

인공지능 객체인식에 관한 파라미터 측정 연구*

최 병 관**

A Study On Parameter Measurement for Artificial Intelligence Object Recognition

Choi, Byung Kwan

〈Abstract〉

Artificial intelligence is evolving rapidly in the ICT field, smart convergence media system and content industry through the fourth industrial revolution, and it is evolving very rapidly through Big Data. In this paper, we propose a face recognition method based on object recognition based on object recognition through artificial intelligence. In this method, We were experimented and studied through the object recognition technique of artificial intelligence. In the conventional 3D image field, general research on object recognition has been carried out variously, and researches have been conducted on the side effects of visual fatigue and dizziness through 3D image. However, in this study, we tried to solve the problem caused by the quantitative difference between object recognition and object recognition for human factor algorithm that measure visual fatigue through cognitive function, morphological analysis and object recognition. Especially, The new method of computer interaction is presented and the results are shown through experiments.

Key Words : Artificial Intelligence, Object Recognition, Parameters, Human Factor

I. 서론

인공지능은 4차 산업혁명을 통해 객체인식과 패턴 인식, 형태소분석, 이미지프로세싱(Image Processing) 등 빅데이터(Big Data)를 기반으로 중요 핵심기술로 자리를 잡아가고 있다. 특히 포스트휴머니즘을 통한 인문학분야에서의 인공지능은 스마트디바이스 ICT 융,복합 산업과 더불어 다양한 디바이스와 플랫폼등

각기 다양한 인식 기반의 파라미터 연구가 활발히 연구를 되고 있으며, 객체인식 기반의 안면 인식은 사람의 얼굴과 표정 그리고 눈의 동공을 통해서 인식할 수 있을 만큼 다양한 기술이 개발되고 있다. 또한 양안 시차를 이용한 인지 기술은 얼굴 인식을 기반으로 객체 인식을 통해 식별할 수 있도록 연구 개발이 활발히 진행되고 있다. 이러한 연구는 휴먼팩터 인지 파라미터 값의 새로운 알고리즘을 통해 인위적으로 양안시차 각을 발생시켜 트랜스휴머니즘(Trans Humanism)을 기초로 하는 사람의 정신적, 육체적 능력을 인공지능을 통해 인식할 수 있는 인지 기술을

* 본 연구는 과학기술정보통신부 및 정보통신기획평가원의 SW 중심대학지원사업의 연구결과로 수행되었음(2019-0-01219)

** 연세대학교 컴퓨터정보통신공학부

적용, 실험을 통해 보다 다양한 실증연구에 대한 연구 결과 사례들을 보여주고 있다.

인공지능 분야에서 객체인식 연구는 기존의 3D영상을 통해 많은 연구가 진행되고 있었으나 3D영상을 통해 나타나는 시각적 피로감과 어지러움증 발생등 양,안시차를 통해서 나타나는 부작용이 해결되지 않고 연구가 진행됨에 따라 많은 문제가 야기되고 있다. 따라서, 본 연구는 인공지능을 통해 패턴인식과, 객체인식, 휴먼인지 기술등 파라미터의 시각적 요소와 피로감등 양,안시차를 통해 발생되었던 문제점들에 대해 객체인식에 관한 연구와 휴먼팩터를 통한 얼굴인식 요소에서 나타났던 문제점들을 보다 쉽고 사용자가 편리하게 인식하고 접근할 수 있는 방법론과 파라미터 측정에 관한 연구를 실험을 통해 측정값을 확인하고 연구 하였다.

또한, 객체인식 기술은 입체 영상과 무안경 디스플레이를 비롯하여 영화, 방송, 공연, 의료, 게임, 애니메이션 등 첨단 스마트미디어를 기반으로 광범위하게 사용되고 있음으며, 인공지능을 통한 패턴인식 기술이 급속히 성장하고 많은 분야에서 사용되고 있음에 따라 3D와 HCI(Human Computer Interaction) 인간과 컴퓨터 상호작용 입체영상 ICT융합 산업으로 급속히 성장하고 있다. 객체인식 분야는 휴먼팩터의 파라미터 측정을 통한 디스플레이와 양안시차의 좌/우안용 영상을 각각의 눈에 입력해 주고 양안시차 원리를 이용하여 사람의 인지정보 체계에서 3D로 인식하도록 설계를 하였으며, 인공의 시각차를 발생시켜서 좌/우안 영상을 디스플레이에서 분리시켜 두 눈에 전달함으로써 뇌에서 패턴인식을 통한 객체인식 체계를 느끼도록 설계를 하였다[1].

일반적으로 2D영상은 좌/우안의 각도 차이 없이 카메라 리그를 이용하여 획득한 영상데이터를 나타내도록 제작되며 디스플레이에서 좌안/우안용 영상을 좌/우에 분리시키는 장치(특수 안경, 무안경식 모니터)없이 두 눈에 전달하게 된다. 이렇듯 두 눈에 전

달되는 입력 영상은 뇌에 전달되어 사람의 인지, 정보 체계를 통해 인식하게 되는데, 3D 입체영상의 경우 좌/우(안) 양안시차를 통해 나타내는 영상이 서로 바뀌어 전달되거나 두 영상 사이의 깊이(Depth)감이 매끄럽지 못할 경우 입체감 인식이 떨어지거나, 시각적 효과의 불편함으로 시청자가 어지러움과 구토등 부작용이 발생할 수 있다.

II. 휴먼팩터 기술요소

본 연구에서는 인공지능을 통한 휴먼팩터 인식기술을 제대로 반영하기 위해서 HVS(Human Visual System)를 이해하고 3차원 영상이미지 시스템에 대한 이해와 휴먼팩터 측정을 통해서 나타나는 문제의 원인을 정확하게 진단하고 그 해결 방안을 제시해야 할 필요가 있다. 이렇듯 사전에 발생 할 문제점을 미리 예측하고 그 대안으로 해결할 수 있는 요소들은 다음과 같은 세부 연구를 선행적으로 연구가 이루어져야 한다[2].

첫째, 인공지능을 이해하고 패턴인식의 알고리즘을 구현하기 이전 시스템 설계 단계에서 부터 휴먼팩터를 고려해야 된다. 인공지능 객체인식의 기술 요소의 모든 경우에 그런 것은 아니지만 일단 시스템의 하드웨어 개발이 일정수준 궤도에 오르고 나면 추가적으로 휴먼팩터를 고려하기 위한 시스템을 수정하기가 곤란하다. 원인은 휴먼팩터 측정값의 결과 값에 대해 기존 시스템의 구조만으로 해결하기 어려운 근본적인 문제일 때 더욱 그러하다. 따라서 개발 도중에 수정을 하는 것 보다는 설계 단계에서 부터 휴먼팩터 측정 객체인식과 연계하여 시스템 기능이나 사양을 정하는 것이 비용절감 효과를 얻는데 최선이다[1].

둘째, 패턴인식을 통한 파라미터의 감성공학 기술로서 시청자가 시각피로 없이 휴먼팩터를 인식하고 영상인식 기술을 이용할 수 있도록 하기 위한 최적의

<표 1> 휴먼팩터 연구사례

구분	내용
HVS 정보처리	패턴인식 정보를 인식하여 양안시차를 인지할 수 있는 HVS를 설계하여 사람의 시각 변이가 0으로 되는 호립터에 관한 연구, 교차변이에 대한 연구를 말한다.
영상객체 상호관계 연구	주시하고 피사체에 양쪽 눈을 회전시켜 시선을 모으는 주시각 조절과 현재 물체의 위치 지점에 안구의 두께를 조절하여 눈의 초점을 주시하고 있는 물체에 맞추는 초점 조절이 연동되어 이뤄진다. .
시각피로/원 인규명 및 해결책 연구	3차원시스템에서 발생하는 대표적인 시각 피로 유발 현상에는 초점과 주시연동과 CB, Cross 현상 등이있다. 실제물체를 볼 때 주시각 조절이 피곤함을 느끼지 않고 3차원 깊이를 인식한다.
양안시차 변화민감도 연구	사람이 감지할 수 있는 양안시차, 시점변화에 대한 연구를 수행, 민감도 범위 내에서 해상도의 양안영상을 이용하여 영상을 하고, 고해상도를 가지는 영상과 거의 같거나 약간 저하되는 정도로 인식한다.

시청 환경에 대한 연구이다.

셋째, 객체인식과 휴먼인지에 대한 심리학적 평가 기술로서 이상적 시스템을 개발하기 위한 여러 단계를 거치면서 각 단계마다 개발된 시스템의 완성도를 사용자에게 제시하여 개발된 시스템의 심리학적 성능을 효과적으로 평가하기 위한 연구이다. 따라서 이러한 관점에서 본 논문에서는 중요시 하고 평가할 수 있는 핵심 원인과 인공지능 휴먼인식 파라미터 측정을 보다 더 확실히 이해하고, 제시한 세 가지의 기술 요소에 대한 자세한 상세 사항에 대해 알아보기로 한다.

2.1 객체인식 HVS시스템 연구

인공지능 객체인식 분야에서 시각 정보를 인식하고 시스템을 설계하고 처리하는 과정의 연구를 통하여 시스템 사용자가 영상인식 정보를 인지할 수 있는 깊이 와 특성 등을 이해함으로써 시스템의 구조를 사람의 시각 시스템과 유사하게 설계할 수 있다. 이를

위해 변이가 “0”으로 되는 호립터(horoptor)에 관한 연구, 교차 변이(crossed disparity)및 미교차 변이(uncrossed disparity)의 인식 범위 등에 대한 연구를 예로 들 수 있다. 1990년 발표된 Yeh 등의 논문에 의하면 사람이 융합 할 수 있는 변이의 범위는 비 교차 변이의 경우에는 24.0 min arc(minutes of arc: 1/60 degree)이고, 교차 변이의 경우에는 27.0 min arc까지 융합이 가능하다[3].

2.2 휴먼팩터 Depth Cue 연구

휴먼팩터 객체인식을 위한 단서들 양안변이(Binocular Disparity), 운동시차(Motion Parallax) 등의 광학정보(Optical Information)와 사용자의 눈이 행하는 초점 조절(Accommodation), 주시각 조절(Convergence) 등의 안과적 정보(Ocular information)에 대한 동작원리를 이해함으로써, 기존의 입체 영상 시스템에서 구현되지 못한 문제점을 분석하고 실제 인공지능 객체인식에 가까운 시스템을 개발할 수 있다[3].

인공지능을 통한 객체인식은 빛에 의한 반사, 그림자에 의한 효과, 가깝게 있는 물체가 크게 보이는 상대적 크기, 다른 물체에 의한 중첩, 가까운 텍스처가 더 선명하게 보이는 텍스처 변화, 멀리 있는 물체가 흐릿하게 보이는 공간적 원근감, 가까운 물체가 빨리 지나가게 보이는 운동시차, 조망(perspective) 등이 있다. 또한 실제 세계에서 사람이 인공지능 객체인식을 통해 감지 할 수 있는 깊이를 인식하는 과정은 주시하고 있는 물체에 양쪽 눈을 회전시켜 시선(line of sight)을 모으는 주시각 조절과 현재 물체의 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 벗어나는 전경이나 배경에 존재하는 물체들은 블러링(blurring)되어 보여지며, 초점 및 주시각 조절

연동 관계를 확인할 수 있다[3].

<표 2> 객체인식 파라미터 권장범위 값

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

2.3 휴먼팩터 실감영상 피로 연구

양안식 입체 영상 시스템에서 발생하는 대표적 시각적 피로의 유발 현상에는 초점과 주시 연동파괴 (accommodation convergence breakdown), 혼신 (crosstalk) 현상 등이 있다. 초점과 주시 연동 파괴 현상은 앞에서 기술한 바와 같이 실제 세계에서 우리가 물체를 볼 때 초점과 주시각 조절이 연동되어 피곤함을 느끼지 않고 3차원 깊이를 인식한다. 그러나 기존의 양, 안식 입체 영상 디스플레이를 통하여 영상을 보게 될 경우에는 큰 시차에 의해 초점 및 주시 연동 파괴 현상이 발생한다.

즉, 우리 눈은 디스플레이 화면상의 평면(on the plane of the screen)에 초점을 맞추는 반면, 화면상의 시차가 생성하는 3D 상의 위치에 주시(converge)하게 되어 이 둘의 위치가 서로 일치하지 않게 된다. 그리고 디스플레이 되는 영상 중에 사람 눈의 초점 심도 범위를 벗어나는 깊이를 가지는 부분이 있더라도 선명한 상태로 보여지기 때문에, 이 영역의 이중 영상이 눈을 피로감을 더욱 가중시키게 된다[4].

2.4 양안시차 시각적 피로 원인 연구

최근 입체영상 분야의 인식 기술로 개발된 3D객체인식 시스템은 Active Glass를 착용하고 입체영상을 시청할 경우, 시스템에 따라 약간씩 차이는 있으나 대부분 20분을 전후하여 영상 시청자가 눈의 피로도

와 불편할 정도의 시각적 피로, 두통 혹은 어지러움을 경험하게 된다. 이렇게 실제의 3D 세상을 경험할 때에는 나타나지 않았던 시각적 피로, 두통, 어지러움 등이 3D 영상 콘텐츠를 시청할 때 발생하는 원인을 측정하고 규명하며 이를 감소시키기 위한 여러 연구가 이에 속한다[3].

III. 인공지능 객체인식 응용 연구

인공지능을 통한 객체인식 시스템 이미지설계를 지능화 값에 적용하여 파라미터를 정의하고자 최적의 입체 영상 시청환경을 객체구조에 맞추어 설계하는 연구가 필요하다. 이를 위해 적절한 시청환경을 결정하는 휴먼팩터 인지기술의 파라미터 값을 정의하고, 이러한 변화를 통해 사용자들에게 미치는 심리적 영향을 파악해야 한다.

인공지능을 통한 객체인식의 시차는 보통 35min arc이내의 크기를 가지도록 제한해야 만 비교적 편안한 3D 영상을 산업기술에 적용하여 사용할 수 있다. 또한 초점 및 주시 연동 파괴 현상이 일어나지 않을 초점 범위는 $\pm 0.2D$ 이어야 한다. 인공지능 객체인식의 혼신은 양,안식 디스플레이 장치에서 좌,우 영상이 정확하게 분리되지 않음으로서 생기는 현상으로, 셔터링 안경의 영상 전환이 불완전 하거나, CRT 모니터 상에서 발광인자의 잔광 효과로 인해 나타나게 된다. 혼신의 인식 정도는 영상의 대조와 시차가 커짐에 따라 증가하는 경향을 보이며, 100:1 대조와 40 min arc 정도의 시차를 가지는 영상에서 0.3%이하이어야 한다[5].

이 밖에도 실제의 양,안 시차를 통한 입체영상의 깊이를 가지는 물체의 관찰거리가 너무 클 경우 평면으로 보이는 평판효과(cardboard effect) 각 대상이 실제 크기보다 작게 보이는 인형극장 효과(puppet theater effect) 등의 원인을 규명하고, 이를 제거할 방

안에 대한 연구가 이에 속한다. 이러한 연구들이 시 각 피로를 최소화하기 위한 3D 객체영상을 이용한 촬영 및 디스플레이 방법 등에 관한 기술 요소로 적용되고 있다.

3.1 심리학적 객체인식 평가 연구

인문학 관점에서 판단하는 영상객체 인식의 평가 기준은 인간과 기계의 경계를 넘어선 존재 즉 인간을 넘어선 초인간의 포스트휴먼 시대를 대비하기 위한 포스트휴머니즘(Post-Humanism)이 논의되고 있으며, 적용기준은 인공지능 관계소통학, 윤리규범학, 기술 비평학, 데이터해석학, 사회문화학, 그리고 심리학적 평가방법을 적용하여 인공지능을 통한 객체인식 분야의 인문학적 스토리텔링을 적용하여 평가기준을 정한다. 특히, 포스트휴머니즘을 통한 5가지의 아젠다는 인공지능의 핵심 방안으로 중점 연구되고 있다. 인공지능을 통한 객체인식은 3D 영상객체 시스템을 비롯한 일반적 시스템 개발에 있어서 완벽한 시스템을 구현한다는 것으로 거의 불가능하다는 기준을 정하고 있으며, 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[6].

따라서 인공지능 분야에서 객체 인식을 통한 심리학적 평가시스템은 공학적인 측면에서 기술적인 평가를 우선 평가항목으로 확인 할 수도 있지만, 객체 인식을 통해 개발된 시스템의 최종 사용자인 시청자들에게 얼마나 친화적인가에 기반하여 평가를 하여야 한다. 이를 위한 것이 3D 영상 시스템의 평가 방법 연구이며, 이에 는 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 인공지능 객체인식 평가 연구

인공지능을 이용한 객체인식 생리적 평가는 정신 물리학적 측정 방법과 심리검사 방법등 모두 기본적으로 실험 적용이 가능하다. 따라서 이러한 연구 기법은 연구자의 주관적인 관점과 자기보고서(self-report)에 기반을 두고 연구를 하게 된다. 이 방법엔 시청자가 의도적으로 이미지를 처리하고 조작하는 거짓 정보를 측정하고 감별하여 조작된 내용을 찾아내기 위한 기법들이 사용되며, 정신 물리학적 측정방법 및 심리검사 방법을 이용 하더라도 상당히 신뢰성이 높고 타당한 평가방법으로 연구되고 개발 할 수 있다[7].

따라서, 본 연구에서는 인공지능을 통한 객체인식 기술을 휴먼팩터 측정 연구를 통해 시청자의 시청 위치와 양안거리 및 위치를 추적할 수 있는 인공지능 포스트휴머니즘(AI Post-Humanism)실감 인지 기술에 대한 개념을 토대로 실험하고 연구를 하였다. 인문학적인 관점에서 영상객체 인식의 평가기준을 정하는 연구 주체는 인공지능 관계소통학, 윤리규범학, 생리학적 평가방법등을 구분하고 있으며, 특히, 의 과학 분야를 통해 집중적으로 연구되어야 할 뇌파 및 안구 운동의 변화를 3차원 영상을 통해 시청할 때와 시청하지 않을 때에 대한 각각의 측정값을 실험 함으로써 시청자가 시각적 피로와 두통을 느낄 때 뇌파와 안구 운동 변화에 어떤 영향을 주고 변화를 알아낼 수 있는지를 평가할 수 있다.

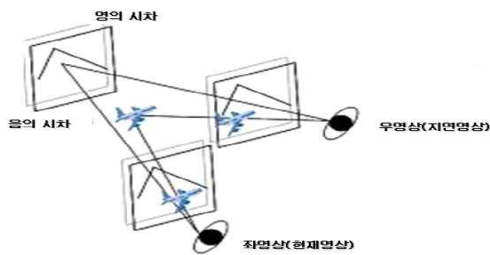
3.3 인공지능 휴먼팩터 인지 기술

인공지능을 통한 3D 휴먼팩터 얼굴인식 기술은 크게 안경식과 무안경식 두 가지의 패턴으로 분류하여 연구를 하게 되며, 3D 객체인식을 통한 Glasses 무안경식의 경우 시역을 형성시키기 위해 사용하는 도구의 종류에 따라 홀로그래피, 체적영상, 광학관식 등으

로 다시 세분화 해서 분류한다. 인공지능 휴먼팩터 인지 입체 영상의 경우는 방식에 따라 특수 안경을 착용하는 것과 착용하지 않고 시청이 가능한 방법이 있다. 이러한 객체인식 이미지 영상을 특수 안경을 착용하지 않고 시청할 수 있게 하기 위해서는 시청자가 응시하는 방향을 추적하는 기술이 필요하다. 따라서 양안 추적 기술은 휴먼팩터 측정을 위한 중요한 기술로 사용된다[7].

인공지능을 통한 3D영상 객체를 위한 양안의 응시 방향을 추적하는 기술은 헤드 트래킹을 할 수 있는 헤드셋을 착용하는 방법과 시청자의 눈을 실시간으로 추적하는 방법으로 구분할 수 있다.

본 논문에서는 인공지능 파라미터 패턴인식 기술을 적용하여 데이터 해석 알고리즘과 파노라마 카메라를 통해 움직이는 피사체를 촬영하여 사용자 콘텐츠를 인식하고, 시청자의 눈을 실시간으로 추적하는 방법에 대해서 인공지능 패턴인식 기술을 통해 실험을 하였다. <그림 1>의 실험에서는 눈 추적 알고리즘을 통해 눈의 위치에 대한 특징점을 추적하고 눈 동공을 파악한 후 눈 추적 알고리즘을 통해 눈동자의 좌표를 출력하는 방식을 적용하여 실험을 한다[8].



<그림 1> 인공지능 패턴인식 양안시차 원리

본 실험에서는 인공지능의 효과 값을 찾아내고 패턴인식을 통해 눈의 추적 알고리즘을 적용하여 눈 구조에서 나타나는 양의시차를 분석하여 동공 한 가운데에 위치한 홍채(Iris)와 흰자위(Sclera)의 경계선을

얼굴 영상에서 검출한 후 추적하는 방식을 사용한다.

또한, 본 실험을 통해서 눈이 응시하는 각도를 정확히 파악하기 어려울 경우, 눈 추적 각도가 작은 경우의 수를 파악하고 정확한 실험을 할 수 없는 문제를 가정하여 눈의 응시각도를 정확히 파악할 수 있도록 음의 시차와 양의시차를 각각의 알고리즘을 통해 얼굴 전체의 특징점을 찾아 각도를 측정하는 방식을 활용하였으며, 본 실험의 방법과 결과는 다음장 “**IV 인공지능패턴인식 얼굴검출**”에서 연구를 통해 실험 결과를 확인 할 수 있다.

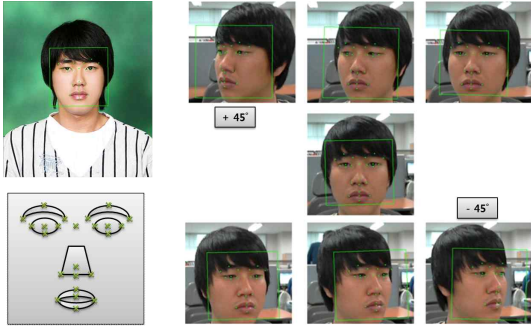
IV. 인공지능 패턴인식 얼굴검출

인공지능을 이용한 패턴인식은 얼굴검출(Face Detection)과 인지모듈을 통해 형태소 분석과 패턴인식의 지능화된 검출 효과의 정보를 얻어 낼 수 있으며, 특히 얼굴검출에 따른 형태소 분석과 패턴인식을 이용한 얼굴의 윤곽과 얼굴모형의 Landmark를 정함으로써, 이를 통해 휴먼인지 객체인식의 결과값을 실험결과로 확인할 수 있다.

우선, 인공지능 객체인식을 구성하기 위해서는 패턴인식을 통해 얼굴인식 시스템의 가장 중요한 요소로 분류할 수 있는 랜드마크 포인트(Landmark Point)를 정하고 이러한 분류를 통해 단계별 형태소 분석 기법을 적용함으로써 영상인식을 위해 얼굴이 있는 위치를 알아 낼 수 있도록 얼굴검출 모듈을 설계하고, 검출된 데이터를 실시간으로 인지해서 얼굴 검출기에 데이터 값을 등록 적용한다[8].

아래의 <그림2>는 AI를 통한 얼굴인식의 패턴 흐름을 찾아내는 인식시스템의 검출요소를 모듈별로 구성하여 찾아내는 방법을 나타내고 있다.

본 연구를 통해 실험하고자 하는 과정을 통틀어 “패턴인식 얼굴검출” 이라고 하는데 최근 실시간 얼굴 인식을 목적으로 다양하게 개발되고 있음에 따라



<그림 2> 인공지능 객체인식 검출 요소

실시간 응용 분야에서의 얼굴 영역을 인식하고 검출하는 과정이 더욱 더 중요하게 활용되고 있다. 사람의 얼굴은 응시하는 방향에 따라 정면 혹은 측면의 각도 고개를 좌,우로 기울이는 정도와 다양한 표정과 카메라와의 거리에 따른 얼굴 영상의 크기에 따라 같은 형태적 변화와 조명에 따른 얼굴 내에서의 발기 정도의 차이가 현격하게 차이가 있으며, 특히 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[9].

이러한 한계 요소로 얼굴 인식 분야에서 얼굴 영역 및 성분 검출 연구가 상당히 중요한 요소로 다루어지고 있으며, 초기 얼굴 인식 연구에서 전처리 단계로 취급되던 검출 부분이 하나의 독립적인 연구 분야로 매우 중요시 다루어지고 있는 실정이다.

4.1 인공지능 객체인식 알고리즘 구현

본 실험은 Adaboost face detection을 이용한 얼굴 인식 영상측정 흐름에 관한 프로토타입을 측정 실험을 통해 영상분류 값을 찾아내었다.

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

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

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

3) 반복 $t = 1, \dots, T$

(a) 비중의 표준화

$$w_{(t,i)} = \frac{w_{(t,i)}}{\sum_{j=1}^n w_{(t,j)}} \text{----- ②}$$

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

(b) 약한 분류기 (h_j)의 에러 ϵ_j

$$\epsilon_j = \sum_i w_i |h_j(x_i) - y_i| \text{----- ③}$$

(c) 분류기의 선택

가장 낮은 에러율을 ϵ_t 라 하고, ϵ_t 를 가지는 약한 분류기 h_j 를 선택.

(d) 비중 업데이트

$$w_{(t+1,i)} = w_t \beta_t^{1-e} \text{----- ④}$$

여기서, 입력영상 x_i 가 h_j 의 값에 따라

올바르게 분류가 된 경우 $e_i = 0$
 올바르게 되지 않은 경우 $e_i = 1$,

$$\text{그리고 } \beta_t = \frac{\epsilon_t}{1 - \epsilon_t} \text{ ----- ⑤}$$

4) 최종 강한 분류기

$$h(x) = \begin{cases} 1 & \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \text{ ----- ⑥} \\ 0 & \text{otherwise} \end{cases}$$

$$\text{여기서, } \alpha_t = \log \frac{1}{\beta_t} \text{ ----- ⑦}$$

최적의 3D 객체인식의 환경을 인지 하기 위해서는 시정 환경에 대한 별도 연구가 필요하다. 이를 위해 적절한 시정환경을 결정하는 파라미터 값을 분석 정의하고, 이의 변화가 시청자에게 미치는 심리적 영향을 파악해야 한다.

4.2 AI 얼굴검출 및 3D모델링

인공지능을 위한 객체인식은 얼굴 검출을 3D모델링을 통해 설계를 하게 되며, 영상객체를 통한 디자인 구성 요소는 다음 그림과 같다. 2개의 카메라를 이용하여 각각의 위상차를 이용한 깊이 정보를 계산하고 이 과정을 통하여 생성하게 되는데, 본 논문에서는 패턴인식을 통해 형태소분석과 얼굴검출등 인공지능을 통해 적용할 수 있는 얼굴검출의 패턴을 Landmark를 적용하여 단일 카메라 입력을 통해 얻어진 각각의 얼굴 영역의 객체 값을 얼굴의 정면 모델로 적용하며, 3D 변환행렬을 구성하여 얼굴 객체 모델에 매핑하여 3D 좌표를 얻어내는 방법을 사용하여 실험에 적용하였다[9].

Adaboost를 이용한 2차원 얼굴영역 검출

정면얼굴모델을 이용한 3D 변환행렬 구성

정면얼굴모델 필터를 이용하여 정면얼굴 정합

특징점 별 Adaboost에 의한 위치 확정

정합률에 기초한 유효성 판정

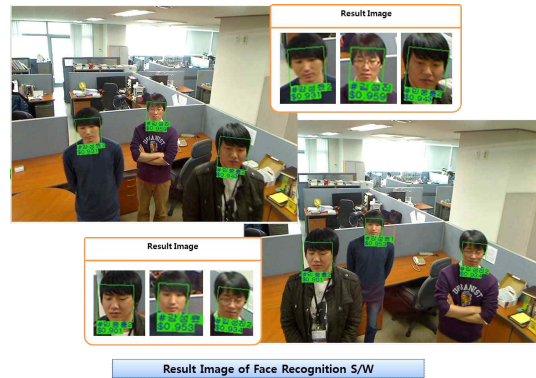
입 모델을 이용한 입술특징점 검출

눈썹 모델을 이용한 눈썹특징점 검출

NelderMead 비선형 최량화기를 이용한 최량정합 진행

<그림 3> 얼굴검출 3D 형태소분석 모델

<그림 3>은 얼굴검출을 위한 패턴인식과 형태소 분석 알고리즘을 적용하여 3D모델링의 분석 모델을 얼굴검출 데이터 추출을 하기 위해 구성하였으며, 얼굴의 특징점 추출을 위한 정합 물과 입과 눈썹등 얼굴 정면의 필터 학습 알고리즘을 얼굴의 특징점으로 Landmark 구성하여 얼굴의 패턴을 측정하고, 얼굴의 모형을 검출하여 다양한 입력 스트림에 대한 얼굴 검출을 표현하는 기법을 분석하여 Adaboost 알고리즘을 적용하여 본 실험을 적용하고 연구를 하였다[9].



<그림 4> 3D 패턴인식 검출 모델링

본 연구에서는 인공지능에 적용 될 객체 이미지 추출을 위해 <그림4> 3D패턴인식 검출 모델링과 사람의 얼굴 특징과 표정을 확인하기 위한 랜드마크를 적용하여 형태 변화의 통계적 모델링 값을 얻기 위해, 3D이미지 추출과 얼굴 특징점 추출 표본 데이터를 얻어 시험에 적용하였다[9].



<그림 5> 3D 이미지 추출 파노라마 촬영

<그림5>는 3D 이미지 추출을 위해 분당/연세웨이 과병원에서 파노라마 x-RAY 촬영을 통해 다양한 얼굴의 표정 데이터 값을 얻기 위한 실험을 과정이다. 이러한 실험 연구는 영상추출을 위해 입력 스트림 값이 들어오면 스트림이미지를 피부색과 타원 얼굴 모델로부터 얼굴 후보 영역을 구하게 된다. 얼굴 후보 영역에서 24x24 크기의 부분 이미지를 추출해내고 다음으로 앞서 설명한 학습을 통해서 얻어진 약분류기를 직접이미지에 대입해서 얼굴임을 판독하도록 설계를 하였다.

<그림6>은 ETRI 융합센싱SoC연구팀에서 얼굴인식을 위한 연구 실험과정을 나타내는 보고서를 인용하여 표기한 내용임.

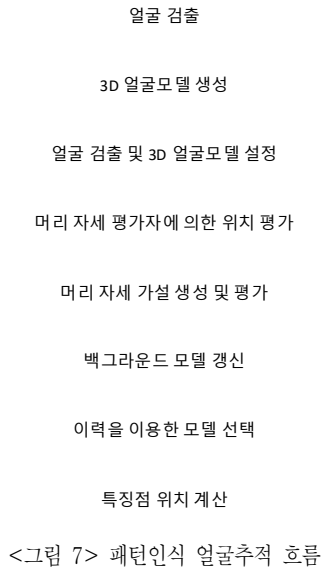
본 연구에서는 표현되는 인공지능 패턴인식과 객체인식의 영상 휘도치 분포값을 다단계 임계값에 적용하여 특징점을 검출 하였으며, 특히 휘도치 분포값은 다단계값의 임계치로 변환하는 방법을 알고리즘 구현에 적용시켜 나타 내도록 구성을 하였다. 임의의 대상 영상들로부터 획득된 확률·통계적 분석에 의해 다단계 임계화 구간을 수직/수평으로 투영하여 얼굴부위에 해당하는 휘도치 분포의 경계 값을



<그림 6> Adaboost 얼굴패턴 검출 모델링

결정하도록 정의하였다. 이 결정된 임계값에 따라 얼굴 이외의 성분으로부터 얼굴의 분할된 부위를 안면 요소의 특징점(눈, 코, 입)으로 구분하여 검출을 하였다[10].

<그림7>은 얼굴검출을 위한 패턴인식의 체계 순서를 순서대로 표기한 내용임. 3D 이미지 추출 파노라마 촬영분 논문에서는 패턴인식을 통해 얼굴의 특징점 검출을 나타내고 영상의 휘도치 분포를 다단계로 임계화하는 방법을 사용한다. 특히 형태소 분석을 통한 얼굴 패턴의 임계치 값은 임의의 대상 영상들로부터 획득된 확률·통계적 분석에 의해 다단계 임계화 된 구간을 수직·수평으로 투영하여 얼굴부위에 해당하는 휘도치 분포의 경계값을 결정한다. 이 결정된 임계값에 따라 얼굴이외의 성분으로부터 얼굴을 분할하고, 분할된 얼굴부위를 바탕으로 안면 요소 특징점(눈,코,입)등을 추출한다. 생성된 3D 얼굴 모델은 8



개(양눈끝에 4개 코에 2개, 입양끝에 2개)의 기본적인 포인트를 포함하게 된다. 눈썹과 입술의 특징점까지 포함하는 경우는 16개의 포인트를 포함한 3D객체 모델링으로 구성하게 된다.

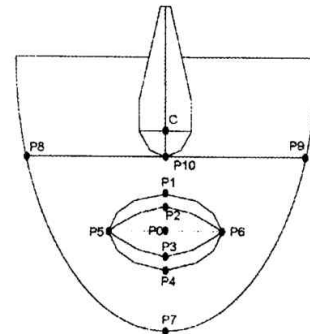
4.3 AI 얼굴추적 특징점 추출

인공지능을 통한 객체인식 얼굴모델 추출은 얼굴 움직임에 대한 추적을 통해 현재의 위치를 기준으로 상하 좌우로 움직일 수 있다는 가설을 세우게 되고, 이러한 가설을 기반으로 머리 자세 평가자를 통해 위치를 설정하게 된다. 또한 새로 입력되는 스트림 영상에 의해 백그라운드 모델이 갱신되게 되며, 이력에 있는 모델의 정보에 따라 새로운 특징점의 위치를 계산하고 3D 얼굴 모델의 추적을 하게 된다. 눈썹 추적의 경우 Eyebrow PCA Shape Model을 생성 Eyebrow Correspondences 계산 후 Eyebrow PCA Minimizer를 이용하여 패턴인식을 통한 얼굴추적을 진행한다[10].

객체인식을 통한 얼굴 포인트 검출은 눈썹과 입술

추적 그리고 입술의 움직임을 통해 입이 벌어진 정도, 턱의 움직임등 입술의 돌출, 뺨의 움직임등과 같은 복합적인 움직임을 감지하여 표정인식과 패턴인식을 병행하여 검출을 하고 이 과정을 통해서 이들의 복합적인 움직임을 포착하고 객체검출을 하게된다.

하지만 입술의 돌출과 뺨의 움직임 등은 2차원적 접근 방법으로는 실험을 통해 정확한 데이터 값을 얻기는 어렵다. 따라서 본 연구에서는 입력된 동영상과 정합된 3차원모델로부터 3차원 특징 벡터를 얻을 수 있도록 데이터 분석을 하고, 아래의 그림과 같이 12개의특징점을 사용하여 4개의 객체와 특징점의 벡터 값을 추출하여 본 실험의 결과를 확인하게 된다. 아래의 <그림8>은 얼굴추적을 위한 특징 추출 Landmark 중요 Point를 표기하여 나타내고 있다.



<그림 8> AI, 얼굴추적 특징점 추출

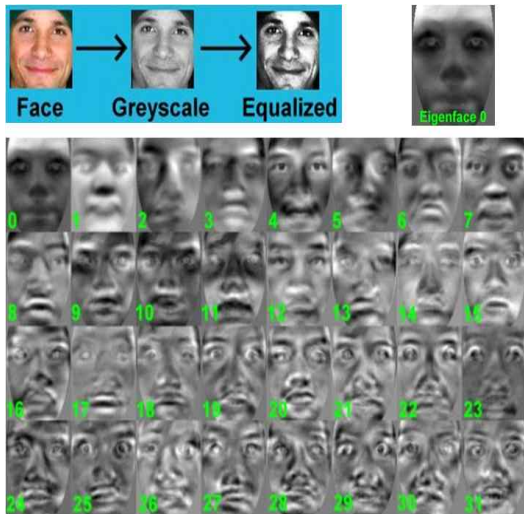
- 윗 입술 움직임 벡터(Up) : $UL = P2-P0$
- 아래 입술 움직임벡터(Down) : $DL=P3-P0$
- 입의 벌어진 정도 : $ML = P5-P6$
- 턱의 상하 움직임 : $JM = P7-P10$

따라서, 인공지능 객체인식의 패러다임을 적용하기 위해서는 같은 움직임을 적용하기 위한 벡터값을 정량적으로 취급하고 개인차를 흡수할 수 있는 기본량의 특징점 추출이 필요하다. 본 실험에서는 입의

종횡폭을 이용하여 기준량으로 정한다. 움직임 벡터의 프레임당 이동변위는 특징점의 이동량을 기준량으로 정규화하여 입술을 추적한다[11].

4.4 AI 분석방법

AI를 이용한 얼굴 인식에서는 얼굴을 가장 잘 표현할 수 있는 얼굴 공간(Face space)을 만드는 것이 중요하다.



<얼굴 인식 모듈(Principal component analysis)>

<그림 9> AI, 얼굴 특징점 분석

이러한 객체인식을 통한 얼굴인식 모듈은 공간들의 벡터가 주성분(principal component)이 되고, 전체 영상공간에서 얼굴을 가장 잘 표현할 수 있는 벡터 값을 찾는 것이 중요하다.

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].

다음에 나타낸 수식은 AI 주성분 분석을 이용한 얼굴 인식 과정을 분석하여 나타낸 것이다.

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.

$$\Psi = \frac{1}{K} \sum_{i=0}^{n^2} \sum_{k=1}^K a_{ik} \text{-----} \textcircled{8}$$

각 얼굴 영상의 차영상 x_{ik} 을 평균 영상으로부터 구한다.

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

$$X = [x_{ik}] \text{-----} \textcircled{9}$$

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

$$C = XX^T \text{-----} \textcircled{10}$$

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.

$$Q = X^T X \text{-----} \textcircled{11}$$

Q 의 고유벡터를 V_k 와 고유값 λ_k 을 구한다.

$$X^T X V_k = \lambda_k V_k \text{-----} \textcircled{12}$$

양변에 X 를 곱해 원래 데이터 집합의 고유벡터를 구한다.

$$XX^T XV_k = \lambda_k XV_k \text{-----} \textcircled{13}$$

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

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

$$E_{im} = \sum_{m=1}^m \sum_{k=1}^K x_{ik} V_{km} \text{-----} \textcircled{14}$$

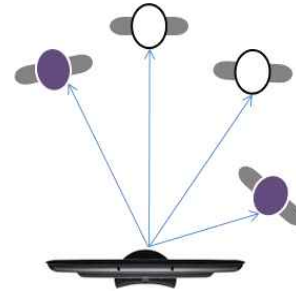
각 학습 얼굴 영상은 고유벡터 공간의 선형결합으로 표시한다.

$$W_{mk} = \sum_{k=1}^K \sum_{m=1}^M E_{im}^T x_{ik} \text{-----} \textcircled{15}$$

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

4.5 객체인식 3D 휴먼팩터 얼굴추적

본 실험은 인공지능을 통해 추진되는 객체인식 기반의 얼굴 추적을 휴먼팩터 측정에 대한 인식 기술을 적용하고 이를 통해 3D기반의 인공지능 패턴인식과 형태소 분석등 얼굴 검출 모델링으로 가장 잘 표현할 수 있는 측정법을 제시하고 이를 통해 나타내는 다양한 입체영상 객체 모델링을 Platform 환경과 연계하여 측정하도록 환경구성을 하여 실험표준(Test bed) 통해 구성요소로 나타내었다.



<그림 10> 3D모델링 객체인식 추적 실험

인공지능의 객체인식 구성은 패턴인식을 통해서 3D모델링의 얼굴추출 측정값을 찾아내고 이를 통해 휴먼팩터 측정에 적용되는 객체인식의 실험을 통해 얼굴모델링의 패턴인식 값을 얻어내며 이러한 시험을 측정을 위해서 다음과 같은 개발 Platform을 구성하여 연구 환경을 구축하여 실험을 하였다.

- 개발언어 : JAVA, PHP, VisualStudio 2019
- 운영체제 : MicroSoft/Windows10
- 입력장치 : 웹카메라(HD_1080P/200만화소)
1/4.5inch, CMOS센서 탑재.



<그림 11> 파라미터 3D특징점 추출/과정

본 연구는 인공지능을 기반으로 설계되고 휴먼팩터를 통한 휴먼인지 실험을 통해 나타나는 다양한 인식표현에 대한 데이터 추출 값을 프로그램을 통해 패턴성능의 실시간 인지와 얼굴인식 기반으로 추출되는 추적상태의 인식 가능한 표현들을 나타내도록 최

소한의 얼굴 픽셀값과 크기, 그리고 최대 인지값의 표현과 임계치의 각도에 따른 Error율을 검색하고 오류데이터를 추출하여 CPU의 로드등을 확인하고 그 연구 값을 얼굴표정과 Landmark적용을 통한 실험값에 적용하여 <그림11> 과 같이 파노라마 촬영을 통해 얼굴의 패턴인식 3D 특징점을 x-RAY 촬영을 통해 실험을 하였다.

본 실험을 통해서 확인된 결과는 인공지능으로 표현될 수 있는 얼굴 포인트의 추출과 객체인식의 얼굴 검출 과정등 시청자의 얼굴 포인트를 패턴인식을 통해 3D인식 기반의 영상처리 기법으로 표현하고, 이를 통해 3D 와이어 프레임과 오버레이등 휴먼팩터 측정으로 표현되는 실험 결과의 값을 확인할 수 있었으며, 그 결과 값은 휴먼팩터 측정을 통한 인공지능 객체인식에 관한 파라미터 측정의 확인값으로 인식하고 인공지능으로 표현할 수 있는 다양한 패턴인식의 휴먼인지 테스트를 실험하고 그 결과 값을 실험결과로 표현하여 나타 내었다.

V. 결 론

4차 산업혁명은 인공지능(AI)과 빅데이터(Big Data), 지능형자동차(EC), 사물인터넷(IoT)등 첨단산업 기술이 더욱 지능화되고 인공지능을 통해 스스로 판단하고 인지하는 능력을 갖춘 새로운 디바이스 산업을 말한다. 따라서 객체인식을 통한 파라미터 측정 연구와 휴먼팩터 인식을 통한 얼굴인식등 실시간 3D 모델링 이미지 프로세싱(Image Processing)기술과 연계하여 더욱 더 확대되어 개발되는 이러한 실시간 인공지능 객체인식 기술은 인문학을 기초로 한 포스트 휴머니즘을 기초로 설계하고 휴먼팩터 인식 기술을 플랫폼과 서비스와 연계하여 파라미터 측정을 통해 다양한 연구를 추진할 수 있다.

본 연구의 핵심은 객체인식 기반의 휴먼팩터를 통

해 다양한 파라미터 측정에 대한 객체인식 기술을 인공지능을 통해 객체인식 분야의 측정 연구를 사람의 양안 시차에서 느끼는 눈의 피로감과 휴먼팩터 핵심 기술인 패턴인식과 형태소 분석등 인공지능에서 필수 요소로 사용되는 인간의 공간과 지각 사이의 함수 관계를 규명하는 것을 목적으로 연구를 하였고 실험을 하였다.

따라서, 본 연구의 향후 목표는 가상화(VR) 공간에서 3D 실감 미디어와 연계된 완전한 인공지능 기반의 휴머니드 연구와 결합 함으로써 스마트디바이스 Plat form을 보다 안정화된 환경으로 적용하고 연구해 보는 것이 연구자로서의 가장 큰 바램이고 목표이다.

참고문헌

- [1] 3차원 영상연구회, 3차원 영상기술 연구 개발 계획 보고서, 전자공학회, 2017.
- [2] 한국통신학회, 얼굴 및 눈 위치 추적을 통한 3DTV 화면 인터페이스제어에 관한 연구, 2018.
- [3] Y.Y. Yeh and L.D. Silverstein, "Limits of fusion and depth judgment in stereoscopic color displays," Human Factors, Vol.32, No.1, Feb. 2014, pp.45-60.
- [4] S. Pastoor and M. Wöpking, "3-D Displays: A review of current technologies," Displays, Vol.17, 2016, pp.7-26, pp.100-110.
- [5] J.O. Merrit, "Stereoscopic display applications issues-Part 1: Human factor Issues," Course Notes, IS&T/SPIE Symposium on Electronic Imaging Science and Technology, 2013.
- [6] S. Shiwa, K. Omura, and F. Kishino, "Proposal for a 3-D display with accommodative compensation: 3DDAC," Journal of SID, Vol.4, 2018,

pp.255-261.

[7] M. Wöpking, "Viewing comfort with stereoscopic pictures: An experimental study on the subjective effects of disparity magnitude and depth of focus," *Journal of the SID*, Vol.3, No.3, 2014, pp.101-103.

[8] 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, 2015, pp.307-313.

[9] P.Engeldrum, *Psychometric Scaling* Imcotek Press, Winchester, Massachusetts, USA.,2014.

[10] N.Hiruma and T. Fukuda, "Accommodation response to binocular stereoscopic TV images and their viewing conditions," *SMPTE Journal*, Dec. 2018, pp.1137-1144.

[11] 최종명 · 김기원, "모바일 플랫폼을 이용한 상황인지 멀티미디어 서비스 시스템 분석," *디지털산업정보학회 논문지*, 제2권 제3호, 2006, pp.95-102.

[12] 정도영 · 홍기천, "얼굴 특징 정보를 이용한 향상된 눈동자 추적을 통한 줄임운전 경보 시스템 구현," *디지털산업정보학회 논문지*, 제5권 제2호, 2009, pp.167-176.

■ 저자소개 ■



최 병 관
(Choi Byungkwon)

2019년 9월~현재
연세대학교/전임교원
과학기술대학컴퓨터정보통신공학부

2019년 8월~ 중앙대학교/전임교원
ICT융합공학 인공지능학전공

2003년 12월 ㈜LG-CAP ICT 첨단기술연구소
기술중앙연구소장/상무

1993년 04월 ㈜한화/한화정보통신
기술중앙연구소/수석연구원

1981년 04월 한국과학기술원 시스템공학연구소

관심분야 : 인공지능, 로봇공학, 지능형자동차,
HCI, AR/VR, SoC 3D융합미디어,
E-mail : zcbk@yonsei.ac.kr

논문접수일	: 2019년 5월 22일
수정일	: 2019년 8월 24일
게재확정일	: 2019년 8월 28일