

# XR 기반 모션 가이드 시스템의 시점 디자인이 현실 세계 동작 학습과 사용자 경험에 미치는 영향

Effect of user perspective designs on

XR-based motion guidance systems for real-world motor skills

오민우

Minwoo Oh

광주과학기술원

Gwangju Institute of Science  
and Technology

alsdn4435@gm.gist.ac.kr

성민우

Minwoo Seong

광주과학기술원

Gwangju Institute of Science  
and Technology

seongminwoo@gm.gist.ac.kr

김승준

SeungJun Kim

광주과학기술원

Gwangju Institute of Science  
and Technology

seungjun@gist.ac.kr

## 요약문

동작을 학습할 때 전문가에게 직접 지도를 받는 방식이 가장 대표적이지만, 시·공간적 제약과 비용 부담이 발생한다는 문제가 있다. 따라서 최근에는 XR 기술을 이용해 3 차원의 동작 정보를 제공하는 XR 기반 모션 가이드 시스템이 새로운 동작 학습법으로 주목받고 있다. XR 기반 모션 가이드 시스템의 더 높은 학습 효율을 위해선 상황에 맞는 디자인을 적용해야 한다. 그중에서도, 시점 디자인이 시스템 성능에 유의미한 영향을 끼칠 수 있음이 이전 연구를 통해 입증되었다. 그러나, 이들의 경우 현실 세계에 존재하지 않는 동작을 사용했다는 한계가 존재한다. 따라서 본 연구는 현실 세계 동작을 위한 XR 기반 모션 가이드 시스템의 시점 디자인이 동작 수행 능력과 사용자 경험에 미치는 영향을 분석하였다. 실험은 태극권 동작을 위한 모션 가이드 시스템에 1 인칭(1PP), 3 인칭(3PP), 거울 시점(MPP), Multi-3PP 시점을 적용한 후 가상환경에서 동작을 따라하는 방식으로 진행되었다. 이때 동작의 오차, 주관적 난이도, 인지 부하를 측정해 비교하였다. 분석 결과 MPP 시점에서 높은 동작 오차와 인지 부하를 보였고, 1PP는 낮은 동작 오차에 비해 높은 주관적 난이도와 인지부하를 발생시키는 것으로 나타났다. 연구 결과는 현실 세계 동작을 위한 XR 기반 모션 가이드 시스템 설계에 있어 시점 디자인이 사용자 경험과 동작 수행 능력에 미치는 영향을 시사한다.

## 주제어

확장현실, 가상현실, 모션 가이드 시스템, 시점 디자인, 현실 기반 동작

## 1 서론

스포츠, 재활 훈련, 춤 등 다양한 분야의 동작을 학습하기 위해선 전문적인 동작 지도를 받아야 한다 [3]. 가장 대표적인 방법은 전문가에게 직접 지도를 받는 방식이지만, 시·공간적 제약과 비용 부담을 유발한다는 단점이 있다 [2]. 대안으로 비디오 튜토리얼을 활용한 동작 학습 방법이 제시되었는데 [4], 이 또한 2D 화면으로 전달되는 불완전한 정보로 인해 잘못된 자세를 학습할 수 있을 뿐만 아니라, 이에 대한 피드백 또한 받기 어렵다는 문제가 남아있다. 따라서, 최근 깊이 정보 등의 3 차원 동작 정보와 교정 피드백을 제공할 수 있는 XR 기반 모션 가이드 시스템(XR-based motion guidance system)이 새로운 동작 학습 방법으로 각광받고 있다 [4][5].

XR 기반 모션 가이드 시스템의 학습 효율을 최적화하기 위해선 상황에 맞게 적절한 디자인을 선택해야 한다 [1][7]. 특히, 사용자 시점 디자인은 시스템의 성능 및 사용자 경험(UX)에 유의미한 차이를 발생시킬 수 있다 [2][3]. 예를 들어, Yu et al. (2020)은 간단한 팔 동작 과제에서 1 인칭 시점과 3 인칭 시점을 사용했을 때 동작 수행 능력과 사용자 경험에 차이가 있는지 관찰한 결과, 1 인칭 시점이 더 높은 동작 정확도를 보인데 반해, 사람들은 양팔을 동시에 보기 쉬웠다는 이유로 3 인칭 시점을 더 선호하는 경향을 보였다 [2].

그러나 기존 시점 디자인 연구는 스포츠, 무술, 재활운동 등 특정 도메인에 기반한 동작이 아닌, 실험을 위해 임의로 생성된 동작이나 비현실적인 동작만을 사용한 한계가 있다. 예를 들어, 측면 공간만을 활용하는 주변 동작(peripheral motion), 정면 공간만을 활용하는 전방 동작(frontal motion) 등 가동 범위에 제약을 둔 동작을 분류하여 실험하거나, 실제로 존재하지 않는 동작을 사용

하는 경우도 있었다 [2][3]. 이러한 임의로 생성된 동작은 XR 기반 모션 가이드 시스템의 궁극적인 목표인 특정 도메인 동작 학습과 거리가 있다는 문제가 존재한다 [1]. 이는 기존 연구의 결과가 XR 기반 모션 가이드 시스템의 실제 적용 가능성을 충분히 반영하지 못한다는 한계를 초래할 수 있다.

따라서 본 연구는 태극권 동작을 활용하여 XR 기반 모션 가이드 시스템의 사용자 시점 디자인을 탐구하는 것을 목표로 한다. 1 인칭(1PP), 3 인칭(3PP), 거울 뷰(MPP), Multi-3PP 시점에 대한 비교를 진행했으며 [1][2][3][8], 실험 환경은 VR 및 웨어러블 센서를 기반으로 구현된 모션 가이드 시스템에 활용하여 구축하였다. 실험 참가자는 각 시점에서 제공된 가이드를 따라 동작을 수행하는 과제를 진행했으며, 이 과정에서 각 시점이 동작 오차, 주관적 난이도, 인지 부하에 미치는 영향을 관찰하였다. 본 연구 결과를 바탕으로, XR 기반 모션 가이드 시스템에서 효과적으로 활용할 수 있는 시점 디자인에 대한 시사점을 제안한다.

## 2 선행 연구

XR 기반 모션 가이드 시스템은 원하는 동작의 입체적인 가이드를 제공하기 때문에, 전문가의 도움 없이 홀로 동작을 학습해야 하는 상황에서 유용하게 활용될 수 있다. 이러한 장점을 기반으로 XR 기반 모션 가이드 시스템은 스포츠[5], 재활 운동[4], 무술[17], 발레[8] 등 다양한 분야에 널리 적용되고 있다.

XR 기반 모션 가이드 시스템의 시점 디자인은 사용자가 가이드의 움직임을 인지하고 이해하는데 영향을 끼치는 요소 중 하나이다 [1]. Yu et al. (2024)은 기존 XR 기반 모션 가이드 시스템에 사용된 시점 디자인을 3가지로 분류하였다. 먼저 1 인칭 시점(1PP)은 가상환경의 시점을 실제 사용자의 시점과 같은 곳에 위치시켜 더 자연스러운 동작 이해를 돕는 방법이다 [4][5][7][17]. 3 인칭 시점(3PP)은 가이드와 사용자 아바타의 후방에서 동작을 관찰할 수 있게 만든 디자인을 의미한다 [4][5]. 거울 시점(MPP)은 사용자와 가이드의 모습을 거울에 반사된 것처럼 보여주는 방법이다 [8]. 추가로, 한 연구에서는 3PP 시점에 측면 시점을 추가해 더 디테일한 동작 정보를 제공하는 Multi-3PP 시점을 제안했다 [3].

XR 기반 모션 가이드 시스템의 시점 디자인 간 비교를 진행한 연구도 이루어졌다. Yu et al. (2020)은 1PP와 3PP 시점을 적용해 각각 정면 공간과 측면 공간을 사용하는 팔동작을 수행했을 때 동작 오차, 주관적 난이도, 선호도 등을 비교하는 실험을 진행했다 [2]. 그 결과, 모든 동작에서 1PP가 3PP보다 낮은 동작 오차를 보였지만,

유의미하진 않았다. 반면 주관적 난이도와 선호도의 측면에서는 1PP가 유의미하게 부정적인 결과를 도출했다. 저자는 1PP로 보이는 화면이 3PP에 비해 실제와 더 비슷하기 때문에 깊이 정보를 파악하기 수월하지만, 양팔을 동시에 확인할 수 없다는 점이 사용자 경험에 부정적인 영향을 끼쳤을 것이라고 설명했다. 다만, 팔을 제한된 구역으로 움직이는 동작이 현실의 시나리오와 부합하지 않기 때문에, 추후 연구에선 더 복잡한 동작을 활용해 시점 간 비교를 수행할 것을 제안했다. 또, Elsayed et al. (2022)는 시점을 1PP, 3PP, Multi-3PP로 나누고, 동작의 방향을 측면, 정면, 후면으로 나누는 등 동작 정확도에 영향을 끼칠 수 있는 요소 간의 비교를 진행했다 [3]. 그 결과, Multi-3PP가 다른 시점 디자인에 비해 유의미하게 낮은 동작 오차를 보였다.

이처럼 기존 연구를 통해 XR 기반 모션 가이드 시스템의 시점 디자인이 시스템 성능에 유의미한 영향을 끼친다는 점을 알 수 있지만, 동작이 특정 공간에 한정되어 있거나 가상의 동작을 사용하는 등 현실과 동떨어진 동작 선택을 했다는 부분은 연구의 한계로 지적된다 [2]. 이러한 한계를 극복하고자 본 연구는 가동 범위가 자유롭고, 기존 XR 기반 모션 가이드 시스템에 자주 적용된 태극권 동작을 사용하여 시점 디자인과 동작 정확도 및 사용자 경험 간의 상관관계를 분석하였다.

## 3 VR 기반 모션 가이드 시스템 구축

본 연구의 인간 대상 실험에서 피험자는 XR 기기를 통해 시각화된 본인의 실시간 움직임과 전문가의 움직임을 관찰하며 특정 동작을 따라 하는 작업을 수행한다. 이를 위해 웨어러블 센서와 VR 기반의 모션 가이드 시스템 구축했다.

실시간 모션 트래킹을 위한 웨어러블 센서는 Noitom사의 Perception Neuron Studio (PNS)를 사용하였다. 17개의 트래커로 구성되어 있고, 각 트래커의 IMU를 기반으로 신체 관절의 위치와 각도를 추정할 수 있다. HMD는 최대 90Hz의 주사율과 시선 추적 기능 등을 제공하는 Meta Quest Pro를 사용하였다.

### 3.1 실시간 사용자 모션 시각화

그림 1(a)은 PNS 센서 데이터를 실시간으로 Unity 환경에 시각화하는 과정이다. 먼저 PNS 센서 전용 소프트웨어인 Axis Studio를 이용해 센서로부터 원시 데이터를 수신한다. 이를 모션 캡처 데이터 포맷인 BVH 파일로 변환해, 소켓 UDP 통신을 통해 Python API로 전송한다.

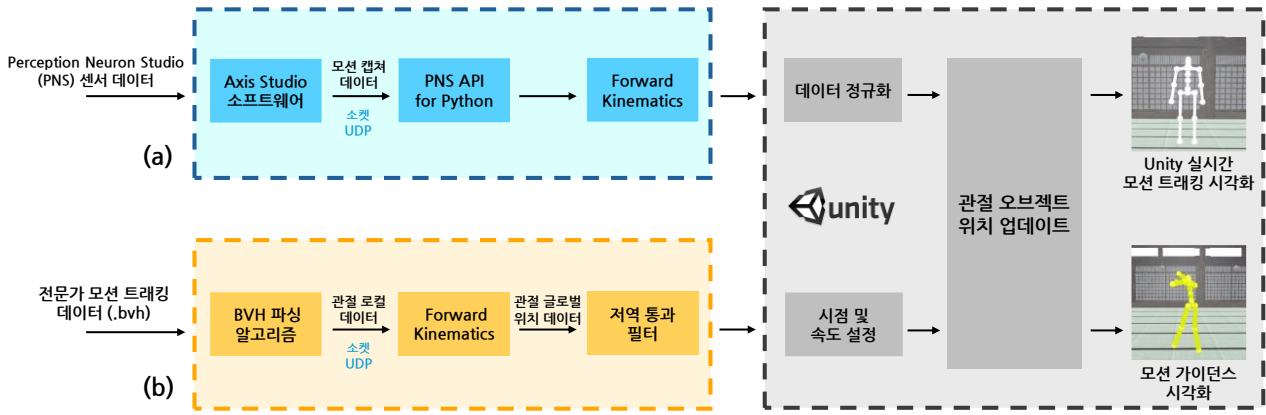


그림 1. VR & 웨어러블 센서 기반 모션 가이드 시스템 파이프라인

PNS Python API는 센서 데이터를 시각화가 가능한 형태로 가공해 Unity로 전송하는 역할을 담당한다 [15]. 우선 BVH 포맷의 모션 캡처 데이터에서 로컬 쿼터니온 각도와 로컬 위치를 추출해 파이썬 리스트 형태로 저장한다. 이를 Forward Kinematics (FK) 알고리즘에 입력해 글로벌 위치 데이터로 변환한다. FK 알고리즘은 관절의 로컬 위치 및 각도 데이터를 이용해 관절 말단의 글로벌 위치 데이터를 추정하는 방법으로, 로봇학과 애니메이션 분야에 주로 사용된다. 마지막으로 소켓 UDP 통신을 이용해 Unity로 결과 데이터를 전송한다

Unity 상의 동작 시각화는 각 관절에 해당하는 오브젝트를 프레임마다 업데이트하는 방법을 사용한다. 수신한 위치 데이터로 각 관절 오브젝트의 위치를 설정하고, 그 사이를 Unity LineRenderer로 연결해서 스켈레톤 형태의 사용자 아바타를 구현한다. 이로써 사용자는 VR 환경 내에서 본인의 움직임을 실시간으로 관찰할 수 있다.

### 3.2 가이드 시스템 모션 시각화

그림 1(b)은 전문가 모션 트래킹 데이터를 Unity 상에 시각화하는 과정을 보여준다. 전문가 모션 데이터는 오픈 소스 데이터셋에서 추출한 관절 위치 데이터를 CSV 파일로 저장해 사용한다. 먼저 BVH 포맷으로 저장된 로컬 관절 위치, 각도 데이터를 파이썬 리스트의 형태로 파싱하고, FK 알고리즘에 통과시켜 글로벌 데이터로 변환한다. 마지막으로 저장 및 필터링에 데이터를 통과시켜 CSV 파일 형태로 저장한다. Unity 시각화 방법은 사용자 데이터에서 사용한 방법과 동일하지만, CSV 형태의 데이터를 원하는 속도 및 시점으로 재생할 수 있다는 점이 다르다.

가이드 시스템에는 action segmentation을 위해 수집된 태극권 모션 캡처 데이터셋을 사용했다 [6]. 두 명의 피험자를 대상으로 10 개의 태극권 동작의 모션 캡처 데이터가 연속으로 수집되었다. Perception Neuron 센서를 이용해, 125Hz의 샘플링 속도로 수집된 로컬 관절 각도 및 위치 데이터를 BVH 포맷으로 제공한다.

## 4 실험 설계 및 결과

현실 세계 동작을 위한 XR 기반 모션 가이드 시스템의 시점 디자인이 동작 오차 및 사용자 경험에 미치는 영향을 탐구하기 위해 인간 대상 실험을 진행하였다. 본 실험에서 목표로 하는 연구 질문은 다음과 같다

- 연구 질문 1: XR 기반 모션 가이드 시스템을 통해 현실 세계 동작을 따라 할 때 시점 디자인에 의해 동작 오차는 어떻게 변화하는가?
- 연구 질문 2: XR 기반 모션 가이드 시스템을 통해 현실 세계 동작을 따라 할 때 시점 디자인에 의해 사용자 경험은 어떻게 변화하는가?

### 4.1 실험 설계

#### 4.1.1 동작 선정



그림 2. “Hold The Lute” 태극권 동작 시퀀스 예시

실험에 사용한 동작은 태극권의 “Hold The Lute” 동작이다 (그림 2). 태극권 종목을 선택한 이유는 태극권이 기존 XR 기반 모션 가이드 시스템에 주로 사용되는 현실 기반 동작이기 때문이다 [17]. 그 중 “Hold The Lute” 동작은 다른 태극권 동작에 비해 가동 범위가 넓고 주변 공간을 고루 사용한다는 점에서 기존 연구에 사용된 동작이 가진 가동 범위의 한계를 극복할 수 있다 [2].

#### 4.1.2 시점 디자인

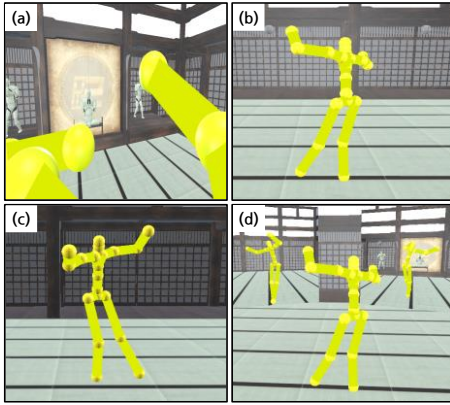


그림 3. 실험에 사용한 사용자 시점 디자인.  
(a) 1PP, (b) 3PP, (c) MPP, (d) Multi-3PP

그림 3은 실험에 사용한 4가지 시점 디자인을 보여준다. 사용자 자신의 실제 시점에서 가이던스를 관찰할 수 있는 1인칭 시점 (1PP) [4][5][7][17], 사용자가 가이던스의 뒷모습을 보고 동작을 따라 할 수 있는 3인칭 시점 (3PP) [4][5], 가이던스의 앞모습을 거울 형태로 제공하는 거울 시점 (MPP) [8], 마지막으로 3PP에 측면 시점을 추가한 Multi-3PP [3]를 사용하였다.

#### 4.1.3 동작 오차 및 사용자 경험 측정 방법

##### 관절 위치 오차

사용자와 전문가 사이의 동작 오차 지표로 두 관절 글로벌 위치 데이터의 차이를 계산한 관절 위치 오차를 사용하였다 [2][7][8]. 피험자 데이터는 힙 관절을 중심으로 정규화해 가이던스 데이터와 같은 위치에서 시작하도록 조정했다. 두 위치 데이터 간의 오차는 모든 프레임에 대해서 각 관절의 3차원 유클리드 거리를 계산한 후 평균을 내었다.

##### Single Ease Question (SEQ)

Single Ease Question (SEQ)은 주어진 작업의 주관적 난이도를 평가하는 사후 설문지이다 [9][18]. 단일 질문으로 구성되어 피험자가 직관적으로 응답할 수 있는 동시에, 다수의 질문으로 이루어진 설문지와 유사한 평가 효과를 제공한다는 장점이 있다 [9]. 본 연구에서는 피험자가 가상환경 UI를 통해 “주어진 시점에서 동작을 따라 하기 얼마나 어려웠나요?” [2]라는 질문에 7-Likert Scale (1: 전혀 어렵지 않음, 7: 매우 어려움)을 이용해 답하는 방법으로 결과를 수집하였다.

##### 시선고정 지표를 사용한 인지 부하 측정

시선 추적 데이터는 시각 자극이 제공되는 작업을 수행할 때, 인지 부하를 측정할 수 있는 대표적인 생체신호 지표

이다 [10]. 그중 시선고정(fixation) 지표는 특정 임계 시간 이상 시선이 한곳에 머물러 있는 순간을 의미한다 [11]. 인지 부하와 유의미한 상관관계를 가진 시선고정 지표로 시선고정 횟수 [14], 시선고정 평균 지속시간 [12][13], 시선고정 최대 지속시간 [13]을 선정하였다.

시선고정 지표를 수집하는 방법으로는 Meta Quest Pro의 시선 추적 기능을 활용하였다. 이를 이용해 가이던스가 위치한 평면과 시선이 교차하는 지점의 3D 위치 데이터를 추출하고, 평면의 중앙을 원점으로 2D 위치 좌표를 계산했다. 그리고 Dispersion-Threshold Identification (I-DT) 알고리즘을 적용해 2D 시선 위치 시계열 데이터에서 시선고정을 검출했다 [11]. I-DT 알고리즘은 최소 시선고정 지속 시간과 시선 분산 임계값을 입력 파라미터로 가진다. 본 시스템에선 이전 연구에서 제안한 최적값을 참고하여 각각 100ms와  $1^\circ$ 로 설정하였다 [11]. 다만 시선 분산 임계값은 데이터의 단위(m)를 고려해 0.05m로 변환하였다.

#### 4.2 실험 진행

##### 4.2.1 참가자

실험에는 보행 및 동작 수행에 문제가 없는 피험자 8명 (남자: 6, 여자: 2)이 참여했다. 피험자의 연령은 19세에서 25세 사이 ( $M: 23.625$ ,  $SD: 2$ )였다. 피험자 중 2명은 VR을 전혀 사용해 본 적이 없었고, 4명은 1년 미만의 사용 경험, 나머지 두 명은 1년 이상의 사용 경험을 가지고 있었다. 모든 참가자는 실험에서 사용된 태극권 동작에 대한 사전 경험이 없었다.

##### 4.2.2 절차

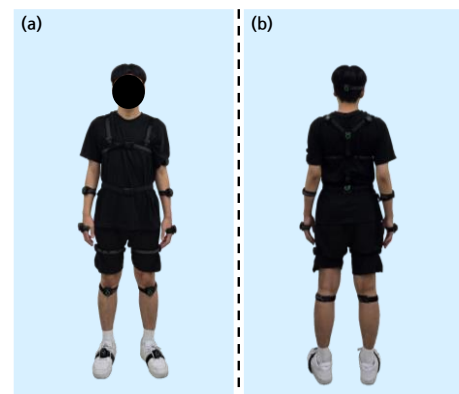


그림 4. PNS 센서 착용 모습. (a) 앞, (b) 뒤

참가자들은 실험 내용을 소개받고, PNS 센서를 착용했다 (그림 4). 그리고 센서의 측정 오차를 최소화하는 과정인 센서 캘리브레이션 [19]과 참가자의 신체 크기를 고려해 가이던스의 크기를 조정하는 스케일링 과정을 진행했다. 스케일링에서는 양팔을 옆으로 들어 올리는 T-pose를

통해 참가자의 신체 크기를 수집하고, 이를 바탕으로 가이던스의 각 관절 사이의 거리를 계층적으로 조정했다 [11].

실험은 두 번의 동작 관찰과 각 시점 디자인을 적용한 연습 수행 1 번, 실전 수행 3 번으로 구성된다. 시점의 순서는 balanced Latin Square 로 counterbalancing 했다. 한 조건이 끝날 때마다 컨트롤러를 이용해 가상환경에 UI 형태로 제공되는 SEQ 설문에 답한다. 이를 각 시점 조건에 대해 총 4 번 반복한다.

모든 동작 수행을 마무리한 후 사후 인터뷰를 진행하였다. 인터뷰는 약 15 분 간 진행되었고, 각 시점 디자인과 실험에 사용된 동작의 주관적인 선호도 및 난이도에 대한 의견을 수집하였다. 총 실험 시간으로 약 1 시간이 소요되었고, 참가자에게 소정의 참가비가 지급되었다.

### 4.3 결과

8 명의 참가자의 관절 위치 오차, SEQ, 시선고정 횟수, 평균 및 최대 시선고정 지속시간에 대한 통계 분석을 진행하였다. 모든 데이터에 대하여 Shapiro-Wilk test 를 수행한 결과 정규성을 보이지 않아 비모수적 통계 분석 방법인 Friedman test 를 수행하였다. 또한 통계적 유의성을 띠는 결과에 대해선 Conover's Post Hoc test 에 Holm 보정을 적용해 추가 분석을 진행하였다.

#### 4.3.1 관절 위치 오차

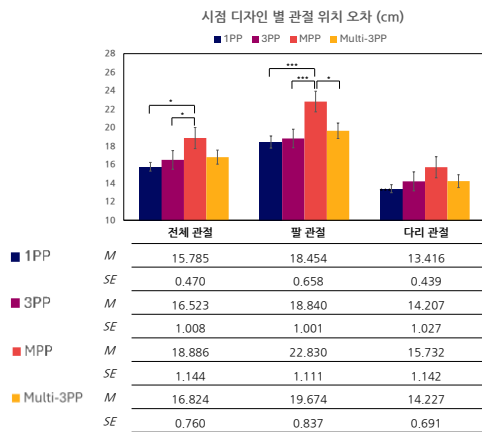


그림 5. 각 시점 디자인 별 관절 위치 오차 결과 (\*:  $p < 0.05$ , \*\*\*:  $p < 0.001$ )

각 시점 디자인 간 전체 관절 위치 오차를 Friedman test 를 통해 비교한 결과 유의미한 차이가 발견되었다 (그림 5) ( $\chi^2(2) = 10.050$ ,  $p = 0.018$ , Kendall's  $W = 0.140$ ). 또한, 팔 관절만을 비교했을 때도 유의미한 차이가 발견되었다 ( $\chi^2(2) = 19.950$ ,  $p < .001$ , Kendall's  $W = 0.277$ ). Conover's Post Hoc test 에 Holm 보정을 적용해 사후 검정한 결과, 전체 관절과 팔 관절 모두에서

MPP 시점과 나머지 시점 디자인 사이의 유의한 차이가 존재함을 확인했다. 반면, 다리 관절에 대한 시점별 관절 위치 오차는 유의미한 차이를 보이지 않았다 ( $\chi^2(2) = 4.2$ ,  $p = 0.241$ , Kendall's  $W = 0.058$ ). 이를 통해 MPP 시점을 사용했을 때, 다른 시점을 사용했을 때보다 동작을 정확히 따라 하기 더 어려웠다는 것을 알 수 있다.

#### 4.3.2 Single Ease Question (SEQ)

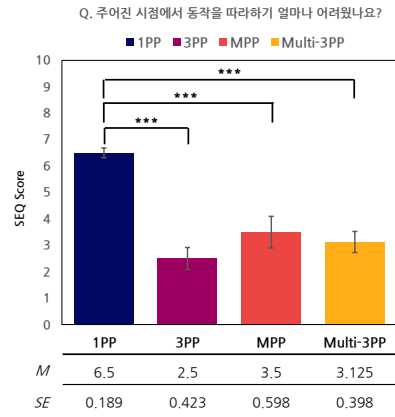


그림 6. 각 시점 디자인 별 SEQ 결과 (\*\*\*:  $p < 0.001$ )

각 시점 디자인 사이의 SEQ 결과를 Friedman test 를 통해 비교한 결과 유의미한 차이가 발견되었다 (그림 6) ( $\chi^2(2) = 17.080$ ,  $p < .001$ , Kendall's  $W = 0.712$ ). Conover's Post Hoc test 에 Holm 보정을 적용해 사후 검정한 결과 1PP 시점과 다른 시점 디자인 사이에 유의한 차이를 확인했다. 이를 통해 참가자들이 다른 조건보다 1PP 를 이용한 동작 수행에 더 많은 어려움을 느꼈다는 것을 확인할 수 있다.

#### 4.3.3 시선고정 지표를 이용한 인지 부하

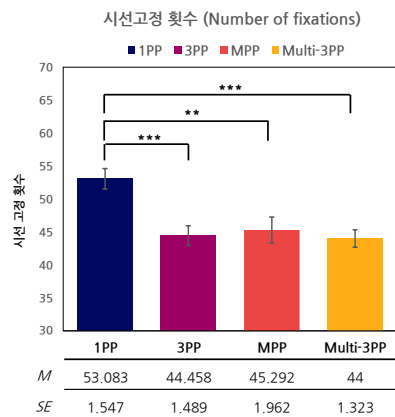


그림 7. 각 시점 디자인 별 시선고정 횟수 (\*\*:  $p < 0.01$ , \*\*\*:  $p < 0.001$ )

각 시점 디자인을 사용했을 때의 시선고정 횟수를 Friedman test 를 통해 비교한 결과 유의미한 차이가 발

견되었다 (그림 7) ( $\chi^2(2) = 20.496, p < .001, Kendall's W = 0.285$ ). Conover's Post Hoc test 에 Holm 보정을 적용해 사후 검정한 결과, 1PP 시점과 다른 시점 디자인 사이의 유의한 차이를 관측하였다.

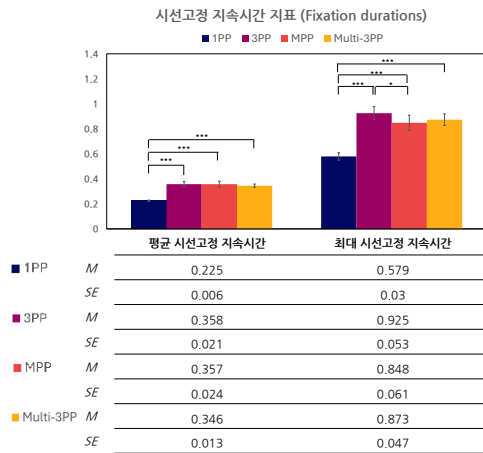


그림 8. 각 시점 디자인 별 시선고정 지속시간  
(\*:  $p < 0.05$ , \*\*:  $p < 0.001$ )

한편, 각 시점 디자인의 평균 시선고정 지속시간을 Friedman test 를 통해 비교한 결과 유의미한 차이가 발견되었다 (그림 8) ( $\chi^2(2) = 43.650, p < .001, Kendall's W = 0.606$ ). 최대 시선고정 지속시간에서도 마찬가지로 유의미한 차이가 발견되었다 ( $\chi^2(2) = 33.650, p < .001, Kendall's W = 0.467$ ). 두 지표에 대해 Conover's Post Hoc test 에 Holm 보정을 적용해 사후 검정한 결과, 평균과 최대 시선고정 지속시간 모두 1PP 시점과 다른 시점 디자인 사이의 유의한 차이가 있음을 확인했다. 추가로, 최대 시선고정 지속시간에서는 3PP와 MPP 조건이 유의미한 차이를 보였다.

이전 연구를 통해 인지 부하가 시선고정 횟수 지표와는 양의 상관관계를 가지고 [14], 시선고정 지속시간 지표들과는 음의 상관관계를 가진다는 점을 알 수 있다 [12][13]. 즉, 결과는 1PP 시점이 다른 시점에 비해 유의미하게 높은 인지 부하를 발생시킨다는 점을 보여준다.

## 5 고찰

먼저 관절 위치 오차에서 거울 시점 (MPP)이 다른 시점 디자인에 비해 유의미하게 높은 동작 오차를 보인 것을 확인할 수 있다. 이는 MPP 시점이 앞뒤 방향, 즉 z 축 방향으로 반전된 모습을 보여주는 것이 원인일 수 있다. 태극권 동작처럼 가동 범위가 넓고 많은 속도 전환이 발생하는 동작에 대해선 이러한 좌표계 변환이 불편함을 느끼게 만들 수 있다 [1]. 또한 사후 인터뷰에서 절반의 피험자로부터 앞뒤가 반전된 모습으로 인해 혼란스러움을 느꼈다는 답변을 받았다 (P1, P2, P3, P7).

혹은 다른 시점 디자인에 비해 MPP 에서 더 낮은 체화 감각 (Sense of Embodiment)을 느꼈을 가능성도 존재한다. Döllinger et al. (2023) [16]은 가상환경에서 3PP 와 MPP 시점으로 스쿼트 등 신체 인식 운동을 수행했을 때 느끼는 체화 감각의 차이를 관찰했을 때, 사람들이 MPP 에서 더 낮은 체화 감각을 느끼는 것을 확인했다. 이를 근거로, VR 속 아바타에게 더 낮은 체화 감각을 느낀 피험자들이 신체를 제어하는 것에 어려움을 느꼈고, 이것이 더 높은 동작 오차를 발생시켰다고 추측할 수 있다. 후속 연구에서 체화 감각 관련 정량 지표를 추가로 측정하여 시점 디자인과 체화 감각 간의 더 명확한 상관관계를 확인할 것을 제안한다.

다음으로, 비록 유의미하진 않지만 1 인칭 시점 (1PP)이 다른 시점에 비해 낮은 동작 오차를 가지고 있는 데 반해, 사용자 경험 지표에서 유의미하게 부정적인 결과가 도출되었다. 동작 오차의 경우, 1PP 가 다른 시점에 비해 더 많은 동작 깊이 정보를 얻을 수 있다는 점이 동작 오차를 줄이는 데 도움이 되었을 것이다 [2]. 반면, 동작 정보가 한 눈에 들어오지 않는다는 점이 사용자 경험을 저하하는 요소로 작용한 것으로 보인다 [2]. 사후 인터뷰에서도 모든 참가자가 1PP 에 대해 전신 움직임을 한 번에 확인할 수 없고, 특히 다리 움직임을 보기 불편했다고 답변했다 (P1~8). 또한, 좁은 시야로 인해 양팔을 번갈아 봐야 했고, 이 과정에서 높은 인지 부하가 발생했을 것으로 추측한다.

마지막으로 Multi-3PP 조건에서 측면 시점이라는 추가적인 동작 정보를 제공했음에도 동작 수행 능력과 사용자 경험 지표 모두에서 3PP 와 유의미한 차이가 발견되지 않았다 [3]. 사후 인터뷰를 통해 이것이 기존 연구에 비해 더욱 복잡해진 동작 때문이라는 것을 알 수 있었다. 절반의 피험자로부터 동작이 복잡하고 생소해서 추가적인 정보를 참고하기 어려웠다고 답변했다 (P3, P4, P5, P8). 그러나 대다수의 피험자가 익숙한 동작을 학습할 시 Multi-3PP 시점을 더 선호한다고 밝힌 점을 고려하면, 동작의 복잡성보다 생소함이 결과에 더 영향을 미쳤을 것으로 추측할 수 있다. 따라서, 복잡하지만 익숙한 동작을 사용했을 때 Multi-3PP 와 3PP 시점 사이의 동작 정확도 및 사용자 경험 지표에서 변화가 발생하는지 추후 연구를 통해 확인할 것을 제안한다.

## 6 한계 및 결론

본 연구는 현실 세계 동작을 위한 XR 기반 모션 가이드 시스템의 사용자 시점 디자인이 동작 수행 능력과 사용자 경험에 미치는 영향을 분석하였다. 1 인칭, 3 인칭, 거울 시점, Multi-3PP 등 다양한 시점 디자인을 비교하여 그



영향을 분석한 결과, 다음과 같은 경향성을 확인할 수 있었다. 먼저 거울 시점(MPP)은 다른 시점에 비해 높은 동작 오차를 보였는데, 이는 좌표계 반전과 낮은 체화 감각 때문으로 추정된다. 1 인칭 시점(1PP)은 높은 인지 부하와 주관적 난이도를 보였다. 이는 사후 인터뷰를 통해 사용자의 시야에 모든 신체 부위가 포함되지 않기 때문으로 드러났다. 한편, Multi-3PP 시점은 추가적인 측면 정보를 제공했음에도 동작 수행 능력과 사용자 경험에서 3 인칭 시점(3PP)과 유의미한 차이를 보이지 않았다. 사후 인터뷰를 통해, 복잡하고 생소한 동작이 추가 정보를 활용하기 어렵게 만들었다는 것이 드러났다. 이러한 결과는 현실 세계 동작을 위한 XR 기반 시스템 설계에서도 시점 디자인이 사용자 경험과 동작 수행에 중대한 영향을 미침을 보여준다.

다만 연구를 진행하는 과정에서 다음과 같은 한계점을 발견할 수 있었다. 먼저 전신을 사용하는 동작을 선정했으나, 팔에 비해 다리의 참여가 크게 요구되지 않았다는 점이다. 실제로 그림 5를 통해 다리 관절 동작 오차가 팔 관절 오차에 비해 작을 뿐만 아니라, 시점 디자인 간 유의미한 차이를 보이지 않았다는 것을 알 수 있다. 이로 인해 팔 동작만을 사용한 기존 연구의 1PP 시점 결과와 본 연구의 결과 사이의 유의미한 차이가 발생하지 않았다고 추정할 수 있다. 따라서, 후속 연구를 통해 다리를 주로 사용하는 동작에 대한 시점 디자인 별 동작 오차는 팔을 사용했을 때와 어떠한 차이가 발생하는지 탐구할 것을 제안한다.

또, 동작 시각화 방법으로 스켈레톤 모델을 사용했다는 점도 한계로 지적된다. 스켈레톤 형태의 동작 시각화는 이전 XR 기반 모션 가이드 시스템 연구에서 종종 사용되었지만 [4][5][6][8], 스켈레톤처럼 추상적인 형태의 시각화는 동작을 정확히 따라하는 데 방해가 될 수 있다는 연구 결과가 존재한다 [7]. 따라서 후속 연구에서는 휴머노이드 [19]나 SMPL 메쉬 모델 [20]같은 인간 형태의 모델을 활용하는 방향을 고려해야 한다.

마지막으로 전문가 동작 데이터를 사용자의 신체 크기에 맞춰 스케일링을 할 때 전문가 동작의 각 관절 사이 길이를 계층적으로 조정하는 방식을 사용하였다. 이는 두 동작을 같은 크기로 보정하는 가장 간단한 방법이지만, 관절 사이 길이가 급격히 조정될 경우 동작이 왜곡될 가능성이 존재한다. 최근 여러 연구에서 AI 모델을 활용해 두 동작 데이터의 중간 동작을 생성하는 방법이 제시되었으므로 [21][22][23], 후속 연구에서는 이를 활용해 더욱 정확한 가이드를 제공할 것을 제안한다.

## 사사의 글

본 논문은 2025 년도 광주과학기술원 GIST-MIT 공동연구사업과 연구개발특구진흥재단의 '지역의 미래를 여는 과학기술 프로젝트' 프로그램(사업번호: 2025-DD-UP-0312)과 한국연구재단을 통해 과학기술정보통신부의 RS-2024-00343397 사업으로부터 지원받아 수행되었음.

## 참고 문헌

1. Yu, X., Lee, B., & Sedlmair, M. (2024, May). Design Space of Visual Feedforward And Corrective Feedback in XR-Based Motion Guidance Systems. In Proceedings of the CHI Conference on Human Factors in Computing Systems (pp. 1-15).
2. Yu, X., Angerbauer, K., Mohr, P., Kalkofen, D., & Sedlmair, M. (2020, November). Perspective matters: Design implications for motion guidance in mixed reality. In 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 577-587). IEEE.
3. Elsayed, H., Kartono, K., Schön, D., Schmitz, M., Mühlhäuser, M., & Weigel, M. (2022, November). Understanding Perspectives for Single-and Multi-Limb Movement Guidance in Virtual 3D Environments. In Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology (pp. 1-10).
4. Hoang, T. N., Reinoso, M., Vetere, F., & Tanin, E. (2016, October). Onebody: remote posture guidance system using first person view in virtual environment. In Proceedings of the 9th Nordic Conference on Human-Computer Interaction (pp. 1-10).
5. Ma, D., Hu, X., Shi, J., Patel, M., Jain, R., Liu, Z., ... & Ramani, K. (2024, October). avaTTAR: Table Tennis Stroke Training with Embodied and Detached Visualization in Augmented Reality. In Proceedings of the 37th Annual ACM Symposium on User Interface Software and Technology (pp. 1-16).
6. Xu, L., Wang, Q., Lin, X., Yuan, L., & Ma, X. (2023). Skeleton-based tai chi action segmentation using trajectory primitives and content. Neural Computing and Applications, 35(13), 9549-9566.
7. Dürr, M., Weber, R., Pfeil, U., & Reiterer, H. (2020, February). EGuide: Investigating different visual appearances and guidance techniques for egocentric guidance visualizations. In Proceedings

of the Fourteenth International Conference on Tangible, Embedded, and Embodied Interaction (pp. 311-322).

8. Anderson, F., Grossman, T., Matejka, J., & Fitzmaurice, G. (2013, October). YouMove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th annual ACM symposium on User interface software and technology* (pp. 311-320).
9. Sauro, Jeff, and Joseph S. Dumas. "Comparison of three one-question, post-task usability questionnaires." *Proceedings of the SIGCHI conference on human factors in computing systems*. 2009.
10. Skaramagkas, V., Giannakakis, G., Ktistakis, E., Manousos, D., Karatzanis, I., Tachos, N. S., ... & Tsiknakis, M. (2021). Review of eye tracking metrics involved in emotional and cognitive processes. *IEEE Reviews in Biomedical Engineering*, 16, 260-277.
11. Blignaut, P. (2009). Fixation identification: The optimum threshold for a dispersion algorithm. *Attention, Perception, & Psychophysics*, 71, 881-895.
12. De Rivecourt, M., Kuperus, M. N., Post, W. J., & Mulder, L. J. M. (2008). Cardiovascular and eye activity measures as indices for momentary changes in mental effort during simulated flight. *Ergonomics*, 51(9), 1295-1319.
13. Borys, M., Tokovarov, M., Wawrzyk, M., Wesołowska, K., Plechawska-Wójcik, M., Dmytruk, R., & Kaczorowska, M. (2017, March). An analysis of eye-tracking and electroencephalography data for cognitive load measurement during arithmetic tasks. In *2017 10th International Symposium on Advanced Topics in Electrical Engineering (ATEE)* (pp. 287-292). IEEE.
14. Menekse Dalveren, G. G., & Cagiltay, N. E. (2018). Insights from surgeons' eye-movement data in a virtual simulation surgical training environment: effect of experience level and hand conditions. *Behaviour & Information Technology*, 37(5), 517-537.
15. Seong, M., Kim, G., Yeo, D., Kang, Y., Yang, H., DelPreto, J., ... & Kim, S. (2024). MultiSenseBadminton: Wearable Sensor-Based Biomechanical Dataset for Evaluation of Badminton Performance. *Scientific Data*, 11(1), 343.
16. Döllinger, N., Wolf, E., Botsch, M., Latoschik, M. E., & Wienrich, C. (2023, April). Are embodied avatars harmful to our self-experience? the impact of virtual embodiment on body awareness. In *Proceedings of the 2023 CHI conference on human factors in computing systems* (pp. 1-14).
17. Han, P. H., Chen, K. W., Hsieh, C. H., Huang, Y. J., & Hung, Y. P. (2016, February). Ar-arm: Augmented visualization for guiding arm movement in the first-person perspective. In *Proceedings of the 7th Augmented Human International Conference 2016* (pp. 1-4).
18. Piumsomboon, T., Lee, G. A., Hart, J. D., Ens, B., Lindeman, R. W., Thomas, B. H., & Billinghamurst, M. (2018, April). Mini-me: An adaptive avatar for mixed reality remote collaboration. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1-13).
19. Hwang, S., Oh, J., Kang, S., Seong, M., Elsharkawy, A. I. A. M., & Kim, S. (2024, May). Ergopulse: Electrifying your lower body with biomechanical simulation-based electrical muscle stimulation haptic system in virtual reality. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1-21).
20. Ikeda, A., Hwang, D. H., Koike, H., Bruder, G., Yoshimoto, S., & Cobb, S. (2018). AR based Self-sports Learning System using Decayed Dynamic Time Warping Algorithm. In *ICAT-EGVE* (pp. 171-174).
21. 장준혁, 성민우, & 김승준. (2024). 딥러닝 기반 인코더-디코더 아키텍처를 이용한 배드민턴 스윙 자세 분석과 시각화. *한국정보과학회 학술발표논문집*, 2370-2372.
22. Seong, M., Kim, G., Kang, Y., Jang, J., DelPreto, J., & Kim, S. (2024). Counterfactual Explanation-Based Badminton Motion Guidance Generation Using Wearable Sensors. *arXiv preprint arXiv:2405.11802*.
23. Liao, C. C., Hwang, D. H., & Koike, H. (2022). Ai golf: Golf swing analysis tool for self-training. *IEEE Access*, 10, 106286-106295.