

3DoF+ 영상에 대한 객관적 영상 품질 평가 기법 성능 비교에 대한 연구

정종범¹⁾, 이순빈¹⁾, 김인애¹⁾, 류일웅²⁾, 김성빈¹⁾, 장동민¹⁾, 류은석¹⁾

1) 성균관대학교 컴퓨터교육과, 2) 가천대학교 컴퓨터공학과
uof4949@skku.edu, soonbinlee@skku.edu, inaelk@skku.edu, dflnd96@gc.gachon.ac.kr,
beencooke@skku.edu, jdm360@skku.edu, esryu@skku.edu

A Study of Video Quality Assessment Methods Performance Comparison for 3DoF+ Video

Jong-Beom Jeong¹⁾, Soonbin Lee¹⁾, Inae Kim¹⁾, Il-Woong Ryu²⁾, Sungbin Kim¹⁾,
Dongmin Jang¹⁾, Eun-Seok Ryu¹⁾

1) Department of Computer Education, Sungkyunkwan University,
2) Department of Computer Engineering, Gachon University

요 약

가상 현실 공간에서 실사 영상 및 그래픽 환경에 대해 사용자 움직임에 대응하는 영상을 제공하는 three degrees of freedom plus (3DoF+), 6DoF 시스템은 사용자의 불편함을 최소화하기 위해 다수의 초고화질 및 고해상도의 영상 전송이 요구된다. 다수의 영상을 사용해 사용자 시점 영상을 생성할 때, 기존의 영상 압축 코덱에서 발생했던 것과는 다른 형태의 품질 저하가 발생한다. 해당 품질 저하를 잘 검출하는 평가 기법을 찾기 위해 본 논문은 3DoF+ 환경에서 합성된 사용자 시점 영상에 대한 다양한 객관적 영상 품질 평가 기법에 대한 성능 평가를 소개한다. 제안한 결과를 통하여 사용자 경험과 잘 부합하는 객관적 영상 품질 평가 기법을 찾기 위한 방향을 제시한다.

1. 서 론

최근 가상 현실(virtual reality; VR)에 대한 수요가 늘어감에 따라, 몰입형 미디어를 지원하기 위한 기술 및 장비의 중요성이 강조되고 있다. Moving picture experts group(MPEG)에서는 MPEG-Immersive (MPEG-I) 그룹을 설립하여 몰입형 미디어의 표준화를 위해 가상 현실에서의 사용자 자유도를 3단계로 구분하였다. 그 중 하나가 three degrees of freedom(3DoF)로, 고정된 시점에서 3 방향으로의 회전을 포함한 움직임을 지원하며, 2020년 현재 표준화가 완료되었다. 6DoF는 회전뿐만 아니라 사용자의 시점 움직임도 완벽히 지원하여 가상 현실 공간에서 사용자의 모든 움직임을 지원하나, MPEG-I에서는 높은 연산량, 대역폭 등의 이유로 2022년까지 점진적으로 표준화를 진행할 예정이다. 이에 3DoF와 6DoF 중간 단계로 3DoF+가 등장하였는데, 이는 가상 현실 공간에서 의자에 앉은 사용자의 머리 움직임 정도를 지원하여 제한된 6DoF로서의 기능을 수행한다. 3DoF+ 및 6DoF는 자유로운 사용자 움직임을 지원하므로, 이를 위해 여러 위치에서 취득된 영상들을 동시에 압축 및 전송하여 사용자 시점의 영상을 합성하는데, 이는 높은 대역폭과 연산량을 요구한다.

이에 2019년 1월, MPEG-I는 3DoF+에 대한 call for proposals(CfP)를 공표하였고[1], 2019년 4월에 philips, technicolor 등 7개의 기업 및 연구소가 5개의 call for response(CfR)[2]를 제출하였다. 또한 효율적인 3DoF+ 영상 스트리밍을 위해 비대칭 다운샘플링[3], 전처리 단계에서의 시점 간 중복성 제거 및 병합[4], 프로젝션 변경 기법[5] 등의 방법이 제시되었다.

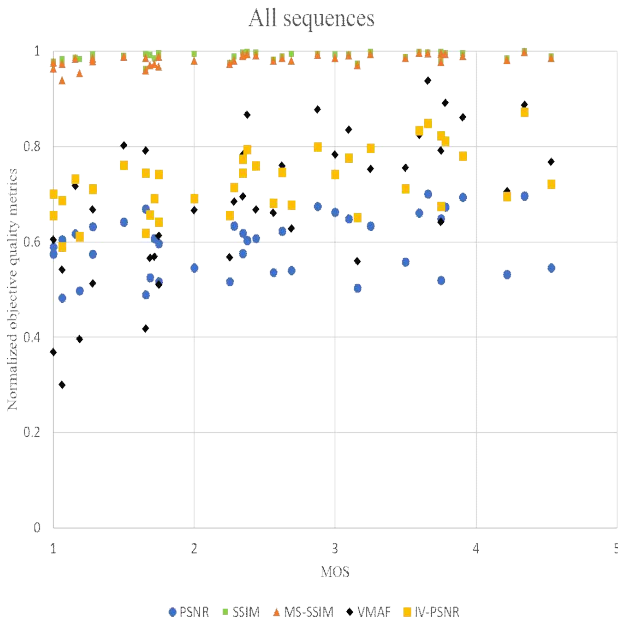
3DoF+ CfR 성능 평가 시 전통적인 객관적 영상 품질 평가 지표인 peak signal-to-noise ratio(PSNR)가 사용된 바 있으나, 3DoF+ 환경에서 발생하는 품질 저하를 검출하기엔 적합하지 않다. 뷰 합성 기술을 이용하여 여러 시점의 영상을 이용해 합성된 영상은 물체의 가장자리 영역에서 노이즈가 발생하는데, 이는 기존의 영상 품질 평가에서 검출하기 어려워 사용자 평가에 부합하는 객관적 평가를 도출하기 어렵다.

본 연구는 최신 객관적 영상 품질 평가 기법들을 선정하여 3DoF+ 영상에 대한 주관적 품질 평가와의 상관관계를 도출한다. 객관적 품질 평가 기법으로 PSNR, structural similarity (SSIM), multi-scale SSIM (MS-SSIM), video multimethod assessment fusion (VMAF), immersive video PSNR (IV-PSNR)을 선정하였다.

본 논문의 구성은 다음과 같다. 2 절에서는 3DoF+ 영상에 대한 객관적 품질 평가 지표와 주관적 평가 결과 및 상관관계를 설명한다. 3 절에서는 본 논문의 결론을 서술한다.

2. 3DoF+ 영상에 대한 객관적 영상 품질 평가 기법과 주관적 평가와의 비교

본 절은 3DoF+ CfR에 대한 주관적 평가에 기반하여 객관적 품질 평가지표와의 비교를 소개한다. 주관적 평가 지표로는 mean opinion score(MOS)를 사용하였고, 해당 데이터는 2019년 3월 3DoF+ CfR에 대해 진행된 주관적 평가에서 취득되었다[6]. MPEG-I의 common test condition(CTC)[7]에 기반하여 ClassroomVideo,



(그림 1) MOS와 객관적 영상 품질 평가 기법과의 상관관계, 가로축: MOS, 세로축: 정규화된 객관적 영상 품질 평가 기법

(표 1) MOS와 객관적 영상 품질 평가 기법과의 상관 계수

품질 평가 기법	MOS와의 상관관계
PSNR	0.3012
SSIM	0.3932
MS-SSIM	0.5694
VMAF	0.6601
IV-PSNR	0.5539

TechnicolorMuseum, TechnicolorHijack, TechnicolorPainter, IntelKermit의 총 5개의 테스트 시퀀스가 사용되었다. 영상은 high-efficiency video coding (HEVC)를 사용하여 부호화 및 복호화되었고, 영상 모두를 선택하여 부호화하는 집합인 P01 anchor와 영상 중 일부를 선택하는 P02 anchor가 사용되었다. 객관적 영상 품질 평가 기법으로는 PSNR, SSIM, MS-SSIM, VMAF, IV-PSNR이 선택되었고, 3DoF+ 뷰 합성에는 reference view synthesizer (RVS)[8]가 사용되었다.

그림 1은 MOS와 객관적 영상 품질 평가 기법 간 상관 관계를 나타낸 그래프이다. 가로축은 MOS, 세로축은 0과 1 사이로 정규화된 객관적 영상 품질 평가 기법을 나타낸다. PSNR과 IV-PSNR은 최소값을 0, 최대값을 60으로 설정하여 정규화를 진행하였다. 전통적인 품질 평가 기법인 PSNR은 MOS와의 연관성을 보여주지 못하고 분산되어 있는 반면, 주관적 영상 품질 평가를 전제하여 개발된 VMAF는 MOS와의 연관성을 잘 보여준다. 표 1은 MOS와 객관적 영상 품질 평가 기법과의 상관 계수를 나타낸다. VMAF가 가장 높은 값을 보여주었고, MS-SSIM, IV-PSNR, SSIM, PSNR 순으로 좋은 결과를 보여주었다. 원본 영상과 부호화 후 복호화된 영상 간 구조적 유사성

을 측정하는 SSIM은 5개 기법 중 4위를 기록하였지만, 이를 다양한 크기의 영상에서 측정하여 결과를 취합하는 MS-SSIM은 2위를 기록하였다. 뷰 합성 시 물체 가장자리에서 발생하는 사용자가 인식하지 못할 정도의 노이즈를 무시하는 IV-PSNR은 3위를 기록하였다.

3. 결론

본 논문은 3DoF+ 및 6DoF 환경에서 제공되는 영상이 전통적인 객관적 영상 품질 평가에 있어서 가지는 한계점을 지적하고 주관적 평가를 기반으로 최근 제안된 객관적 영상 품질 평가 기법들과의 비교를 수행하였다. 해당 비교를 통해 VMAF, MS-SSIM, IV-PSNR 등의 품질 평가 기법이 가상 현실 몰입형 미디어 시스템 평가에 사용될 수 있으며, 추후 각 품질 평가 기법을 분석하여 주관적 평가에 더 부합하는 평가 기법을 개발할 계획이다.

ACKNOWLEDGMENT

본 연구는 한국전력공사의 2016년 선정 기초연구개발과제 연구비에 의해 지원되었음 (과제번호 : R17XA05-68). 이 논문은 또한 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1A2C1010476).

참고 문헌

- [1] ISO/IEC JTC1/SC29/WG11, "Call for Proposals on 3DoF+ Visual", 125th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/n18145, 2019.
- [2] Bart Kroon, Bart Sonneveldt, "Philips response to CfP on 3DoF+", 126th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG2019/m47179, 2019.
- [3] JongBeom Jeong, Dongmin Jang, Jangwoo Son, Eun-Seok Ryu, "3DoF+ 360 Video Location based Asymmetric Down-sampling for View Synthesis to Immersive VR Video Streaming", Sensors, 18(9):3148, 2018.
- [4] Jong-Beom Jeong, Soonbin Lee, Dongmin Jang, Eun-Seok Ryu, "Towards 3DoF+ 360 Video Streaming System for Immersive Media", IEEE Access, Vol. 7, pp. 136399-136408, 2019.
- [5] 정종범, 이순빈, 장동민, 김성빈, 이상순, 류은석, "A Study on Projection Conversion for Efficient 3DoF+ 360-Degree Video Streaming", 방송공학회논문지, 제24권 제7호, pp. 1209-1220, 2019.
- [6] Vittorio Baroncini, Giacomo Baroncini, "Report of the evaluation of the responses to the CfP on 3DoF+ Visual", 126th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG2019/m47979, 2019.
- [7] Joel Jung, Bart Kroon, Renaud Doré, Gauthier Lafruit, Jill Boyce, "Common Test Conditions on 3DoF+ and Windowed 6DoF", 124th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/n18089, 2018.
- [8] ISO/IEC JTC1/SC29/WG11, "Reference View Synthesizer (RVS) manual", 124th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/n18068, 2018.