

가상 현실 서비스에서의 객관적 품질 측정

김성빈, 류영일, 정종범, 김인애, 이순빈, 김승환, 류은석

성균관대학교 컴퓨터교육과

beencooke@skku.edu, yeongilyoo@gmail.com, uof4949@skku.edu, inaelk@skku.edu,

soonbinlee@skku.edu, whitekomani@skku.edu, esryu@skku.edu

Objective Quality Metrics for Virtual Reality Services

Sungbin Kim, Yeongil Ryu, Jong-Beom Jeong, Inae Kim, Soonbin Lee, Seunghwan

Kim, Eun-Seok Ryu

Department of Computer Education, Sungkyunkwan University

요약

가상 현실 시스템은 전방위 영상을 획득하고 가상 시점을 생성하여 사용한다는 점에서 기존 2차원 영상과 차이가 있다. 따라서 기존 2차원 영상의 품질을 측정하기 위해 개발된 기법을 이용하여 가상 현실 영상의 품질을 측정할 때 주관적 품질과의 상관성이 낮아지는 문제가 발생한다. 본 논문은 기존의 객관적 품질 측정 기법들을 이용하여 주관적 품질과의 상관성을 높이는 방법을 소개한다.

1. 서론

가상 현실 서비스에 대한 관심이 증가하면서 360도 카메라로 획득된 전방위 영상을 이용하는 가상 현실 콘텐츠의 생산과 수요가 증가하고 있다. Moving picture experts group(MPEG)에서는 고정된 위치에서의 360도 영상을 제공하는 3 degrees of freedom(3DoF)과 이에 약간의 자유도를 추가한 3DoF+에 대한 표준화를 진행하였으며, 현재는 완전한 자유도를 제공하는 6DoF 시스템을 위한 표준화 과제를 진행하고 있다.

양질의 서비스를 제공하기 위해서는 사용자가 경험할 영상의 품질을 예측하는 과정이 필요한데, 영상의 주관적인 품질을 가장 정확하게 알 수 있는 방법은 당연하게도 사람이 직접 품질을 평가하는 것이다. 하지만 생산되는 모든 영상에 대하여 사람이 직접 평가하는 것은 시간적, 경제적 제약이 따르기 때문에 품질 측정 알고리즘을 이용하여 주관적 품질을 예측한다.

품질 측정 알고리즘으로 영상의 품질을 측정하는 것을 객관적 품질 측정이라 하며, peak signal-to-noise ratio(PSNR)와 structural similarity(SSIM)가 오랜 기간 보편적으로 사용되어 왔지만, 비교적 최근 개발된 multi-scale SSIM(MS-SSIM), immersive video PSNR(IV-PSNR), video multimethod assessment fusion(VMAF) 등 사람의 시각 체계를 고려하는 측정 알고리즘이 개발되면서 주관적 품질을 보다 정확하게

예측할 수 있게 되었다[1]-[5].

기존과 마찬가지로 가상 현실 콘텐츠 또한 사용자가 경험할 영상의 품질을 예측해야 하는데, 6DoF 영상은 가상 시점을 생성하여 사용한다는 점에서 기존의 영상과 차이가 있다. 따라서 2차원 영상의 품질을 측정하는 기법으로 가상 현실 영상의 품질을 측정했을 때 주관적 품질과 상관성이 낮아지는 문제가 발생한다. 본 논문에서는 기존의 여러 가지 객관적 품질 측정 기법을 이용하여 가상 현실 영상의 주관적 품질과의 상관성을 높이는 방법을 소개한다.

2. 관련 연구: 객관적 품질 측정 기법

본 절에서는 객관적 품질 측정 기법 중 본 논문이 제시하는 방법에 사용된 기법들에 대하여 간략히 설명한다.

2.1 PSNR

PSNR은 전통적인 영상 품질 측정 기법으로, 최대 신호와 잡음의 비율로 품질을 측정한다. 잡음이란 원본 영상과 손실 영상의 차이를 의미하며 mean squared error(MSE)를 사용하여 값을 구한다. 단순히 원본 영상의 픽셀과 손실 영상의 픽셀의 차이만을 고려하기 때문에 사람의 시각 체계를 고려하지 않는다는 단점이 있지만, 직관적이고 계산이 쉽다는 장점 때문에 여전히

보편적으로 사용되는 측정 기법이다.

2.2 SSIM

SSIM은 PSNR 기법이 사람의 시각 체계를 고려하지 않는다는 단점을 보완하기 위해 개발된 측정 기법으로, 영상의 구조 정보를 반영하여 품질을 평가하는 시각 체계의 특성을 고려한다. 측정값은 원본 영상과 손실 영상의 휘도, 명암, 구조 정보를 비교하여 계산하며 PSNR과 함께 보편적으로 사용되는 측정 기법이다.

2.3 MS-SSIM

MS-SSIM은 SSIM과 동일하게 영상의 휘도, 명암, 구조 정보를 비교하지만, 디스플레이 해상도와 시청 거리를 고려한다는 점에서 차이가 있다. 다운샘플링과 저역 통과 필터를 통해 다양한 해상도와 품질의 영상을 생성하여 SSIM 값을 계산하고 시청 환경에 적응적으로 가중치를 합하여 품질을 측정하는 기법이다.

2.4 IV-PSNR

IV-PSNR은 360도 영상과 같은 몰입형 비디오의 품질을 측정하기 위해 개발된 기법이다. 가상 시점 생성 과정에서 발생하는 오류가 대부분 사물의 경계에서 발생하는데, 사람의 시각 체계는 이러한 왜곡을 화질에 크게 반영하지 않는다는 특성을 고려한다. IV-PSNR은 사물의 경계에서 오류가 발생하면 블록 단위로 PSNR을 계산하는데, 이 방식을 이용하면 픽셀과 픽셀의 차이만을 비교하는 PSNR에 비해 측정 값의 하락 폭을 줄임으로써 주관적 품질과의 차이를 줄이는 측정 기법이다.

2.5 VMAF

VMAF는 기존의 품질 측정 기법들을 결합한 측정 기법으로, 주관적 평가 점수의 원본 영상과 손실 영상의 정보량을 비교하는 visual information fidelity(VIF) 기법과 품질에 영향을 주는 세부 정보의 손실을 측정하는 detail loss metric(DLM) 기법을 support vector machine(SVM)으로 학습 후 가중합하여 값을 계산하는 머신 러닝 기반 품질 측정 기법이다.

3. 가중치 할당 기법

본 논문에서 제안하는 가중치 할당 기법은 객관적 품질 측정 기법과 주관적 품질 평가 지표인 mean opinion score(MOS)의 상관성을 높이기 위해 역전파 알고리즘을 이용한다. 측정 과정은 여러 테스트 영상에 대하여 미리 측정된 5개의 객관적 품질 측정 기법 점수에 역전파 알고리즘을 이용하여 각각의 가중치를 구한 뒤 가중합한 결과로 영상의 품질을 측정한다.

4. 실험 및 평가분석

가중치 할당 기법과 PSNR, SSIM, MS-SSIM, VMAF, IV-PSNR의 성능을 비교하기 위한 실험을 진행하였다. 실험에 사용된 테스트 영상은 MPEG-I의 common test condition(CTC)에 기반하여 ClassroomVideo, TechnicolorMuseum, TechnicolorHijack, TechnicolorPainter, IntelKermit으로 총 5개가 사용되었다[6]. 또한 테스트 영상은 high-efficiency video coding(HEVC) 기술로 인코딩 및 디코딩되었고, P01 anchor와 P02 anchor가 사용되었다.

역전파 알고리즘을 이용하여 얻어낸 가중치는 PSNR, SSIM, MS-SSIM, VMAF, IV-PSNR이 각각 -0.21289 , -0.41435 , -0.61026 , 0.03103 , 0.22695 이며, bias는 -0.66383 이다. 가중치를 적용하여 계산한 MOS와의 상관계수는 0.7525 이며, 위의 5가지 품질 측정 기법과 MOS와의 상관계수를 비교한 표는 표 1과 같다.

표 1. 객관적 영상 품질 측정 기법 및 제안하는 기법과 MOS와의 상관계수

품질 측정 기법	MOS와의 상관계수
PSNR	0.3012
SSIM	0.3932
MS-SSIM	0.5694
VMAF	0.6601
IV-PSNR	0.5539
가중치 할당 기법	0.7525

4. 결 론

실험 결과는 표 1과 같이 제안하는 기법과 MOS의 상관관계수가 기존의 5가지 품질 측정 기법보다 높은 것을 알 수 있다. 하지만 가상 현실 영상에 대한 MOS 데이터가 한정적이기 때문에 결과의 신뢰성이 높지 않다. 가상 현실 영상에 대한 MOS 데이터를 충분히 확보하여 결과의 신뢰성을 높이기 위한 후속 연구가 필요하다.

사 사 문 구

이 논문은 2020년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-00920,(세부2)중대형 공간용 초고해상도 비정형 플렌옵틱 영상 저장/압축/전송 기술 개발)

참 고 문 헌

[1] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," in The Handbook of Video Databases: Design and Applications (B. Furht and O. Marques, eds.), pp. 1041-1078, CRC Press, Sept. 2003.

[2] Z. Wang, E. P. Simoncelli, and A. C. Bovik. "Multi-scale structural similarity for image quality assessment," In IEEE Asilomar Conference on Signals, Systems and Computers, volume 2, pages 9-13, 2003.

[3] ISO/IEC JTC1/SC29/WG11, "Objective quality metric for immersive video", 127th MPEG meeting of ISO/IEC JTC1/SC29/WG11, MPEG/m48093, 2019

[4] T.-J. Liu, J. Y. Lin, W. Lin, and C.-C. J. Kuo, "Visual Quality Assessment: Recent Developments, Coding Applications and Future Trends," APSIPA Transactions on Signal and Information Processing, 2013.

[5] J. Y. Lin, T.-J. Liu, E. C.-H. Wu, and C.-C. J. Kuo, "A Fusion-based Video Quality Assessment (FVQA)

Index," APSIPA Transactions on Signal and Information Processing, 2014.

[6] Jong-Beom Jeong, Soonbin Lee, Inae Kim, Il-Woong Ryu, Sungbin Kim, Dongmin Jang, Eun-Seok Ryu, "A Study of Video Quality Assessment Methods Performance Comparison for 3DoF+ Video", Korean Society for Internet Information (KSII) Spring Conference, May. 22-23, 2020.