

깊이 맵을 활용한 2D 영상의 3D 재구성 기술에서 품질 향상을 위한 부가 정보 활용에 대한 연구

이주현*, 김성훈, 정경훈*, 강동욱*

*국민대학교, 한국전자통신연구원

*andrew8849@kookmin.ac.kr, steve-kim@etri.re.kr

A Study on the Utilizing Additional Information for Quality Enhancement in 3D Reconstruction of 2D Videos Based on Depth Map

Lee Ju Hyeon*, Kim Sung Hoon, Jung Kyeong Hoon*, Kang Dong*

*Kookmin Univ., Electronics & Telecommunications Research Institute.

요약

본 논문은 깊이 맵 기반 2D to 3D 변환 과정에서 발생하는 화질 저하 문제를 해결하기 위해, Video Enhancement Information(VEI)을 활용한 3 단계 품질 개선 구조를 제안한다. 제안된 구조는 입력 영상의 화질 열화(VEI-1), 깊이 맵 정밀도 저하(VEI-2), 시점 영상의 시각적 결함(VEI-3) 등 주요 품질 저하 요인에 단계적으로 대응하며, 각 단계별로 보조 정보를 활용해 경계 흐림, 깊이 왜곡, 가려짐 영역 등을 보완한다. 이를 통해 다양한 네트워크 환경에서도 고품질 S3D 콘텐츠 생성을 가능하게 한다. 제안된 구조는 인코딩 시점에서 추출한 부가 정보를 및 주변 환경의 정보를 효율적으로 활용함으로써 고정형 디스플레이 환경은 물론 이동통신 기반의 저해상도 스트리밍 환경에서도 안정적으로 고품질 S3D 콘텐츠 생성을 가능하게 함을 목적으로 한다.

I. 서론

실감형 미디어 기술의 발전과 함께 몰입도 높은 사용자 경험을 제공하는 스테레오스코픽 3D(S3D: Stereoscopic 3D) 영상에 대한 수요가 꾸준히 증가하고 있다. 특히 TV, 영화, 가상현실(VR), 증강현실(AR) 등 다양한 분야에서 입체 영상 콘텐츠의 활용이 확대되고 있으나 고품질 3D 콘텐츠의 제작은 여전히 많은 비용과 시간 그리고 복잡한 촬영 시스템을 필요로 한다. 이로 인해 3D 콘텐츠의 절대적인 부족 현상이 발생하고 있으며 이는 실감형 미디어 서비스 확산의 주요한 제약 요인으로 작용하고 있다.

이를 해결하기 위한 대안으로 기존의 2D 영상을 3D로 변환하는 2D to 3D 변환 기술이 주목받고 있다. 해당 기술은 별도의 스테레오 촬영 없이 3D 콘텐츠를 생성할 수 있어 콘텐츠 제작의 효율성과 경제성을 크게 향상시킨다. 현재까지 제안된 2D to 3D 변환 기술은 크게 학습 기반(Deep Learning) 변환, Depth map 기반 변환으로 구분된다. [1, 2]

이 중에서도 본 논문에서는 Depth map 기반의 변환 방식에 주목한다. [3] 해당 방식은 단안 영상으로부터 장면 내 각 픽셀의 상대적인 깊이 정보를 추정한 후 이를 기반으로 새로운 시점의 영상을 생성함으로써 입체감을 구현하는 방식이다. 깊이 정보를 이용함으로써 시점 이동(parallax)을 수학적으로 명시적으로 모델링 할 수 있어 구조적 왜곡을 예측 가능하게 하며 기존의 영상 렌더링 및 후처리 기술과의 호환성이 높아 시스템 통합 및 구현이 용이하다는 점에서도 실용적인 이점을 갖는다.

그러나 현실에서는 깊이 맵 추정의 정확도 저하, 압축 영상으로 인한 입력 화질 열화, 그리고 novel view 생성 과정에서의 가려짐(occlusion) 및 왜곡 발생 등으로 인해

S3D 영상의 품질이 크게 저하될 수 있다. 따라서 본 논문에서는 Video Enhancement Information(VEI) 기반의 3 단계 부가 정보 활용 기법을 제안한다. 제안하는 방식은 첫째, 압축된 2D 영상의 화질을 사전 보정하고, 둘째, 깊이 맵의 정밀도를 향상시키며, 셋째, 생성된 타 시점 영상의 시각적 품질을 개선하는 일련의 프로세스를 기반으로 한다. 이와 같은 접근은 고해상도 스트리밍이 원활한 고정형 디스플레이 환경뿐만 아니라 이동통신 기반의 저해상도 스트리밍 환경에서도 안정적으로 적용 가능하므로 다양한 미디어 소비 환경에서 고품질 S3D 미디어 서비스 실현에 기여할 수 있을 것으로 기대된다.

II. 본론

본 논문에서는 2D 영상을 깊이 맵을 활용하여 3D로 재구성하기 위한 구조를 그림 1과 같이 제안한다. 제안하는 구조는 크게 세단계로 구성되며 각 단계는 입력 stream의 디코딩과정, 깊이 맵 생성 및 새로운 시점 영상 생성으로 이루어져 있으며 최종 시점 영상의 시각적인 완성도를 높이기 위해서 Video Enhancement Information(VEI)기반의 부가 정보가 단계별로 활용된다.

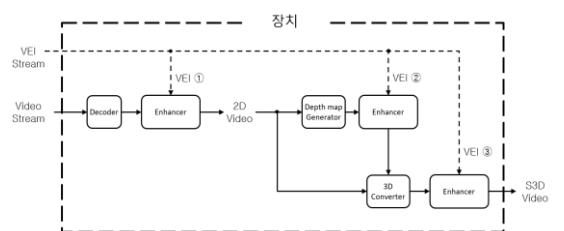


그림 1. 깊이 맵을 활용한 2D to 3D 변환 구조
Fig 1. Depth Map-based 2D-to-3D Conversion Framework

각 단계는 변환 전반에 걸쳐 최종적으로 생성되는 시점 영상의 품질을 결정하는 요인을 단계적으로 대응하기 위해 구성된다. 첫번째로 입력 영상은 최종적으로 사용자에게 전달되는 정보이며 현대의 전송 기술 환경에서는 압축 손실이 불가피하게 발생한다. 이로 인해 영상 내 경계가 흐려지고 세부 구조가 손실되며 이러한 화질 저하는 깊이 맵 추론 과정에서 경계 검출 실패나 세그멘테이션 오류를 유발하기 때문에 결과적으로 부정확한 깊이 맵 생성을 초래한다. 이에 따라 VEI-1 을 활용한 사전 화질 보정(pre-enhancement)을 통해 경계의 명확도와 색상 정확도 향상 등을 실현함으로써 안정적인 깊이 추정을 지원할 수 있다. 더 나아가 VEI-1 은 인코딩 단계에서 사전 처리된 보정 정보를 함께 전송하는 방식 외에도 수신부의 입장에서는 네트워크 상에 존재하는 주변 고해상도 영상 스트림 또는 보조 채널로부터 특정 영역의 화질을 대체하거나 보완하는 형태로도 동작할 수 있다. 이러한 방식은 고해상도 주변 스트림을 활용한 능동적 보정을 가능하게 하고 수신 측에서 영상 품질을 보다 유동적으로 향상시킬 수 있는 환경을 구축할 수 있다. 두 번째로 깊이 맵은 시점 이동(parallax generation)의 기반이 되는 핵심 정보로서 그 정밀도는 S3D 입체감 형성에 직접적인 영향을 미친다. 그러나 압축 영상의 열화는 깊이 추정 과정에서 깊이의 연속성과 구조적 안정성을 저하시킬 수 있으며 이는 그림 2 에서 확인할 수 있듯이 깊이 맵의 품질 차이가 발생한다. 이를 해결하기 위해 VEI-2 를 활용하여 원본 영상에서 추출된 경계 정보, 텍스처 및 지역 대비 정보 등을 보조 신호로 제공함으로써 디코더 단에서도 깊이 맵의 정밀도를 유지하고 시각적으로 일관된 깊이 구조를 복원할 수 있다.

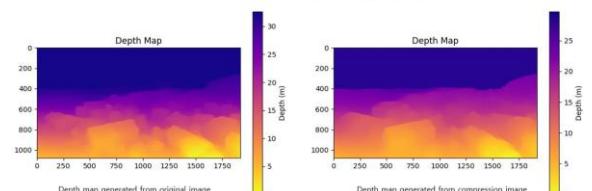


Fig 2 left) Depth map generated from original image, right) Depth map generated from compression image

마지막으로 생성된 시점 영상은 깊이 맵을 기반으로 시차 보상을 통해 렌더링되지만 이 과정에서 가려짐(occlusion) 영역의 발생, 깊이 불연속 경계의 왜곡 및 텍스처 왜곡과 같은 다양한 시각적 결함이 수반된다. 특히 occlusion 은 원본 영상에는 존재하지 않던 영역을 새롭게 생성해야 하는 문제로, 복원되지 않은 픽셀 구간이나 불규칙한 경계 왜곡을 유발하며 이는 입체감의 불균형과 시청 피로를 초래할 수 있다. 이러한 시각적 오류는 단순히 깊이 맵 정밀도만으로는 해결이 어려우며 생성된 영상에 대한 후처리 보정(post-enhancement)이 필수적이다. 따라서 VEI-3 에서는 원본 영상 기반으로 inpainting, 경계 보정을 하기 위한 경계 마스크, 공간의 미적 영역 정보, 인접 프레임 참조 텍스처 등을 부가 정보로 제공함으로써 시점 영상 생성 이후 발생하는 occlusion 영역의 복원 경계 선명화 노이즈 제거 등을 가능하게 한다. 이러한 방식은 생성된 S3D 영상의 시각

적 안정성과 입체감의 일관성을 유지하는 데 핵심적으로 작용하며, 전체 변환 구조의 마지막 단계에서 품질을 결정짓는 중요한 역할을 수행한다.

III. 결론

본 논문에서는 깊이 맵을 기반으로 2D 영상을 3D로 변환하는 과정에서 발생할 수 있는 화질 저하 문제에 대응하기 위한 가능성 있는 접근으로 Video Enhancement Information(VEI)을 활용한 3 단계 품질 향상 구조를 제안하였다. 제안한 구조는 입력 영상의 사전 보정(VEI-1), 깊이 맵 정밀도 향상(VEI-2), 그리고 시점 영상의 시각적 결합 보정(VEI-3)으로 구성하여 세 가지 주요 품질 저하 요인을 단계적으로 대응함으로써 전체적인 S3D 영상 품질을 개선하는 것을 목표로 한다.

VEI-1 은 인코딩된 보정 정보 외에도 수신 측에서 고해상도 스트리밍이나 보조 채널을 활용해 특정 영역의 화질을 능동적으로 보완할 수 있어 다양한 환경에서 유연한 품질 향상을 가능하게 한다. VEI-2 는 경계, 텍스처, 지역 대비 정보를 보조 신호로 제공함으로써 깊이 맵의 정밀도와 구조적 안정성을 높이고 VEI-3 은 inpainting 및 경계보정을 위해 경계 마스크, 의미 기반 영역 정보, 인접 프레임 텍스처 등을 활용한 후처리로 가려짐과 왜곡을 보정하여 시각적 일관성과 입체감을 유지하는 데 기여한다. 이러한 방식은 고정형 디스플레이 환경뿐만 아니라 네트워크 품질이 불안정한 모바일 스트리밍 환경에서도 적용 가능성이 높으며 다양한 미디어 소비 환경에서 안정적인 고품질 S3D 콘텐츠 제공을 가능하게 할 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 한국전자통신연구원 연구운영비지원사업의 일환으로 수행되었었음. [25ZC1110, 초실감 입체공간 미디어·콘텐츠 원천기술 연구].

참 고 문 헌

- [1] Xie, Junyuan, Ross Girshick, and Ali Farhadi. "Deep3d: Fully automatic 2d-to-3d video conversion with deep convolutional neural networks." European conference on computer vision. Cham: Springer International Publishing, 2016.
- [2] Sisi, Li, Wang Fei, and Liu Wei. "The overview of 2D to 3D conversion system." 2010 IEEE 11th International Conference on Computer-Aided Industrial Design & Conceptual Design 1. Vol. 2. IEEE, 2010.
- [3] Fehn, Christoph. "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV." Stereoscopic displays and virtual reality systems XI. Vol. 5291. SPIE, 2004.