



3D 콘텐츠 피로도 예측을 위한 데이터베이스 구축 및 3D 콘텐츠 분석 방법

+ 이상훈 연세대학교 교수

서론

3차원 영상 처리 기술 및 3차원 디스플레이의 발전은 3차원 영상 시장의 빠른 발전을 가져왔으며 차세대 방송기술로 인식되고 있으며, 최근 3D 영화관 및 3DTV 등의 다양한 3D 콘텐츠의 등장으로 3D 콘텐츠에 대한 화질 및 피로도 분석 기술이 각광받고 있다. 현재 3D 입체 영상은 양안 시차를 가지는 두 장의 영상을 이용한 스테레오 방식으로써 시청자에게 깊이감을 제공하며 편광 및 셔터글라스 방식 등 다양한 방법을 이용하고 있다. 또한, 무안경식 3D 입체 디스플레이에 대한 연구도 함께 진행되고 있으며, 보다 현실감 있는 깊이감 제공을 위한 가상현실 분야 또한 활발히 연구가 진행되고 있다. 3D 콘텐츠는 기존의 2D 영상과는 달리 깊이감이라는 새로운 사용자 경험을 제공하지만, 이에 따른 시각 피로 및 영상 멀미 등의 인체 안전성 문제를 내포한다. 스테레오 3D는 양안 시차를 강제로 이용하기 때문에, 눈에서 일어나는 제어 시스템에 문제를 일으켜 피로를 유발하며, 개인차에 따라 심하게는 극심한 어지러움과 두통, 메스꺼움을 느끼기도 한다 [1][2]. 이에 따라, 현재 3D 영상기술에서 시청자의 피로도의 원인 규명을 위한 3D 인체안전성 연구가 활발하게 이루어지고 있지만, 실제 촬영된 3D 콘텐츠를 분석하여 시청자의 인체안전성에 미칠 영향을 예측하기 위한 피로도 예측 모델 정립에 관한 연구는 부족한 실정이다. 안전한 3D 콘텐츠 제작 및 제공을 위해서는 피로도 원인 규명뿐 아니라 시청자의 주관적/객관적 피로도를 측정하고 예측하는 연구가 반드시 선행되어야 한다. 이러한 연구 흐름에 따라, 현재 3D 콘텐츠를 위한 피로도 예측 기법에 대한 표준화 작업이 IEEE-SA(Institute of Electrical and Electronics Engineers Standard Association)에서 진행되고 있다. IEEE-SA에서는 3D 휴먼팩터 워킹그룹이 신설되어 3D 콘텐츠의 피로도 분석을 위한 데이터베이스 구축 및 3D 콘텐츠 분석을 통한 3D 시각 피로도의 주관적/객관적 평가 방식에 대한 논의가 활발히 진행 중이다. 본 글에서는 현재 IEEE-SA 3D 휴먼팩터 워킹그룹에서 이루어지고 있는 3D 데이터베이스 구축 및 특성, 주관적 화질 평가 기법 등에 대해 살펴본 후, 이를 바탕으로 피로도 분석을 위한 3D 콘텐츠 분석 방법에 대해 알아볼 것이다.

3D 콘텐츠 피로도 분석을 위한 3D 데이터베이스 구축 및 특성

A. 데이터베이스 구축 (Acquisition of Database)

IEEE-SA 3D 휴먼팩터 그룹에서는 스테레오 3D 콘텐츠 데이터베이스 구축을 위해서, 두 가지 타입의 카메라 시스템을 사용하였다. 먼저 첫 번째는 일체형 3D 양안 카메라(Twin-lens) 시스템으로, 이는 두 개의 카메라가 내장되어 있으므로, 카메라 사이의 거리(Inter Ocular Distance, IOD)는 고정되어 있지만, 수렴(Vergence)의 조절을 통해, 깊이의 범위(Range of Depth)를 조절하여 촬영이 가능하다. 두 번째 카메라 타입은 두 개의 단안(Monocular) 카메라를 리그(Rig)에 연결하여 사용하는 Rig-type 2-camera 시스템이므로, 이는 리그를 활용하여 카메라 사이의 IOD를 조절할 수 있다. 이를 통해 깊이의 범위 및 깊이감 조절이 양안 카메라 시스템보다 용이하다. 하지만, 분리된

카메라 사이에 기하학적, 광학적, 백색 균형 불일치를 유발하는 좌/우 불일치 문제가 발생할 수 있다. 따라서 이러한 문제를 고려하여 촬영을 진행해야 한다. [그림 1]은 Rig-system의 카메라 세팅 및 실제 촬영 현장을 보여준다.



그림 1. Rig-system의 카메라 세팅(왼쪽)과 실제 촬영 현장(오른쪽)

이렇게 두 가지 종류의 카메라 시스템을 통해 획득된 3D 콘텐츠는 아래와 같은 특성을 고려하여 분류된다.

- 실내(Indoor)/실외(Outdoor) : 비디오의 깊이 규모는 촬영 환경에 의해 결정된다. 일반적으로, 실외 영상은 실내 영상보다 더 큰 규모의 깊이를 가진다.
- 카메라 움직임(Ego-motion)/카메라 고정(Non Ego-motion) : 카메라 움직임은 비디오 영상의 전체적인 움직임 특성을 결정짓는 데 중요한 요소이다. 카메라의 움직임은 3D 콘텐츠의 화질 및 시각적 피로에 중요한 영향을 미치는 요소이므로, 데이터베이스 구축 시 반드시 고려해야 한다.

B. 데이터베이스 후처리 과정 (Post-processing)

3D 카메라로 촬영된 3D 콘텐츠는 수직 정렬오차(VERTICAL Misalignment)나 색 부조화(Color Mismatch)와 같은 왜곡이 있을 수 있다. 이러한 왜곡을 최소화하기 위해, 수직 정렬이나 색 일치 과정을 수행하는 후처리 과정은 데이터베이스 구축을 위해 반드시 필요하다고 할 수 있다.

- 수직 정렬(VERTICAL Alignment) : 일반적인 방법으로 제작된 스테레오 3D 카메라의 리그 시스템 및 교정(Calibration)이 정확하다고 해도, 의도치 않게 좌/우 불일치가 발생할 수 있다. 이러한 왜곡은 눈의 피로와 시각적 불편함을 유발하는 잠재적 원인으로 알려져 있다. 이런 이유로, 수직 정렬은 내부적인 카메라 왜곡을 보정하는 것에 초점을 맞춘다.
- 색 일치(Color Matching) : 인간시각체계(Human Visual System, HVS)는 색 성분을 민감하게 받아들이므로, 백색 균형과 노출(Exposure)이 조절되었다고 해도, 좌/우안 영상 사이에 다른 색이 인지될 수 있다. 색 일치 과정의 목적은 두 영상의 색 차이를 보정하는 것이다.

C. 데이터베이스 분석 (Analysis of Database)

3D 콘텐츠의 피로도를 효과적으로 분석하기 위해서는 이에 적합한 데이터베이스 구축이 반드시 필요하다. 이에, IEEE-SA 휴먼팩터 그룹에서는 데이터베이스 구축 시, 다양한 깊이감(Disparity)과 움직임(Motion)을 갖는 콘텐츠를 촬영하였고, 이를 통해 다양한 측면에 대한 3D 콘텐츠 피로도 분석을 수행하였다.

- 깊이감(Depth) : [그림 2(a)]는 구축된 데이터베이스의 깊이 정보의 다양성을 나타낸다. 이는 깊이감을 의미하는 영상 내 시차

(Disparity)를 정규화(Normalization)한 평균값의 분포를 나타내고 있다. 이를 살펴보면, 데이터베이스 영상(Sequence)들의 시차값이 다양하고 넓은 분포를 갖는다는 것을 확인할 수 있다. 이는 인간시각체계의 다양한 공간적 특성을 분석하기 위한 데이터베이스가 잘 구축되었다는 것을 의미한다.

- 움직임(Motion) : [그림 2(b)]는 구축되어 있는 데이터베이스의 움직임 정보의 다양성을 나타낸다. 이는 XY 평면과 Z축 움직임의 크기에 대한 정규화 평균값을 나타내고 있다. 이를 살펴보면, 데이터베이스로 구축된 다양한 콘텐츠들의 움직임 정보가 다양하고 넓은 분포를 갖는다는 것을 확인할 수 있다. 즉, 3D 콘텐츠의 움직임 정보에 따른 인간시각체계의 다양한 시간적 특성을 분석하기 위한 데이터베이스가 잘 구축되어 있다는 것을 나타낸다.

3D 콘텐츠 피로도 분석을 위한 주관적 평가 방법 (Subjective Assessment Methodology)

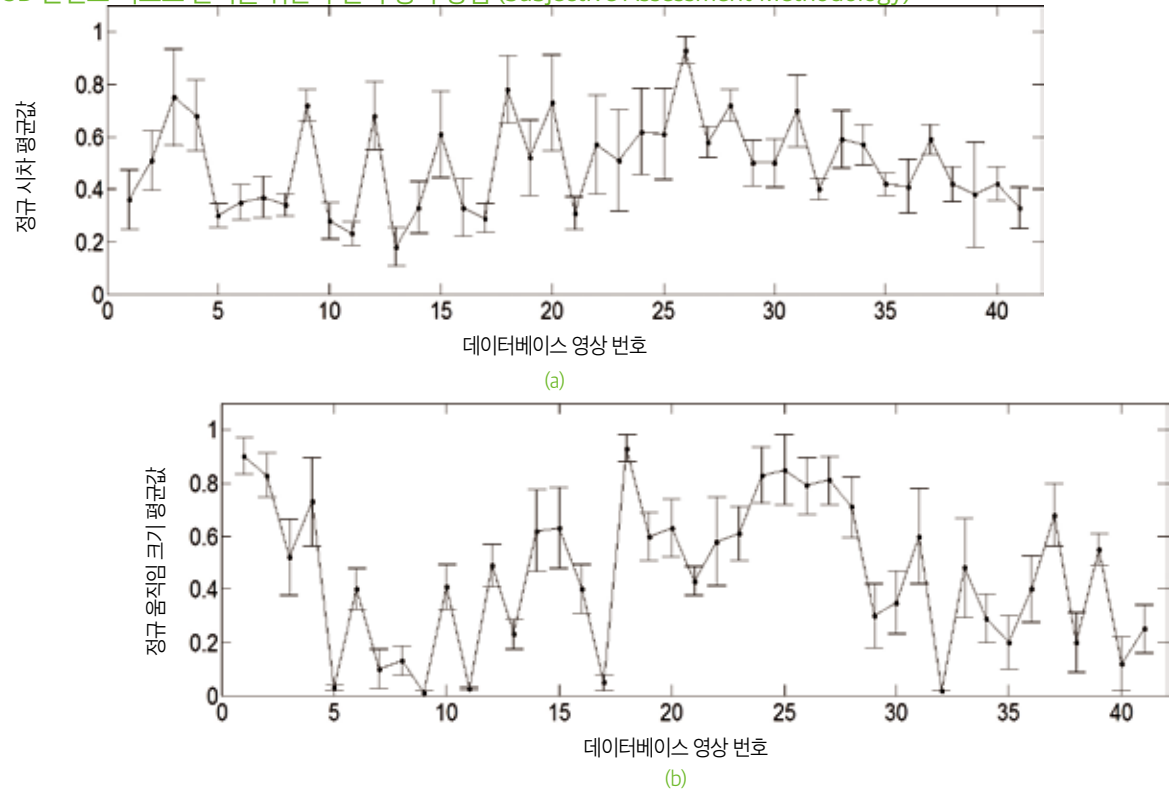


그림 2. 데이터베이스분석 결과. (a) 정규화 된 시차. (b) 정규화 된 움직임 크기

3D 입체영상은 기존의 2D 영상과는 달리 다양한 원인의 복합적인 영향을 받아 시각적 피로를 유발하게 된다. 이러한 피로현상을 분석하기 위해서 3D 영상에 대한 주관적 평가 연구가 활발히 이루어지고 있다. 특히, 2D 영상과는 다른 3D 입체 영상의 특성 및 평가 환경 등이 고려된 신뢰도 높은 주관적 평가 방법이 반드시 필요하다. 3D 입체영상의 주관적 평가 방법에 대해서는 휴먼팩터가 시간에 따라 크게 변하는 특성 때문에 연속적인 평가(Continuous Assessment) 방법이 많이 사용되고 있으며, 기존에는 단일자극연속화질평가(Single Stimulus Continuous Quality Evaluation, SSCQE) 방식을 주로 사용하였다 [3]. 하지만, 이는 3D 디스플레이 시청 시 시청자의 몰입감을 방해할 수 있으며, 집중력 저하 및 다양한 문제가 나타날 수 있다. 따라서 IEEE-SA 휴먼팩터 그룹에서는 시청자, 화질 평

가 도구, 3D 디스플레이 사이의 상호작용을 통해 효과적인 주관적 3D 피로도 평가를 수행하는 새로운 방식인 Multi-modal Interaction Continuous Scoring of Quality (MICSQ)를 제안하였다 [4]. MICSQ는 3D 디스플레이와 (가능하다면 여러 개의)화질 평가 도구간의 무선 통신을 이용하여, 한 명 혹은 여러 명의 시청자가 교실이나 영화관 같은 환경에서 동일한 비디오를 평가하는 실험에 동시에 참여할 수 있도록 한다.



그림 3. MICSQ 프로토콜

[그림 3]은 MICSQ 프로토콜을 나타낸다. MICSQ는 주관적 화질평가 과정을 분리된 상호작용으로 나눈다;

- 1) 평가 도구로 사용되는 태블릿과 3D 디스플레이 사이의 상호작용(기기 간 상호작용), 이 경우에는 무선 네트워크 프로토콜을 이용함.
 - 2) 태블릿과 시청자 사이의 상호작용(인간의 상호작용), 이 경우에는 효과적으로 주관적 결과를 기록하기 위해 촉각과 청각을 이용함.
- 또한, 3D 디스플레이 서버와 화질 평가 도구(태블릿) 간의 상호 작용 과정은 실시간으로 이루어진다. 서버는 각 3D 비디오의 재생과 화면 표시, 그리고 시청자 평가 점수의 저장을 다룬다. 인간의 상호작용 과정에서는 태블릿에 기반한 다중-시청자 인터페이스에 대한 새로운 아이디어를 제시한다. 앞서 언급한 집중력 저하 문제는 연속적인 평가 수행에 대한 방해나 몰입감 문제로 이어질 수 있다. 따라서 이러한 문제를 다루기 위해 다양한 모드의 프로토콜을 제안한다. 시청자가 평가를 하는 처음부터 끝까지 평가 도구는 청각과 촉각적인 신호를 전달한다. [표 1]에서는 이러한 신호들을 나타낸다.

	집중력 저하 문제	몰입감 문제
청각 신호	주기적 '뽀' 소리	무작위의 '뽀' 소리 / 실시간 점수 알림
촉각 신호	주기적 진동	무작위의 진동

표 1. 다중모드 평가 도구 신호

3D 시각피로 객관화를 위한 콘텐츠 분석 방법

A 인간 시각 인지 분석 (Human Visual Perception Analysis)

IEEE-SA 휴먼팩터 그룹에서는 3D 시각피로 객관화를 위한 콘텐츠 분석을 위해 우선적으로 인간의 시각 인지 및 주의(Human Visual

Perception and Attention)를 분석하였다. 인간의 시각 인지 특성은 콘텐츠 분석을 위해서 반드시 선행되어야 하므로, 콘텐츠 내의 시각적 관심영역 추출 기법을 연구하였다. [그림 4]는 관심 영역 추출 알고리즘의 전체적인 구조도이다 [7].

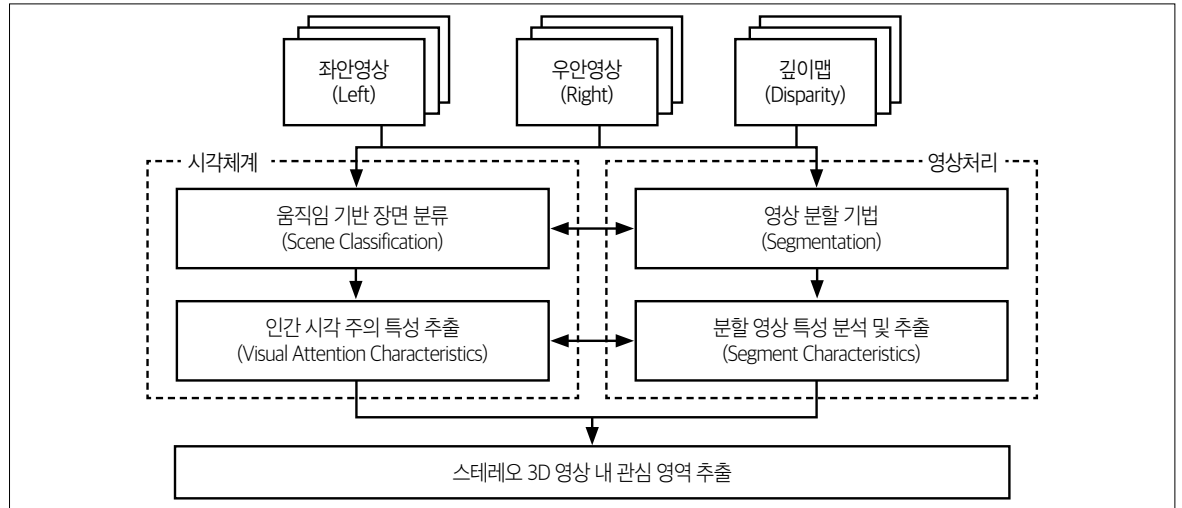


그림 4. 3D 비디오와 시차 맵의 입력을 통한 관심영역 추출

비디오 영상과 시차 맵은 시각체계 모듈과 영상처리 모듈의 입력으로 들어가게 된다. 시각 체계 모듈에서는, 움직임 정보에 기반한 장면 분류(Scene Classification)를 통해 인간의 시각적 관심영역 특성을 적응적으로 추출하였다. 각 장면은 총 5가지의 종류로 구분되며, 각각의 장면에 따른 시선의 특성은 다음과 같이 나타난다.

- 1) 카메라 정지 : 움직이는 물체에 주로 시선이 감,
- 2) 움직이는 물체를 추적하는 카메라 움직임 : 추적하는 물체에 주로 시선이 감,
- 3) 2D 카메라 움직임 : 주위와 상이한 움직임을 갖는 물체에 주로 시선이 감,
- 4)/5) 카메라 줌인/줌아웃 : 영상의 가운데 부분에 주로 시선이 감.

또한, 모든 장면에 대해서 앞으로 튀어나와 입체감을 주는 물체(음시차)에 시선이 많이 가는 것을 확인할 수 있다. 이러한 특성을 활용하여, 영상처리 모듈에서는, 장면별로 시각적으로 관심을 갖는 물체를 추출하고 분할된 영역의 특성을 얻기 위해 분할(Segmentation)을 수행한다. 두 가지 모듈에서 얻은 결과를 이용하여, 3D 입체 영상 내 관심영역을 자동으로 검출할 수 있다. [그림 5]는 관심영역 추출 결과 및 눈동자 추적(Eye-tracker) 결과를 나타낸다. 이를 살펴보면, 실제 눈동자 추적결과와 자동으로 추출된 관심영역이 거의 유사하게 나타나는 것을 확인할 수 있다.



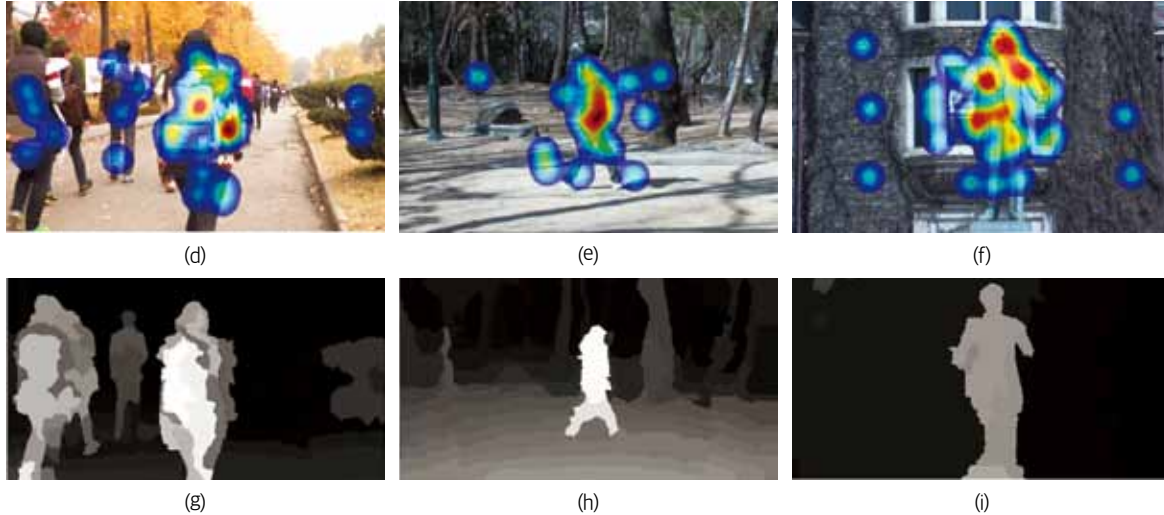


그림 5. 관심영역추출 결과 (a)~(c) 왼쪽 원본 좌안 이미지 : "Marathon 1"의 42번째 프레임 (장면 1), "Walking-person 1"의 58번째 프레임 (장면 2), "Statue 3"의 109번째 프레임 (장면 3). (d)~(f) 아이트래커 실험 결과. (g)~(i) 관심영역 추출 결과.

B. 3D 콘텐츠 피로도 분석을 위한 시각적 복잡도 분석 (Visual Activity Analysis)

3D 콘텐츠의 피로도 분석을 위해서는 우선적으로 3D 콘텐츠가 포함하고 있는 다양한 성분(깊이감, 움직임, 색, 텍스처 등)에 대한 분석이 선행되어야 한다. 따라서 이에 IEEE-SA 휴먼팩터 그룹에서는 3D 비디오 콘텐츠의 통계적 특성을 3D 복잡도(Visual Activity) 관점에서 분석하였다 [8]. 3D 복잡도는 3D 공간에서의 시차와 움직임의 자유 정도(Randomness)에 기반하여 콘텐츠의 성분을 정량화한다. 3D 물체는 다양한 지점에서 다르게 움직이고 위치할 수 있기 때문에, 3D 물체의 역동성은 시각적 피로도, 관심영역, 그리고 다른 요소들과 긴밀하게 연관되어 있다. 게다가, 2D 이미지에서 3D 공간으로의 확장을 통해, 물체는 더 많고 다양한 방향으로 움직일 수 있다. 특히, 깊이 움직임(z축으로의 움직임)은 시각적 피로와 불편함의 원인이 되므로 이에 대한 분석이 반드시 필요하다고 할 수 있다.

3D 복잡도를 얻기 위해 우선적으로 콘텐츠의 깊이감, 움직임, 색, 텍스처 맵을 추출하여, 각 요소의 분포를 측정한다. 각각의 맵을 통계적 모델에 적용하기 위해 우선적으로 웨이블릿 변환(Wavelet Transform)을 수행하여, 공간, 주파수 성분으로 3D 콘텐츠를 변환한다.

이러한 콘텐츠의 통계적 성분에 대한 정량화를 위해 자연 영상 통계(Natural Scene Statistics, NSS) 분석을 수행하였다. 이는 웨이블릿 계수의 정량적 분석에 기반 하며, 웨이블릿 계수의 개수를 크기에 따라 측정하는 방법을 이용하여 NSS를 분석한다. 웨이블릿 계수의 히스토그램을 구하기 위해 확률 질량 함수(Probability Mass Function, PMF)를 모델링한다.

다음 단계로, 각 요소의 웨이블릿 계수를 통계적 모델에 적용하고 그것의 모양 변수(Shape Parameter)를 얻는다. 보통의 일반 영상의 웨이블릿 계수 분포는 일반적으로 두꺼운 꼬리와 높은 최상치를 갖는 것으로 알려져 있다. 많은 연구의 결과, 웨이블릿 계수의 분포는 일반화된 가우시안으로 나타내는 것이 적합하다[9]. 가우시안으로 나타낸 웨이블릿 계수를 관찰했을 때, 고주파 대역에서의 웨이블릿 계수 분포가 증가하면, 모양 변수의 값은 감소한다. 따라서 웨이블릿 계수가 넓고 무작위로 분포할수록, 모양계수의 값은 작아진다. 이러한 관찰에 기반하여, 모양변수를 3D 복잡도를 측정하는 데 주요한 변수로 사용하며, 이를 정량화함으로써, 3D 복잡도를 계산할 수 있다. 이러한 3D 복잡도를 통해 3D 콘텐츠를 효과적으로 분석할 수 있고, 통계적 분산과 자유도를 정량화할 수 있다. [그림 6]은 "Car 1"과 "University 1" 영상의 3D 복잡도를 시간 축으로 나타낸 그림이다. 이를 살펴보면, 시간 축에 대해 영상 내 포함된 복잡도 정도가 다양하게 나타나는 것을 확인할 수 있다. 이는 추후 3D 콘텐츠 분석에 사용될 수 있을 것으로 예상되며, 이러한 복잡도와 같은 다양한 영상 분석 요소를 활용한 객관적 3D 콘텐츠 피로도 예측 모델을 구축할 수 있을 것으로 예상된다.

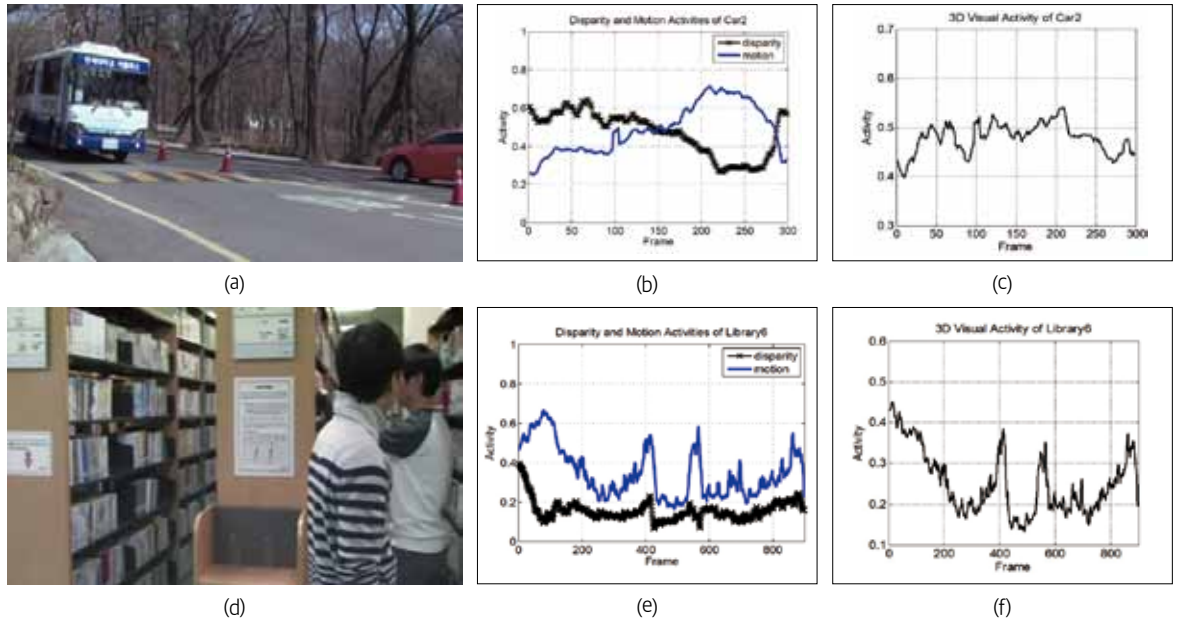


그림 6. 3D 복집도 측정 결과. (a) "Car2". (b) "Car2"의 시차, 움직임 복집도. (c) "Car2"의 3D 복집도.
(d) "Library6". (e) "Library6"의 시차와 움직임 복집도. (f) "Library6"의 3D 복집도

맺음말

3D 콘텐츠의 피로도 분석을 위한 데이터베이스 구축 및 분석 방법, 주관적 피로도 평가 방법 및 콘텐츠 분석 기법에 대한 현재 IEEE-SA 휴먼팩터 그룹의 진행 상황에 대해 간략히 개관해보았다. 본 고에서 언급하지 않은 피로도 분석을 위한 뇌 신경계의 피로 모델이나 실제 객관적 피로도 매트릭에 대한 표준화 작업도 현재 함께 진행 중에 있다. 앞서 언급한 것처럼, 3D 콘텐츠의 원활한 제공을 위해서는 인체 안전성이 우선적으로 보장되어야 한다. 지금까지 활발히 연구되어온 다양한 연구를 통해 3D 콘텐츠의 인체안전성이 보장되는 기준이 확립될 것 기대하며, 이를 통해 안전하고 생동감 넘치는 3D 콘텐츠 서비스가 제공될 것 기대한다. 

참고문헌

- [1] S. Pastoor, "Human factors of 3DTV: an overview of current research at Heinrich-Hertz-Institut Berlin", Stereoscopic Television, IEE Colloquium on., pp. 11/1 - 11/4, 1992.
- [2] M. Lambooi and W. IJsselstein, "Visual discomfort and visual fatigue of stereoscopic displays: a review," Journal of Imaging Science and Technology, 53(3) 030201(2009).
- [3] W. IJsselstein, H. de Ridder, R. Hamberg, D. Bouwhuis and J. Freeman, "Perceived depth and the feeling of presence in 3DTV," Displays, vol. 18, no. 4, pp. 207-214, 1998.
- [4] T. Kim, J. Kang, S. Lee and A. C. Bovik, "Multimodal interactive continuous scoring of subjective 3D video quality of experience," IEEE Trans. on Multimedia, vol. 16, no. 2, pp. 387-402, 2014.
- [5] IEEE Standards Association Stereoscopic (3-D imaging) Database : <http://grouper.ieee.org/groups/3dhf/>
- [6] L. Goldmann, F. De Simone, and T. Ebrahimi, "A comprehensive database and subjective evaluation methodology for quality of experience in stereoscopic video," Electron. Imag., 3D Image Process. Applicat., 2010.
- [7] H. Kim, S. Lee and A. C. Bovik, "Saliency prediction on stereoscopic videos," IEEE Trans. on Image Processing, vol. 23, no. 4, pp. 1476-1490, 2014.
- [8] K. Lee, A. K. H. Moorthy, S. Lee and A. C. Bovik, "3D visual activity assessment based on natural scene statistics," IEEE Trans. on Image Processing, vol. 23, no. 1, pp. 450-465, 2014.
- [9] Y. Liu, L. K. Cormack and A. C. Bovik, "Statistical Modeling of 3-D Natural Scenes With Application to Bayesian Stereopsis," IEEE Trans. on Image Processing, vol. 20, no. 9, pp. 2515-2530, Sep. 2011.