research have. In this paper, we first propose the necessary technical elements of the human factor technology trends and look at the human object recognition based spFACS (Smile Progress Facial Action Coding System) for detecting smiles study plan of the image recognition technology recognizes objects. Study scheme 1). ASM algorithm. By suggesting ways to effectively evaluate psychological research skills through the image object 2). By applying the result via the face recognition object to the tooth area it is detected in accordance with the recognized facial expression recognition of a person demonstrated the effect of extracting the feature points.

Key Words : ASM Algorithm, Imaging Process, 3D Image, spFACS, Object Recognition

I. 서론

영상객체 분야는 스마트-폰 기술의 발전과 더불어 개인 휴대용 단말기기로 자리 잡고 있다. 영상객

체 기술은 특수안경 착용을 통하여 입체 효과를 느꼈던 기존 방식에서 무안경 디스플레이 영상기술로 진화되고 있음에 따라 콘텐츠와 더불어 개발속도가 급속히 발전하고 있다.

영상인식을 비롯한 일반적인 시스템 개발에 있

* 가톨릭대학교 미디어기술콘텐츠학과 부교수

어 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 사람 시청자에게 얼마나 친화적인가에 기반하여 평가를 하여야 한다. 이를 위한 것이 객체영상 기술의 평가 방법 연구이며, 이에에는 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.

영상객체 spFACS ASM의 알고리즘은 양안시차의 입체효과를 발생시켜 좌/우 영상을 디스플레이에서 분리시켜 두 눈에 전달함으로써 뇌에서 양안시차를 통한 입체감을 느끼게 하는 것이다. 특히 영상인식 기술에 있어서 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[1].

II. 휴먼팩터 영상인식 기술요소

입체영상 시스템은 실제로 Image 공간에서 깊이와 대상의 크기가 실제대로 지각되고 피로를 유발하지 않으며 실감이 있어야 한다. 특히 영상인식을 통한 Image Processing처리를 위한 spFACE(Smile Progress Facial Action Coding System), ASM

(Active Shape Model)에 있어서는 영상인식 기술을 통해 실현하기 위한 휴먼팩터 기술에 대한 연구가 활발히 연구 개발되고 있으며, 이러한 기술 요소들을 통해 본 연구에서 중요사항으로 인식하고 있는 휴먼인지 기술을 이해하기 위하여 이들 세 기술 요소에 대하여 좀더 상세하게 알아보기로 한다.

첫째, 시스템 설계 단계에서부터 휴먼팩터를 고려하여 설계 디자인을 해야된다. 모든 경우에 시스템의 하드웨어 개발이 일정 수준의 궤도에 오르고 나면, 추가적으로 휴먼팩터를 고려하기 위해 시스템을 수정하기가 곤란하다.

둘째, 영상객체를 통한 방송시청 환경 파라미터 규명 및 이에 대한 감성 공학 기술로서 시청자가 편안하게 시각 피로 없이 영상인식을 할 수 있도록 하기 위한 최적의 시청 환경에 대한 연구이다.

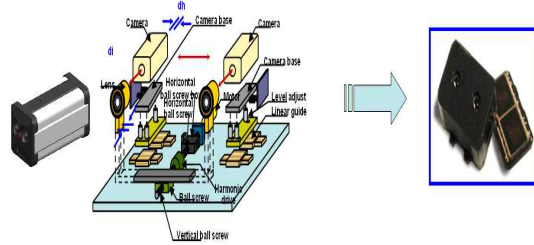
셋째, 객체인식 시스템에 대한 심리학적 평가 기술로서 이상적 시스템을 개발하기 위한 여러 단계를 거치면서 각 단계마다 개발된 시스템의 완성도를 시청자에게 제시하여 개발된 시스템의 심리학적 성능을 효과적으로 평가하기 위한 연구이다.

특히, 고려해야 할 휴먼팩터가 기존 시스템의 구조만으로 해결하기 어려운 근본적인 문제일 때 더욱 그러하다. 따라서 개발 도중에 수정을 하는 것보다는 설계 단계에서부터 휴먼팩터를 측정하거나 고려하여 시스템 기능이나 사양을 정하는 것이 비용절감 효과를 얻을 수 있다[2].

<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 설계할 수 있다.
영상인식을 위한 단서 및 상호관계에 대한 연구	휴먼팩터를 통한 주시각 조절과 현재 물체의 3D 위치 지점에 안구의 두께를 조절하여 눈의 초점을 주시하고 있는 물체에 맞추는 초점 조절이 연동되어 이뤄진다.

구분	내용
시각피로의 원 인규명 및 해결 책에 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 시 점 변화에 대한 시청자의 민감 도연구	영상의 해상도 차이와 공간적 해상도의 경우 50%의 해상도 차이에 대해 20%의 상위치 오차가 일어나며, 3D 영상의 해상도는 고훈상도를 가지는 영상과 거의 같거나 약간 저하되는 정도로 인식된다.



<그림 1> 입체영상 카메라 Module

<표 2> 시청환경 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

2.1 객체인식 휴먼팩터 연구

영상객체를 통한 휴먼팩터 연구 기술을 제대로 반영하기 위해서는 휴먼팩터 관점에서 문체의 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[2].

2.2 시청환경의 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 물리학적적으로 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 MITD(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[3].

일반적으로 사람이 사물을 바라볼때 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[3].

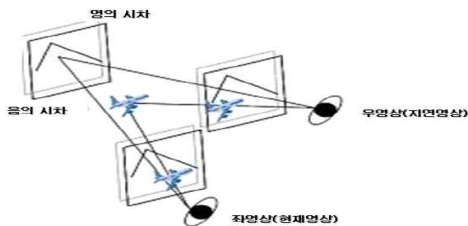
2.3 시각적 피로유발 원인 연구

현재까지의 기술로 개발된 입체영상 시스템은 영

상을 시청할 경우, 시스템에 따라 약간씩 차이는 있으나, 대부분 20분을 전후하여 시청자들이 불편할 정도의 시각적 피로나 두통 혹은 어지러움을 경험하게 된다. 이렇게 실제의 입체영상시스템은 3D 세상을 경험할 때에는 나타나지 않던 시각적 피로, 두통, 어지러움이 3D 방송 영상을 시청할 때 발생하는 정도 및 원인을 측정하고 규명하며, 이를 감소시키기 위한 여러 연구가 이에 속한다[4].

2.4 입장감에 대한 감성공학적인 연구

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



<그림 2> Ross현기반의 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

로 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[5].

III. 심리학적 평가방법 측정 연구

방송미디어 기술을 비롯한 영상처리 기술은 일반적인 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[5].

시스템의 평가라는 것은 공학적인 측면에서 평가할 수도 있지만, 개발된 시스템의 최종수요자인 사람 시청자에게 얼마나 친화적인가에 기반하여 평가를 하여야 한다. 이를 위한 것이 3D 방송 시스템의 평가 방법 연구이며, 이에 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.

3.1 평가항목 실험 환경 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 있

다. 따라서, 이들 평가 방법에 대한 표준화가 필요하다. 이 때 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[6].

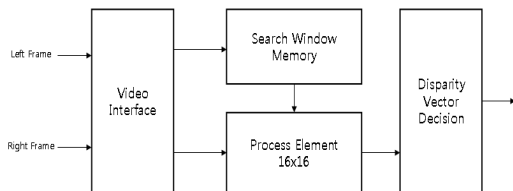
3.2 생리학적 평가방법 개발 연구

정신물리학적 측정방법과 심리검사방법 모두 기본적으로 시청자의 주관적인 자기보고서(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[7].

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[7].

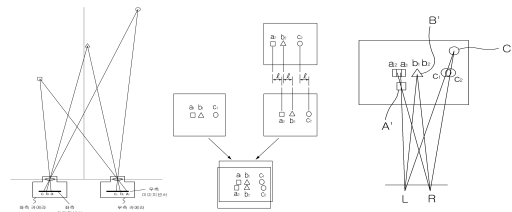
본 논문에서는 휴먼팩터 측정 및 연구에 필요한 시청자의 시청 위치와 양안의 거리 및 위치를 추적할 수 있는 휴먼인지 기술 개발에 대해 기술하였다.

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



<그림 3> 수평방향 변이

무안경 다시점 디스플레이는 2D plus depth의 영상을 사용함으로써 높은 전송 효율을 가질 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 피로도를 예측하는 방법이 제안되었다. 깊이 영상만을 사용함으로써 계산의 복잡도를 낮추고, 편안한 영상을 위한 깊이 값 조절도 용이하다[7].



<그림 4> 3D영상 입체 표현기법

3.3 얼굴인식 휴먼인지 기법

객체인식을 통해 검출된 얼굴 영역의 특징점들을 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. 본 논문에서 시청자 인식을 위하여 개발된 기술은 입력되는 한 개의 카메라 영상으로부터 얼굴의 검출 후 3D 모델링을 통한 얼굴의 3D 모델을 추출하여 시청자의 위치와 응시 방향 및 각도를 추출하는 3D 얼굴추적 기반 휴먼인지 기술이다[5].

카메라 입력 얼굴검출 3D 얼굴 모델링 3D 얼굴 추적 눈썹 및 입술 추적 3D 데이터 출력

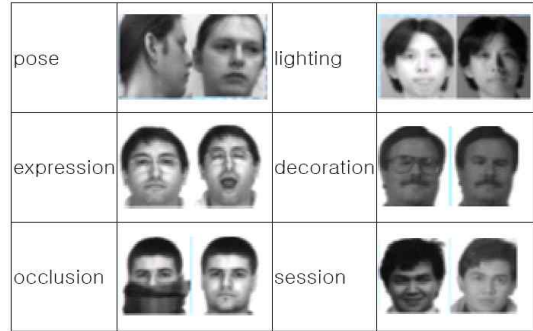
<그림 5> 3D 얼굴추적 기반 휴먼인지 흐름도

IV. 얼굴인식 모듈 검출방법

4.1 검출모듈 및 실증연구 기법

얼굴 인식 시스템에서 중요한 첫 번째 단계는 영상처리 기술에서 얼굴이 있는 위치를 정확히 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 <그림 6> 는 얼굴 인식 시스템에서 취급되는 한계 요소를 나타내고 있다.



<그림 6> 얼굴인식을 어렵게 만드는 한계 요소들

4.2 얼굴인식 검출 실증연구

연구방법(Algorithm detection): 본 연구에서는 다면적 영상 기법을 적용하여 실시간 얼굴인식 검출방법을 나타내고 있으며, 얼굴인식에 관한 분석 시스템은 얼굴형태 모델 학습, 얼굴 특징 검출과 표정분석 및 평가 모듈로 구성되어 있다. 얼굴의 특징을 검출하기 위해 입력된 영상으로부터 Adaboost 알고리즘을 이용하여 얼굴을 검출한다. spFACE(Smile Progress Facial Action Coding System) 패턴으로 검출된 얼굴 영상에 대하여 ASM(Active Shape Model) 기반으로 학습된 얼굴 형태 모델을 이용하여 얼굴의 특징(얼굴윤곽, 눈, 눈썹, 코, 입, 입술 등)을 검출한다.

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

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

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

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

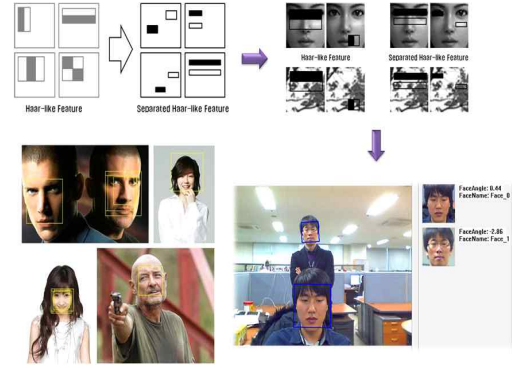
여기서, 입력영상 x_i 가 h_j 의 값에 따라
올바르게 분류가 된 경우 $e_i = 0$
올바르게 되지 않은 경우 $e_i = 1$,

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

4) 최종 강한 분류기

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

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

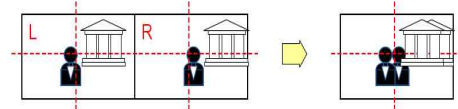


<얼굴 검출 모듈(Adaboosting Face Detection)>

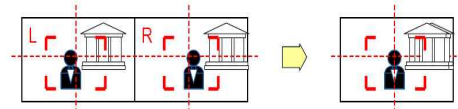
<그림 7> 얼굴검출 모듈의 결과물

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 모델에 이용한 3D 변환행렬을 구성하여 3D 얼굴 모델에 매핑하여 3D 좌표를 얻어내는 방법을 사용하고 있다[3, 8].

- Without Auto Convergence, it caused dizziness and uncomfortable manual operation needed



- With Auto Convergence, it is very easy to feel natural 3D without manual operation

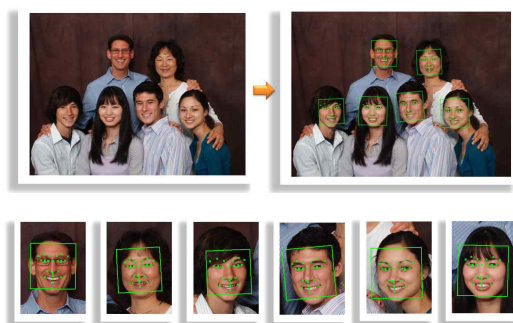


<그림 8> 3D 영상합성 이미지 추출

입력 스트림이 들어오면 이 스트림의 이미지에서 피부색과 타원 얼굴 모델로부터 얼굴 후보 영역을 구하게 된다. 얼굴 후보 영역에서 24x24 크기의 부

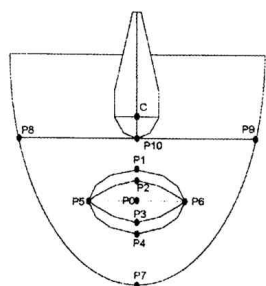
분 이미지를 추출해내고 다음으로 앞서 설명한 학습을 통해서 얻어진 약 분류기를 직접이미지에 대입해서 얼굴임을 판단한다.

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[7].



<그림 9> Adaboost를 이용한 얼굴 검출방법

본 연구에서는 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[8].



<그림 10> 입술의 특징점

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

눈썹 추적의 경우 Eyebrow PCA Shape Model을 생성한 후 Eyebrow Correspondences를 계산한 후 Eyebrow PCA Minimizer를 이용하여 최량화를 진행한다. 눈썹과는 달리 입술의 추적은 입술의 움직임, 입의 벌어진 정도, 턱의 움직임, 입술의 돌출, 뺨의 움직임등과 같은 복합적인 움직임으로 이루어진다. 그러므로 이들의 복합적인 움직임을 고려하여야 하지만 입술의 돌출과 뺨의 움직임 등은 2차원적인 접근방법으로는 어렵다. 따라서 여기에서는 입력된 동영상에 정합된 3차원모델로부터 3차원 특징 벡터를 얻을 수 있다. 본 논문에서는 아래 그림과 같이 12개의 특징점을 사용하여 4개의 특징 벡터를 추출한다. 4개의 특징벡터는 다음과 같다[5, 8].

- 윗입술 움직임 벡터(UP): $UL = P2 - P0$
- 아랫입술 움직임 벡터(Down): $DL = P3 - P0$
- 입의 벌어진 정도: $ML = P5 - P6$
- 턱의 상하 움직임: $JM = P7 - P10$

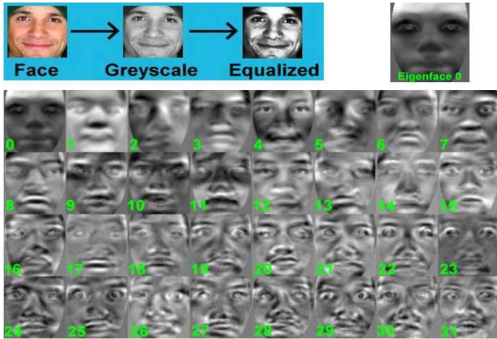
이와 같은 움직임 벡터를 정량적으로 취급하기 위해서는 개인차를 흡수할 수 있는 기본량이 필요하다. 여기에서는 입의 중횡폭을 이용하여 기준량으로 정한다. 움직임 벡터의 프레임당 이동변위는 특징점의 이동량을 기준량으로 정규화하여 입술을 추적한다[9].

4.3 얼굴인식 분석방법

PCA를 이용한 얼굴 인식에서는 얼굴을 가장 잘 표현할 수 있는 얼굴 공간(Face Space)을 만드는 것으로 가능해진다. 이러한 얼굴 공간들의 벡터가 주 성분(principal component)이 되고, 전체 영상공간에

서 얼굴을 가장 잘 표현할 수 있는 벡터를 찾는 것이 목적이다[10].

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[9].



<그림 11> 얼굴인식 분석 모듈

4.4 영상인식 주성분 분석 처리기술

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^TX 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.

$$Q = X^TX$$

Q 의 고유벡터를 V_k 와 고유값 λ_k 을 구한다.

$$X^TXV_k = \lambda_k V_k$$

양변에 X 를 곱해 원래 데이터 집합의 고유벡터를 구한다.

$$XX^TXV_k = \lambda_k XV_k$$

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.

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

각 학습 얼굴 영상은 고유벡터 공간의 선형결합으로 표시한다.

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

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[11].

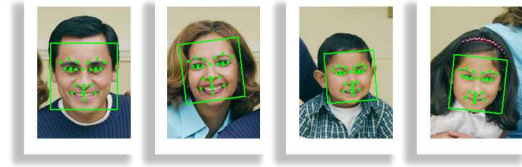


<그림 12> spFACS 얼굴표정 영상추출

4.5 객체인식 휴먼팩터 구현 결과

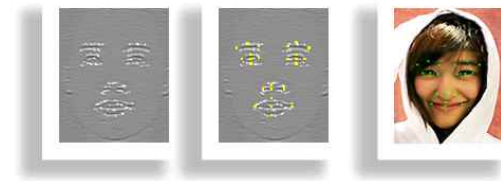
3D 얼굴인식 기반 휴먼인지 기술 개발환경은 다음과 같다.

- 개발언어 : Visual C++, JAVA
- 운영체제 : Windows10
- DataBase: MS_SQL, Oracle
- 입력장치: 웹 카메라(HD 720P/500만화소 CMOS)
- 렌즈: F1 8f=45mm 5G렌즈 / 접사 = 30Cm



<그림 13> 객체인식을 통한 3D얼굴 얼굴모델링

다음 그림은 개발된 프로그램으로 시청자의 얼굴을 3D로 인식하여 원 얼굴에 3D 와이어 프레임으로 오버레이하여 제대로 인지되었는지 확인하기 위한 휴먼인지 테스트 결과 화면이다[5].



<그림 14> 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[8].

V. 결론 및 향후 연구방향

현재 연구된 영상객체 spFACS ASM 휴먼팩터 측정기술을 실용화 또는 상용화를 앞당기기 위해서는 해당분야의 연구개발이 활성화 되어야 될 것으로 보이며, 특히 휴먼팩터에 대한 연구와 휴먼인지 기

술이 보다 활발히 개발되어 상용화되는 전체를 기반으로 개발 되어야 한다[11].

다면기법 영상 인식을 통한 ASM(Active Shape Model)알고리즘 기반의 spFACS(Smile Progress Facial Action Coding System) 휴먼인식 얼굴 검출 인식시스템 기술이 객체인식 기반에서 실시간으로 설계되고 특히 입체영상을 통한 실증 Platform연구가 활발히 개발되어 얼굴 객체인식을 통한 휴먼팩터 측정 연구와 실증연구가 다양하게 연구되어져야 된다고 생각한다.

또한, 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 처리 기술의 가능성을 확인할 수 있었다[8, 10].

따라서 향후에는 보다 다양한 얼굴검출 방법안을 연구하여 고도화된 압축기술과 MPEG7, MPEG21, H.264(Part 10)을 적용한 UHD 4K Platform을 적용한 영상 압축기술이 개발되어져야 할 것으로 생각된다[12].

참고문헌

- [1] Pavlovic, A.Garg, M. Rehg. Multimodel speaker detection using error feedback dynamic Bayesian network, in: Proc. IEEE Internet. Conf. Computer Vision and Pattern Recognition, 2010.
- [2] C.Padgett,G.Cottrell, "Representing faceimages for emotion classification," Advances in Neural Information Processing System. Vol9, MIT Press. 2011.
- [3] P. Ekman, W.V. Friesen. J.C. Hager, Facial Coding System: Manual, CD Rom, Sanfrancisco, CA, 2011.
- [4] Jeek "Adaboost and circular multitarget classification technique based on the function Rox," thesis paper 2010.03, 2010-47CI-3-3.
- [5] S. Pastoor and M. Wöpking, "3-D Displays: A review of current technologies," Displays, Vol.17, 2012, pp. 100-110.
- [6] Kipo "3-dimensional imaging technology," 2001 Technology Trends survey report, electrical / electronic field Article 1, 2009. 11, pp. 80-82.
- [7] N. Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," SMPTE Journal, Dec. 2008, pp. 1137-1144.
- [8] J.O. Merrit, "Stereoscopic display applications issues-Part 1: Human factor Issues," Course Notes, IS&T/SPIE Symposium Electronic Imaging ScienceandTechnology, Feb.2007.
- [9] S. Pastoor, "Human factors of 3D displays in advanced image communications," Displays, Vol. 14, No.3, 2013, pp. 150-157.
- [10] P.Engeldrum, Psychometric Scaling Imcotek Press, Winchester, Massachusetts, USA., 2011.
- [11] W. Blohm, I.P. Bleddie, K. Schenke, K. Fazel, S. Pastoor, "Stereoscopic image representation with synthetic depth of field," Journal of the SID, Vol.5, No.3, 2009, pp. 307-313.
- [12] N. Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images

- and their viewing conditions,” SMPTE Journal, Dec. 2010, pp. 1137-1144.
- [13] W.A. Ijsselsteijn, 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. 2012, pp. 43-57.
- [14] ETRI, CRC, “Binocular vision system, reduce fatigue Techniques,” Final Report of the International Joint, 2013. 7.
- [15] S. Yano and M. Emoto, “Two factors in visual fatigue caused by stereoscopic HDTV images,” Proceedings of SPIE, Vol. 4864, Aug. 2010.

■ 저자소개 ■



최 병 관
(Choi Byungkwon)

2015년 8월~현재, 가톨릭대학교 미디어기술콘텐츠학과 / 부교수
2004년 10월 ㈜LG-NORTEL ICT사업본부장
1993년 04월 ㈜한화/한화정보통신 기술중앙연구소
1981년 04월 한국과학기술원 시스템공학연구소

관심분야 : 디지털미디어공학, 증강현실, 로봇공학
SoC응용설계, 3D융합미디어, 객체인식,
지능형자동차, IoT사물인터넷

E-mail : seri@catholic.ac.kr

논문접수일 : 2016년 9월 07일
수정일 : 2016년 11월 17일(1차), 11월 25일(2차)
게재확정일 : 2016년 11월 29일