

논문 2020-57-11-6

가상현실 콘텐츠의 시공간적 인지 특징 기반 정량적 멀미 예측

(Predicting the Level of Cybersickness via Spatiotemporal Perceptual Features Extracted from Virtual Reality Content)

오 희 석*

(Heeseok Oh[©])

요 약

머리장착형 디스플레이(Head-Mounted Display: HMD) 기반의 가상현실 체험은 시각 정보로 하여금 전정 및 체성 감각에 앞서 신체 균형 감각의 지배를 강요하므로, 시청자로 하여금 이른바 생체인지적 감각불일치로 인한 극심한 가상현실 멀미를 유발한다. 가상현실 콘텐츠의 보급과 시장 확대를 위해서는 제작자와 서비스 제공자로 하여금 멀미 수준에 대한 사전 예측을 통해 콘텐츠를 적절히 제어함으로써 이용자의 시청안전성을 보장함이 요구되나, 이를 위한 기존 가상현실 멀미에 관한 연구는 각 요소에 대해 Ad-hoc 관점에서 연구되어 왔다. 본 논문에서는 가상현실 멀미의 정량적 예측을 위해 지도학습 기반의 멀미 예측 모델을 학습하였다. 선행 연구에서 밝혀진 멀미 관련 콘텐츠 요소를 고려하여 52종의 가상현실 장면을 생성하였고, 150명의 피험자를 동원하여 대규모 임상실험을 수행함으로써 주관적 가상현실 멀미 점수를 획득하였다. 시-전정 감각불일치를 대표하는 특징 추출을 위한 Raw 데이터의 가공과, 시각적 민감도/복잡도를 반영하는 RGB 영상으로부터의 특징 정의, 깊이맵을 활용한 특징 정의를 통해 정량적 가상현실 멀미 예측에 강인한 표현력을 갖는 17종의 공간적 인지 특징을 추출하였다. 또한, 추출된 공간적 특징을 인간의 시각 인지 체계를 반영한 4가지 풀링(Pooling) 기법을 도입함으로써 시간적 인지 특징으로 변환하였다. 임상실험을 통해 획득한 피험자의 주관적 멀미 점수를 레이블로 하여 제안한 특징들을 학습시킨 결과, 객관적/정량적 가상현실 멀미 수준 예측 모델을 구축할 수 있었으며, 예측 성능은 실제 피험자의 의견 대비 76%의 상관도를 달성하였다.

Abstract

Although an exploring virtual reality (VR) based on a head-mounted display (HMD) delivers the extended visual experience to a viewer, whose artificially forced stimuli cause inter- and intra- sensory conflicts between visual, vestibular and somatosensory proprioceptor. Such abnormal neurological interaction might provoke a deleterious side-effect, which is called cybersickness. In order to guarantee viewer's safety, producer and provider have to predict the degree of cybersickness prior to regulating the related attribute of a VR content. Unfortunately, previous researches on cybersickness have only focused on the mutually independent factors. To cope with this problem, we propose a novel scheme for prediction of the level of cybersickness by employing the supervised learning. Towards this, 52 synthetic VR scenes are newly generated, and then the subjective opinions are collected through a large-scale clinical evaluation with participation of 150 volunteers. To capture the robust representations of cybersickness, 17 spatial perceptual features are extracted including the processed raw data which represent the visual-vestibular sensory conflict, the estimated visual sensitivity and complexity from RGB images, and the depth information. Moreover, the extracted spatial perceptual features are transformed to the temporal perceptual features by deploying four pooling methods which reflects the human's visual characteristics. The extracted features are regressed onto the ground-truth (i.e., the rated score), thus the prediction of quantified level of cybersickness can be achieved by using the trained model, and whose predictive performance has ~76% correlated relationship with the subjective opinions.

Keywords : Virtual reality, Cybersickness prediction, Spatiotemporal perceptual feature, Supervised learning

* 정회원, 한성대학교 IT융합공학부(Department of IT Convergence Engineering, Hansung University)

© Corresponding Author(E-mail : ohhs@hansung.ac.kr)

※ 본 연구는 한성대학교 교내학술연구비 지원과제임.

Received ; July 27, 2020

Revised ; October 4, 2020

Accepted ; October 12, 2020

I. 서 론

시각 정보 기반의 실감성 향상에 대한 시청자의 수요 확대에 따라 영상콘텐츠는 평판 디스플레이를 필두로 한 8K UHD, IMAX 같은 초고화질 하드웨어 디바이스 개발과 더불어 양안식 3차원 영상을 비롯한 혼합, 증강, 가상현실 콘텐츠 보급을 위한 원천 기술 발전을 촉진하고 있다. 이 중, 가상현실은 머리장착형 디스플레이(Head-Mounted Display: HMD)를 활용하여 그래픽 엔진에 의해 생성된 실감형 콘텐츠의 자유로운 제어가 가능하다는 이점이 있다. 이는 서비스 제공자로 하여금 의료, 교육, 테마파크 등의 실감콘텐츠 응용 사례 제시를 가능케 함으로써, 근래에는 이외의 다양한 분야에서 그 활용 가능성에 대한 기대감이 증폭되고 있다. 하지만 관련 기술의 눈부신 발전에도 불구하고 대중들의 반응과 시장의 확대는 기대에 미치지 못하는 것이 사실인데, 무엇보다 가상현실 시청 시 발생하는 피로감과 멀미로 인한 사용자의 거부감이 시장의 성장을 방해하는 주요한 요인으로 분석된다. 연구에 따르면 80% 이상의 시청자가 가상현실 환경에 노출되었을 시 10분 이내에 극심한 멀미를 경험한다^[1].

과거 신경과학 및 생체심리학 분야에서는 가상현실 시청 시 발생하는 멀미의 원인을 규명하고자 다양한 이론이 제시되었으며, 그 중 시각 정보와 전정 기관 및 체성 감각과의 불일치를 근거로 멀미 기전을 설명하는 감각불일치 이론(Sensory conflict theory)이 멀미 발생에 대해 대중적으로 통용되는 이론으로써 자리 잡았다^[2]. Oman은 감각불일치 이론을 발전시켜 시각 기관, 전정 기관, 그리고 체성 감각 기관 간의 관계를 수학적으로 설명하고 동적 모델로 표현하였다^[3]. Bos 등은 Oman 모델로 설명하지 못하는 회전 시각성 운동 및 중력 변화에 따른 멀미 유발에 관련해 그림 1과 같이 신경 저장소의 개념을 도입하여 과거 경험과의 감각 차이를 추가적으로 반영한 새로운 멀미 유발 모델을 정립하였다^[4]. 상기 이론을 근거로, HMD를 이용한 가상현실 시청에 따른 멀미 발생은 눈을 통해 획득하는 시각 정보가 시청자의 다른 균형 감각을 지배함으로써 감각 간 불일치를 유발하는 현상으로 설명 가능하다^[5].

이러한 이론적 해석의 존재에도 불구하고, 가상현실 멀미 수준에 대한 정량적인 분석과 예측에 관련한 연구는 미진한 상태이다. 현재 가상현실 멀미 저감을 위한 방법론은 경험에 따른 콘텐츠 제작자들의 주관적인 노하우만으로 존재하고 있으며, 제작 단계에서 시청자가 경험할 가상현실 멀미 수준을 객관적으로 예측하고 이를 제어하

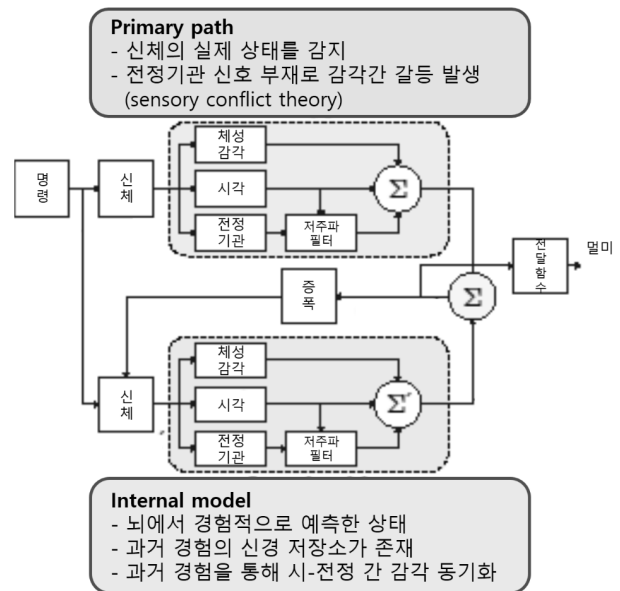


그림 1. 감각불일치 이론 기반 Bos의 멀미 모델^[4]

Fig. 1. Bos model describes motion sickness regarding psychophysical mechanism based on the sensory conflict theory^[4].

려는 시도는 연구된 바가 없다. 따라서 본 연구에서는 상기 멀미 이론에 근거하여 가상현실 콘텐츠의 영상 성분을 분석하고, 기계학습을 통해 정량적인 가상현실 멀미 수준을 예측하는 기술을 개발하였다.

본 연구는 콘텐츠 입력만으로 시청자가 경험할 가상현실 멀미 수준을 정량적으로 예측함으로써 가상현실 콘텐츠 제작자 및 서비스 제공자로 하여금 별도의 피드백 과정 없이 시청 안전성을 보장하는 가상현실 환경 구축을 가능하게 하여 향후 가상현실 콘텐츠 보급 및 시장의 확대에 도움을 주고자 한다. 이를 위한 구체적 개발 방법으로 가상현실 멀미 측정을 위한 레퍼런스 콘텐츠를 제작하고 임상 실험 프로토콜을 수립하여 다수의 피험자를 동원해 주관적 의견을 수집한 뒤, 멀미 수준에 강인한 콘텐츠 요소 및 시간적, 공간적 인지 특징을 정의하고 추출하여 기계학습 모델을 학습시킨다. 학습된 멀미 예측 모델은 임의의 가상현실 콘텐츠가 입력되었을 시, 정량적인 멀미 수준을 출력함으로써 제작 과정에서 시청 안전성 보장을 위한 가이드라인의 역할이 가능하다. 나아가 궁극적으로는 본 연구 결과가 멀미 저감을 위한 각종 기술 개발의 표준 지표로 응용될 수 있을 것으로 예상된다.

II. 선행 연구 및 한계

1. 가상현실 멀미와 하드웨어 요소

기존 가상현실 멀미에 관한 연구는 체험에 요구되는 하드웨어 사양 관점에서 주로 분석되어왔다. Mittelstaedt 등은 해상도와 시야각이 상이한 이중의 HMD 성능에 따른 멀미 발생 여부를 통계적으로 분석하였으나 유의한 결과를 도출하지는 못하였다^[6]. Moss 등은 Motion-to-Photon 지연시간 관점에서 영상콘텐츠의 디스플레이 렌더링 속도에 따른 가상현실 멀미를 분석하였으며, 지연시간이 짧을수록 멀미를 적게 발생함을 확인하였다^[7]. Sharples 등은 HMD 기반의 가상현실과 평판 및 프로젝션 형태의 디스플레이가 유발하는 시야발 멀미 발생의 차이에 대해 실험하였다^[8]. 이와 같이 하드웨어 사양에 따라 달라지는 가상현실 멀미는 시청단에서 활용하는 장비의 종류에 의존적이기 때문에 실제 콘텐츠 제작 단계에서는 고려하기 어려운 요소들이다. 또한 근래의 HMD의 사양 역시 급격히 발전하고 있으므로 하드웨어 성능 저하로 인해 야기되는 멀미는 점차 개선되고 있다고 볼 수 있으나, 근본적으로 콘텐츠 자체가 유발하는 멀미 요소에 관한 규명과 이에 관한 정량적 분석은 여전히 난제로 남아있다.

2. 생체신호를 통한 가상현실 멀미 분석

하드웨어 성능에 따른 멀미 유발의 차이에서 나아가 심리학 분야에서는 인간의 생체 반응을 이용해 가상현실 멀미 수준을 정량화하고자 하였다. Lin 등은 프로젝션 형태의 가상현실 체험 공간에서의 자동차 주행 시뮬레이션 시 발생하는 멀미의 수준을 뇌파와 심전도 스펙트럼 분석을 통해 예측하였다^[9]. Rahnev 등은 뇌파를 분석하여 뇌의 전전두엽 영역이 멀미 유발과 관련한 감각불일치에 연관되어있음을 밝혔다^[10]. Guna와 Gersak은 360° 가상현실 영상 시청 후 멀미 발생에 따른 심전도와 피부전도도의 변화가 통계적으로 유의함을 실험적으로 증명하였다^[11]. 이 같은 생체신호 기반의 가상현실 멀미 분석은 인간의 생체인지적 멀미 기전의 해석에 큰 의의가 있으나, 실제 가상현실 콘텐츠 제작자에게 있어서는 현업에서 활용하기 어려운 정보이다. 또한 제작 단계에서 각종 생체반응 측정 장비를 동원하여 그 결과를 검증하며 작업하는 과정이 사실상 불가능하므로, 제어 가능한 콘텐츠 요소 변화에 따라 달라지는 멀미 수준의 자동적 예측 기술이 요구된다.

3. 가상현실 멀미 유발 콘텐츠 요소

콘텐츠 요소에 따른 가상현실 멀미 유발은 개별 요소 단위로 분석되어왔다. HMD 착용 상태에서 가상의 코를 생성해 화면 내 배치함으로써 프레임 이동의 기준으로 작용해 가상현실 멀미가 저감된다는 결과가 발표되었다^[12]. Singla 등은 가상현실 콘텐츠의 카메라의 이동 속도가 빠를수록 더욱 빈번하게 멀미가 유발됨을 실험적으로 증명하였다^[13]. Keshavarz 등은 일정한 카메라의 움직임이 아닌, 가속과 감속이 혼재되어 이동하는 상황에서 시청자가 더욱 극심한 멀미를 경험함을 보였다^[14]. Joseph과 Griffin은 여러 가지 카메라 움직임 중 roll과 pitch 회전이 멀미 유발과 통계적 유의성을 가짐을 확인하였다^[15]. 기존 수행된 콘텐츠 요소에 따른 가상현실 멀미 연구는 주로 발생 여부에 초점을 맞추어 진행되었다. 따라서 시청자가 경험하는 멀미에 대한 정량적인 지표로서의 활용에는 한계가 있으며, 객관적인 멀미 수준의 예측은 불가능하다.

III. 가상현실 멀미 예측 모델

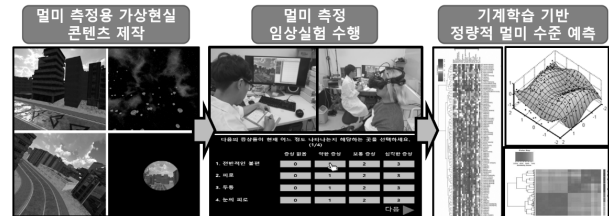


그림 2. 본 연구의 가상현실 멀미 예측을 위한 수행 절차
Fig. 2. Overall framework for prediction of cybersickness arising from virtual reality.

본 연구는 가상현실 멀미 예측 기술 개발을 위하여 그림 2와 같이 크게 세 부분을 수행하였다. 첫 번째로 그래픽 엔진을 이용하여 각종 콘텐츠 요소를 단계별로 제어함으로써 가상현실 멀미를 신뢰도 높게 측정 가능한 콘텐츠를 제작하였다. 두 번째로 실제 피험자를 동원한 임상실험을 통해 HMD를 착용한 상태에서 레퍼런스 콘텐츠를 시청한 후 각자가 느끼는 주관적 멀미 점수를 기록하였다. 세 번째로 가상현실 콘텐츠를 요소별로 분석하여 가상현실 멀미 표현에 강인한 특징을 추출한 후, 주관적 점수와 지도학습을 통해 정량적 가상현실 멀미 예측 모델을 구축하였다.

1. 가상현실 멀미 측정용 콘텐츠 제작

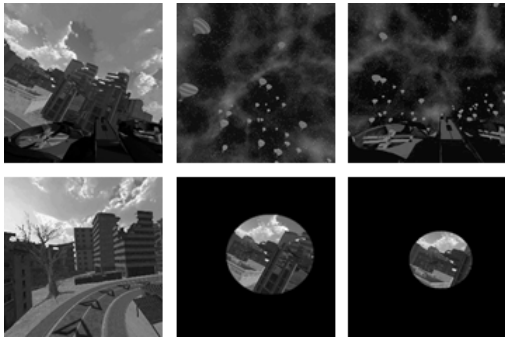


그림 3. 가상현실 멀미 측정을 위해 제작된 콘텐츠 장면의 예
Fig. 3. Examples of the generated scenes by employing Unity3D for measuring cybersickness.

기존 가상현실 멀미 측정에 이용된 콘텐츠는 실사 기반의 360° 파노라마 영상이 주를 이루었다. Padmanaban 등은 360° 영상에서의 움직임 정보를 추출하여 벡터의 크기를 토대로 정량적인 멀미 예측을 수행하였다[16]. 하지만 360° 영상은 특성상 각종 콘텐츠 요소를 독립적으로 제어하기 어려우며 이로 인해 통제변인의 변화에 따른 멀미 수준의 단계별 평가가 불가능하다는 한계가 존재한다. 본 논문에서는 이와 같은 단점을 극복하고자 Unity3D를 활용하여 콘텐츠 요소 조절을 통해 다양한 가상현실 장면을 생성하였다. 가상현실 멀미 측정용 콘텐츠 제작을 위해 다음과 같은 8종의 멀미와 관련되어 있다고 알려진 콘텐츠 요소를 고려하였다. ① 콘텐츠 배경 (도시공간, 우주공간), ② 카메라 움직임 복잡도 (단축 운동, 다축 운동), ③ 카메라 이동 속도(4m/s, 9m/s), ④ 카메라 이동 가속도 (등속, 가속), ⑤ 시야각 (30°, 45°, 90°), ⑥ 프레임 레퍼런스 (레퍼런스 유무), ⑦ 콘텐츠 길이(13초, 24초), ⑧ 사용자 조작 여부(조작 가능, 불가능). 8종의 요소들을 조합하여 총 52개의 가상현실 장면을 제작하였으며, 그림 3에서는 제작된 멀미 측정용 콘텐츠 장면의 예제를 보여준다 (좌상단부터 도시공간 배경 프레임 레퍼런스 존재 장면, 우주공간 단축 회전 장면, 우주공간 프레임 레퍼런스 존재 장면, 45° 시야각 장면, 90° 시야각 장면, 사용자 조작 가능 장면).

2. 가상현실 멀미 측정 임상 실험

제작된 52개의 가상현실 멀미 측정용 콘텐츠 장면을 이용하여 실제 피험자를 동원해 주관적 멀미 점수를 획득하고 데이터베이스를 구축하였다. 총 150명의 피험자가 임상실험에 참여하였으며, 그림 4와 같은 임상실험



그림 4. 가상현실 멀미 측정을 위한 임상실험 프로토콜
Fig. 4. Subjective evaluation protocol for obtaining cybersickness score.

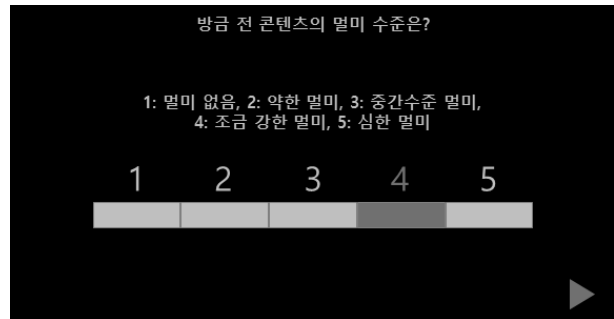


그림 5. 피험자의 가상현실 멀미 수준 평가 화면
Fig. 5. Graphical user interface for scoring the level of cybersickness.

프로토콜 설계를 통해 연구 대상자의 인체 안전성 보장을 위한 IRB(Institutional Review Board) 승인을 득하였다. 피험자는 14세에서 59세까지 정상 시력을 가진 가상현실 관련 비전문가 남녀(각 75명) 집단으로 구성되었다. 실험을 위한 HMD는 HTC VIVE를 이용하였다. 각 피험자는 실험 전 실험 개요에 관한 안내를 받음과 더불어 동의서를 작성하고, 이 과정에서 실험 결과에 영향을 미칠 수 있는 가설에 관한 내용 설명은 배제되었다. HMD를 착용 후, 컨트롤러를 비롯한 가상현실 기기 사용법 설명을 위한 사전 학습 세션을 수행하였다. 사전 학습 세션을 위해 실제 52개 장면과 중복되지 않는 6개의 가상현실 장면을 별도로 제작하였다. 사전 학습 세션 이후 3분의 휴식^[17]을 취한 뒤, 본 실험 세션을 시작하였다.

본 실험 세션에서는 52개의 가상현실 장면이 콘텐츠 요소 특성에 따라 무작위로 디스플레이되며, 피험자는 해당 장면에 대해 1-5단계의 Likert 척도 기반으로 주관적 멀미 체험에 관한 점수를 부여하게 된다. 기존 가상현실 멀미 연구에서 활용되었던 설문지 기반의 멀미 점수 부여^[18]와는 달리, 본 연구에서는 HMD를 착용한 상태에서 주관적 평가가 가능하도록 그림 5와 같이 멀미 점수 획득을 위한 Graphical User Interface(GUI)를 별도로 구성하여 장비 탈부착에 따른 부가적인 피로 발생을 차단하였다. 더불어 장시간 가상현실 시청에 따른 멀미 누적을 방지하기 위해 실험 세션은 3번에 걸쳐 나누어 수행되었으며, 각 실험 세션 사이에는 3분의 휴식시

간이 주어졌다. 하드웨어 지연에 따른 멀미 발생을 최소화하고자 콘텐츠의 재생 속도는 90fps를 유지하였다.

피험자가 시청하는 가상현실 장면은 RGB 영상 데이터로 기록하였으며, 동시에 깊이맵을 함께 저장하였다. 가상현실 멀미 측정 임상 실험을 통해 총 7,800개(150명 x 52개 장면)의 데이터-정답 쌍을 획득하였으며, 콘텐츠에서 출력하는 Raw 데이터 목록은 표 1과 같다.

표 1. 가상현실 멀미 측정 임상실험 획득 데이터
Table1. Raw data obtained from subjective evaluation.

카메라 움직임	HMD 움직임
x축 위치 좌표 (x_{tp}^c)	x축 위치 좌표 (x_{tp}^h)
y축 위치 좌표 (y_{tp}^c)	y축 위치 좌표 (y_{tp}^h)
z축 위치 좌표 (z_{tp}^c)	z축 위치 좌표 (z_{tp}^h)
x축 이동 속도 (x_{tv}^c)	x축 이동 속도 (x_{tv}^h)
y축 이동 속도 (y_{tv}^c)	y축 이동 속도 (y_{tv}^h)
z축 이동 속도 (z_{tv}^c)	z축 이동 속도 (z_{tv}^h)
x축 이동 가속도 (x_{ta}^c)	x축 이동 가속도 (x_{ta}^h)
y축 이동 가속도 (y_{ta}^c)	y축 이동 가속도 (y_{ta}^h)
z축 이동 가속도 (z_{ta}^c)	z축 이동 가속도 (z_{ta}^h)
x축 회전 각 (x_{rp}^c)	x축 회전 각 (x_{rp}^h)
y축 회전 각 (y_{rp}^c)	y축 회전 각 (y_{rp}^h)
z축 회전 각 (z_{rp}^c)	z축 회전 각 (z_{rp}^h)
x축 회전 각속도 (x_{rv}^c)	x축 회전 각속도 (x_{rv}^h)
y축 회전 각속도 (y_{rv}^c)	y축 회전 각속도 (y_{rv}^h)
z축 회전 각속도 (z_{rv}^c)	z축 회전 각속도 (z_{rv}^h)
x축 회전 각가속도 (x_{ra}^c)	x축 회전 각가속도 (x_{ra}^h)
y축 회전 각가속도 (y_{ra}^c)	y축 회전 각가속도 (y_{ra}^h)
z축 회전 각가속도 (z_{ra}^c)	z축 회전 각가속도 (z_{ra}^h)

3. 가상현실 멀미 예측 모델 학습

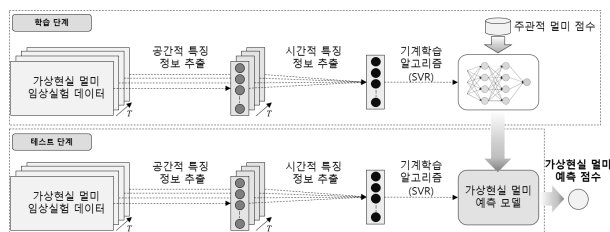


그림 6. 공간적/시간적 특징 추출과 지도학습을 통한 가상현실 멀미 예측 모델 학습

Fig. 6. Spatiotemporal feature extraction and cybersickness predictor training scheme in supervised manner.

그림 6은 임상실험을 통해 획득한 데이터를 활용한 정량적 가상현실 멀미 예측 모델의 학습을 위한 방법론

을 도시하였다. 각각의 프레임에서 가상현실 멀미에 대한 표현력을 가지는 시각 인지 특성을 공간적 인지 특징이라고 정의하였고, 가상현실 장면 내 모든 프레임에 대한 공간적 특징의 풀링(Pooling) 결과를 시간적 인지 특징으로 정의한다. 추출된 특징들은 임상실험에서 획득한 피험자의 주관적 멀미 점수를 레이블(Label)로 하여 지도학습을 통해 회귀 모델을 학습하며, 학습을 위한 알고리즘은 Radial Basis Function(RBF) 커널 기반의 비선형 Support Vector Machine(SVM)을 이용하였다.

가. 공간적 인지 특징 추출

매 n 번째 프레임에서 추출하는 공간적 인지 특징을 데이터에 따라 다음과 같은 3개의 카테고리로 분류하였다. ① Raw 데이터 기반 특징, ② RGB 영상 기반 특징, ③ 깊이맵 기반 특징.

(1) Raw 데이터 기반 특징

표 1과 같이 획득한 Raw 데이터를 이용하여 특징을 추출한다. 앞서 Bos 모델^[4]에서 서술한 바와 같이, 가상현실 멀미 유발의 주요 원인은 시각 정보와 전정 감각의 불일치이다. 불일치 정도를 특징으로 반영하기 위해, 시각 정보를 대표하는 카메라 움직임과 전정 감각을 대표하는 HMD 움직임을 Raw 데이터를 활용함으로써 계산하였다.

$$\{f_{1,n}, f_{2,n}, f_{3,n}\} = \{|p_{tv}^c - p_{tv}^h| | p \in (x, y, z)\}, \quad (1)$$

$$\{f_{4,n}, f_{5,n}, f_{6,n}\} = \{|p_{ta}^c - p_{ta}^h| | p \in (x, y, z)\}, \quad (2)$$

$$\{f_{7,n}, f_{8,n}, f_{9,n}\} = \{|p_{rv}^c - p_{rv}^h| | p \in (x, y, z)\}, \quad (3)$$

$$\{f_{10,n}, f_{11,n}, f_{12,n}\} = \{|p_{ra}^c - p_{ra}^h| | p \in (x, y, z)\}. \quad (4)$$

식 (1)과 (2)는 선형 이동 속도 및 가속도에 대한 시-전정 감각정보의 불일치를 대표하며, (3)과 (4)는 회전 속도 및 가속도에 대한 시-전정 감각정보의 불일치를 표현하는 특징이다.

(2) RGB 영상 데이터 기반 특징

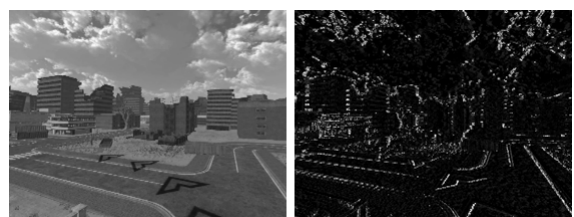


그림 7. 좌: RGB 영상, 우: 시각적 민감도 계산
Fig. 7. Left: RGB image, Right: contrast sensitivity map.

HMD 착용 시, 실제 시청자에게 전달되는 시각정보는 가상현실 콘텐츠의 RGB 영상이다. Raw 데이터와 더불어, 보다 정확한 시각 정보의 계산을 위해 RGB 영상 처리를 통한 공간적 인지 특징을 추출한다.

RGB 영상을 이용한 첫 번째 특징으로 시각적 민감도를 계산한다. 인간의 시각 체계는 공간적/시간적으로 특정 주파수 대역에 보다 민감하게 반응한다. 신경과학 분야에서는 이를 대비민감도함수(Contrast Sensitivity Function: CSF)로 정의하였다^[19].

$$CSF(\rho, \nu) = kc_0c_2\nu(c_12\pi\rho)\exp\left(-\frac{c_14\pi\rho}{\rho_{\max}}\right), \quad (5)$$

$$k = s_1 + s_2 \left| \log\left(\frac{c_2\nu}{3}\right) \right|^3, \quad (6)$$

$$\rho_{\max} = \frac{p_1}{c_2\nu + 2}. \quad (7)$$

여기서 ρ 와 ν 는 공간주파수와 움직임 속도를 의미한다. 나머지 파라미터들은 상수로써, $c_0 = 1.14$, $c_1 = 0.67$, $c_2 = 1.7$, $s_1 = 6.1$, $s_2 = 7.3$, $p_1 = 45.9$ 로 세팅하였다^[19]. 공간주파수의 계산은 이산 코사인 변환 도메인에서 수행하였으며, 움직임 속도는 optical flow 알고리즘을 통하여 계산하였다. 그림 7에서는 RGB 영상에 대해 CSF를 이용하여 계산한 시각적 민감도를 가시화하였다. 시각적 민감도를 가상현실 멀미 예측의 특징으로 활용하기 위하여 전체 영상 영역에 대한 평균과 분산을 이용하였다.

$$f_{13,n} = \text{mean}(CSF(\rho, \nu)), \quad (8)$$

$$f_{14,n} = \text{var}(CSF(\rho, \nu)). \quad (9)$$

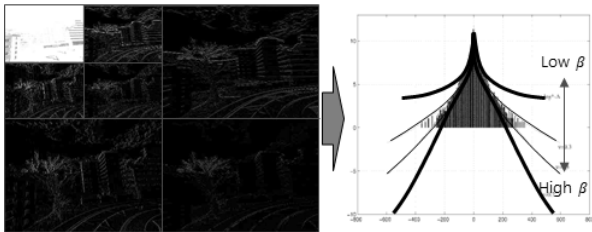


그림 8. 좌: RGB 영상의 웨이블릿 변환, 우: 웨이블릿 계수의 GGD fitting

Fig. 8. Left: Wavelet transform of RGB image, Right: GGD fitting of wavelet coefficients.

두 번째 RGB 영상으로부터의 가상현실 멀미 예측을 위한 특징으로 시각적 복잡도를 계산하였다. Lee et al.은 영상 성분의 복잡도를 visual activity라 정의하고, visual activity가 클수록 시각 피로가 증가함을 보였다^[20]. 시각

적 복잡도는 웨이블릿 변환 도메인에서 계산한다. 저주파 영역을 제외한 고주파 영역에서의 웨이블릿 계수 v 는 Generalized Gaussian Distribution(GGD)을 따르며, GGD는 다음과 같이 표현된다.

$$GGD(v; \eta, \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} \exp(-(|x - \eta| / \alpha)^\beta) \quad (10)$$

여기서 $\Gamma(\cdot)$ 는 감마함수를 의미하며, η , α , 그리고 β 는 각각 분포의 location, scale, 그리고 shape를 표현하는 파라미터이다. 그림 8에서는 RGB 영상의 웨이블릿 변환과 GGD로의 fitting을 보여준다. Visual activity는 shape 파라미터인 β 를 통해 표현되며, 낮은 β 는 고주파 성분이 넓게 분포해 있음을 나타내므로 높은 시각적 복잡도를 의미한다. 따라서 GGD의 shape 파라미터 β 는 가상현실 멀미 예측을 위한 표현력 강한 특징이 될 수 있다.

$$f_{15,n} = \beta. \quad (11)$$

(3) 깊이맵 기반 특징

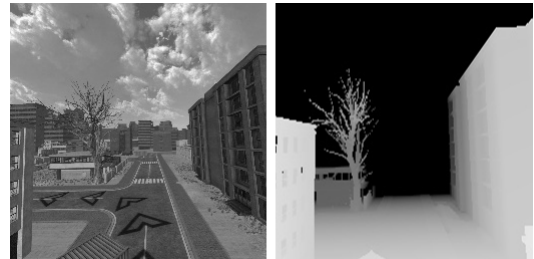


그림 9. 좌: RGB 영상, 우: 뎀스맵

Fig. 9. Left: RGB image, Right: depth map.

HMD 기반의 가상현실 콘텐츠는 양안 시차를 통해 깊이 정보를 강제함으로써 보다 높은 실감성을 제공한다. 하지만 이는 강한 시각 피로를 유발하며 가상현실 멀미의 한 원인이 된다^[21, 22]. 본 논문에서는 이에 대한 반응을 위해 가상현실 콘텐츠에서 깊이맵을 함께 추출하여 이를 가상현실 멀미 예측을 위한 특징 정보로 활용한다. 그림 9에서는 RGB 영상에 대해 함께 추출된 깊이맵을 가시화하였다. (i, j) 번째 픽셀에서의 깊이값을 $d(i, j)$ 라고 할 때, 깊이맵을 이용한 특징은 다음과 같이 계산된다.

$$f_{16,n} = \frac{1}{K} \sum_{i,j} d(i, j), \quad (12)$$

$$f_{17,n} = \frac{1}{K} \sum_{i,j} (f_{16,n} - d(i, j))^2. \quad (13)$$

여기서 K 는 전체 픽셀 개수를 의미하므로, $f_{16,n}$ 과 $f_{17,n}$ 는 각각 한 프레임에서 전체적인 깊이 수준과 그 분포를 나타낸다.

나. 시간적 인지 특징 추출

n 번째 프레임에서 17개의 공간적 인지 특징을 추출하였다. 제작된 가상현실 장면은 길이가 서로 다른 비디오 시퀀스이므로 이를 특징으로 활용해 멀미 예측 모델을 학습시키기 위해서는 하나의 가상현실 장면 당 모든 N 개의 프레임에 대해 시간축을 따라 폴링하여 전체 장면을 대표하는 특징으로 가공하여야 한다. 본 연구에서는 가상현실 장면 내 모든 N 프레임에 대해 m 번째 공간적 특징의 평균과 분산을 이용해 폴링하였다.

$$\mu_m = \frac{1}{N} \sum_{n=1}^N f_{m,n}, \quad (14)$$

$$\sigma_m = \frac{1}{N} \sum_{n=1}^N (f_{m,n} - \mu_m), \quad (15)$$

인간의 시각 체계는 시간축으로 입력되는 시각 정보에 대해 평균과 분산으로 대표되는 전체적인 강도 및 분산과 더불어, 특정 강하거나 약한 대표적 자극에 보다 민감한 특성이 있다. 기존 시각 경험을 정량화 하는 연구들에서는 이러한 생체심리학적 특성 반응을 위해 p -percentile 폴링 방법을 활용하였다^[22, 23].

$$\mu_m^h = \frac{1}{\psi_{m,\max}} \left(\frac{1}{N \cdot p/100} \right)_{n < N(100-p)/100} \sum_{n=1}^N \psi_{m,n}, \quad (16)$$

$$\mu_m^l = \frac{1}{\psi_{m,\max}} \left(\frac{1}{N \cdot p/100} \right)_{n < N \cdot p/100} \sum_{n=1}^N \psi_{m,n} \quad (17)$$

여기서 ψ 는 m 번째 특징을 프레임별 값의 크기대로 정렬한 데이터를 의미하며, 이에 따라 μ_m^h 과 μ_m^l 은 각각 상위 및 하위 p -percentile 데이터의 평균을 의미한다. 본 연구에서는 $p = 10\%$ 로 세팅하였다^[22]. (14)–(17)의 과정에 의해 17개의 특징은 각각 4개의 시간적 인지 특징으로 폴링되며, 최종적으로 68개의 특징을 추출하여 각 특징에 대해 최소-최대 정규화 과정을 거친 뒤, 가상현실 멀미의 정량적 예측에 이용한다.

IV. 가상현실 멀미 예측 성능

총 150명의 피험자가 52개의 가상현실 장면을 시청함으로써 7,800개의 데이터-정답 쌍에 대한 SVM 기반

학습이 가능하다. 본 연구에서는 전체 피험자의 80%인 120명을 학습 데이터로 이용 하였으며 20%인 30명에 대해 예측된 가상현실 멀미 수준을 테스트하였다. 학습 및 테스트 데이터는 시도 별 무작위로 추출되었으며, 학습 모델의 일반화 성능 검증을 위하여 시도마다 다른 학습 및 테스트 데이터를 구성하여 200번의 학습 및 검증 수행하였고 성능의 평균을 측정하였다. 즉, 매번 다른 구성의 5-fold 교차 검증을 40번 수행하였다. 성능 측정은 학습된 모델의 예측 값과 실제 피험자가 평가한 주관적 멀미 점수 간 상관성을 측정하였으며, 단조성을 평가하기 위한 지표인 SROCC(Spearman's Rank Order Correlation Coefficient)와 선형성을 평가하기 위한 지표인 PLCC(Pearson's Linear Correlation Coefficient)를 이용하였다. 각각의 상관계수는 1에 가까울수록 정량적 가상현실 멀미를 정확히 예측하였음을 의미한다.

표 2. 주관적 멀미 점수와 예측 멀미 수준 간 상관계수를 통한 검증 (200회 평균)

Table2. SROCC and PLCC for 200 trials of randomly chosen train and test sets.

공간적 인지 특징 종류	SROCC	PLCC
Raw 데이터 기반 특징 ($f_1 \sim f_{12}$)	0.712	0.722
RGB 영상 데이터 기반 특징 ($f_{13} \sim f_{15}$)	0.652	0.617
깊이맵 기반 특징 (f_{16}, f_{17})	0.437	0.397
전체 특징 (SVM)	0.763	0.751
전체 특징 (MLP)	0.738	0.722

표 2에서는 200회의 학습 및 테스트 데이터의 무작위 추출을 통한 SROCC와 PLCC의 평균을 나타내었다. III장에서 정의하고 추출한 공간적 인지 특징별 성능을 Raw 데이터 기반 특징, RGB 영상 데이터 기반 특징, 깊이맵 기반 특징으로 분류해 측정하였다. 시-전정 불일치를 반영한 Raw 데이터 기반 특징이 피험자의 주관적 의견과 SROCC 기준 약 71%의 상관도를 가짐으로써, 각각 65%와 43%의 상관도를 갖는 RGB 영상 데이터 기반 특징과 깊이맵 기반 특징보다 가상현실 멀미 예측에 있어 강인한 표현력을 지니는 특징이라고 설명할 수 있다. 모든 특징을 이용하여 가상현실 멀미를 예측하였을 경우, 주관적 의견과의 상관도는 약 76%를 나타내었다. 더불어, 멀미 예측 모델의 학습 방법에 따른 비교를 위해 SVM 뿐 아니라 한 개의 은닉층(100개의 은닉 노드)을

포함한 다층 퍼셉트론(Multi-Layer Perceptron; MLP)을 이용해 같은 방법으로 학습 및 테스트를 통한 성능 확인을 수행하였으며, 이 경우 주관적 의견과의 상관도는 약 73%를 나타내었다. 이는 Raw 데이터 기반 특징과 RGB 및 깊이맵 기반의 특징들이 상보적으로 작용하며, 가상현실 멀미의 예측에 있어 효과적이고 신뢰성 있는 특징이라고 해석할 수 있다.

본 논문에서는 전체 특징이 이용되어 가상현실 멀미를 예측함에 있어 특징별 가중치를 계산해보았다. 예측 모델 학습에 이용된 SVM 알고리즘 특성 상, 특징에 대한 가중치의 개별적 계산이 불가하므로, 최소자승추정(Least Square Estimation)을 통한 정규화된 특징 별 예측 중요도를 분석하였다. 앞서 추출한 m 개의 특징 행렬을 F , 예측 값의 벡터를 P 로 두었을 때, 각 특징에 대한 가중치를 나타내는 m 차원의 벡터 W 는 다음과 같이 계산 가능하다.

$$W = (F^T F)^{-1} F^T P. \quad (18)$$

표 3. 최소자승추정을 통한 특징별 예측 가중치
Table3. Weight of each feature for cybersickness prediction estimated by least square estimation.

공간적 인지 특징 종류	특징정보 별 가중치
Raw 데이터 기반 특징 ($f_1 \sim f_{12}$)	10.3%
RGB 영상 데이터 기반 특징 ($f_{13} \sim f_{15}$)	74.4%
깊이맵 기반 특징 (f_{16}, f_{17})	15.3%

표 3에서는 (18)을 통해 추정된 특징 별 가상현실 멀미 예측 가중치를 나타내었다. 특징 공간에서 가상현실

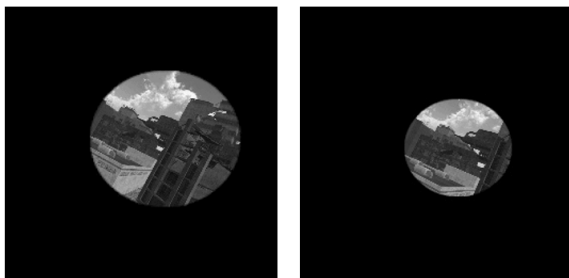


그림 10. 좌: 90° 시야각 장면 (예측 멀미도: 1.562), 우: 45° 시야각 장면 (예측 멀미도: 1.415)
Fig. 10. Left: 45° field of view (predicted sickness level: 1.562), Right: 90° field of view (predicted sickness level: 1.415).

멀미 예측을 위해 RGB 영상에서의 시각적 민감도 및 복잡도가 중요한 특징으로 작용함을 확인할 수 있다.

다만 이러한 RGB 영상 기반 특징의 의존도가 높음으로 인해 그림 10과 같이 대부분의 RGB 정보가 제한적인 서로 다른 시야각을 가진 유사한 장면에 대해서는 예측을 통한 멀미 수준 구분이 어려운 한계를 보인다.

V. 결 론

가상현실 콘텐츠의 다양성과 시장의 확대를 위해서는 가상현실 시청 시 발생하는 멀미를 예측하고 멀미 발생 요소를 저감함으로써 시청자의 안전성을 보장함이 최우선 요소이다. 이는 제작자 혹은 서비스 제공자 입장에서 사용자의 피드백 없이 사전에 정량적인 멀미 수준의 예측이 요구됨을 의미한다. 본 연구에서는 이를 위해 다양한 멀미 유발 요소를 포함한 가상현실 콘텐츠 장면을 생성하고, 150명의 피험자를 동원한 대규모의 임상실험을 통해 데이터를 축적하였다. 이 과정에서 다른 멀미 변인을 통제하고자 새롭게 고안한 임상실험 프로토콜을 설계하였으며, 주관적 멀미 수준에 대한 의견을 수집하였다. 기존 멀미 이론과 영상처리 기법을 활용하여 가상현실 멀미 예측에 효과적인 68개의 시공간적 인지 특징을 추출하였으며, 피험자의 주관적 점수를 레이블로 하여 지도학습을 통해 정량적 가상현실 멀미 예측 모델을 완성하여 실제 임상 실험의 결과와 76%의 상관도를 달성하였다. 본 연구를 통해 도출된 가상현실 멀미 예측 방법은 콘텐츠 제작의 가이드라인으로 활용될 수 있으며, 추후 실감성과 시청 안전성을 보장하는 새로운 가상현실 서비스 창출에 막대한 도움이 될 것으로 예상된다. 추후 획득 데이터의 보완을 통해 심층신경망을 활용한 보다 높은 성능의 가상현실 멀미 예측 모델 학습을 기대할 수 있으며, 더불어 연령별, 성별 및 개인 특성에 기반한 다양한 분석이 가능할 것으로 예상된다. 나아가 제안하는 기술의 프레임워크는 가상현실 멀미뿐만이 아닌 실감성, 현실성, 몰입감 등 다양한 체험품질(Quality of Experience: QoE) 정량화에 응용할 수 있다.

REFERENCES

- [1] S. V. G. Cobb, "Measurement of postural stability before and after immersion in a virtual environment," *Applied Ergonomics*, vol. 30, no.1, pp. 47-57, Feb. 1999.

- [2] J. T. Reason and J. J. Brand, "Motion sickness," Academic Press, London, pp. 83-101, 1975.
- [3] C. M. Oman, "A heuristic mathematical model for dynamics of sensory conflict and motion sickness hearing in classical musicians," *Acta Oto-Laryngologica*, vol. 94, no. sup392, 1982.
- [4] J. E. Bos, W. Bles and E. L. Groen, "A theory on visually induced motion sickness," *Displays*, vol. 29, no. 2, pp. 47-57, March 2008.
- [5] Y. Y. Kim, H. J. Kim, E. N. Kim, H. D. Ko and H. Y. Kim, "Characteristic changes in the physiological components of cybersickness," *Psychophysiol.*, vol. 42, no. 5, pp. 616-625, Sept. 2005.
- [6] J. Mittelstaedt, J. Wacker and D. Stelling, "Effects of display type and motion control on cybersickness in a virtual bike simulator," *Displays*, vol. 51, pp. 43-50, Feb. 2018.
- [7] J. D. Moss, J. Austin, J. Salley, J. Coats, K. Williams and E. R. Muth, "The effects of display delay on simulator sickness," *Displays*, vol. 32, no. 4, pp. 159-168, Oct. 2011.
- [8] S. Sharples, S. Cobb, A. Moody and J. R. Wilson, "Virtual reality induced symptoms and effects (VRISE): comparison of head mounted display (HMD), desktop and projection display systems," *Displays*, vol. 29, pp. 58-69, Oct. 2007.
- [9] Y. T. Lin, Y. Y. Chien, H. H. Wang, F. C. Lin and Y. P. Huang, "The quantization of cybersickness level using EEG and ECG for virtual reality head-mounted display," *SID Symposium Digest of Tech. Paper*, May 2018.
- [10] D. Rahnev, H. Lau and F. P. de Lange, "Prior expectation modulates the interaction between sensory and prefrontal regions in the human brain," *J. Neuroscience*, vol. 3, no. 29, pp. 10741-10748, July 2011.
- [11] J. Guna, G. Gersak, "Influence of video content type on user's virtual reality sickness perception and physiological response," *Future Generation Comput. Syst.*, vol. 91, pp. 263-276, Feb. 2019.
- [12] D. M. Whittinghill, B. Ziegler, J. Moore and T. Case, "Nasum virtualis: a simple technique for reducing simulator sickness in head mounted VR," *Game Developers Conf. (GDC)*, 2015.
- [13] A. Singla, S. Fremerey, W. Robitza and A. Raake, "Measuring and comparing QoE and simulator sickness of omnidirectional videos in different head mounted displays," *9th Int'l Conf. Quality of Multimedia Experience (QoMEX)*, May 2017.
- [14] B. Keshavarz, B. E. Riecke, L. J. Hettinger and J. L. Campos, "Vection and visually induced motion sickness: how are they related?," *Front. Psychol.*, vol. 6, no. 472, pp. 1-11, April 2015.
- [15] J. A. Joseph and M. J. Griffin, "Motion sickness: effect of the magnitude of roll and pitch oscillation," *Aviat. Space Environ. Med.*, vol. 79, no. 4, pp. 390-396, April 2008.
- [16] N. Padmanaban, T. Ruban, V. Sitzmann, A. M. Norcia and G. Wetzstein, "Towards a machine learning approach for sickness prediction in 360 stereoscopic video," *IEEE Trans. Visualizations and Comput. Graph.*, vol. 24, no. 4, pp. 1594-1603, April 2018.
- [17] H. Ujiaki, T. Yokoi and S. Saida, "Effects of virtual body motion on visually-induced motion sickness," in *Proc. 26th Annual Int'l Conf. IEEE Eng. Med. Biology Soc. (EMBS)*, Sept. 2004.
- [18] R. S. Kennedy, N. E. Lane, K. S. Berbaum and M. G. Lilienthal, "Simulator sickness questionnaire: an enhanced method for quantifying simulator sickness," *Int'l J. Aviat. Psychol.*, vol. 3, no. 3, pp. 203-220, Nov. 2009.
- [19] C. W. Tang, "Spatiotemporal visual considerations for video coding," *IEEE Trans. Multimedia*, vol. 9, no. 2, pp. 231-238, Feb. 2007.
- [20] K. Lee, A. K. Moorthy, S. Lee, and A. C. Bovik, "3D visual activity based on natural scene statistics," *IEEE Trans. Image Process.*, vol. 23, no. 1, pp. 450-465, Jan. 2014.

저 자 소 개



오 희 석(정회원)

2017년 연세대학교 전기전자공학과 박사 졸업
2017년~2017년 삼성전자 DMC 연구소 책임연구원
2017년~2020년 한국전자통신연구원 선임연구원
2020년~현재 한성대학교 IT융합공학부 조교수

<주관심분야: 영상처리, 컴퓨터비전, 혼합현실, 심층생성모델 등>