

# 혼합현실 환경 기반 직접 볼륨 렌더링을 위한 상호작용 시스템 실효성 연구

## Interactive System for Direct Volume Rendering in Mixed Reality Environments: An Effectiveness Study

김예진<sup>1</sup>, 강지윤<sup>2</sup>, 최준서<sup>3</sup>, 추동혁<sup>4</sup>, 정윤현<sup>5\*</sup>

Yejin Kim<sup>1</sup>, Jiyun Kang<sup>2</sup>, Junseo Choi<sup>3</sup>, Donghyuk Choo<sup>4</sup>, Younhyun Jung<sup>5\*</sup>

### 요약

Direct Volume Rendering(DVR)은 Transfer Function(TF)을 조절하여 3차원 볼륨 데이터의 내부 구조를 추출하고 관심 영역을 효과적으로 시각화하는 기법이다. 최근 컴퓨팅 환경이 확장됨에 따라 Head-Mounted Display를 활용한 혼합현실(MR) 환경에서도 볼륨 데이터와 직관적으로 상호작용하려는 연구가 활발하다. 그러나 기존 연구들은 사전에 정의된 TF를 사용하거나 손 제스처와 같은 단일 입력 방식에 의존하여, 사용자가 물입 환경에서 TF를 자유롭게 정밀하게 조정하는 데 한계가 있었다. 이에 본 연구에서는 손 제스처, 시선 추적, 그리고 대규모 언어 모델 기반의 음성 인식을 통합한 다중 모달 인터페이스를 제안하여 MR 환경에서의 제어 유연성을 극대화하고자 하였다. 본 연구는 HoloLens 2 환경에서 제안된 시스템을 구현하고, 머리, 척추 등 의료 데이터와 시뮬레이션 데이터를 활용하여 DVR의 실효성을 분석하였다. 실험 결과, 하드웨어적 제약으로 인해 평균 10 FPS의 렌더링 성능과 1차원 TF의 표현적 한계가 관찰되었으나, 다중 인터페이스를 통해 복잡한 메뉴 조작 없이도 직관적인 데이터 탐색과 시각화 조정이 가능함을 확인하였다. 본 연구의 결과는 향후 수술 내비게이션과 같은 전문 의료 및 산업 분야에서 MR 기반 시각화 인터페이스가 나아가야 할 방향성과 실질적인 발전 가능성을 제시한다.

핵심어 : 볼륨 렌더링, 혼합현실, 전달 함수, 시각화

1 School of Computing, Gachon University, Seongnam, Korea [Undergraduate Student]

e-mail: vvvstfhk@gachon.ac.kr

2 School of Computing, Gachon University, Seongnam, Korea [Undergraduate Student]

e-mail: jiyun1215@gachon.ac.kr

3 School of Computing, Gachon University, Seongnam, Korea [Undergraduate Student]

e-mail: wnstl1234@gachon.ac.kr

4 School of Computing, Gachon University, Seongnam, Korea [Undergraduate Student]

e-mail: choodh105@gachon.ac.kr

5 School of Computing, Gachon University, Seongnam, Korea [Professor]

e-mail: younhyun.jung@gachon.ac.kr (Corresponding author)

\* 이 논문은 한국연구재단 우수신진연구사업(과제명 : 수술 환경에서의 체적 의료영상 재구성과 물리적 표면 재구성을 위한 NeRF · GS 알고리즘 연구 및 수술 증강 시각화에서의 검증, 과제번호 : RS-2025-00554526, 기여율 : 100%)으로 수행되었음.

Received(September 1, 2025), Review Result(1st: September 23, 2025), Accepted(December 12, 2025), Published(December 31, 2025)



© 2025 The Authors. Published by NCISS.

This is an open access article licensed under the Creative Commons Attribution-NonCommercial 4.0 International License.

To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## Abstract

Direct Volume Rendering (DVR) is a technique that extracts internal structures of 3D volumetric data and effectively visualizes regions of interest by adjusting Transfer Function (TF). Recently, as computing environments expand, research on intuitively interacting with volumetric data in Mixed Reality (MR) environments using Head-Mounted Displays (HMD) has become active. However, existing studies have relied on predefined TF or single input methods such as hand gestures, which limited users in freely and precisely adjusting TF in immersive environments. Therefore, this study aimed to maximize control flexibility in MR environments by proposing a multi-modal interface that integrates hand gestures, eye tracking, and Large Language Model (LLM)-based voice recognition. We implemented the proposed system in the HoloLens 2 environment and analyzed the effectiveness of DVR using medical data, such as head and spine, and simulation data. Experimental results showed that while an average rendering performance of 10 FPS and expressive limitations of 1D TF were observed due to hardware constraints, intuitive data exploration and visualization adjustment were possible without complex menu manipulations through multi-modal interfaces. The results of this study present the direction and practical development potential of MR-based visualization interfaces in specialized medical and industrial fields such as surgical navigation.

Keyword : Direct Volume Rendering, Mixed Reality, Transfer Function, Visualization

## 1. 서론

혼합현실(Mixed Reality, MR)은 실제 현실과 가상 객체를 융합하여 현실의 사용자가 상호작용 할 수 있는 기술이다. 사용자는 Head Mounted Display(HMD)와 같은 MR 기기를 통해 실제 현실에 가상의 물체를 투영하고, 손 제스처, 시선 추적, 음성 인식 등의 기능을 사용하여 가상 물체와 직접적인 상호작용을 할 수 있다.

직접 볼륨 렌더링(Direct Volume Rendering, DVR) [1]은 3D 볼륨 데이터의 내부 구조를 탐색하고 분석할 때 사용되는 대표적인 시각화 기법이다. DVR은 등가면 추출이나 메시 생성과 같은 기하학적 전처리 과정 없이, 복셀 단위의 데이터를 레이 캐스팅하고 일정하게 샘플링하여 2D 영상으로 투영한다. 이로 인해 데이터의 손실을 최소화하면서 전체 구조를 시각화할 수 있으며, 불투명도 조절을 통해 표면뿐 아니라 내부의 세부 구조까지 동시에 관찰할 수 있다.

전이함수(Transfer Function, TF)는 DVR 환경에서 사용자가 관심 영역(Region of Interest, ROI)을 강조하기 위해 볼륨의 복셀에 색상과 불투명도를 할당하는 매핑 함수이다. 사용자는 데이터의 특성에 따라 TF를 조정함으로써 원하는 영역을 강조하고 효과적으로 시각적 정보를 획득할 수 있다. 초기에는 복셀의 강도(intensity) 값만을 기반으로 색상과 불투명도를 부여하는 1차원 TF가 제안되었다 [1]. 이후 연구에서는 1차원 TF를 CT 데이터에 적용하여 볼륨의 영역을 분류하는 방법을 제시하기도 하였다 [2]. 1차원 TF는 조작이 간단하고 직관적이라는 장점이 있으나, intensity 값만을 고려하기 때문에 동일한 intensity를 가지더라도 서로 다른 위치에 있는 복셀을 구분할 수 없으며, 복셀 간 경계 영역에서는 intensity 값이 혼합되어 구조 분리가 어렵다는 한계가 존재한다. 이러한 한계를 보완하기 위해 이후 연구에서는 intensity 외의 다양한 특징을 함께 고려하는 다차원 TF가

제한되었다. 다차원 TF는 intensity 값을 바탕으로 기울기 크기 [2], 2차 미분 [2], 곡률(curvature) [3][4], 질감(texture) [5] 등의 속성을 함께 활용하여 볼륨 데이터의 구조적 특징을 더욱 정교하게 표현한다. 그러나, 차원이 증가함에 따라 사용자 인터페이스의 복잡성이 높아지고 TF의 직관적 조정이 어려워지는 한계가 나타났다 [2][6].

이전의 TF 조작 방식은 주로 컴퓨팅 환경에서 연구됐으나, 최근 들어 손 제스처, 시선 추적, 음성 인식과 같은 자유로운 상호작용을 중심으로 MR 환경에서 DVR을 적용하려는 시도가 의료 및 산업 분야에서 활발히 이루어지고 있다 [7][8]. 그러나, MR 기기는 컴퓨팅 환경보다 연산량과 메모리 용량이 부족하여 복잡한 다차원 TF보다는 연산 효율성이 높은 1차원 TF가 대안으로 선택된다.

최근 연구들은 MR 환경에서 TF 조정을 구현하기 위한 다양한 접근을 시도하였다. 예를 들어, Song 등 [9]은 시선 추적과 손 제스처를 활용하여 MR 환경에서 ROI를 자동으로 최적화하는 방법을 제안하였다. 그러나 이 방식은 사전 정의된 TF를 사용하기 때문에 사용자의 직접적인 제어가 불가능하고 실시간 조정이 어렵다는 한계가 있다. Choi 등 [10]은 손 제스처를 이용하여 실시간 TF 편집이 가능한 TF editor를 구현하였으나, 시선 추적이나 음성 인식 등 MR 환경의 다양한 상호작용을 반영하지 못하였다. Kim 등 [11]은 이러한 한계를 극복하기 위해 시선 추적, 손 제스처, 음성 인식 등 다양한 MR 상호작용을 통합하고, 여기에 대규모 언어 모델(ChatGPT)을 결합하여 사용자가 자연어 음성 명령을 통해 TF를 직관적으로 조작할 수 있는 애플리케이션을 제안하였다. 해당 연구는 MR 환경의 상호작용 기법과 DVR의 특성을 잘 반영한 애플리케이션을 제안하였지만, 다양한 환경에서의 실험이 제한적이라는 한계가 있었다.

본 연구는 해당 애플리케이션이 MR 환경과 DVR을 결합한 다양한 환경에서 유연하게 작동하는지를 분석하여 이 애플리케이션의 실효성을 검증하였다.

## 2. 애플리케이션

### 2.1 개발 환경

MR 환경에서 해당 애플리케이션의 실효성을 검증하기 위해 본 연구는 [11]에서 개발된 애플리케이션을 사용했다. MR 기기는 퀄컴 스냅드래곤 850 프로세서, 4GB의 DRAM을 장착하고 있는 마이크로소프트의 홀로렌즈 2(HoloLens 2)를 사용하였고, Unity 게임 엔진과 Mixed Reality Toolkit(MRTK)를 활용하여 개발하였다. 음성 인식 기능 향상을 위해 Whisper API와 자연어 해석을 위한 GPT-3.5 Turbo API를 사용하였다.

### 2.2 상호작용 기능

해당 애플리케이션은 이전 연구 [10]에서 구현된 1차원 TF 기반으로 구현되었다. MR 환경의 손 제스처를 활용한 노드 생성, 노드 삭제, 노드 색상 변경, 슬라이더 기반 불투명도 제어 기능을 제공한다. 사용자가 노드 생성 기능을 활성화한 후 원하는 위치를 터치하면 해당 위치에 노드가 생성되고, 노드 삭제 기능을 활성화한 후 삭제를 원하는 노드를 터치하면 해당 노드가 삭제된다. 노드 색상 변경 기능은 원하는 노드를 터치하면 바꿀 수 있는 대표적인 9개 색상(빨강, 파랑, 초록, 검정, 하양, 노랑, 보라, 민트, 갈색)이 뜨고 원하는 색상을 터치하면 노드의 색상이 바뀐다. 정교한 불투명도 조절이 필요한 경우에는 슬라이더 기능을 사용하여 세밀하게 불투명도 값을 조정할 수 있다. 이러한 기능들을 통해 볼륨 데이터의 ROI를 직관적으로 관찰할 수 있다. 그러나, 손 제스처를 이용한 기존 TF 조작 방식은 MR 기기 특유의 다양한 상호작용 방식을 충분히 반영하지 못한다는 한계가 있다. 이를 개선하기 위하여 본 애플리케이션은 시선 추적 기능과 음성 인식 기능을 추가하여 시각 및 음성을 이용한 TF 조작 기능을 추가했다.

시선 추적 기능의 직관성을 높이기 위해 사용자가 응시하는 곳에 빨간 구체를 표시하고, 음성 기반 상호작용을 위해 “클릭(Click)” 키워드를 설정했다. 사용자가 특정 객체를 바라보고 “Click”을 말하면 해당 객체가 선택되도록 구현했다. 이를 통해 사용자는 손을 사용하지 못하는 상황에서도 TF를 편집할 수 있다. 하지만, 손 제스처나 시선 처리만을 활용하여 TF를 조정하려면 여러 단계의 상호작용이 필요하다. 예를 들어, 한 노드의 색상을 변경하기 위해서는 Color 변경 기능을 활성화한 뒤, 변경하고자 하는 노드를 선택하고 색상 팔레트 중 원하는 색상을 클릭하는 번거로움이 존재한다. 또한, MRTK는 영어만 지원하여 다국어 사용에 제약이 있으며, 이를 보완할 수 있는 MRTK의 키워드 인식 기능 또한 단순한 키워드만 인식하기 때문에 사용자의 복잡한 지시를 처리하기 어렵다.

이러한 한계점을 극복하기 위해 해당 애플리케이션은 대규모 언어 모델을 활용한 GPT 기반의 음성 인터페이스를 도입했다. 사용자가 각국의 언어로 음성 명령을 말하면 Whisper API가 이를 녹음하여 텍스트로 변환한다. 변환된 텍스트는 사전 정의된 프롬프트 엔지니어링을 거친 GPT 3.5 turbo API로 전달된다. GPT는 입력 문장을 분석하여 MRTK가 처리할 수 있는 형식으로 명령어를 반환한다. 예를 들어, 사용자가 “2번 노드 삭제”라고 명령하면 GPT API는 해당 텍스트를 “Delete 2” 형식으로 반환한다. 반환된 명령어는 HoloLens 2 내에서 원하는 기능을 수행하도록 연결된다. 이러한 접근은 음성 기반의 직관적 조작을 통해 사용자의 편의성을 높이고 다국어 명령 처리를 가능하게 하여 MRTK의 기존 한계를 극복한다.

### 3. 실험 및 분석

제안된 애플리케이션을 바탕으로 MR 환경과 DVR을 결합한 다양한 분야에서 실효성을 평가하

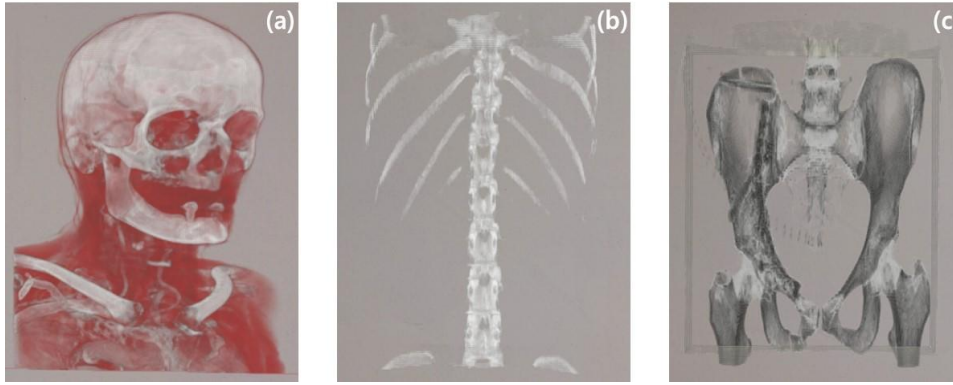
기 위하여 Head, Calix, Pelvis, Engine으로 구성된 의료 이미지 및 시뮬레이션 데이터를 활용하여 실험을 진행하였다. 의료 데이터의 경우 다양한 내부 조직이 내포되어 있어 MR 기기에서 원본 데이터를 렌더링하기에는 연산량이 과도하여, 오픈소스 볼륨 렌더링 엔진 프레임워크인 Voreen [12]을 활용하여 각 볼륨 데이터에 대해 다운샘플링을 적용해 데이터 크기를 줄이고 MR 기기의 연산 부담을 완화하였다. 볼륨 데이터는 (x, y, z) 차원 순서로 구성되어 있으며 Head는 (256, 256, 256) 차원, Calix는 (200, 200, 134) 차원, Pelvis는 (256, 256, 113) 차원, Engine은 (256, 256, 256) 차원으로 구성하였고, 샘플링 비율은 1.2로 설정하였다.

[그림 1]은 애플리케이션을 바탕으로 Engine 데이터에 DVR을 적용한 렌더링 이미지이다. Engine 내부의 실린더와 개스킷의 정보를 잘 표현하고 있지만, 내부의 굴곡진 공간적 구조 정보를 충분히 제공하지 못했다.



[그림 1] Engine TF editor 및 DVR 결과  
[Fig. 1] Engine TF editor and DVR result

[그림 2]는 애플리케이션을 활용하여 Head(a), Calix(b), Pelvis(c)에 대해 각각 서로 다른 TF를 적용한 DVR 결과이다. Head(a)는 뼈와 근육의 정보를 함께 보여주고 있으나, 피부에 대한 시각적 정보를 충분히 제공하지 못했다. Calix(b)는 척추뼈와 갈비뼈의 정보를 잘 보여주고 있으나, 콩팥에 대한 정보를 렌더링하지 못했다. Pelvis(c)는 골반에 대한 형태 정보를 잘 제공하고 있으나, 뼈 주변의 시각형 노이즈를 제거하지 못했다. 이는 해당 애플리케이션의 DVR에 적용된 1차원 TF의 한계로 서로 다른 위치에 있지만 동일한 intensity를 지닌 복셀의 경우 분리할 수 없는 한계점을 명확히 보여준다.



[그림 2] Head(a), Calix(b), Pelvis(c)의 DVR 결과

[Fig. 2] Head(a), Calix(b), Pelvis(c) DVR result

또한, MR 기기의 평균 메모리 사용량은 3.4GB, 평균 렌더링 프레임은 10 FPS로 측정되었다. 이는 컴퓨팅 환경과 비교해 MR 기기의 적은 메모리 용량과 연산량에 의한 단점을 보여주는 결과이다.

#### 4. 결론

본 연구에서는 제안된 애플리케이션과 의료 이미지 및 시뮬레이션 볼륨 데이터를 활용하여 다양한 데이터 환경을 바탕으로 MR 환경에서 DVR의 시각화 결과 및 MR 기기의 성능을 통해 애플리케이션의 실효성을 분석했다. 손 제스처, 시선 추적, 음성 인식과 같은 다양한 상호작용 기법을 이용하여 HMD에서 서로 다양한 데이터의 내부 구조 모습을 시각화하였다. 이를 통해 MR 기기에서도 데이터의 특징을 잘 표현할 수 있고 1차원 TF 기반의 DVR 연산이 가능하다는 것을 확인했다.

다만, MR 기기의 낮은 메모리 용량과 연산량에 의해 평균 메모리 사용량은 3.4GB, 평균 렌더링 프레임은 10 FPS로 측정되었다. 또한, 1차원 TF에 의해 같은 intensity의 복셀이 다른 위치에 있는 경우 데이터를 명확하게 구분하지 못하는 한계가 존재한다. 이러한 이유로 데이터의 노이즈를 제거하지 못하거나, 높은 샘플링 비율을 연산하지 못해 시각화 품질이 떨어지는 단점이 있다.

해당 애플리케이션을 통해 수술 내비게이션 [13-15]과 같은 의료 분야는 물론 산업 등 실제 응용 분야에서의 MR 기반 DVR 시각화 기술 발전 가능성을 확인할 수 있었다.

향후 연구에서는 해당 애플리케이션을 발전시켜 볼륨 데이터의 내부 특성을 더 잘 표현할 수 있는 2차원 TF를 접목하고, 연산 효율성을 향상해 용량이 큰 시뮬레이션 데이터에 적용할 수 있는 새로운 알고리즘 개발이 필요할 것이다.

## References

- [1] R. A. Drebin, L. Carpenter, P. Hanrahan, "Volume rendering", *SIGGRAPH Computer Graphics*, vol. 22, no. 4, August 1988, pp. 65-74, doi: 10.1145/378456.378484.
- [2] J. Kniss, G. Kindlmann, C. Hansen, "Multidimensional transfer functions for interactive volume rendering", *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 3, July 2002, pp. 270-285, doi: 10.1109/TVCG.2002.1021579.
- [3] G. Kindlmann, J. W. Durkin, "Semi-automatic generation of transfer functions for direct volume rendering", *IEEE Symposium on Volume Visualization*, October 19-20, 1998, Research Triangle Park, NC, USA, pp. 79-86, doi: 10.1109/SVV.1998.729588.
- [4] G. Kindlmann, R. Whitaker, T. Tasdizen, T. Moller, "Curvature-based transfer functions for direct volume rendering: methods and applications", *IEEE Visualization 2003*, October 19-24, 2003, Seattle, WA, USA, pp. 513-520, doi: 10.1109/VISUAL.2003.1250414.
- [5] J. J. Caban, P. Rheingans, "Texture-based Transfer Functions for Direct Volume Rendering", *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, November 2008, pp. 1364-1371, doi: 10.1109/TVCG.2008.169.
- [6] C. R. Salama, M. Keller, P. Kohlmann, "High-Level User Interfaces for Transfer Function Design with Semantics", *IEEE Transactions on Visualization and Computer Graphics*, vol. 12, no. 5, September 2006, pp. 1021-1028, doi: 10.1109/TVCG.2006.148.
- [7] K. C. Wong, E. Y. Sun, I. O. L. Wong, S. M. Kumta, "Mixed Reality Improves 3D Visualization and Spatial Awareness of Bone Tumors for Surgical Planning in Orthopaedic Oncology: A Proof of Concept Study", *Orthopedic Research and Reviews*, vol. 15, July 2023, pp. 139-149, doi: 10.2147/ORR.S421077.
- [8] S. Lee, H. Jung, E. Lee, Y. Jung, S. T. Kim, "A Preliminary Work: Mixed Reality-Integrated Computer-Aided Surgical Navigation System for Paranasal Sinus Surgery Using Microsoft HoloLens 2", *Computer Graphics International Conference (CGI 2021)*, September 6-10, 2021, Geneva, Switzerland, pp. 633-641, doi: 10.1007/978-3-030-89029-2\_47.
- [9] M. Song, S. Kim, Y. Jung, "An intuitive and semi-automated transfer function design for interactive region of interest-based direct volume rendering in mixed reality head mounted devices", *Virtual Reality*, vol. 29, no. 53, March 2025, doi: 10.1007/s10055-025-01121-4.
- [10] J. Choi, H. Kim, H. An, Y. Jung, "Real-Time Transfer Function Editor for Direct Volume Rendering in Mixed Reality", *SIGGRAPH Asia 2024 Posters*, December 3-6, 2024, Tokyo, Japan, Article 38, pp. 1-2, doi: 10.1145/3681756.3697953.
- [11] Y. Kim, S. Kim, Y. Jung, "Instinctual-TF: Feel Free to Manipulate Transfer Function in Mixed Reality Environments", *2025 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct)*, October 8-12, 2025, Daejeon, Republic of Korea, pp. 715-716, doi: 10.1109/ISMAR-Adjunct68609.2025.00154.
- [12] J. Meyer-Spradow, T. Ropinski, J. Mensmann, K. Hinrichs, "Voreen: A rapid-prototyping environment for ray-casting-based volume visualizations", *IEEE Computer Graphics and Applications*, vol. 29, no. 6,

November 2009, pp. 6-13, doi: 10.1109/MCG.2009.130.

- [13] S. M. Heining, V. Raykov, O. Wolff, H. Alkadhi, H. C. Pape, G. A. Wanner, “Augmented reality-based surgical navigation of pelvic screw placement: an ex-vivo experimental feasibility study”, *Patient Safety in Surgery*, vol. 18, no 3, January 2024, doi: 10.1186/s13037-023-00385-6.
- [14] T. Ito, Y. Kawashima, A. Yamazaki, T. Tsutsumi, “Application of a virtual and mixed reality-navigation system using commercially available devices to the lateral temporal bone resection”, *Annals of Medicine and Surgery*, vol. 72, no. 103063, December 2021, doi: 10.1016/j.amsu.2021.103063.
- [15] C. T. Morley, D. M. Arreola, L. Qian, A. L. Lynn, Z. P. Veigulis, T. F. Osborne, “Mixed Reality Surgical Navigation System; Positional Accuracy Based on Food and Drug Administration Standard”, *Surgical Innovation*, vol. 31, no. 1, February 2024, pp. 48-57, doi: 10.1177/15533506231217620.