

Heterogeneous Sensor Coordinate System Calibration Technique for AR Whole Body Interaction

Hangkee Kim[†] · Daehwan Kim^{††} · Dongchun Lee^{†††} · Kisuk Lee^{††††} · Nakhoon Baek^{†††††}

ABSTRACT

A simple and accurate whole body rehabilitation interaction technology using immersive digital content is needed for elderly patients with steadily increasing age-related diseases. In this study, we introduce whole-body interaction technology using HoloLens and Kinect for this purpose. To achieve this, we propose three coordinate transformation methods: mesh feature point-based transformation, AR marker-based transformation, and body recognition-based transformation. The mesh feature point-based transformation aligns the coordinate system by designating three feature points on the spatial mesh and using a transform matrix. This method requires manual work and has lower usability, but has relatively high accuracy of 8.5mm. The AR marker-based method uses AR and QR markers recognized by HoloLens and Kinect simultaneously to achieve a compliant accuracy of 11.2mm. The body recognition-based transformation aligns the coordinate system by using the position of the head or HMD recognized by both devices and the position of both hands or controllers. This method has lower accuracy, but does not require additional tools or manual work, making it more user-friendly. Additionally, we reduced the error by more than 10% using RANSAC as a post-processing technique. These three methods can be selectively applied depending on the usability and accuracy required for the content. In this study, we validated this technology by applying it to the "Thunder Punch" and rehabilitation therapy content.

Keywords : AR, Rehabilitation, Contents, Whole-Body, Interaction

AR 전신 상호작용을 위한 이중 센서 간 좌표계 보정 기법

김 항 기[†] · 김 대 환^{††} · 이 동 춘^{†††} · 이 기 석^{††††} · 백 낙 훈^{†††††}

요 약

꾸준히 증가하는 노인성 질환 환자 대상으로 몰입형 디지털 콘텐츠를 이용한 쉽고 정확한 전신 재활 상호작용 기술이 필요하며, 본 연구에서는 이를 위해 홀로렌즈와 키넥트를 사용한 전신 상호작용 기술을 소개한다. 이를 위해 메쉬 특징점 기반 변환, AR 마커 기반 변환, 신체 인식 기반 변환 방법의 3가지 좌표 변환 방식을 제안한다. 메쉬 특징점 기반 변환은 공간 메쉬에 3개의 특징점을 지정하고 변환 행렬을 이용하여 좌표계를 일치시킨다. 이 방법은 수작업이 필요하여 사용성이 떨어지지만 정확도가 8.5mm로 상대적으로 높다. AR 마커 기반 방식은 홀로렌즈와 키넥트가 동시에 인식하는 AR, QR마커를 사용하여 평균오차 11.2mm로 준수한 정확도를 달성했다. 신체 인식 기반 변환 방법은 두 기기가 동시에 인식하는 머리 혹은 HMD위치와 양 손 혹은 컨트롤러 위치를 이용하여 좌표계를 일치시킨다. 이 방법은 정확도가 떨어지지만 부가적인 도구나 수작업이 필요 없으므로 사용성이 좋다. 또한 후처리 기술로 RANSAC을 적용함으로써 오차를 10% 이상 줄였다. 이러한 3가지 방식은 콘텐츠가 필요한 사용 편의성과 정확도에 따라서 선택적으로 적용할 수 있다. 본 연구에서는 이 기술을 이용해서 '번서펀치'와 재활 치료 콘텐츠에 적용하여 검증하였다.

키워드 : 증강현실, 재활, 콘텐츠, 전신, 상호작용

※ 이 논문은 2022년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임(No.2021-0-00742, 의료수준 전신 메디컬트윈 핵심기술 개발).
※ 이 논문은 2022년 한국정보처리학회 정보처리학회지 6월호에 "AR 몰입형 전신 재활치료를 위한 신체 상호작용에 관한 연구"의 제목으로 발표된 논문을 확장한 것이다.
† 정 회 원 : 한국전자통신연구원 VR/AR콘텐츠연구실 책임연구원
†† 정 회 원 : 울산대학교 IT융합전공 교수
††† 비 회 원 : 한국전자통신연구원 VR/AR콘텐츠연구실 책임연구원
†††† 비 회 원 : 한국전자통신연구원 VR/AR 콘텐츠연구실 실장
††††† 중신회원 : 경북대학교 컴퓨터학부 교수
Manuscript Received : November 23, 2022
First Revision : April 4, 2023
Second Revision : May 8, 2023
Accepted : June 5, 2023
*Corresponding Author : Nakhoon Baek(oceanru@gmail.com)

1. 서 론

한국은 2018년까지 저출산(연평균 3.1% 감소), 고령화(연평균 3.3% 증가) 속도가 OECD 37개국 중 가장 빠르다 [1]. 이러한 고령화의 영향으로 노인성 질환인 치매, 파킨슨, 뇌졸중 등의 환자가 꾸준히 증가하고 있다[2]. 본 연구는 노인성 질환 중 하나인 뇌졸중 환자를 대상으로 몰입형 디지털 콘텐츠를 이용한 전신 재활 상호작용 기술에 대한 것이다.

뇌졸중 환자에 대한 재활 콘텐츠는 여러 가지 장치 및 효과성을 대상으로 꾸준히 연구되어 왔다. 치료목적인 디지털 프로그램으로 ADHD 환자를 대상으로 하는 EndeavorRx[3]

와 같은 기술들이 효과가 입증되고 있다. 또한 뇌졸중의 디지털 치료와 관련하여 집에서 뇌졸중 환자 대상 원격 재활치료의 효과성에 관한 연구가 진행된 바 있으며[4], 여러 가지 재활 연구자료를 취합하여 15개의 운동 재활에 대한 기준을 제시하기도 하였다[5].

상지/하지/전신 재활 치료를 수행할 때 몰입형 재활 콘텐츠 장치인 가상현실(VR) 혹은 증강현실(AR) HMD를 사용할 수 있다[6]. 이 중 기존에 많이 시도되었던 VR HMD 장비(예: Oculus)를 이용한 재활 콘텐츠의 경우 HMD가 시야를 차단하여 바깥이 보이지 않으므로 거동이 불편한 환자는 치료사가 옆에서 보조를 해주지 않으면 다치거나 사고가 발생할 위험이 있다. 반면에 광학식/비디오 외부 투과(Optical / Video See Through) 방식의 AR HMD(예: 홀로렌즈)의 경우 주위 환경이 보이는 상황에서 그래픽 콘텐츠를 체험할 수 있으므로 HMD 착용 중 사고의 위험성을 낮추면서도 사용자의 몰입을 높일 수 있다. 하지만 AR HMD는 VR장치 보다 일반적으로 시야각이 좁고 야외나 혹은 밝은 공간에서 영상이 어둡고 뚜렷하지 않다. 또한, 별도의 컨트롤러 없이 손을 이용하여 상호작용해야 하는 경우가 많다.

본 연구는 이러한 AR HMD의 특징인 사용자의 머리 위치와 방향 추적, 손 상호작용밖에 할 수 없다는 제약을 극복하고 전신을 이용하기 위해서 외부 센서를 연동하여 재활이 가능하도록 지원한다. AR 콘텐츠에서 전신 상호작용을 위해서 IMU, 마커 기반 광학 인식(예: 옵티트랙), 깊이기반 전신 동작 인식 센서(예: 키넥트) 등이 사용 가능하다.

본 연구에서는 이들 중 홀로렌즈(HoloLens, Microsoft)와 키넥트(Kinect, Microsoft)를 이용하여 전신 재활을 위한 상호작용 기술을 제안한다. 이 경우 홀로렌즈와 키넥트에서 인식된 공간의 좌표계가 서로 다른 지역(Local) 좌표계를 사용하므로 전역 좌표계를 지정 후 좌표계 변환(Transform)을 통한 일치가 필요하다[7].

본 연구가 기여하는 바는 첫째, AR HMD와 외부 RGB-D 센서를 함께 사용할 때, 일치하는 정점 위치를 이용하여 서로 다른 로컬 좌표계를 글로벌 좌표계로 일치시키는 3가지 방식을 사용성과 정확도의 관점에서 제안하는 것이며, 둘째, AR HMD에서 외부 이미지 및 대응하는 깊이 정보를 가져올 수 없어서 나머지 방법이 불가능할 때 사용하는 신체 부위를 이용한 방법을 제안하는 것에 있다.

2. 관련 연구

2.1 마커를 이용한 좌표계 변환

AR HMD와 외부 RGB-D 센서의 좌표계를 일치시키는 여러 가지 방법 중, 실시간으로 좌표계 변환을 통한 전역 좌표계 일치에는 마커를 이용한 좌표계 변환 기술이 사용 가능하다. 이 경우 좌표계 일치를 위한 기존 기술로는 QR 코드 혹은 AR

마커 등을 사용한 변환 기술[8]이 소개되었다. HMD와 RGB-D 센서가 함께 볼 수 있는 위치에 인식 가능한 마커를 배치하고 각각의 지역 좌표계에서 RGB 영상을 이용하여 마커를 인식한다. 이 때 마커가 너무 작거나 너무 멀리 떨어져 있거나 대비가 좋지 않을 경우 인식이 떨어져서 부정확할 수 있으므로 미리 고려해야 한다. 그 후 대응되는 깊이 데이터를 기반으로 위치와 방향을 추적한다. 추적된 각 좌표계를 비교 후 강제 변환행렬(3x3 회전행렬과 1x3 이동 벡터)을 계산하여 일치시킨다. 한 번 변환행렬이 계산되면 마커는 제거되어도 무관하다. 또한 양 기기의 추적 데이터 사이에 약간의 대기 시간 오차가 발생할 수 있어 마커를 움직여서 사용할 경우 사용자가 너무 빨리 움직이면 부정확성을 초래할 수 있다. 따라서 추적 알고리즘의 최적화를 통하여 대기 시간을 최소화하는 것이 중요하다. 또한 마커가 아닌, 체커보드를 이용하여 RealSense와 홀로렌즈 장치의 좌표계를 일치시키는 경우[9] 혹은 DSLR 카메라와의 좌표계를 일치하는 경우[10] 등이 소개된 바 있다.

본 연구는 이 방식 중 QR 코드를 마커로 사용하여 좌표계 변환하는 방식을 제안하고 있으며, 이 때 여러 개의 고정적인 마커 집합을 사용하는 방법이 아닌, 한 개의 마커를 손으로 이동시키며 대응 데이터를 모아 변환 행렬을 계산한다.

2.2 특징점을 이용한 좌표계 변환

서로 다른 좌표계를 가진 메쉬를 통합하는 방법은 특징점 (Feature Point) 정합, Iterative Closest Point(ICP), Sample Consensus(SaC) 등이 있다. ICP는 메쉬의 모든 점을 기준으로 최소의 오차가 발생하도록 변환하므로 정확도를 높일 수 있으나[11], 많은 CPU 부하와 연산시간이 필요하며, GPU를 통한 가속으로 실시간 사용이 가능하나 홀로렌즈와 같은 모바일 장치에서 실시간으로 처리하기에는 부적합하다. 또한 홀로렌즈가 스캔하는 공간 메쉬는 세밀함이 떨어지므로 이를 이용한 메쉬 매핑은 정확도가 더욱 떨어진다.

SaC는 체험 공간의 메쉬의 형태가 도형 Primitive(구, 평면, 실린더 등)와 유사할 경우에만 적용할 수 있다. 따라서 특징점을 이용한 좌표계 변환이 적합한 해결책이 될 수 있다. 이 때 다양한 방식의 특징점 추적 기법이 사용 가능하다. 특징점은 각 센서의 RGB영상에서 컴퓨터 비전 알고리즘을 이용하여 추출하거나 깊이로부터 변환된 3D Point Cloud를 이용하여 추출이 가능하다. 예를 들어 KLT 알고리즘을 이용한 SIFT 특징점 추적 기법[12]이 사용되기도 하였으며, 이러한 특징점 기반 방법은 자동화된 방식도 있지만 수동 개입으로 지정하기도 한다.

본 연구는 이러한 특징점을 지정하는 방식에 있어서 메쉬의 공통적인 위치 3곳을 수작업으로 지정 후 이를 기준으로 변환 행렬을 계산하는 방식과 공통적으로 인식하는 신체의 일부를 이용하여 변환하는 방식을 제안하고, 각 기술의 장단점 및 정확도를 분석하였다.

2.3 Hybrid 기법을 이용한 좌표계 변환

Hybrid 기법은 여러 가지 기법의 조합을 이용해서 더욱 강인하고 정확한 변환을 추구하는 방법이다. 이 때 조합 방식이 잘못될 경우 각각이 서로 잡음(Noise)으로 작용하여 단일 기술을 사용하는 것보다 예러가 커지는 경우도 있다. 또한 단일 기법을 사용하는 것보다 많은 계산량이 요구되고, 홀로렌즈의 경우 CPU 성능에 제약이 있으므로 조합에 사용할 요소와 융합 방식에 주의하여야 한다. 이러한 방식의 예는 I-ICP와 AR Code 인식, 특징점 인식을 병용해서 사용하거나[13], 스마트폰의 Plane인식 및 IMU를 병용하여 AR HMD의 컨트롤러로 사용[14]하는 등 다양한 조합이 소개된 바 있다. 본 연구는 RANSAC을 후처리에 포함시킴으로써 오차를 줄이도록 하였다.

3. 다중 센서를 이용한 AR 몰입 콘텐츠 구조

본 연구는 다중 센서를 이용하여 AR 몰입 콘텐츠의 상호작용을 처리하며 그 구조는 Fig. 1과 같다. 사람의 신체 동작을 인식하기 위한 센서 장비로 홀로렌즈와 키넥트를 이용한다. 이때 홀로렌즈는 동작 인식뿐만 아니라 몰입형 AR 콘텐츠 가시화와 재활콘텐츠 전반적인 진행을 수행한다. 홀로렌즈에서 동작 인식이 가능한 신체 부위는 HMD를 착용하고 있는 머리 부분과 양 손이다. 머리 부위는 홀로렌즈의 가시화를 위한 카메라 위치, 방향 정보를 활용한다. 손 추적은 홀로렌즈 HMD에 부착된 손 추적 카메라의 시야각 이내에 손이 보일 경우에만 추적이 가능하다. 시야각을 벗어나면 추적할 수 없다. 키넥트는 깊이 카메라에서 분석한 데이터를 사용하여 전신 관절을 추적한다. 홀로렌즈를 착용하고 있는 경우에도 키넥트는 머리 부위를 대부분 올바르게 찾아내므로 두 장치를 함께 사용하는 데 무리가 없다.

전신 재활 콘텐츠가 홀로렌즈에서 동작하여 머리, 손만 추적하기 때문에 부족한 전신 추적 기능을 보완하기 위해서 키넥트를 활용한다. 키넥트는 Personal Computer(PC)에 연결되어 구동되고, 키넥트를 통해 얻은 전신 추적 데이터는 무선통신(WiFi)을 통해서 홀로렌즈로 전달된다.

홀로렌즈와 키넥트가 사용하는 좌표계의 불일치를 해소하기 위해서 전달받은 데이터에 대해 좌표계 변환을 처리한다. 이러한 공간 좌표계 변환은 상호작용을 위해서 중요한 단계이지만 여러 가지 센서의 누적 오차로 인해서 정확한 변환이 쉽지 않다. 따라서 본 논문에서는 상황에 따라 사용성 및 정확성을 고려해 선택 가능한 3가지 좌표계 변환 기술을 제안하며, 추가로 RANSAC을 기반으로 한 정확도 보정을 통해서 오차를 추가로 감소시킨다. 이렇게 오차를 줄여 좌표계 변환이 된 전신 관절 데이터를 통해서 몸체, 다리, 발 등을 사용하는 AR 몰입형 전신 재활 콘텐츠 상호작용을 수행한다.

4. 전신 관절 인식을 위한 공간 좌표계 변환

PC에서 전달받은 키넥트의 전신 데이터는 홀로렌즈와는 다른 원점, 좌표축을 사용하므로, 홀로렌즈에서 유지하고 있는 3차원 좌표계와 일치하지 않는다. 따라서 동일한 공간 상에서 같은 물체를 보더라도 홀로렌즈와 키넥트가 사용하는 위치와 방향 수치값은 다르다. 이런 불일치를 해소하기 위해서 좌표계 변환을 처리한다. 좌표계 변환은 용도 및 상황에 따라 메쉬 특징점 기반 변환, AR 마커 기반 변환, 신체 인식 기반 변환 방법 중 한 방법을 선택하여 활용한다.

4.1 메쉬 특징점 기반 공간 좌표계 변환

홀로렌즈와 키넥트 모두 깊이 카메라 센서를 가지고 있어서 이를 이용하여 공간 인식 및 메쉬 생성이 가능하다. 이 구조를 이용해서 서로 다른 좌표계의 융합이 가능하다.

본 연구에서 첫 번째 제안하는 방식은 메쉬에서 공통으로 뽑아낼 수 있는 특징적인 위치를 3개를 지정하고, 이를 기준으로 좌표계 변환을 수행하는 것이다. 동일한 위치를 홀로렌즈와 키넥트가 가진 각각의 다른 센서가 자신만의 좌표계로 보았을 때 이 위치들의 변환 정보를 찾아내고, 이렇게 한 변환에서 다른 변환으로 변환되도록 변환행렬을 구함[15]으로써 좌표계 변환이 가능하다. 이를 위해서 특징점은 3D 공간에서 삼각형을 만들어 낼 수 있도록 최소 3개 이상이 필요하다.

특징점은 사용자가 3D Pointing 혹은 3D Modeling 도

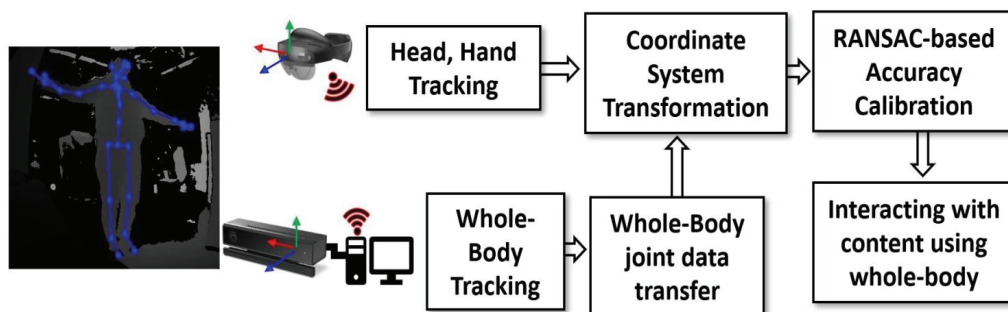


Fig. 1. Structure of AR Immersive Content Interaction Using Multiple Sensors

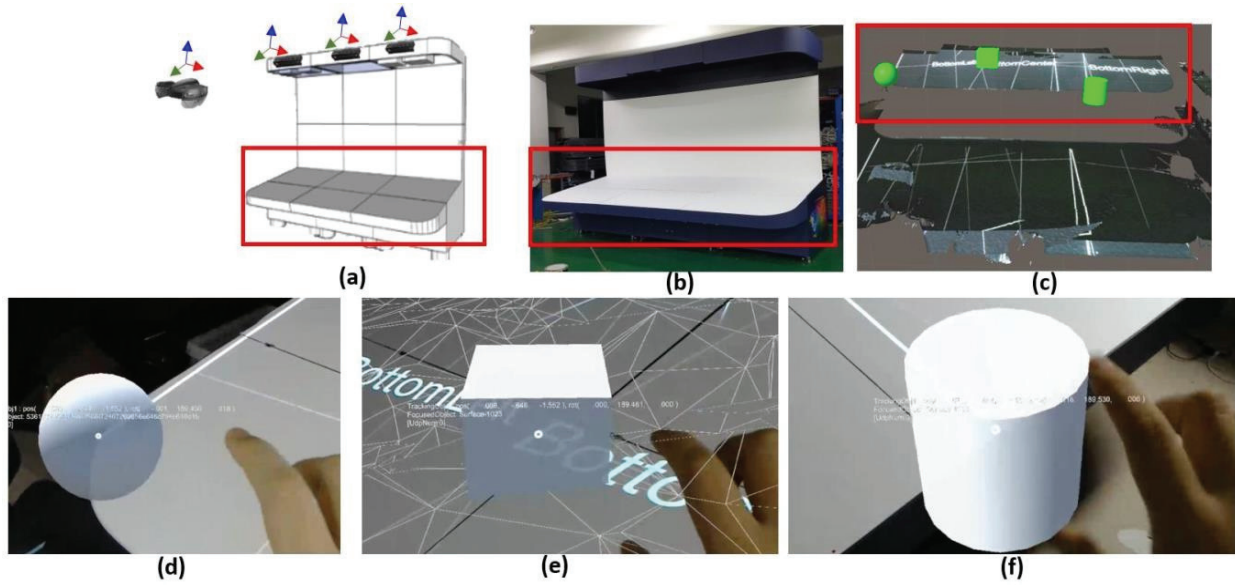


Fig. 2. An Example of Using Feature Points of Kinect and HoloLens on the Spatial Mesh of Content

구를 이용하여 수작업으로 지정하며, 홀로렌즈에서와 키넥트 각각에서 특징점을 지정할 때 동일한 위치를 지정하기 용이하도록 뾰족하게 튀어나오거나, 움푹 들어가거나, Texture의 컬러값이 쉽게 구분할 수 있는 것을 지정한다.

키넥트로부터 구성한 메시는 홀로렌즈보다 조금 더 많은 Vertex를 가져서 세밀하다. 따라서 특징점을 지정할 경우 상대적으로 세밀하지 않은 홀로렌즈의 메시에서 쉽게 구분할 수 있는 위치를 우선시한다. 키넥트에서 얻어낸 메시 데이터에서 특징점을 사용자가 지정하기 위해서는 Blender와 같은 3D 편집 도구 혹은 Unity3D와 같은 3D 콘텐츠 제작 도구를 이용할 수 있다. 이때 서로 구분하기 쉽게 하기 위해서 Fig. 2와 같이 콘텐츠 체형 공간(a)에 메시 특징점 기반 공간 좌표계 변환용 키넥트 메시에 설정한 특징점(c, 연두색)과 홀로렌즈에서 손 상호작용으로 대응되는 특징점 위치에 서로 다른 도형이 배치된다. 현재 사용하고 특징점 도형은 구(d), 정육면체(e), 실린더(f)를 사용했다.

홀로렌즈는 재부팅이 되거나 애플리케이션을 다시 실행할 경우 기존의 위치와 같은 위치일지라도 다른 좌표 값으로 표현된다. 즉, 매번 좌표계가 변한다. 따라서 배치된 특징점은 Space Anchor라고 하는 홀로렌즈의 기능을 이용해서 프로그램이 홀로렌즈가 다시 켜졌을 때도 계속 같은 위치에 유지가 될 수 있도록 저장한다.

제안한 첫번째 좌표계 변환 방법은 사람이 직접 3D 공간 상에 3개의 조절점을 지정하고, 깊이 데이터만 사용하므로, 다음에 서술할 다른 2개의 방법보다 더 정확할 수는 있지만, 각 장치별로 사람의 수작업이 필요하고, 수작업의 난이도와 숙련도에 따라 차이가 커지며, 이 때 상황에 따라 다른 방식보다 오히려 오차가 커질 수 있다.

4.2 AR 마커 기반 공간 좌표계 변환

홀로렌즈와 키넥트는 둘 다 컬러 카메라 센서를 포함하고 있다. 이 컬러 센서를 이용한 AR 마커가 중간 매개체가 될 수 있다. 이 때 동일한 마커를 키넥트와 홀로렌즈에서 인식해서 AR 마커의 3차원 좌표를 기준으로 Fig. 3과 같이 좌표계를 융합할 수 있다.

키넥트의 AR 마커 좌표를 분석은 먼저 컬러 센서로부터 AR 마커를 인식하고 인식된 각 4개 외곽선 코너의 위치를 깊이 공간 좌표(2D)로 변환한 후, 그것을 다시 신체 인식하는 공간인 카메라 공간 좌표(3D)로 변환하여 신체 관절 좌표계와 동일한 좌표계에서의 키넥트 AR 마커 변환을 계산할 수 있다. 홀로렌즈에서도 마찬가지로 동일한 AR 마커를 컬러 센서에서 하고, 홀로렌즈의 3차원 공간인식 좌표계로 맵핑한 후 코너 3D 좌표를 이용하여 홀로렌즈의 AR 마커 변환을 계산할 수 있다. 이렇게 동일한 AR 마커를 이용하여 각각 장치별 인식 후 계산된 변환 정보를 이용하여 한 기기의 변환에서

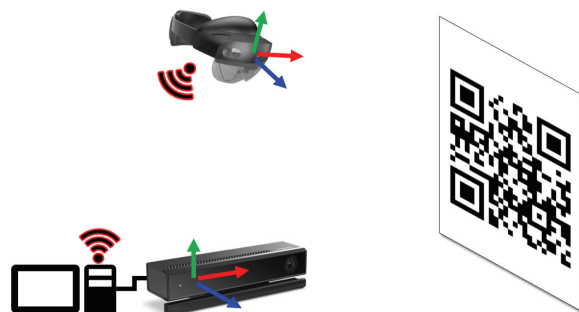


Fig. 3. Transformation of the Coordinate System in Space based on AR Marker

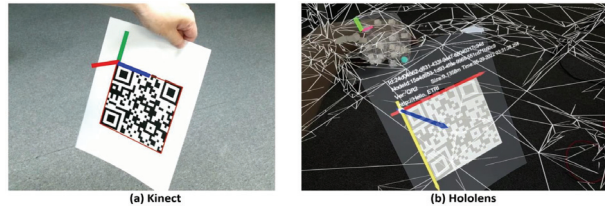


Fig. 4. Transform Analysis using QR Code Recognition

다른 기기의 변환으로 변환하는 행렬을 구할 수 있다. 사용 가능한 AR 마커는 ArUco[16], ARToolKit[17], QR 코드 [18] 등이 있다.

이 방법은 AR 마커의 자동 인식 기술을 이용하므로 수작업이 필요하지 않지만, 키넥트와 홀로렌즈가 동시에 볼 수 있는 위치에 AR 마커를 배치해야 하고, 양 센서에서 위치, 방향을 인식할 수 있을 정도로 AR 마커의 크기가 커야 한다. 또한 AR 마커는 기본적으로 컬러 카메라에서 획득한 영상을 사용하고 그 영상과 대응되는 깊이 데이터를 연동해서 사용하게 되는데 이때 키넥트와 홀로렌즈의 컬러 이미지와 깊이 이미지의 맵핑 관계가 오차를 포함하고 있다. 이러한 각 단계별 인식 오차 및 공간 맵핑 오차가 누적되어 최종 오차는 커진다.

본 연구에서는 여러 가지 AR 마커 중 QR 코드를 인식하고, 이를 기준으로 두 기기 간의 공간 좌표계 변환을 수행하였다. 이 때 여러 개의 고정적인 마커 집합을 사용하는 방법이 아닌, 하나의 마커를 손으로 이동시키며 대응 데이터를 모아 변환 행렬을 계산한다. 기기별 QR 코드 인식을 통한 변환 분석의 예시는 Fig. 4와 같다. (a)는 키넥트에서 인식한 QR 코드 변환이고, (b)는 홀로렌즈에서 인식한 변환이다.

4.3 신체 인식 기반 공간 좌표계 변환

홀로렌즈와 키넥트는 깊이 카메라의 데이터를 기반으로 신체 정보를 인식하는 기능을 제공한다. Fig. 5와 같이 키넥트는 전신 관절을 인식할 수 있고 홀로렌즈는 머리와 양손을 포함한 3개의 신체부위를 인식할 수 있다. 서로 인식하는 위치가 약간의 차이는 있을지 모르지만 두 장치가 각각 대응되는 부

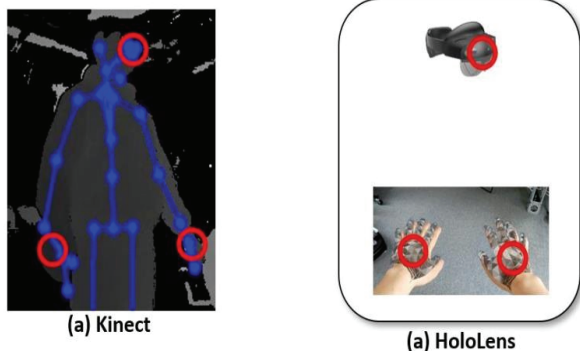


Fig. 5. Body Parts Corresponding to Coordinate System Transformation Based on Body Recognition

분을 인식한다고 볼 수 있다.

이렇게 같은 위치 3개를 각각의 기기에서 인식하므로 이를 이용하여 앞서 언급한 좌표계 변환을 위한 3개의 공통 특징점으로 활용할 수 있다. 즉, 인식된 신체 3개의 신체 부위를 이용해서 좌표계를 변환할 수 있다. 본 연구에서는 대응되는 신체 관절 부위로 키넥트의 Head, HandLeft, HandRight 관절 정보를 사용하였고, 홀로렌즈의 경우 HMD의 위치, Left Palm, Right Palm 관절 정보를 사용하였다.

해당 방법은 별도의 수작업, 또는 마커 등 추가적 장치가 필요 없이 정합할 수 있으므로 사용이 편리하다는 장점이 있다. 하지만, 신체 추적기술 자체가 오차를 포함하고 많은 떨림 현상을 수반하고 있다. 또한 키넥트에서 사용하는 머리와 손, 그리고 홀로렌즈에서 사용하는 머리와 손의 위치가 인식 방법의 차이로 인해 일치하지 않기 때문에 맵핑 시 오차는 더욱 커지게 된다. 이렇게 기본적인 신체 부위 인식 차이와 실시간 트래킹 에러로 인해서 좌표계 변환 오차가 앞의 두가지 좌표계 변환 방법보다 더 크다. 또한 AR HMD에서 외부 이미지 및 대응하는 깊이 정보를 가져올 수 없어서 앞서 소개한 두 가지 방법이 불가능할 때 적용 가능하다.

5. RANSAC 기반 좌표계 변환 정확도 보정

위와 같이 키넥트의 전신 관절 정보를 홀로렌즈에서 전달 받아 좌표계 변환을 통해 일치시킨 것을 홀로렌즈에서 가시화한 결과는 Fig. 6과 같다. 홀로렌즈의 기본 기능인 손(a) 인식 결과인 분홍색 구(Sphere) 이외에 팔(b), 다리(c)와 같이 녹색 선으로 표현되는 전신 관절이 사용자의 움직임에 따라 연동되어 올바르게 관절위치가 추적됨을 확인할 수 있다. Fig. 7은 RANSAC을 이용하여 좌표계 변환 행렬을 구하는 알고리즘이다.

AR 마커기반 좌표계 변환과 신체 인식기반 좌표계 변환 방법은 카메라 등 여러 가지 오차가 중첩되어 오차가 커질 수 있다. 이러한 오차를 보정하기 위하여 본 연구에서는 RANSAC [19]을 적용하였다. 일정 시간 동안 마커 혹은 신체 데이터를 누적하여, 이 중 이상치(Outlier)를 제외한 데이터의 오차를 최소가 되도록 좌표계 변환행렬을 구한다. 신체 인식기반의 경우 각 관절의 인식 신뢰도가 나오므로 인식 신뢰도를 고려하여 신뢰도가 높은 신체 부분의 경우 가중치를 높게 주고, 신뢰도가 낮은 경우 가중치를 낮게 주어 오차 최소화 기법을 적용한다.

이러한 좌표계 변환행렬 계산은 많은 시스템 자원을 소모한다. 키넥트의 위치, 방향이 바뀌지 않고, 홀로렌즈의 공간추적이 리셋되지 않는 이상 각각의 좌표계는 유지되며, 따라서 좌표계의 변환 행렬은 유지된다. 따라서 초기 일정 시간 동안 최적화 후 안정화되면 중단하고 그때까지의 좌표계 변환행렬을 고정하여 사용함으로써 시스템 자원을 절약할 수 있다.

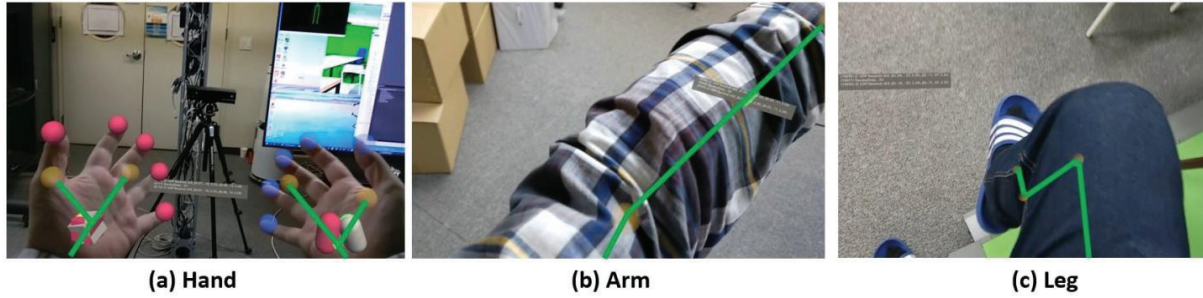


Fig. 6. The Result of Mapping to Whole Body Joint Information Seen from HoloLens (green line)

Algorithm 1 Calculation of Coordinate Transform using RANSAC

Input: $S = \{ (x_{h,l}, x_{k,l}), \dots, (x_{h,m}, x_{k,m}) \}$, A set of matched pairs of three vertex positions on HMD and Kinect

Output: T_m , Transform matrix from Kinect to HMD coordinate

```

1:  $I_{max} \leftarrow 0$ 
2: for each round  $t = 1$  to  $N$  do
3:    $X_t \leftarrow$  set of  $m$  point pairs randomly sampled from  $S$ 
4:    $T_{h,t} \leftarrow$  Compute Transform Matrix from  $X_t$  on HMD
5:    $T_{k,t} \leftarrow$  Compute Transform Matrix from  $X_t$  on Kinect
6:    $T_{m,t} \leftarrow T_{h,t} \times (T_{k,t})^{-1}$ 
7:    $X_{k,m} \leftarrow$  Positions on Kinect from  $S$  transformed by  $T_{m,t}$ 
8:    $I_t \leftarrow$  Count the Inlier Number of  $X_{k,m}$ 
9:   if  $I_t > I_{max}$  then
10:      $I_{max} \leftarrow I_t$ 
11:      $T_m \leftarrow T_{m,t}$ 
12:   end if
13: end for
    
```

Fig. 7. Pseudo Code of Coordinate Transform Calculation

6. 실험 결과 및 분석

본 연구에서 몰입형 AR HMD 장치로 홀로렌즈 2를 사용했다. 전신 동작인식을 위해 키넥트 2를 PC에 연결하였다. PC는 Intel Core i9-11900K, RAM 32G를 장착하였으며 운영체제로 Windows 10를 사용하였다. 홀로렌즈 구동 및 콘텐츠 개발 도구로 Unity3D, 키넥트 구동 및 Open CV 구동용으로 MS Visual Studio를 컴파일에 사용하였다. QR 코드를 검출하기 위해서 PC에서는 OpenCV를, 홀로렌즈에서는 MS Mixed Reality QR SDK를 활용하였다.

추적된 전신 관절 데이터로 상호작용하기 위해 관절 부위에 Fig. 8과 같이 Collider를 부착하였다. 머리와 손은 구(Sphere) 형태를 Collider로 사용하였고, 몸, 상완, 전완, 허벅지, 정강이, 발은 캡슐(Capsule) 형태의 기본형(Primitive)를 가진 Collider를 사용하였다. 홀로렌즈의 CPU 계산 부담을 덜기 위해 몸 전체를 둘러싸는 구 Collider를 두어 세부

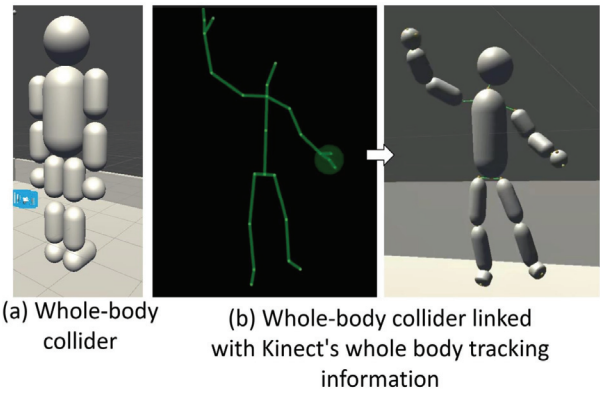


Fig. 8. Collider for Collision Detection Linked with Whole- Body Joint Information

관절 Collider 충돌체크 이전에 수행하여 대상 물체가 충돌된 경우에만 세부 관절 충돌 체크를 수행하였다. 캡슐은 신체의 길이와 굵기에 따라서 모양이 결정된다. 홀로렌즈와 PC의 통신을 위해 무선통신(WiFi)를 사용하였으며, 이때 매 프레임마다 생성되는 전신동작 관절 정보는 속도를 위해 UDP, 콘텐츠 관련 정보의 교환은 안정성을 위해 TCP를 사용하였다.

이러한 홀로렌즈와 키넥트를 연동한 상호작용 실험을 위해 2개의 콘텐츠를 적용하였다. 첫 번째는 대형 다중 프로젝터 연상을 이용하여 몰입감을 주고 키넥트와 홀로렌즈를 이용하여 입체감을 느낄 수 있는 “썬더펀치”[20,21] 콘텐츠를 Fig. 9와 같이 적용하였다. 프로젝터 영상을 출력하기 위한 대형 L자형 디스플레이 장치(a)를 사용하였으며, 이러한 대형 장치형 콘텐츠는 초기 설치 후 이동이 없으므로 수작업으로 미리 세팅하는 것이 유리하다. 따라서 썬더펀치에서 이들 장치 간의 좌표계 연동은 메쉬 특징점 기반 공간 좌표계 변환 방법을 사용하였다. 이렇게 함으로써 손에서 쏘아지는 레이저(b)를 입체감 있게 가시화할 수 있다.

두 번째는 뇌졸중 재활 치료용 콘텐츠로써 전신 운동을 재할하기 위한 기능성 콘텐츠에 Fig. 10과 같이 적용하였다. 해당 콘텐츠는 여러 방향에서 몬스터가 나타나 재활 환자로 다가오면 환자는 이때 몬스터를 피하거나 잡는 동작을 하여 운동 기능을 향상시킨다. 이 콘텐츠는 AR 마커 방식 및 신체인

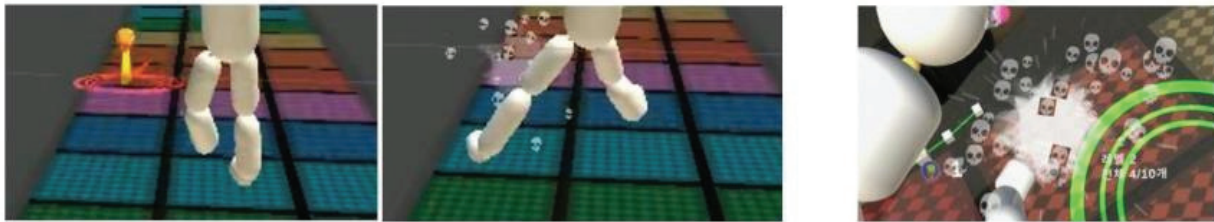


(a) “Thunder Punch” content using multi-projector, HoloLens, and Kinect



(b) An example of laser shooting from the hand of the HoloLens point of view

Fig. 9. The Result of Applying AR Whole Body Interaction Technology to Thunder Punch Contents



(a) Examples of monster interaction using leg collision detection

(b) Interactions from a HoloLens perspective

Fig. 10. Results of AR Rehabilitation Content Interaction Using Whole Body Motions

Table 1. Comparison of Features of 3 Methods

	Usability	Accuracy	External image Requirement
Feature Point	Low	High	Yes
AR Marker	Mid	Mid	Yes
Body Skeleton	High	Low	No

Table 2. Average Errors of 3 Methods

	Feature Point	AR Marker	Body Skeleton
Error (mm)	8.519	11.296	58.385

Table 3. Average Errors based on RANSAC

		AR Marker	Body Skeleton
Error (mm)	with RANSAC	11.296	58.385
	without RANSAC	12.647	67.287

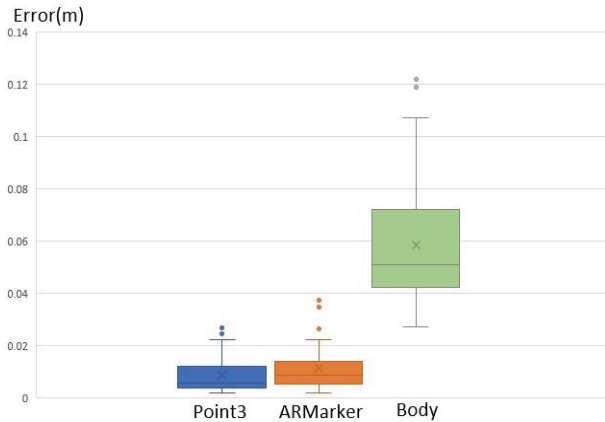


Fig. 11. Box-and-whisker Plot of Errors

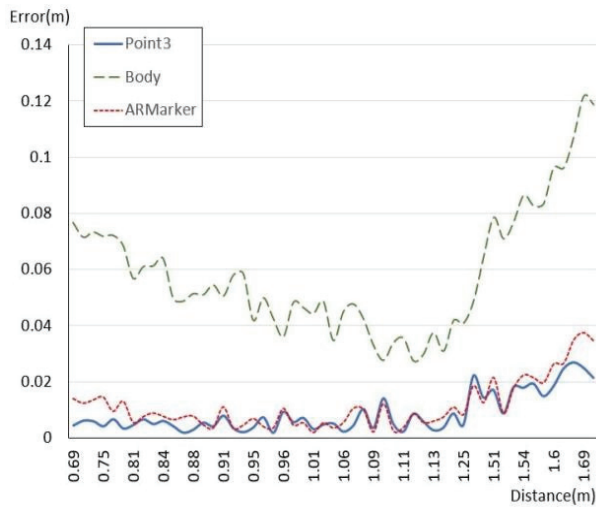


Fig. 12. Errors by Distance

식 기반 방식을 사용하여 좌표계를 변환하였다. 사용의 편의성이 요구될 때는 신체인식기반 방식을, 좀 더 정확한 좌표계 변환이 필요하면 AR 마커방식을 사용하였다. 홀로렌즈는 손을 이용한 상호작용밖에 되지 않으므로 키넥트와 연동하여 전신 동작을 인식함으로써 전신 재활 치료가 가능하였다.

Table 1은 각 방식을 사용성, 정확도, 외부 카메라 영상 필요성에 따라서 비교한 것이다. Table 2는 각 방식의 평균 오차를 비교한 것이다. 메쉬 특징점 기반 방식(Feature Point)이 약 8.5mm의 오차로 가장 오차가 작았고, AR 마커 방식은 약간 더 오차가 높았으며, 신체 인식 기반 방식은 약 58.4mm로 오차가 가장 크게 나타났다. 이처럼 신체 인식 기반 방식은 큰 오차를 보여주었다. 하지만 신체 인식 기반 방식은 추가적인 도구가 필요 없고 조정 단계가 간단하므로, 콘텐츠에서 필요한 정확도와 편의성을 고려하여 사용할 수 있다. Fig. 11은 각 방식 에러의 최댓값, 최소값, 사분위수를 구하여 표현한 그림이다.

Table 3은 RANSAC을 적용한 경우와 적용하지 않은 경우

의 오차를 비교한 것이다. RANSAC을 적용한 경우 오차가 줄어들었으며, 측정할 때 이상치(Outlier)가 많을수록 RANSAC을 적용하는 경우가 더 유리하다. 메쉬 특징점 기반 방식은 미리 스캔된 메쉬에 사용자가 수동으로 지정하여 RANSAC을 사용하지 않으므로 비교 대상에서 제외하였다.

Fig. 12는 3가지 방법에 대해 거리에 따른 오차를 측정할 그래프이다. 1m 거리에서 각 기준 값을 측정하여 변환 정보를 구한 후, 거리를 변화시켜 주며 오차를 측정할 값이다. 측정할 기준 거리인 1m에서 멀어질수록 오차가 증가하는 것을 확인할 수 있다. 이 때 메쉬 특징점 기반 방식과 AR 마커 방식은 거리에 따른 오차가 크지 않은 반면 신체 인식 기반 방식은 거리에 따라서 오차가 크게 차이가 났다.

7. 결 론

본 연구는 뇌졸중 환자의 재활 치료에 있어서 AR 몰입형 콘텐츠를 전신 상호작용할 수 있는 기술을 소개하였다.

AR 몰입 장치인 홀로렌즈와 전신 동작 인식 장치인 키넥트를 연동하여 전신 상호작용을 제공하는 데에 있어서 각 기기가 서로 다른 지역 공간 좌표계를 사용하므로 이를 전역 좌표계로 일치시키기 위해서 메쉬 특징점 기반 변환, AR 마커 기반 변환, 신체 인식 기반 변환 등 세 가지 방법을 제시하였다. 사용자는 이러한 세 가지 방법 중 콘텐츠의 체험 환경을 고려하여 한 가지를 선택하여 활용할 수 있다. 이때 발생하는 오차를 줄이기 위해서 신체 인식의 신뢰도 정보와 RANSAC 기반 최적화 기법을 활용하였다. 이러한 좌표계 변환 기술을 이용하여 '썬더펀치'와 뇌졸중 재활 치료 콘텐츠에 적용함으로써 실용성이 있음을 보였다.

본 연구에서 제시한 기술을 통해 보다 안전하고 몰입감 있는 상호작용을 제공하는 AR 전신 재활 콘텐츠를 서비스할 수 있을 것으로 기대한다.

References

- [1] KERI, International Comparison of Low Fertility and Aging Trends and Implications, 2021.
- [2] S. H. Shin, "Use of rehabilitation medical services for stroke patients," *HIRA*, pp.72-81, 2019.
- [3] S. H. Kollins et al., "A novel digital intervention for actively reducing severity of paediatric ADHD (STARS-ADHD): A randomised controlled trial," *The Lancet Digital Health*, Vol.2, No.4, pp.168-178, 2020.
- [4] S. C. Cramer et al., "Efficacy of home-based telerehabilitation vs in-clinic therapy for adults after stroke: A randomized clinical trial," *JAMA neurology*, Vol.76, No.9, pp.1079-1087, 2019.

- [5] M. Maier, B. R. Ballester, and P. F. Verschure, "Principles of neurorehabilitation after stroke based on motor learning and brain plasticity mechanisms," *Frontiers in Systems Neuroscience*, 2019.
- [6] D. Y. Kim et al., "Clinical practice guideline for stroke rehabilitation in Korea 2016," *Brain & Neurorehabilitation*, 13:74, 2016.
- [7] B. Liu, F. Zhang, X. Qu, and X. Shi, "A rapid coordinate transformation method applied in industrial robot calibration based on characteristic line coincidence," *Sensors*, Vol.16, No.2, pp.239, 2016.
- [8] M. De Cecco et al., "Sharing augmented reality between a patient and a clinician for assessment and rehabilitation in daily living activities," *Information*, Vol.14, No.4, pp.204, 2023.
- [9] M. Garon, P.-O. Boulet, J.-P. Doiron, L. Beaulieu, and J.-F. Lalonde, "Real-time high resolution 3D data on the HoloLens," *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, pp.189-191, 2016.
- [10] J. Smisek, M. Jancosek, and T. Pajdla, "3D with Kinect," *Consumer Depth Cameras for Computer Vision: Research Topics and Applications*, pp.3-25, 2013.
- [11] S. H. Choi et al., "An integrated mixed reality system for safety-aware human-robot collaboration using deep learning and digital twin generation," *Robotics and Computer-Integrated Manufacturing*, Vol.73, 2022.
- [12] C.-H. Hsieh, J.-D. Lee, and C.-T. Wu, "A Kinect-based medical augmented reality system for craniofacial applications using image-to-patient registration," *Neuropsychiatry*, Vol.7, No.6, pp.927-939, 2017.
- [13] M.-L. Wu, J.-C. Chien, C.-T. Wu, and J.-D. Lee, "An augmented reality system using improved-iterative closest point algorithm for on-patient medical image visualization," *Sensors*, Vol.18, No.8, pp.2505, 2018.
- [14] K. Hattori and T. Hirai, "Inside-out tracking controller for VR/AR HMD using image recognition with smartphones," *ACM SIGGRAPH 2020 Posters*, 2020.
- [15] C. T. Russell, "Geophysical coordinate transformations," *Cosmic Electrodynamics*, Vol.2, No.2, pp.184-196, 1971.
- [16] S. Garrido-Jurado, R. Munoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marin-Jimenez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, Vol.47, No.6, pp.2280-2292, 2014.
- [17] H. Kato and M. Billinghurst, "Marker tracking and hmd calibration for a video-based augmented reality conferencing system," *Proceedings 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR'99)*, IEEE, pp.85-94, 1999.
- [18] Denso Wave Corporation, QR Code features [Internet]. <http://www.qrcode.com.2000>.
- [19] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Communications of the ACM*, Vol.24, No.6, pp.381-395, 1981.
- [20] S. Park et al., "Development of Projection based Virtual Experience System using Punching," *Korea HCI Conference*, pp.847-849, 2017.
- [21] D. H. Kim, H. M. Kim, H. K. Kim, S. R. Park, K. S. Lee, and K. H. Kim, "ThunderPunch: A bare-hand, gesture-based, large interactive display interface with upper-body-part detection in a top view," *IEEE Computer Graphics and Applications*, Vol.38, No.5, pp.100-111, 2018.



김 항 기

<https://orcid.org/0000-0002-3998-636X>

e-mail : hangkee@etri.re.kr

1999년 경북대학교 컴퓨터과학과(학사)

2001년 경북대학교 컴퓨터과학과(석사)

2001년 ~ 현 재 한국전자통신연구원

VR/AR콘텐츠연구실 책임연구원

관심분야 : 컴퓨터 그래픽스, VR/AR, Bio, 의료



김 대 환

<https://orcid.org/0000-0001-8239-3503>

e-mail : daehwankim@ulsan.ac.kr

2002년 동국대학교 컴퓨터멀티미디어

공학과(학사)

2004년 POSTECH 컴퓨터공학과(석사)

2011년 POSTECH 컴퓨터공학과(박사)

2012년 ~ 2022년 한국전자통신연구원 선임연구원

2022년 ~ 현 재 울산대학교 IT융합전공 교수

관심분야 : 컴퓨터비전, 머신러닝, 인공지능, HCI, 가상현실,

증강현실



이 동 춘

<https://orcid.org/0000-0003-4917-4346>

e-mail : bluepine@etri.re.kr

1999년 경북대학교 컴퓨터공학과(학사)

2001년 경북대학교 컴퓨터공학과(석사)

2001년 ~ 현 재 한국전자통신연구원

VR/AR콘텐츠연구실 책임연구원

관심분야 : 컴퓨터 그래픽스, VR/AR, Bio, 의료



이 기 석

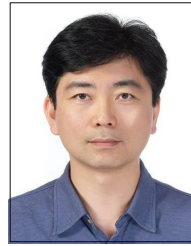
<https://orcid.org/0000-0002-3843-5339>

e-mail : mvr_lks@etri.re.kr

1999년 성균관대학교 제어계측공학과
(학사)

2001년 성균관대학교 전기전자 및
컴퓨터공학과(석사)

2001년 ~ 현 재 한국전자통신연구원 VR/AR 콘텐츠연구실 실장
관심분야: 컴퓨터 그래픽스, VR/AR, Bio, 의료



백 낙 훈

<https://orcid.org/0000-0003-2136-843X>

e-mail : oceancru@gmail.com

1990년 한국과학기술원 전산학과(학사)

1992년 한국과학기술원 전산학과(석사)

1997년 한국과학기술원 전산학과(박사)

1997년 George Washington

University, Visiting Scholar

1998년 경북대학교 전자전기공학부 초빙교수

2004년 ~ 현 재 경북대학교 컴퓨터학부 교수

관심분야: 모바일 그래픽스, 병렬컴퓨팅, 실시간 그래픽 및
특수효과