디지털산업정보학회 논문지 제11권 제3호-2015년 9월
http://dx.doi.org/10.17662/ksdim.2015.11.3.099

다면기법 SPFACS 영상객체를 이용한 AAM 알고리즘 적용 미소검출 설계 분석*

최 병 관**

Using a Multi-Faced Technique SPFACS Video Object Design Analysis of The AAM Algorithm Applies Smile Detection

Choi Byungkwan

〈Abstract〉

Digital imaging technology has advanced beyond the limits of the multimedia industry IT convergence, and to develop a complex industry, particularly in the field of object recognition, face smart-phones associated with various Application technology are being actively researched. Recently, face recognition technology is evolving into an intelligent object recognition through image recognition technology, detection technology, the detection object recognition through image recognition processing techniques applied technology is applied to the IP camera through the 3D image object recognition technology Face Recognition been actively studied. In this paper, we first look at the essential human factor, technical factors and trends about the technology of the human object recognition based SPFACS(Smile Progress Facial Action Coding System)study measures the smile detection technology recognizes multi-faceted object recognition. Study Method: 1)Human cognitive skills necessary to analyze the 3D object imaging system was designed. 2)3D object recognition, face detection parameter identification and optimal measurement method using the AAM algorithm inside the proposals and 3)Face recognition objects (Face recognition Technology) to apply the result to the recognition of the person's teeth area detecting expression recognition demonstrated by the effect of extracting the feature points.

Key Words : 3D Visual Object Recognition, Face Recognition, Smile Recognition

Ⅰ. 서론

다면적 영상기술은 정보통신, 방송, 의료, 영화, 게임, 애니메이션 등과 같은 기존의 모든 산업 제품개발에 광범위하게 응용되는 핵심기술로 폭 넓게 자리잡고 있으며, 최근에는 3D 응용기술을 통한 정보기술이 기존 산업과 융합되면서 새로운 패러다임을 창출하는 신(新)성장 동력산업으로 급 부상하고 있다. 특히, 사람의 얼굴인지를 통한 인식기술이 활발히 연구

* 본 연구는 2015년 가톨릭대학교 교비연구비의 지원으로 이루어졌음.
** 가톨릭대학교 미디어기술콘텐츠학과 부교수

되면서 양안시차 원리를 통한 사람의 인지, 정보 체계에서 3D 영상 인식기법을 적용한 인식처리 기술로 발전하고 있다. AAM(Active Appearance Models)알고리즘을 통한 얼굴특징 추출 자료분석과 미소검출을 통한 치아영역에 대한 외곽선 추출 방법론등의 실험자를 대상으로 한 연구설계 방안을 제시하여 각각의 상호관계에 있는 내용을 실험, 분석하여 제안한다.

3D 영상인식 기반을 통한 SPFACS 얼굴 검출 및 얼굴 모델링 랜드마크를 통한 실험을 하였으며, 미소인식에 대한 평가와 3D 얼굴 검출 기술을 이용한 변화량과 분석을 통하여 얼굴의 핵심이 되는 미소이미지의 대칭점 추출과 얼굴인식의 파라미터값을 측정하여 3D적용을 위한 측정값에 본 실험값을 대입하여 이론적이고 수치적인 통계값을 연구를 통하여 본 논문에 실험결과를 고찰 하였다[3].

마지막으로 AAM(Active Appearance Models)알고리즘을 이용한 3D영상 SPFACS의 휴먼팩터 측정을 위한 미소검출의 이론적, 기술적 인식기술 연구를 제안하며, 본 연구를 통하여 기존의 3D 얼굴인식을 통한 기술의 차별성과 객체인식을 통한 얼굴검출과 파라미터 측정기술에 대한 연구를 구현하고 증명하는 것으로 본 연구를 마무리 하였다.

Ⅱ. 다면기법 영상인식 이론적 배경

다면기법 영상을 시청할 때에 시청자가 시각적 피로나 두통, 어지러움증, 멀미등을 경험한다면 이는 3D 정보처리 매커니즘에 제시되는 3D자극이나 시청환경이 적절하지 못해서 발생하는 현상에서 나타난다. 시청자가 어지러움증과 구토등의 불편등을 해소하기 위해서는 3D자극의 특성이나 제시 방식이 바뀌어야 한다. 이는 크게 3가지 측면에서 시청자의 3D 시각적 피로에 영향을 미치는 물리적 파라미터로 볼 수 있다[2].

<표 1> 다면적 입체영상 처리 과정

과 정	기 술
기 술	형상 기술
	운동 기술
	변형 처리
해 석	물체의 형상 해석
	물체의 움직임 해석
	물체의 장면 해석
작 성	트레이싱
	평행, 회전운동
	시차계산
처 리	입체영역 개선
	2차원 영상처리
	3차원 영상처리

첫째, 3D 디스플레이와 관련된 파라미터를 들 수 있다. 이는 다른 변이들이 모두 적절 하더라도 디스플레이 자체가 문제를 안고 있다면 시청자에게 불편함을 유발한다. Cross- talk, 좌안/우안 영상제시 방식에 따른 좌안/우안 영상의 수직시차, 디스플레이의 resolution, 좌/우안 영상의 시간상에서의 제시 차이시간, 좌/우안 영상의 밝기차이, 색상 차이등이 디스플레이 요인과 관련된 물리적 파라미터들이다[4].

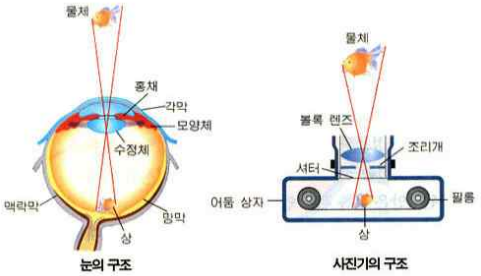
둘째, 컨텐츠 요인이 있다. 3D 디스플레이 시스템이 완벽하다 할지라도 디스플레이에 제시되는 컨텐츠가 시청자에게 문제를 유발할 수 있다. <표 1> 3D 입체영상 처리 과정에서 보이는 바와 같이 동영상의 경우까지 고려한다면 시간에 따른 화면시차의 변화량이 매우 중요할 수 있다. 디스플레이 자체가 좌, 우안 영상 차이를 유발할 수도 있겠지만 컨텐츠 자체가 다양한 좌, 우안 영상차이를 유발 수 도 있다. 컨텐츠 요인과 밀접하게 관련된 요인으로 영상획득시의 카메라 요인이 있다. 두대의 카메라 배열방법과 주 시각 제어방법 그리고 초점 거리가 획득된 영상에 영향을 미치고 이는 궁극적으로 시청자의 시청피로에 영

향을 미칠 수 있다[7].

셋째, 수렴과 더불어 3D지각에 중요한 역할을 하는 것은 렌즈의 조절(accommodation) 작용이다. 조절은 망막에 깨끗한 상을 맺게하여 궁극적으로 흐릿한 상이 아닌 선명한 상을 지각하게 하는 역할을 한다.

<표 2> 사람의 눈과 카메라의 원리 비교

	빛의굴절	빛의 양조절	원근조절	상이맺힘	빛의 차단	산란방지
눈	수정체	홍 채	모 양 체	망 막	눈꺼풀	맥 락 막
카메라	렌 즈	조 리 개	거 리 개	필 름	셔 터	어둠상자



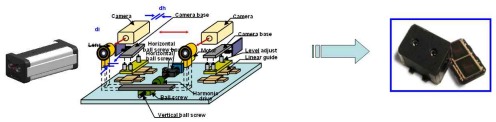
<그림 1>사람의 눈과 카메라의 조절작용

사람의 랜즈는 망막에 깨끗한 상을 맺기 위하여 대상이 멀리 있을 때에는 <그림 1>에서 보이는 바와 같이 렌즈의 굴절률을 위해 렌즈를 얇게 하고 가까이 있을 때에는 굴절률을 높이기 위하여 렌즈를 두껍게 한다. 일반적으로 3D환경에서는 수렴과 조절 작용이 서로 협응하고 렌즈의 초점거리(focal distance)와 수렴거리는 대부분의 경우에 크게 차이가 나지 않는다[7].

하지만, 3D 디스플레이에 컨텐츠가 제시되는 상황에서는 수렴과 조절이 분리되는 경향을 보이는데 이렇게 조절의 불일치가 발생하는 이슈의 수렴은 양안시차가 과도한 경우에 이중상이 경험되는 것을 방지하기 위하여 디스플레이로 부터 떨어진 곳으로 나타나게 되는데, 이에 반하여 렌즈의 조절은 여전히 광학자극이 발생하는 디스플레이로 맞춰지기 때문이다. 화면시차가 크면 클수록 조절과 수렴의 불일치양은 증가하게 되고, 이러한 수렴과 조절의 불일치가 3D 시각피로의 중요한 요인으로 간주되고 있다[14].

2.1 3D 객체인식 영상 변환

3D 콘텐츠에 대한 수요와 입체감을 느낄 수 있는 영상을 제작하기 위해서는 양안에 해당하는 두 대의 카메라시스템을 사용하는 것이 일반적이다. 하지만 3D 콘텐츠를 제작하는 경우 아직도 많은 비용과 노력을 필요로 하고 있기 때문에 기존에 2D로 제작된 영상을 3D 입체 영상으로 변환하여 <그림 2>에서 보이는 바와 같이 3D 콘텐츠를 제작하는 방법으로 사용되고 있다[5]. 2D 영상을 3D입체 영상으로 변환하는 방법은 여러 가지가 있는데, 그중 실시간 서비스를 위해 자동으로 변환하는 방법이 필요한데, 이 경우 변환된 3D 입체영상의 품질은 두 대의 카메라를 직접 이용하여 제작한 경우에 비해 현저히 떨어지게 된다[3].



<그림 2> 3D 입체영상 카메라 Module

가장 기본적인 3D 입체 변환 기법으로는 MTD(modified time differnce)를 이용한 방법과 Garcia의 시간, 공간, 보간(spatial-temporal interpolation)을 이용한 방법, 그리고 Matsumoto의 영상의 깊이 정보를 활용한 방법등이 있다[3]. 하지만 기존 대부분의 방법들은 영상 변형을 통해 입체 영상의 화질이 떨어지고, 움직이는 물체의 속도와 방향을

결정해야 하며, 수직, 수평운동 방향을 구분하는 과정을 거쳐야 한다.

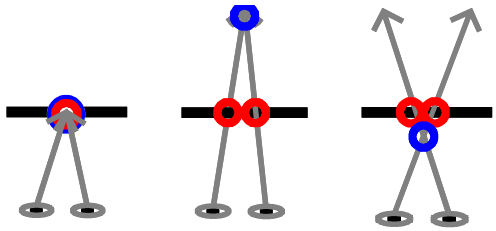
일반적으로 사람이 사물을 바라볼때 움직임이 없는 배경보다는 움직임이 있는 객체에 대한 입체감을 더 느낀다. 따라서 객체의 움직임 방향에 영향을 받지 않고 객체를 구분하여 깊이감을 부여 함으로써 입체 변환시 부자연스러움을 최대한으로 줄여주는 변환 기법들이 연구되어 개발되고 있다[9].

2.2 3D 입체영상 원리

3D 입체영상의 변환 원리는 양의시차. 음의시차, 영의시차등 크게 세 종류의 양안시차(binocular disparrity)를 제공함으로서 시청자에게 영상의 깊이감을 느낄 수 있게 한다[7]. 시차의 종류는 인간이 스크린을 통해 입체영상을 보았을 때 수렴점이 어느 위치에 생기느냐에 따라 결정된다. <그림3>에서 보이는 것처럼 영의 시차는 스크린 상에, 양의 시차는 스크린의 뒤쪽에, 음의 시차는 스크린의 앞쪽에 수렴점이 존재한다[11].

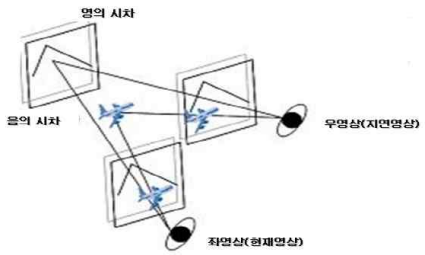
좌 영상으로는 현재 영상을 사용하고 우영상은 지연영상을 사용한다. 영상내의 비행기는 좌에서 우로 이동하고 있으며, 배경이 되는 산은 정지하고 있다. <그림4>는 Ross현상을 이용한 입체영상의 원리를 보여주고 있다. 좌안과 우안은 모두 배경이 되는 산과 비행기를 바로 보게 되며 그때 좌 영상과 우영상의 차이인 시차가 발생하게 된다. 따라서 비행기는 스크린보다 앞쪽에 수렴점이 생기는 음의 시차를 갖게 되고, 산은 스크린에 수렴점이 생기게 되어 영의 시차를 가진다[10].

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



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

<그림 3> 양안 시차의 종류



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

<그림 4> Ross현상기반의 3D 변환원리

디스플레이 장치을 통해서 재현하고 상영하는 영상들은 2D 영상이다. 이러한 2D 영상을 가지고 3D효과를 내기 위하여 좌 영상과 우 영상으로 구성되는 3D 입체 영상을 생성하는 것이 3D 입체 변환의 기본 원리이다.

Ⅲ. 3차원 영상의 피로도 측정기술

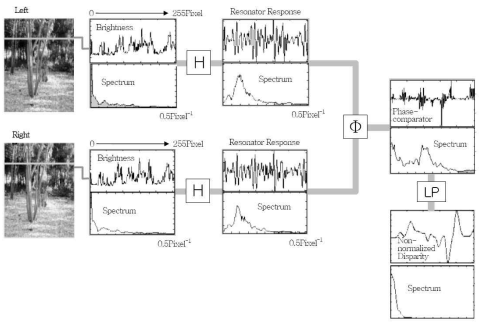
스테레오 영상과 비디오의 인지 화질을 측정하는 연구들이 많이 진행되고 있다. ITU는 주관적 화질 평가를 표준화하기 위해 모노_스코픽 텔레비전과 스테레오 텔레비전 영상을 위한 권고안을 가지고 있다. 스테레오 영상의 평가 방법, 촬영환경, 시청환경과 테스트 자료를 포함하는 평가 조건을 제시하였다. 이러

한 주관적 영상 화질 평가 방법은 시각 피로의 메커니즘은 아직 완벽히 밝혀진 바 없지만, 시각 피로 증상은 일반적으로 눈의 피로, 눈물, 안구의 압박감, 눈 주위의 통증, 눈을 깜박일 때의 불편함, 안구의 뜨거움, 초점 맞추기 어려움, 흐린시야, 굳은 어깨, 그리고 두통 등이 있다. 또한, 객관적인 피로도 측정 방법으로서 조절 능력, 시력, 동공의 직경, 중요 합성 주파수, 안구 이동 속도, 시각피로도의 주관적 등급 그리고 작업능력과 같은 시각 피로도를 측정하는 방법들이 있다[14].

뇌파와 ERP 신호를 분석하여 시각 피로를 객관적으로 평가하는 방법안을 측정영상 그래프로 보여주고 있다.

카메라 간 거리와 수직 변이를 파라미터로 사용하는 영상을 이용하여 주관적 화질 평가를 시행하여 얻은 결과를 분석하고, 새로운 영상이 입력되었을 때 주관적 화질 평가를 수행하지 않고, 영상 분석을 통하여 정확하게 예측할 수 있는 객관적 피로감 예측이다. 카메라 간 거리가 조절 되었으며, 카메라의 수직 정렬이 정확하게 이루어지지 않은 여러 영상을 만들기 위하여 3D Max를 사용한 실험 영상을 제작 하였다. 카메라간 거리는 등 간격 배열을 하여 4가지 경우의 수로 제작되었으며, 카메라의 수직 정렬 상태는 3가지 경우의 수로 제작하였다[7].

피로도의 주관적 실험 결과에 따르면 카메라 거리와 수직변이의 두 항목은 서로 독립적이며, <그림 3-1>에서 보이는 것처럼 피로감은 주로 카메라간 거리에 영향을 받음을 알 수 있다. 피로감 정량화를 위해 사용되는 영상 분석 결과는 SIFT를 통해 추출된 변이의 분포와 수직변이를 이용한다. 주관적 평가 결과를 이용하여 선형 회귀식을 풀게 되면 변이의 분포와 수직변이의 가중치를 얻을 수 있으며, 새롭게 피로도를 평가 할 영상이 입력되면 획득했던 가중치를 곱하여 피로도를 예측할 수 있다. 스테레오 피로도 측정 방법은 주관적 평가를 거쳐 정량화된 주관적 화질 지수에 대비해 상관도를 산출할 수 있게 된다[4].



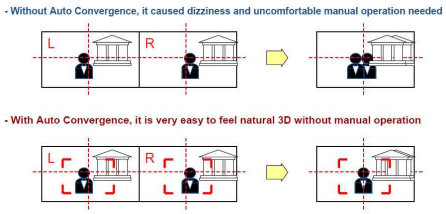
<그림 5> 3D 영상의 피로도 측정영상

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.

3.1 영상합성 모델링 기법

3D 영상합성은 합성방법에 있어 고려해야 할 사항들이 많다. 자신이 원하는 장면을 합성하기 위해 철저한 계획에 따라 진행되어야 하는데 합성의 소스가 3차원이라면 기존의 영상 합성 방법인 마스크 기능을 이용한 레이어의 합성 또는 알파 채널을 이용한 방법이외에 추가적인 방법이 필요하게 된다. 이것은 3D 3차원 공간상에 존재하는 카메라의 움직임 정보들이다[5].

보편적으로 많이 쓰이는 방법은 모션컨트롤 카메라를 이용하여 합성될 영상들의 카메라 모션 정보를 동일하게 촬영하여 합성하는 것이다. 이 방법은 정확한 카메라 모션 정보를 사용하기 때문에 합성시 자연스러운 합성 영상을 얻을 수 있지만 고가의 장비를 사용하기 때문에 많은 비용이 들게 된다. <그림 6>에서 보이는 것은 3D툴을 통해 생성된 3D 모델영상과 디지털 소스 영상의 합성 방법이다[2].



<그림 6> 3D 영상합성 이미지 추출

일반적인 3D 툴을 이용하여 모델링을 하고 디지털 영상 소스의 카메라 모션 정보를 트래킹 기법을 통해 추적한 카메라 정보를 이용하여 랜더링한 영상과 디지털 영상 소스와 합성하게 된다. 이때 기존의 카메라 트래킹 데이터를 추출하는 기법을 사용한다[6].

<그림 7> 3D 디지털 영상 합성 이미지

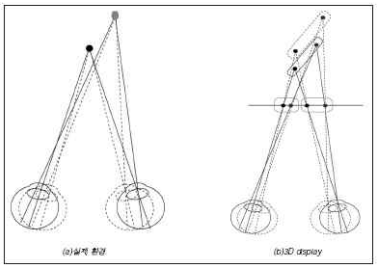
대표적인 트래킹 툴로는 Realvizdml의 Matchmove와 2D의 Boujou가 가장 많이 사용되고 있다. 기존의 방법들은 카메라 모션 정보들이 존재하는 영상들을 사용하기 때문에 합성하기가 쉽다. 그러나 카메라의 모션 정보가 존재하는 디지털 영상 소스와 2D영상의 합성은 개념적으로 불가능하다[14].

3.2 휴먼팩터 연구

이상적인 휴먼팩터 기술은 실제 3D 공간을 볼 때처럼 3D 공간의 깊이와 대상의 크기가 실제대로 지각되고 피로를 유발하지 않으며 실감이 있어야 한다. 이러한 휴먼팩터 기술이 실현되기 위해서는 다음과 같은 3가지 기술 요소들이 선행되어 져야 된다.

첫째, 3D 방송 시스템 설계에 필요한 휴먼팩터 기술로서 이는 휴먼팩터가 기존 시스템의 구조만으로 해결하기 어려운 근본적인 문제일 때 개발 도중에 수정을 하는 것보다는 설계 단계에서부터 휴먼팩터를 측정하거나 고려하여 시스템 기능이나 사양을 정하는 것이 비용도 절감할 수 있다[5].

둘째, 휴먼팩터 환경 파라미터 규명 및 이에 대한 감성 공학 기술로서 시청자가 편안하게 시각 피로 없이 3D 방송을 시청할 수 있도록 하기 위한 최적의 시청 환경에 대한 연구이다.



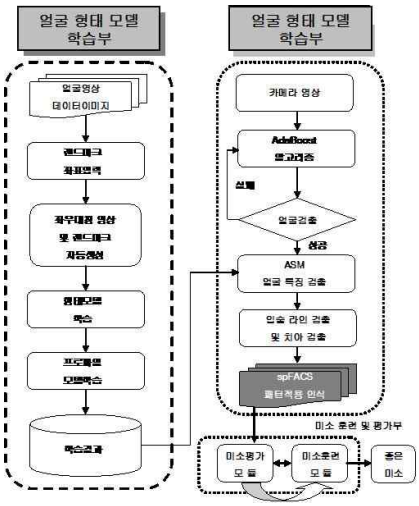
<그림 8> 실제환경과 3D디스플레이 동공간 거리의 영향

셋째, 휴먼팩터 기술에 대한 심리학적 평가 기술로서 이상적 시스템을 개발하기 위한 여러 단계를 거치면서 각 단계마다 개발된 시스템의 완성도를 시청자들에게 제시하여 개발된 시스템의 심리학적 성능을

효과적으로 평가하기 위한 연구이다. 본 논문의 핵심 내용인 3D객체인식의 휴먼인지 기술을 이해하기 위해서는 위의 3가지 기술요소를 적용기술로서 충분히 활용해야 된다.

Ⅳ. 다면적 얼굴검출 SPFACS 설계연구

4.1 설계연구 및 절차

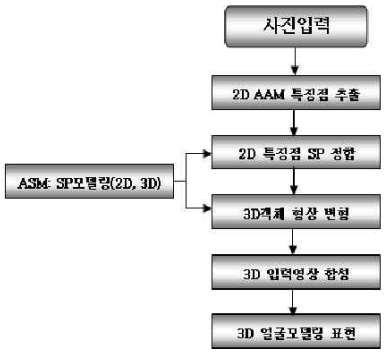


<그림 9> 시스템 구성도

본 연구는 다면적 영상 SPFACS(Smile Progress Facial Action Coding System)객체 인식기반 AAM (Active Appearance Models) 알고리즘을 이용한 미소 검출인식 연구를 통하여 안면근의 복합적인 근 운동에 의해 나타는 즐거움, 기쁨, 찬성, 조소 또는 어떤 다른 감정들을 표현하는 웃음으로써, 눈을 밝게 하고 소리를 내지 않으며 입꼬리를 상방으로 올리는 얼굴 표정의 변화이다.

미소는 눈의 반짝임과 얼굴의 아래에 있는 입술 주위의 근육 작용에 의해 표현된다. 얼굴 표정에는 표정 근육이 많이 있는데, 주로 눈과 입 근처에 몰려 있다. 입은 아주 중요한 표정 표현 기관으로, 매력적인 미소뿐만 아니라 심술을 나타내는 불룩한 입술, 공포를 나타내는 파르르 떠는 입술, 의지를 나타내는 꽉 다문 입술 등 다양하게 자신의 심리상황을 상대방에게 전달 할 수 있는 중요한 표현 수단인 것이다.

본 연구에서는 다면적 얼굴 모델링을 기초로 AAM(Active Appearance Model)알고리즘 SPFACS 얼굴표정 분석 패턴기법을 적용하여 컴퓨터에서 카메라 영상을 입력받아 얼굴의 특징 정보를 이용하여 입술 라인과 개별 치아를 추출하여 미소를 분석한다. 추출된 입술라인과 개별치아의 관계를 이용하여 미소를 분석함으로써 미소가 잘된 미소인지 바로 알 수 있게 자동으로 미소 평점 결과를 보여 주고, 사용자가 스스로 좋은 미소를 짓기 위해 훈련을 할 있는 시스템을 연구하였다[13].



<그림 10> SPFACS 다면적 얼굴모델링 과정

또한 얼굴 표정인식 기술로 많이 쓰고 있는 FACS (Facial Action Coding System) 감정표현 방식에서 미소인식기술을 분석 적용함에 따라 Adaboost와 같은 Boosting기법을 통하여 고화질의 AAM, ASM, LDA, ICA, PCA등 홀리스틱 분석기법의 보다 인식률이 정확하고 고화질의 얼굴인식 연구와 다양한 애플

리케이션이 개발 연구가 필요하다.

4.1 연구내용

본 연구는 AAM 알고리즘을 이용하여 제안된 얼굴 표정 분석 기법으로 얼굴의 정확한 윤곽 데이터를 추출하며, 인체 해부학적 토대로 사람의 얼굴 근육을 분석하여 웃음을 통하여 나타내는 86개의 Landmak Action Unit으로 미소 검출 인식시스템을 설계 하였다.

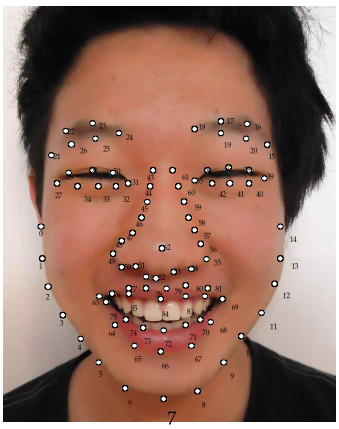
본 연구에서 실험, 분석한 각각의 Action Unit은 객체인식을 통한 미소검출 감정 표현으로 움직이는 얼굴 근육에 대한 모사로 SPFACS의 경우 다수의 얼굴 표정을 표현하고 있는 얼굴 이미지에 대하여 사람이 직접 수동으로 Action Unit을 비교, 분석하여 어떤 표정을 표현하고 결정하는 SPFACS이 자동으로 동작하는 Human Computer Interface상에 적용된 연구사례는 국, 내외에서 현재까지 연구된 적이 없는 것으로 확인 되었다.

본 연구에서는 이러한 부분에 주안점을 두고 SPFACS을 자동화시스템(Automation)에 적용하여 사람의 얼굴전체의 표정을 분석하기 보다는 미소에서 표출되는 입술과 치아의 표정에서 차별성을 두어 검출효과를 얻을 수 있는 시스템을 연구하고자 하는 것이다[13].

4.2 SPFACS 영상 패턴분석

본 연구에서는 다면적 영상 기법을 적용하여 실시간 얼굴 특징 검출방법을 나타내고 있으며, 자동 미소 분석 시스템은 얼굴 형태 모델 학습, 얼굴 특징 검출과 미소 분석 및 평가 모듈로 구성되어 있다. 얼굴의 특징을 검출하기 위해 입력된 영상으로부터 Adaboost 알고리즘을 이용하여 얼굴을 검출한다. SPFACS 패턴으로 검출된 얼굴 영상에 대하여 AAM 기반으로 학습된 얼굴 형태 모델을 이용하여 얼굴의 특징(얼굴윤곽, 눈, 눈썹, 코, 입, 입술 등)을 검출한다[9].

미소 분석과 미소 평가를 하기 위해서 입술영역에서 다시 치아 영역을 검출하고, 검출된 치아 영역과 입술 라인의 관계를 이용하여 미소를 분석하고 평가하는 결과를 사용자에게 실시간으로 보여주게 된다. Landmak의 수는 <그림 11>같이 86개로 설정하였으며 얼굴의 윤곽과 눈, 눈썹, 코, 입술을 표현하였다. 얼굴 특징 형태를 이루는 Landmark 86개의 번호는 얼굴 윤곽 0-14번, 우측눈썹 15-20번, 좌측눈썹 21-26번, 좌측눈 27-35번, 우측눈 36-44번, 코 45-59번, 입과 입술은 60-84번, 앞니 85-86번으로 지정하였다[12-13].



<그림 11> SPFACS Landmark 얼굴모델링

얼굴표정에 대한 Landmark Shape 변화의 표시된 영상들에 대해서는 SPFACS 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[6].

4.3 적용방법 및 절차

본 연구에서 사용되는 SPFACS의 특징기반 형태 모델은 얼굴의 특징 요소에서 Landmark의 연결로 표현된다. 이 방법은 형태 변화를 통계적으로 모델링 하는 방법으로 특징 기반 모델은 얼굴의 형태를 특징 점들의 집합으로 표현되고, 초기 위치를 얼굴 영상에 점진적으로 정합한다. 얼굴 형태 모델을 통계적으로 모델링하기 위해서는 다양한 데이터가 필요하다. 학습 영상들을 평면에 표출하기 쉽게 하기 위해서 유사 변형에 의해 2차원 형태로 주어진다. 각 특징점들은 동일한 공간에서 얼굴 형태의 경계선을 따라 정렬되게 되며 (식 1)와 같이 86개의 랜드마크 좌표 벡터로 표현된다. 여기서 x_n, y_n은 물체를 구성하는 n번째 좌표 값을 나타낸다.

$$X_i = [x_1,\ x_2,\ \cdots,\ x_n,\ y_1,\ y_2,\ \cdots,\ y_n]^T \qquad \text{(식 1)}$$

학습 영상에 있는 얼굴 특징 형태는 크기와 방향등의 새로운 좌표값들의 집합을 갖게 되며 이것은 차원이 커짐에 따라 그 양도 증가한다. 이러한 문제를 해결하기 위해 대표하는 하나의 모델을 구성하여 유사한 크기, 방향을 갖고 무게중심의 좌표가 동일하도록 확대/축소, 이동, 회전 변환하는 선형 기하학적 변환을 통하여 정규화 한다. 정규화된 벡터들의 평균 백터 $\overline{X}$를 (식 2)과 같이 구한다. 평균 벡터를 구한 다음에는 각 훈련 영상에 대한 벡터들이 평균 벡터로부터 얼마나 벗어나 있는지 나타내는 변위 벡터들을 (식 3)와 같이 계산한다.

$$\overline{X} = \frac{1}{N}\sum_{i=1}^{N} X_i \qquad \text{(식 2)}$$

$$dX_i = X_i - \overline{X} \qquad \text{(식 3)}$$

변위 벡터들을 구한 다음에는 이들의 공분산 행렬을 (식 4)와 같이 구한다.

$$S = \frac{1}{N}\sum_{i=1}^{N} dX_i\, dX_i^T \qquad \text{(식 4)}$$

공분산행렬을 구한 다음에는 공분산 행렬에 대하여 (식 5)을 만족하는 고유 백터와 값을 구한다. 이 식 P_i는 고유 벡터이고 λ_i는 고유값이며, $\lambda_i \geq \lambda_{i+1}$이 만족되도록 고유값을 정렬한다.

$$SP_i = \lambda_i P_i \qquad \text{(식 5)}$$

이렇게 정렬된 얼굴 특징 형태는 주성분분석을 이용하여 (식 6)과 같이 나타 낼 수 있다.

$$X \approx \overline{X} + Pw \qquad \text{(식 6)}$$

여기에서, $P = (P_1, P_2,\ \ldots, P_k)$ 는 형태 주성분석을 통해 계산된 고유벡터, 형태 모델을 구성하는 기저벡터들로 이루어진 행렬이며 $w = (w_1, w_2, \ldots, w_k)^T$는 각 고유벡터의 가중치를 나타내는 벡터이다. 형태 파리미터 벡터 ω를 조절함으로써 X 를 합성할 수가 있게 되며, 고유벡터를 통한 합성특성은 정합 알고리즘의 토대가 된다.

4.4 표본추출

본 연구의 얼굴검출 미소분석 실험을 위하여 수도권에 거주하고 있는 40대 미만의 성인 남녀 80명의 모집단을 선정하여 개인별 영상데이터를 촬영하여 데이터베이스 영상으로 구성하였으며, 실험자에 대해 10장씩 800장의 영상을 촬영하였다. 각 학습영상의 크기는 720 x 540으로 설정 하였으며, <그림 13>와

같이 좌측(a), 정면(b), 우측(c)의 세 방향으로 영상 촬영 표본데이터를 얻어 시험에 적용하였다[12].

<그림 12> 표본추출 spFACS 3면 영상 촬영

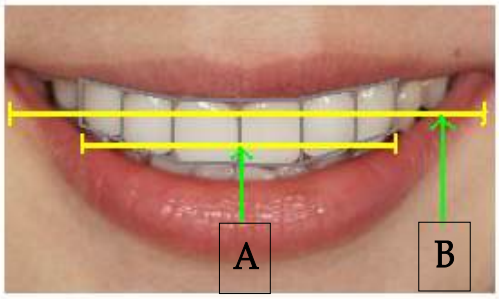
기존의 학습 영상 데이터베이스에는 미소 영상이 충분하게 포함되어 있지 않아 미소 훈련 학습 영상으로 활용하기에는 부적합하였다. 따라서 본 논문에서는 미소 영상을 획득하기 위해 50명을 촬영하여 데이터베이스 영상으로 구성 하였으며 실험자에 대해 10장씩 500장의 영상을 촬영하였다. <그림 13>와 같이 좌측(a), 정면(b), 우측(c)의 세 방향에서 촬영하였다[13].

<그림 13> SPFACS 입술표정 영상추출

<표 3> 표본추출

구분	표본 추출 방법	비고
대상	성인남녀 피험자 80명	회사원(25) 대학생(30) 고교생(5) 치과 외례환자(20)
연령	만 10-40세 연령대별 각 5명씩	
표본영상	실험자 당 10장씩 800장	
영상크기	720 x 540	디카/800만화소(이상)
표본추출영상	좌측, 정면, 우측	(a), (b), (c) 세방향
표정특징	웃음표정, 치아영역	정상치아보유자 틀니착용자 제외
성별 구성비	남녀 25 : 25	연령대별 5 : 5
치과적 특성	정상치아 보유자	충치환자, 보철착용자 틀니착용자 제외

4.5 실험



<그림 14> 검출된 치아영역

4.6 분석방법

다면기법 입체영상 SPFACS 객체인식을 통한 미소검출 실험은 피험자 개인당 실제 미소인식을 위한 영상 촬영 측정에 걸리는 시간은 평균 18분정도 소요되나 실험 개요에서 설명 하였듯 실험자 한 명당 얼굴의 좌측, 정면, 우측 3면촬영등 실험준비 시간을 포함하여 약 25분이 소요 되었다. 미소검출 인식은 추출된 입술라인과 치아 영상데이터 촬영 그리고 웃음표정에서 나타내는 치아의 상태 데이터를 수집하기 때문에 피험자 80명 개개인에게 본 연구의 실험에 대한 충분한 이해와 영상검출 3D회전 파노라마 영상 촬영을 전체피험자 80명에게 하루 10명 5일간 30분씩 총 2시간 30분 동안 영상 촬영시험을 하였다.

<그림 15> SPFACS 미소검출 영상 촬영장면

4.7 실험내용

다면기법 영상 SPFACE X-Ray 인식처리장치 DX-3000장비를 통하여 피험자의 모든 치아를 검출하지 않고 상악 치아 중에서 비교적 검출이 용이한 가운데 8개의 치아를 검출한다. 검출된 치아 영역에서 개별적인 치아를 찾기 위해 먼저 3x3 크기의 수직 경계선 검출용 마스크를 사용하여 가로 방향으로 2차 미분을 수행함으로써 세로 방향 경계선을 구한다. <그림 16>(a)는 2차 미분을 적용한 영상 결과를 보여주고 있다[11].

(a) 가로방향 2차 미분결과 영상 (b) 세로방향 2차 미분결과 영상

<그림 16> 실험으로 측정된 2차 미분 결과 영상

본 연구에서 제안하는 SPFACE 객체인식 미소 분석 애플리케이션은 입술 라인과 드러난 치아 모습을 사용하여 미소 정도를 측정한다. 웃음표현시 치아가 노출될 때 윗 치아들이 상순과 하순 사이에 보이며, 상순은 입꼬리가 위로 올라가거나 직선이며, 상악 전치의 절단연이 하순과 평행하며, 윗 치아 들이 8개 이상 드러나 보이는 경우의 미소가 더욱 매력적인 것으로 평가하고 있다.

<표 4> 미소검출 평가 기준

Land Mark	Point No	입술라인에 따른 항목별 평가 기준	
		평가 기준	점수
CLow / 81		**○ (상순(윗니)하연의 중심점)** 조건: 입술라인과 가운데 치아의 윗경계선과의 사이가 세 화소이하이고 치아의 높이가 너비보다 크면	1.5점
LCh / 84		○ (양 입술 끝점(입꼬리) 조건: 거리가 세 화소 보다 작고 치아 높이가 너비보다 크지 않으면	0.5점
RCh / 78			
CLab / 75		**○ (하순의 중심점)** 조건: 거리가 세 화소 보다 크면 입술라인이 올라가 있어 윗 잇몸이 많이 보여지면	1.0점

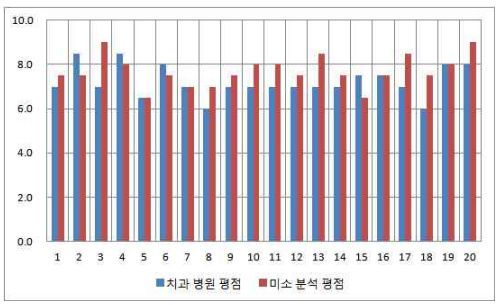
본 연구의 실험대상자는 실험군(N=50)의 연령별로 10대, 20대, 30대 40대 모두 각 25%이며, 청소년, 성인 남녀등 다양한 연령대의 피험자를 대상으로 한 검출 방법을 채택해서 준비했으며, 시험 또한 각 50%의 성별과 연령에서 고르게 분포하고 있다.

(a) 윗방향(1.5점) (b) 중간(1.0점) (c) 아랫방향(0.5점)

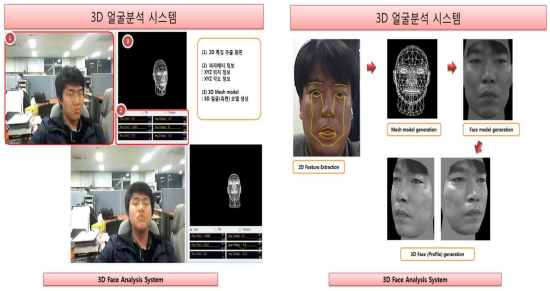
<그림 17> 치아검출 평가 기준

4.8 실험결과

본 실험은 피험자를 80명을 대상으로 시행 하였으며, 미소 분석 시스템을 적용함에 따라 신뢰도 계수는 ±0.892으로 나타났다. <그림 17>는 피험자를 통한 미소분석 시험 평가 결과를 나타낸 것이다[7].



<그림 18> 미소 평점 비교 결과표



<그림 19> 다면적 얼굴분석 시스템

<그림 19>은 다면적 얼굴분석 시스템을 나타낸 것이다. 다면적 기반의 얼굴인식 기술은 인식된 얼굴의 특징 정보를 추출하여 미소 훈련의 단계에서와 같이 미소 짓는 연습을 하게 되면 <그림 20>와 같이 각 단계의 미소 그림이 바뀌게 된다[14].

<그림 20> 객체인식 미소 시스템

실험분석 결과 치과 병원에서 전문가에 의해 수작업으로 평가한 평점과 미소 분석 시스템의 평가 평점의 평균 오차는 0.77점이었다. 만점이 9점이므로 평균 오차는 8.6%로서 치과 병원에서 이루어지는 수작업으로 분석한 결과와 제안된 자동 미소 분석 시스템을 이용하여 미소 분석한 것이 큰 차이를 보이지 않았다.

V. 결론

다면기법 영상 인식을 통한 AAM(Active Appearance Model)알고리즘 기반의 SPFACS(Smile Progress Facial Action Coding System) 휴먼인식 미소 검출 인식시스템 기술은 기존의 고정된 검출방식의 객체인식기술에 비해 실시간으로 움직이는 입체영상 깊이 정보가 추가됨으로써 얼굴 객체인식을 통해 얼굴 특징을 찾은 다음 미소 분석에 필요한 입술라인과 개별치아 영역을 추출하는 방법을 통하여 치아 인식기술을 보다 정확성을 적용한 방법으로 설계를 하였다. 이러한 영상객체 인식기술은 미소의 정도를 분석하기 위해 입술 라인과 치아와의 관계 판별이 필요하며, 얼굴검출에 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. 얼굴인식 기술에 대한 연구는 그동안 많은 분야에서 개발되어져 왔고 현재까지도 다양한 Adaboost알고리즘을 기초로 한 홀리스틱 분석방식을 이용한 연구가 다양한 분야에서 연구개발 되고 있다.

그러나 본 논문에서 연구한 미소의 정도를 분석하고 입술라인과 치아와의 관계 판별에 있어서 객체인식을 기초로한 2차 미분법을 적용하여 연구를 실행한 예가 없기 때문에 입술라인과 개별치아 영역에 대한 체계적이고 깊이 있는 알고리즘의 적용 분석을 통하여 휴먼팩터에 대한 연구와 SPFACS(Smile Progress Facial Action Coding System)의 고화질 기술을 적용한 HD급 실시간 객체인식 기술이 절대적으로 개발되며, 더불어 MPEG-4 기반의 H. 264를 기초로한 압축영상 기술을 탑재한 연구가 이루어 져야 될 것으로 판단된다.

참고문헌

[1] Koni "Improved Face Detection Using Adaboost Algorithm tilted," Papers 2008,06 Article 12 No. 3.

[2] Ieek "Using Adaboost for Iris Recognition in Mobile Environments A Study on Eye Detection," paper, 2008.07 papers 2008-45CI-4-1.

[3] Ieek" AdaBoost and circular multitarget classification technique based on the function Rox "thesis paper 2010.03 2010-47CI-3-3

[4] Journal of Korea Multimedia Society "Adaboost algorithm, real-time face detection and tracking using," 2006,10 pp. 1266-1275.

[5] Stephen Milborrow, Fred nicolls, "Locating Facial Feature with an Extended Active shape Model," Lecture in computer science, 2008.

[6] C. Padgett, G. Cottrell, "Representing faceimages for emotion classification," Advances in Nural Information Processing System, Vol9, MIT Press, 2011.

[7] Mark F. Bear, Barry Connors, and Michael Paradiso, Neuroscience: Exploring the Brain 3E, 2009.

[8] 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, 2010.

[9] P. Ekman, W. V. Friesen. J. C. Hager, Facial Coding System: Manual, CD Rom, Sanfrancisco, CA, 2009.

[10] 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, 2007, pp. 843~848.

[11] Yongmian Zhang, Qiang Ji "Facial Expression Understanding in Image Sequences Using Dynamic and Active Vision," Proceedings of the Ninth IEEE International Conforence on Computer Vision(ICCV 2008), 2-Volume, Set 2008.

[12] Kipo, "3-dimensional imaging technology," 2010 Technology Trends survey report, electrical/ electronic field Article 1, 2010.11, pp. 80~82.

[13] S. Pastoor and M. Wöpking, "3-D Displays: A review of current technologies," Displays, Vol. 17, 2008, pp. 100~110.

[14] N. Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," SMPTE Journal, Dec. 2010, pp. 1137~1144.

■ 저자소개 ■

최 병 관
(Choi Byungkwan)

2015년 8월~현재
가톨릭대학교
미디어기술콘텐츠학과 부교수
2013년 2월 호서대학교 융합미디어학부
(공학박사)
1998년 2월 홍익대학교 컴퓨터공학과
(이학석사)
1991년 2월 서울과학기술대학교 전산학과
(이학사)

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: 2015년 8월 5일
수 정 일: 2015년 8월 18일(1차)
2015년 8월 28일(2차)
게재확정일: 2015년 9월 4일