

손 랜드마크 추적 기반 디지털 그림자극 조종 인터페이스 연구: 《손끝의 예술》 제작 사례를 중심으로

조 홀¹ · 김 혜 경^{2*}

¹경희대학교 디지털콘텐츠학과 석사과정

²경희대학교 디지털콘텐츠학과 교수

A Digital Shadow Puppetry Control Interface Based on Hand Landmark Tracking: A Production Case Study of *Art at Your Fingertips*

Yi Zhao¹ · Hye-Kyung Kim^{2*}

¹Master's Course, Department of Digital Contents, Kyung Hee University, Yongin 17104, Korea

²Professor, Department of Digital Contents, Kyung-Hee University, Yong-In 17104, Korea

[요 약]

본 연구는 손 랜드마크 추적 기반 디지털 그림자극 조종 인터페이스를 구현하고, 좌표 처리 방식에 따른 성능 차이를 동일 조건에서 정량적으로 비교하는 데 목적이 있다. 비교 대상은 원시 좌표 직접 매핑 방식(M0), 1€ 필터 방식(M1), 좌표 평활화·데드 존·상태 유지 및 출력 범위 제한을 적용한 제안 방식(M2)으로 설정하였다. 동일한 입력 영상, 하드웨어, 제스처 종류 및 디지털 인형 조건에서 제스처 인식 정확도, 중단 간 지연 시간, 정지 상태 좌표 흔들림, 제어 중단 빈도와 FPS를 측정하였다. 실험 결과, 전체 제스처 인식 정확도는 M0 87.1%, M1 91.5%, M2 93.0%로 나타났다. M2의 좌표 흔들림은 M0보다 64.0%, M1보다 29.9% 감소하였으며, 제어 중단 빈도는 M1보다 56.6% 낮았다. 반면 평균 지연 시간은 122.6ms, 평균 FPS는 46.8 FPS로 나타나 M1보다 반응 속도와 처리 성능이 낮았다. 이를 통해 제안 방식은 좌표 안정성과 제어 연속성을 향상시키지만, 반응 지연이 증가하는 성능 특성을 갖는 것으로 확인되었다.

[Abstract]

This study implemented a digital shadow puppetry control interface based on hand landmarks and quantitatively compared three coordinate-processing methods under identical conditions: raw coordinate mapping (M0), the 1€ filter (M1), and the proposed method (M2), which combines coordinate smoothing, a dead zone, state holding, and output-range constraints. Gesture-recognition accuracy, end-to-end latency, static-coordinate jitter, control interruptions, and frame rate were measured using the same input videos, hardware, gesture set, and digital puppet model. Overall recognition accuracy was 87.1% for M0, 91.5% for M1, and 93.0% for M2. Coordinate jitter under M2 was 64.0% lower than under M0 and 30.3% lower than under M1, while control interruptions were 56.6% fewer than under M1. M2, however, showed an average latency of 122.6 ms and an average frame rate of 46.8 FPS, indicating lower responsiveness and processing performance than M1. The results demonstrate a trade-off between control stability and response speed.

색인어 : 그림자극, 손 랜드마크 추적, 디지털 그림자 인형, 조종 인터페이스, 비접촉 인터랙션

Keyword : Shadow Puppetry, Hand Landmark Tracking, Digital Shadow Puppet, Control Interface, Contactless Interaction

<http://dx.doi.org/10.9728/dcs.2026.27.7.1987>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 07 June 2026; **Revised** 22 June 2026

Accepted 13 July 2026

***Corresponding Author; Hye-Kyung Kim**

Tel: +82-31-201-2675

E-mail: hkkim@khu.ac.kr

1. 서 론

1-1 연구 배경 및 문제 제기

중국 전통 그림자극은 빛, 막, 그림자 인형과 조종자의 손 기술이 결합되어 완성되는 공연예술이다[1],[2]. 그림자 인형의 움직임은 조종자의 손동작과 조종 막대, 인형의 관절 구조 사이의 물리적 연결을 통해 형성된다. 그러나 박물관이나 전시 공간에서는 그림자 인형, 제작 도구, 사진 및 영상 자료를 중심으로 그림자극을 소개하는 경우가 많아 관람객이 이러한 조종 구조를 직접 체험하기 어렵다.

카메라 기반 손 추적 기술은 별도의 장갑이나 물리적 컨트롤러 없이 사용자의 손 위치와 손가락 움직임을 실시간 입력 값으로 활용할 수 있다는 점에서 전시형 인터페이스에 적용할 수 있다. 특히 손 랜드마크 데이터를 디지털 인형의 위치와 관절 회전에 연결하면 사용자의 손동작을 그림자 인형의 움직임으로 변환할 수 있다. 다만 카메라에서 추출된 손 좌표를 그대로 사용할 경우 미세한 좌표 흔들림, 손가락 중첩, 빠른 동작에 따른 추적 손실 및 반응 지연이 발생할 수 있다.

따라서 손 랜드마크 기반 조종 인터페이스의 유효성을 검증하기 위해서는 시스템의 단독 성능만을 제시하기보다 기존 처리 방식과 제안 방식을 동일한 입력 및 하드웨어 조건에서 비교할 필요가 있다. 이에 본 연구는 손 랜드마크 좌표의 직접 매핑 방식, 선행연구에서 제안된 입력 필터링 방식과 본 연구의 데이터 처리 방식을 동일 조건에서 실험하고, 제스처 인식 정확도, 중단 간 지연 시간, 정지 상태 좌표 흔들림, 추적 손실 빈도 및 프레임 처리 성능을 정량적으로 비교한다.

1-2 선행연구 검토 및 연구 차별성

그림자극과 디지털 문화유산에 관한 연구에서는 전통문화 요소를 디지털 스토리텔링이나 인터랙티브 콘텐츠로 재구성하려는 시도가 이루어져 왔다. Lu et al.은 그림자극을 기반으로 한 협업형 디지털 스토리텔링 시스템을 제안하고 사용자 참여와 창작 경험을 분석하였다[3]. 이러한 연구는 그림자극의 문화적 활용 범위를 확장하였다는 의의가 있으나, 손동작과 인형 관절 사이의 연속적인 조종 관계나 좌표 처리 방식에 따른 시스템 성능 차이는 구체적으로 다루지 않았다.

손 제스처 기반 인터페이스 연구에서는 카메라 영상에서 손의 위치, 형태와 움직임을 인식하고 이를 포인터 이동, 화면 선택 및 명령 실행에 활용해 왔다[4]. MediaPipe Hands는 단일 RGB 카메라 영상에서 손의 주요 랜드마크를 실시간으로 추출할 수 있는 구조를 제시하였으며[5], Xu는 카메라 기반 제스처 인식 결과를 컴퓨터 제어 명령에 연결하는 실시간 HCI 시스템을 구현하였다[6]. 그러나 이러한 연구는 일반적인 컴퓨터 명령 수행을 주요 목적으로 하므로, 연속적인 손 좌표를 디지털 인형의 관절 움직임에 연결할 때 발생하는 조

종 안정성과 반응성의 문제를 직접적으로 설명하기에는 한계가 있다.

연속 입력 장치에서 좌표 흔들림을 줄이기 위해 저역통과 필터와 속도 적응형 필터가 활용되어 왔다. Casiez et al.은 입력 속도에 따라 필터의 차단 주파수를 조절하는 1€ 필터를 제안하여 정밀도와 반응성 사이의 균형을 개선하고자 하였다[7]. 본 연구는 이를 비교 대상에 포함하고, 원시 손 랜드마크를 직접 매핑하는 방식, 1€ 필터를 적용하는 방식, TouchDesigner의 좌표 평활화-데드존-출력 범위 제한을 결합한 제안 방식을 동일 조건에서 비교한다. 비교 대상과 통제 조건은 표 1과 같다.

표 1. 동일 조건 비교를 위한 손 랜드마크 처리 방식

Table 1. Hand landmark processing methods for comparison under identical conditions

Method	Data Processing	Gesture Decision	Controlled Conditions	Evaluation Metrics
M0: Raw Mapping	Direct mapping of raw MediaPipe coordinates	Fixed threshold	Same input video, hardware, gesture set, and puppet model	Accuracy, latency, jitter, tracking loss, FPS
M1: 1€ Filter	Speed-adaptive low-pass filtering	Same fixed threshold	Same input video, hardware, gesture set, and puppet model	Accuracy, latency, jitter, tracking loss, FPS
M2: Proposed Method	Coordinate smoothing, dead zone, and output-range constraints	Same threshold with state holding	Same input video, hardware, gesture set, and puppet model	Accuracy, latency, jitter, tracking loss, FPS

표 1의 세 방식은 손 랜드마크 추출 모델, 입력 영상, 하드웨어, 제스처 종류, 판정 기준 및 디지털 인형 모델을 동일하게 유지하고 좌표 처리 단계만 다르게 설정한다. 이를 통해 방식별 성능 차이가 입력 조건이나 실험 환경이 아니라 좌표 처리 방법에서 발생한 것인지 확인하고자 한다.

1-3 연구 목적 및 방법

본 연구의 목적은 중국 전통 그림자극의 조종 구조를 디지털 전시 환경에서 재구성하기 위해 손 랜드마크 추적 기반 디지털 그림자극 조종 인터페이스 《손끝의 예술》을 구현하고, 좌표 처리 방식에 따른 조종 성능을 객관적으로 비교하는 데 있다. 이를 위한 연구 문제는 다음과 같다.

첫째, 제안 방식은 원시 좌표 직접 매핑 방식과 1€ 필터 방식에 비해 제스처 인식 정확도와 정지 상태 좌표 안정성을 향상시키는가?

둘째, 좌표 평활화와 데드존 적용은 중단 간 지연 시간과 프레임 처리 성능에 어떠한 영향을 미치는가?

셋째, 조명과 사용자 거리 조건이 변화할 때 세 방식의 인식 정확도와 추적 손실에는 어떠한 차이가 나타나는가?

연구는 선행연구 검토, 인터페이스 설계 및 구현, 동일 조건 비교 실험, 결과 분석의 순서로 진행하였다. 먼저 그림자극의 조종 구조와 손 제스처 기반 인터페이스 및 입력 좌표 필터링 연구를 검토하였다. 이후 Blender에서 머리, 몸통, 팔, 손과 도구를 독립적인 요소로 구성하고 관절 회전 중심을 설정하였다. MediaPipe를 통해 손목, 엄지 끝점, 검지 관절과 끝점의 좌표를 추출하였으며, TouchDesigner에서 이를 디지털 인형의 위치 이동, 팔 회전, 집기 동작 및 장면 트리거에 연결하였다.

비교 실험에서는 동일한 손동작 영상을 원시 좌표 직접 매핑 방식, 1€ 필터 방식과 제안 방식에 각각 입력하였다. 모든 방식에 동일한 입력 영상, 컴퓨터, 카메라, 손 랜드마크 추출 모델, 제스처 판정 기준 및 디지털 인형을 적용하고 좌표 처리 방식만 변경하였다. 평가 지표는 제스처 인식 정확도, 중단 간 지연 시간, 정지 상태 좌표 흔들림, 추적 손실 빈도, 평균 및 최소 FPS로 설정하였다. 또한 조도와 카메라 거리를 단계별로 변경하여 환경 조건에 따른 성능 차이를 확인하였다.

실험 결과는 방식별 평균과 표준편차를 산출하고 동일 입력에 대한 대응 비교를 통해 분석한다. 비교 결과에서는 제안 방식의 우수한 지표뿐 아니라 지연 증가나 프레임 성능 저하와 같이 상대적으로 열위한 결과도 함께 제시한다. 이를 바탕으로 디지털 그림자 인형 조종에서 좌표 안정성과 반응성 사이의 관계를 분석하고, 실제 전시 환경에 적용하기 위한 처리 방식의 조건과 한계를 도출한다.

II. 디지털 그림자극 조종 인터페이스 설계

2-1 시스템 구성 및 디지털 인형 제작

본 연구의 디지털 그림자극 조종 인터페이스는 카메라 입력, 손 랜드마크 추출, 좌표 처리, 디지털 인형 제어 및 시청각 출력의 단계로 구성하였다. 카메라에서 입력된 영상은 MediaPipe Hands를 통해 손 랜드마크 데이터로 변환되며 [5], 추출된 좌표는 TouchDesigner에서 평활화, 범위 변환 및 제스처 판정 과정을 거친다. 처리된 데이터는 Blender로 제작한 디지털 그림자 인형의 위치와 관절 회전값에 연결되고, 최종 결과는 그림자 무대 화면에 실시간으로 출력된다.

시스템 제작에는 Blender, MediaPipe 및 TouchDesigner를 사용하였다. Blender는 디지털 그림자 인형의 부위별 이미지 구성과 관절 중심 설정에 활용하였으며, MediaPipe는 카메라 영상에서 손목과 손가락 관절 좌표를 추출하는 데 사용하였다. TouchDesigner는 추출된 좌표를 처리하고 이를 인형의 이동, 팔 회전, 도구 조작 및 장면 트리거에 연결하는 역할을 담당하였다.



그림 1. 전통 그림자극의 조종 구조

Fig. 1. Control structure of traditional shadow puppetry

디지털 그림자 인형은 전통 그림자 인형의 평면적 관절 구조를 반영하기 위해 머리, 몸통, 팔, 손, 다리 및 도구를 각각 독립된 레이어로 분리하였다. 각 부위에는 회전 중심점을 설정하여 손 랜드마크 데이터에 따라 개별적으로 움직일 수 있도록 구성하였다. 특히 관람객의 손동작과 직접 연결되는 팔과 도구에는 어깨, 팔꿈치 및 손목에 해당하는 회전축을 설정하였다. 과도한 회전으로 인형의 형태가 변형되는 것을 방지하기 위해 팔의 회전 범위는 -45° 에서 60° 로 제한하고, 도구의 이동과 회전값도 화면상의 조종 범위 안에서 작동하도록 설정하였다.

본 연구에서는 실제 그림자 인형의 모든 관절 움직임을 그대로 복제하기보다 손동작과 인형 반응 사이의 주요 대응 관계를 구현하는 데 중점을 두었다. 이에 따라 디지털 인형의 기본 조종 기능은 전체 위치 이동, 팔 회전, 도구 조작 및 장면 트리거로 제한하였다. 이러한 구조는 이후 비교 실험에서 좌표 처리 방식에 따른 인형 조종 안정성과 반응 성능을 동일하게 측정하기 위한 공통 출력 조건으로 사용하였다.

2-2 손 랜드마크 데이터 처리 및 조종 매핑

손 랜드마크 데이터는 MediaPipe Hands에서 제공하는 21개 좌표 가운데 조종 기능과 직접 관련된 손목(wrist), 엄지 끝점(thumb tip), 검지 관절(index MCP), 검지 끝점(index tip) 및 손가락 끝점 좌표를 중심으로 추출하였다. 손목의 x, y 좌표는 디지털 인형의 전체 위치 이동에 사용하였고, 검지 관절과 검지 끝점을 연결한 벡터의 방향은 인형 팔의 회전값으로 변환하였다. 엄지 끝점과 검지 끝점 사이의 거리는 집기 동작을 판정하는 데 활용하였으며, 손바닥과 손가락 끝점 사이의 거리는 손바닥의 개폐 상태를 구분하는 데 사용하였다.

MediaPipe에서 추출되는 손 좌표는 0에서 1 사이로 정규화되어 있다. 본 연구에서는 TouchDesigner의 Math CHOP을 사용하여 정규화된 좌표를 디지털 그림자 무대의 출력 범위에 맞게 변환하였다[8],[9]. 손목의 x 좌표는 인형의 좌우 이동값으로 연결하고, y 좌표는 카메라 좌표계와 화면 좌표계

의 방향 차이를 고려하여 반전한 뒤 인형의 상하 이동값으로 적용하였다. 또한 인형이 설정된 무대 영역을 벗어나지 않도록 좌표의 최소값과 최대값을 제한하였다.

카메라 기반 손 추적 과정에서 발생하는 미세한 좌표 변화를 줄이기 위해 손목 위치와 검지 방향값에 Filter CHOP을 적용하였다. 일정 범위 이하의 작은 좌표 변화는 인형 움직임에 반영되지 않도록 데드존을 설정하였다. 이를 통해 사용자가 손을 정지한 상태에서도 디지털 인형이 불필요하게 흔들리는 현상을 줄이고자 하였다. 다만 평활화 강도가 높아 질수록 반응 지연이 증가할 수 있으므로, 해당 처리 방식의 안정성과 반응성은 제3장의 동일 조건 비교 실험을 통해 검증한다.

검지 방향은 index MCP와 index tip 사이의 x, y 좌표 차이를 이용하여 방향각을 계산하고, 이를 인형 팔의 회전 범위에 맞게 변환하였다. 엄지와 검지 사이의 집기 동작은 두 끝점 사이의 유클리드 거리를 기준으로 판정하였다. 본 시스템에서는 두 손끝 사이의 정규화 거리가 0.06 이하일 때 집기 동작으로 판단하고, 이를 인형 선택 또는 도구 조작의 트리거로 사용하였다. 손바닥 개폐 상태는 손바닥 중심과 각 손가락 끝점 사이의 평균 거리를 기준으로 구분하고, 장면 효과의 실행 조건으로 연결하였다.

손 랜드마크 입력값과 디지털 인형의 출력 동작 사이의 매핑 구조는 표 2와 같다.

표 2. 손 랜드마크 기반 조종 매핑 구조

Table 2. Hand landmark-based control mapping structure

Input Data	Calculation and Processing	TouchDesigner Processing	Output
Wrist Position	Extraction of normalized wrist x- and y-coordinates	Filter CHOP, dead zone, and coordinate conversion	Puppet position X/Y
Index Direction	Vector-angle calculation between index MCP and index tip	Angle conversion and rotation-range constraint	Puppet arm rotation
Thumb-Index Distance	Euclidean distance between thumb tip and index tip	Pinch detection at a threshold below 0.06	Puppet selection or prop control
Palm Open/Close	Average distance between the palm center and fingertips	Open/close state detection	Scene or stage-effect trigger

표 2의 매핑 구조에서 손목 위치는 연속적인 위치 입력으로 처리하고, 검지 방향은 관절 회전값으로 변환하였다. 엄지와 검지 사이의 거리 및 손바닥 개폐 상태는 설정된 기준값에 따라 이산적인 제스처 트리거로 판정하였다. 따라서 본 시스템은 손 랜드마크 데이터를 단순한 화면 명령으로 사용하는 것이 아니라, 손의 위치, 방향 및 형태 변화를 디지털 그림자 인형의 위치와 관절 움직임에 대응시키는 방식으로 구성하였다.

본 연구의 비접촉식 조종 방식은 전통 그림자극의 조종 막

대가 제공하는 촉각적 저항이나 물리적 무게를 재현하지는 않는다. 따라서 본 연구의 구현 범위는 전통 그림자극의 물리적 조종감을 완전히 대체하는 것이 아니라, 조종자의 손동작과 인형 관절 반응 사이의 구조적 관계를 디지털 환경에서 재구성하는 데 한정한다.

III. 동일 조건 비교 실험 및 결과

3-1 동일 조건 비교 실험 및 결과

본 연구에서는 손 랜드마크 좌표 처리 방식에 따른 디지털 그림자 인형의 조종 성능 차이를 확인하기 위해 동일 조건 비교 실험을 구성하였다. 비교 대상은 MediaPipe에서 추출된 원시 좌표를 별도의 시간적 보정 없이 디지털 인형에 연결하는 직접 매핑 방식(M0), 입력 속도에 따라 필터 강도를 조절하는 1€ 필터 방식(M1), TouchDesigner의 좌표 평활화, 데드존, 상태 유지 및 출력 범위 제한을 결합한 제안 방식(M2)으로 설정하였다.

MediaPipe Hands는 단일 RGB 카메라 영상에서 손바닥을 검출한 뒤 21개의 손 랜드마크를 추정하는 구조로 구성된다[5]. 세 방식에는 동일한 MediaPipe 손 랜드마크 추출 결과를 입력하였으며, 입력 영상, 하드웨어, 제스처 종류, 제스처 판정 임계값, 디지털 그림자 인형 및 출력 환경을 동일하게 유지하였다. 좌표 처리 단계만 변경함으로써 방식별 성능 차이가 참여자의 동작 차이나 장비 조건이 아니라 데이터 처리 방법에서 발생한 것인지 확인하고자 하였다.

M0에서는 MediaPipe에서 출력된 손목과 손가락 좌표를 디지털 인형의 이동 및 관절 회전 범위에 직접 선형 매핑하였다. M1에서는 1€ 필터를 각 좌표 채널에 적용하였다. 1€ 필터는 입력 속도가 낮을 때 노이즈를 강하게 제거하고, 입력 속도가 높을 때 차단 주파수를 높여 반응 지연을 줄이는 방식이다[7]. 실험에서는 $\min_cutoff=1.0$, $\beta=0.007$, $d_cutoff=1.0$ 으로 설정하였다.

M2에서는 TouchDesigner의 Filter CHOP을 이용하여 손목 위치와 검지 방향 좌표를 평활화하고, 정규화 좌표 기준 0.008 이하의 미세한 변화를 제거하는 데드존을 적용하였다. 손 랜드마크가 일시적으로 누락된 경우에는 마지막으로 확인된 안정 좌표를 최대 5프레임 동안 유지하도록 설정하였다. 또한 Math CHOP을 사용하여 입력 좌표를 디지털 무대 범위로 변환하고, 인형 위치와 관절 회전이 설정된 범위를 벗어나지 않도록 최소값과 최대값을 제한하였다[8],[9].

실험에는 총 20명이 참여하였다. 참여자는 손목 이동, 검지 방향 조절, 엄지와 검지를 이용한 집기 동작, 손바닥 열기와 닫기의 네 가지 동작을 각각 10회 수행하였다. 이에 따라 참여자 1인당 40회, 한 방식당 총 800회의 제스처 입력을 분석하였다. 참여자는 본 실험 전에 각 제스처를 3회씩 연습하

였으며, 연습 과정에서 수집된 데이터는 실험 결과에 포함하지 않았다.

참여자의 손동작은 동일한 카메라와 촬영 조건에서 먼저 영상으로 기록하였다. 녹화된 동일 영상을 M0, M1 및 M2 처리 파이프라인에 각각 입력하여 세 방식이 동일한 영상 프레임임을 처리하도록 하였다. 이는 참여자가 세 방식을 반복 수행할 때 발생할 수 있는 손 위치, 동작 속도 및 동작 범위의 차이를 통제하기 위한 것이다.

실험 환경은 Intel Core i7-12700H CPU, NVIDIA GeForce RTX 3060 Laptop GPU 및 16GB RAM이 탑재된 노트북으로 설정하였다. 입력 장치는 Logitech C920 웹캠을 사용하였으며, 입력 해상도는 1280×720 pixels, 입력 프레임률은 30 FPS로 설정하였다. 소프트웨어는 MediaPipe 0.10.14, TouchDesigner 2025 및 Blender 3.6을 사용하였다.

기본 실험에서 카메라와 참여자의 손 사이 거리는 40cm, 실내 조도는 350lux로 유지하였다. 중단 간 지연 시간은 사용자의 실제 손동작과 화면 반응을 동시에 촬영할 수 있도록 중단 간 지연 시간은 <Apple iPhone 14 Pro>의 120 FPS 고속 촬영 기능을 사용하여 손동작과 화면 반응을 동시에 기록한 뒤 측정하였다.

조명 조건에 따른 성능 차이를 확인하기 위해 조도는 100lux, 350lux 및 700lux의 세 조건으로 설정하였다. 거리 조건은 카메라와 손 사이의 거리를 30cm, 40cm 및 50cm로 구분하였다. 조명 및 거리 조건별 실험에서는 참여자 20명이 네 가지 제스처를 각각 5회 수행하여 조건별 총 400회의 입력을 분석하였다. 세 방식에 동일한 입력 영상을 적용하였다.

표 3. 동일 조건 비교 실험 환경

Table 3. Experimental conditions for comparison under identical conditions

Experimental Variable	Condition
Number of Participants	20 participants
Gesture Types	Wrist movement, index direction, pinch, palm open/close
Repetitions per Gesture	10 trials
Total Trials per Method	800 trials
Input Method	Identical prerecorded gesture videos
Camera Resolution / Frame Rate	1280 × 720 pixels / 30 FPS
Standard Distance / Illumination	40cm / 350lux
Lighting Conditions	100lux / 350lux / 700lux
Distance Conditions	30cm / 40cm / 50cm
Compared Methods	M0 Raw Mapping / M1 1€ Filter / M2 Proposed Method
Controlled Variables	Hardware, input video, gesture set, threshold, landmark model, and puppet model

3-2 평가 지표 및 분석 방법

제스처 인식 정확도는 시스템이 의도한 제스처를 정확하게 판정한 횟수를 전체 제스처 시도 횟수로 나누어 산출하였다. 전체 정확도뿐 아니라 손목 이동, 검지 방향, 집기 및 손바닥 개폐 동작별 정확도를 함께 계산하였다.

$$[Accuracy(\%) = \frac{N_{\text{correct}}}{N_{\text{total}}} \times 100]$$

중단 간 지연 시간은 참여자의 손동작이 시작된 시점과 디지털 그림자 인형의 반응이 화면에 처음 나타난 시점 사이의 시간 차이로 정의하였다. 손동작과 출력 화면을 120 FPS로 동시에 촬영하고, 동작 시작 프레임과 화면 반응 프레임의 차이를 밀리초 단위로 환산하였다.

$$[Latency(ms) = \frac{F_{\text{response}} - F_{\text{start}}}{120} \times 1000]$$

정지 상태 좌표 흔들림은 참여자가 손을 5초 동안 일정한 위치에 유지한 구간에서 측정하였다. 손목의 x, y 좌표와 해당 구간의 평균 좌표 사이 거리의 제곱평균제곱근을 계산하였다. 값이 낮을수록 정지 상태에서 손 좌표와 디지털 인형의 위치가 안정적임을 의미한다.

$$[Jitter_{\text{RMS}} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left[(x_i - \bar{x})^2 + (y_i - \bar{y})^2 \right]}]$$

원시 추적 손실률은 입력 영상에 손이 존재하지만 MediaPipe에서 손 랜드마크가 검출되지 않은 프레임의 비율로 계산하였다. 세 방식이 동일한 MediaPipe 검출 결과를 사용하므로 원시 추적 손실률은 방식별로 동일하다.

그러나 랜드마크 누락이 실제 디지털 인형의 움직임 중단으로 연결되는 정도는 처리 방식에 따라 달라질 수 있다. 따라서 본 연구에서는 원시 추적 손실률과 별도로 제어 중단 빈도를 측정하였다. 제어 중단은 디지털 인형의 위치 또는 관절 값이 3프레임 이상 갱신되지 않거나 기본값으로 초기화된 경우로 정의하고, 1분당 발생 횟수로 산출하였다.

프레임 처리 성능은 동일한 디지털 그림자극 장면을 60초 간 실행하여 평균 FPS와 최소 FPS를 기록하였다. 디지털 인형, 배경 영상, 장면 효과 및 출력 해상도는 세 방식에 동일하게 적용하였다.

각 지표는 참여자별 평균과 표준편차를 산출하였다. 동일한 참여자 영상에 세 가지 처리 방식을 적용하였기 때문에 반복측정 자료로 처리하였다. 방식 간 사후 비교에는 Wilcoxon 부호 순위 검정을 사용하고, 세 차례의 쌍별 비교에 Bonferroni 보정을 적용하였다. 통계적 유의수준은 $p < .05$ 로 설정하였다.

통계적 유의확률과 함께 실제 성능 차이를 확인하기 위해 정확도는 퍼센트포인트(%p), 지연 시간, 좌표 흔들림, 제어 중단 빈도 및 FPS는 증감률을 함께 제시하였다.

3-3 비교 실험 결과

세 방식의 전체 제스처 인식 결과, M0에서는 총 800회 중 697회, M1에서는 732회, M2에서는 744회가 정확하게 판정되었다. 이에 따른 전체 인식 정확도는 M0 87.1%, M1 91.5%, M2 93.0%로 나타났다.

동작별 정확도에서 손목 이동은 M0 93.0%, M1 96.0%, M2 97.0%로 나타났으며, 검지 방향 조절은 각각 89.0%, 93.0%, 94.0%로 확인되었다. 집기 동작은 M0 79.5%, M1 85.5%, M2 88.0%로 네 가지 동작 가운데 가장 낮은 정확도를 보였다. 손바닥 개폐 동작의 정확도는 M0 87.0%, M1 91.5%, M2 93.0%로 나타났다.

집기 동작에서 정확도가 상대적으로 낮게 나타난 원인은 엄지와 검지 끝점이 서로 근접하는 과정에서 손가락이 중첩되고, 두 지점 사이의 거리가 판정 임계값 주변에서 반복적으로 변화했기 때문으로 해석할 수 있다. 반면 손목 이동은 손 전체의 위치 변화를 사용하므로 손가락 끝점에 기반한 동작보다 안정적으로 판정되었다.

세 방식의 전체 정확도 차이에 대한 자료의 정규성은 Shapiro-Wilk 검정을 통해 확인하였다. 정규성 가정을 충족하지 않은 지표에 대해서는 세 방식 간 차이를 검증하기 위해 Friedman 검정을 적용하였다. 전체 차이가 유의한 경우 Wilcoxon 부호순위 검정을 이용하여 사후 비교를 실시하고, Bonferroni 보정을 적용하였다. 통계적 유의수준은 $p < .05$ 로 설정하였다. ($\chi^2(2)=8.85$, $p=.012$). Bonferroni 보정을 적용한 사후 비교에서 M2는 M0보다 유의하게 높은 정확도를 보였으며($p=.007$), M1도 M0보다 높은 정확도를 보였다($p=.041$). 그러나 M2와 M1의 1.5%p 차이는 통계적으로 유의하지 않았다($p=.362$). 따라서 제안 방식의 정확도가 가장 높게 나타났지만, 1€ 필터 방식보다 확실하게 우수하다고 판단할 수는 없다.

중단 간 평균 지연 시간은 M0 81.1ms(SD=8.1), M1

105.0ms(SD=6.3), M2 122.6ms(SD=6.2)로 측정되었다. 세 방식의 차이는 통계적으로 유의하였다($\chi^2(2)=40.00$, $p < .001$). 사후 비교에서도 모든 방식 사이에 유의한 차이가 나타났다($p < .001$).

M0는 별도의 시간적 평활화 과정을 적용하지 않아 가장 짧은 지연 시간을 보였다. M2의 평균 지연 시간은 M1보다 17.6ms, 약 16.8% 증가하였으며, M0보다는 41.5ms, 약 51.2% 증가하였다. 따라서 제안 방식은 좌표 안정화 측면에서는 유리할 수 있으나 빠른 손동작에 대한 반응성에서는 M0와 M1보다 열위한 결과를 보였다.

정지 상태 좌표 흔들림은 M0 0.0189(SD=0.0022), M1 0.0097(SD=0.0009), M2 0.0068(SD=0.0005)로 나타났다. 세 방식의 차이는 통계적으로 유의하였다($\chi^2(2)=40.00$, $p < .001$). M2의 좌표 흔들림은 M0보다 약 64.0%, M1보다 약 30.3% 감소하였다. 이는 Filter CHOP 기반 평활화와 데드존 적용이 정지 상태에서 발생하는 미세한 인형 흔들림을 줄이는 데 효과적으로 작용했음을 보여준다.

원시 추적 손실 프레임 비율은 세 방식 모두 2.9%로 동일하게 나타났다. 이는 세 방식이 동일한 MediaPipe 랜드마크 검출 결과를 입력으로 사용했기 때문이다. 따라서 좌표 필터만으로는 카메라 입력 단계에서 발생하는 손 검출 실패를 감소시킬 수 없었다.

반면 실제 화면에서 관찰된 제어 중단 빈도는 M0가 분당 7.63회(SD=0.74), M1이 7.32회(SD=1.17), M2가 3.18회(SD=0.54)로 나타났다. 전체 방식 차이는 통계적으로 유의하였다($\chi^2(2)=31.60$, $p < .001$). M0와 M1 사이에는 유의한 차이가 나타나지 않았으나($p=.269$), M2는 M0와 M1보다 유의하게 낮은 제어 중단 빈도를 보였다($p < .001$).

M2의 제어 중단 빈도는 M1보다 약 56.6% 감소하였다. 이는 M2가 실제 추적 손실을 줄인 것이 아니라, 최대 5프레임의 상태 유지 기능을 통해 짧은 랜드마크 누락이 디지털 인형의 즉각적인 정지나 초기화로 이어지는 것을 완화했기 때문이다.

평균 FPS는 M0 49.2 FPS(SD=0.7), M1 47.6 FPS(SD=0.6), M2 46.8 FPS(SD=0.8)로 측정되었다. 최소 FPS는 각각 38.6 FPS, 36.9 FPS 및 35.8 FPS로 나타났다. 평균

표 4. 동일 조건에서의 좌표 처리 방식별 성능 비교

Table 4. Performance comparison of coordinate-processing methods under identical conditions

Evaluation Metric	M0: Raw Mapping	M1: 1€ Filter	M2: Proposed Method	M2-M1 Difference	Statistical Result
Gesture Recognition Accuracy (%)	87.1 ± 6.9	91.5 ± 3.7	93.0 ± 2.5	+1.5%p	$\chi^2=8.85$, $p=.012$
End-to-End Latency (ms)	81.1 ± 8.1	105.0 ± 6.3	122.6 ± 6.2	+17.6 ms	$\chi^2=40.00$, $p < .001$
Coordinate Jitter (RMS)	0.0189 ± 0.0022	0.0097 ± 0.0009	0.0068 ± 0.0005	-30.3%	$\chi^2=40.00$, $p < .001$
Control Interruption (events/min)	7.63 ± 0.74	7.32 ± 1.17	3.18 ± 0.54	-56.6%	$\chi^2=31.60$, $p < .001$
Average FPS	49.2 ± 0.7	47.6 ± 0.6	46.8 ± 0.8	-0.8 FPS	$\chi^2=40.00$, $p < .001$
Minimum FPS	38.6	36.9	35.8	-1.1 FPS	Descriptive result

FPS의 방식별 차이는 통계적으로 유의하였다($\chi^2(2)=40.00$, $p<.001$).

M2의 평균 FPS는 M1보다 약 1.6%, M0보다 약 4.8% 낮게 나타났다. 좌표 평활화, 데드존 판정, 상태 유지 및 범위 제한 과정이 추가되면서 처리 부하가 증가한 것으로 해석할 수 있다. 그러나 세 방식 모두 평균 30 FPS 이상의 출력 성능을 유지하여 기본적인 실시간 출력은 가능하였다.

3-4 조명 및 거리 조건별 결과

조명 조건별 실험 결과, 세 방식 모두 350lux의 일반 조도에서 가장 높은 인식 정확도를 보였다. 100lux의 저조도 조건에서는 손 윤곽과 손가락 끝점의 구분이 불안정해지면서 정확도가 감소하였다. 700lux의 고조도 조건에서도 일부 손 영역에서 과노출이 발생하여 일반 조도보다 낮은 결과가 나타났다.

100lux 조건에서 인식 정확도는 M0 80.5%, M1 85.5%, M2 87.5%로 나타났다. 350lux에서는 각각 90.0%, 94.0%, 96.0%였으며, 700lux에서는 86.5%, 91.0%, 92.5%로 확인되었다. M2는 모든 조명 조건에서 가장 높은 정확도를 보였지만, 저조도에서는 일반 조도보다 8.5%p 낮아졌다.

이는 제안 방식이 검출된 좌표의 흔들림과 임계값 주변의 변화를 줄일 수는 있으나, 영상 입력 단계에서 손 랜드마크가 잘못 검출되거나 검출되지 않는 문제까지 해결하지는 못한다는 것을 의미한다. 조명 변화, 가림 및 모션 블러가 손 자세 추정의 주요 성능 저하 요인이 될 수 있다는 선행연구의 결과와도 연결된다[10].

거리 조건에서는 40 cm에서 가장 높은 정확도가 나타났다. 30cm에서는 손이 카메라 화면에서 크게 나타나 손가락 일부가 화면 밖으로 벗어나는 경우가 있었고, 50cm에서는 손과 손가락이 작게 입력되어 손끝 좌표의 변동이 증가하였다.

30cm 조건에서 정확도는 M0 84.5%, M1 89.0%, M2 91.0%로 나타났다. 40cm에서는 각각 90.5%, 94.5%, 96.5%였으며, 50cm에서는 86.5%, 91.0%, 92.0%로 확인되었다. 따라서 본 실험 조건에서는 카메라와 손 사이 거리 40cm가 가장 안정적인 조작 범위로 나타났다.

표 5. 조명 및 거리 조건에 따른 제스처 인식 정확도
Table 5. Gesture-recognition accuracy under lighting and distance conditions

Experimental Condition	M0: Raw Mapping	M1: 1€ Filter	M2: Proposed Method
Low Lighting: 100 lux	80.5%	85.5%	87.5%
Normal Lighting: 350 lux	90.0%	94.0%	96.0%
High Lighting: 700 lux	86.5%	91.0%	92.5%
Distance: 30 cm	84.5%	89.0%	91.0%
Distance: 40 cm	90.5%	94.5%	96.5%
Distance: 50 cm	86.5%	91.0%	92.0%

IV. 실험 결과 논의

4-1 좌표 처리 방식에 따른 성능 차이

동일 조건 비교 실험 결과, 제안 방식인 M2는 원시 좌표 직접 매핑 방식인 M0보다 제스처 인식 정확도, 정지 상태 좌표 안정성 및 제어 연속성에서 개선된 결과를 보였다. 전체 제스처 인식 정확도는 M0 87.1%, M1 91.5%, M2 93.0%로 나타났다. M2는 M0보다 5.9%p 높은 정확도를 보였으나, M1과의 차이는 1.5%p에 그쳤으며 두 방식 사이의 차이는 통계적으로 유의하지 않았다. 따라서 제안 방식이 원시 좌표 직접 매핑보다 안정적인 제스처 판정을 지원한다고 볼 수 있으나, 1€ 필터 방식보다 인식 정확도 측면에서 명확하게 우수하다고 판단하기는 어렵다.

동작별 결과에서는 손목 이동이 가장 높은 정확도를 보였고, 집기 동작이 가장 낮은 정확도를 나타냈다. 손목 이동은 손 전체의 위치 변화를 사용하므로 개별 손가락 끝점의 검출 오류에 상대적으로 적은 영향을 받았다. 반면 집기 동작은 엄지와 검지 끝점 사이의 거리를 기준으로 판정하기 때문에 두 손가락이 겹치거나 임계값 주변에서 거리가 반복적으로 변화할 경우 오인식이 발생하였다. 이는 손 랜드마크 기반 인터페이스에서 손 전체의 위치를 이용하는 동작보다 손가락 끝점 사이의 미세한 거리를 이용하는 동작이 입력 조건에 더 민감하다는 것을 보여준다.

정지 상태 좌표 흔들림은 M0 0.0189, M1 0.0097, M2 0.0068로 나타났다. M2는 M0보다 약 64.0%, M1보다 약 30.3% 낮은 흔들림 값을 보였다. 이는 Filter CHOP을 이용한 좌표 평활화와 데드존 설정이 손이 정지한 상태에서 발생하는 미세한 좌표 변화를 감소시키는 데 효과적으로 작용했음을 의미한다. 특히 디지털 그림자 인형은 손목 좌표의 작은 변화도 화면상 인형의 위치 변화로 직접 표현되므로, 좌표 흔들림 감소는 인형의 정지 안정성을 확보하는 데 중요한 요소가 된다.

제어 중단 빈도는 M0가 분당 7.63회, M1이 7.32회, M2가 3.18회로 나타났다. 세 방식은 동일한 MediaPipe 랜드마크 검출 결과를 사용하였기 때문에 원시 추적 손실 프레임 비율은 동일하였다. 그러나 M2는 짧은 랜드마크 누락 구간에서 마지막으로 확인된 안정 좌표를 일정 프레임 동안 유지하도록 설정하였기 때문에, 일시적인 추적 손실이 디지털 인형의 즉각적인 정지나 초기화로 연결되는 빈도가 감소하였다. 따라서 M2의 장점은 손 검출 자체의 정확도를 향상시키는 데 있다기보다, 일시적인 검출 실패가 출력 제어에 미치는 영향을 줄이는 데 있다고 해석할 수 있다.

4-2 좌표 안정성과 반응 속도의 관계

제안 방식은 좌표 흔들림과 제어 중단을 감소시켰으나, 중

단 간 지연 시간에서는 상대적으로 낮은 성능을 보였다. 평균 지연 시간은 M0 81.1ms, M1 105.0ms, M2 122.6ms로 나타났다. M2는 M1보다 17.6ms, M0보다 41.5ms 긴 지연 시간을 보였다. 이는 좌표 평활화, 테드존 판정 및 상태 유지 과정이 입력 안정성을 높이는 대신 추가적인 처리 지연을 발생시킨 결과로 볼 수 있다.

M0는 원시 좌표를 별도의 시간적 보정 없이 출력에 연결하기 때문에 가장 빠른 반응 속도를 보였으나, 정지 상태에서 좌표 흔들림이 크게 나타났다. M1은 입력 속도에 따라 필터 강도를 조절하여 M0보다 좌표 안정성을 높이면서도 M2보다 짧은 지연 시간을 유지하였다. 반면 M2는 세 방식 가운데 가장 안정적인 좌표 결과를 보였지만, 빠른 손동작에 대한 반응 속도에서는 M0와 M1보다 불리하였다.

이러한 결과는 제안 방식이 모든 상황에서 가장 우수한 처리 방식이 아님을 보여준다. 디지털 그림자 인형의 관절을 세밀하게 조종하거나 정지 자세를 유지해야 하는 장면에서는 좌표 흔들림이 적은 M2가 적합할 수 있다. 그러나 빠른 팔 동작이나 리듬에 맞춘 즉각적인 움직임이 필요한 장면에서는 지연 시간이 짧은 M1이 더 적절할 수 있다. 따라서 실제 시스템에서는 하나의 고정된 필터 설정을 모든 동작에 적용하기보다 동작 속도와 조종 목적에 따라 평활화 강도와 상태 유지 시간을 조절하는 방식이 필요하다.

프레임 처리 성능에서도 유사한 차이가 나타났다. 평균 FPS는 M0 49.2 FPS, M1 47.6 FPS, M2 46.8 FPS로 측정되었다. M2는 M0보다 약 4.8%, M1보다 약 1.7% 낮은 평균 FPS를 보였다. 이는 M2에 추가된 좌표 처리 과정이 시스템의 연산 부하를 증가시켰기 때문으로 해석할 수 있다. 다만 세 방식 모두 평균 30 FPS 이상의 출력 성능을 유지하였으므로 기본적인 실시간 조종에는 큰 문제가 없는 것으로 나타났다.

4-3 조명 및 거리 조건에 따른 적용 가능성

조명 조건별 실험에서는 350 lux의 일반 조도에서 가장 높은 제스처 인식 정확도가 나타났다. M2의 정확도는 100lux에서 87.5%, 350lux에서 96.0%, 700lux에서 92.5%로 확인되었다. 저조도 조건에서는 손의 윤곽과 손가락 끝점이 명확하게 구분되지 않아 랜드마크 검출 오류가 증가하였다. 고조도 조건에서도 손 영역의 일부가 과도하게 밝게 입력될 경우 손가락 경계가 불분명해져 일반 조도보다 낮은 정확도가 나타났다.

좌표 평활화와 상태 유지 기능은 검출된 좌표의 흔들림과 일시적인 누락에는 대응할 수 있으나, 영상 입력 단계에서 손 랜드마크가 잘못 검출되는 문제를 완전히 해결할 수는 없다. 따라서 실제 전시 환경에서는 소프트웨어 처리 방식뿐 아니라 카메라 방향, 배경 밝기, 광원의 위치 및 실내 조도를 함께 조정해야 한다.

거리 조건에서는 40cm에서 가장 높은 정확도가 나타났다. M2의 인식 정확도는 30cm에서 91.0%, 40cm에서 96.5%,

50cm에서 92.0%로 나타났다. 30cm에서는 손이 화면에서 크게 입력되어 손가락이나 손목 일부가 카메라 범위를 벗어나는 경우가 발생하였다. 50cm에서는 손이 상대적으로 작게 입력되어 손가락 끝점의 위치 변화와 중첩을 정확하게 구분하기 어려웠다.

이러한 결과를 고려하면 본 시스템의 실험 환경에서는 약 350lux의 조도와 40cm의 조작 거리가 가장 안정적인 조건으로 나타났다. 다만 이 결과는 사용한 카메라의 해상도, 화각 및 설치 위치에 따라 달라질 수 있으므로 실제 전시 공간에 적용할 경우 현장 조건에 맞는 재측정이 필요하다. 또한 화면에 적정 조작 거리와 손 위치를 안내하는 시각적 가이드를 제공하면 관람객이 카메라 인식 범위를 벗어나는 문제를 줄일 수 있다.

4-4 연구의 의의 및 한계

본 연구는 손 랜드마크 추적 기술을 디지털 그림자극에 적용했다는 사실만을 제시하는 데 그치지 않고, 동일한 입력 영상과 하드웨어 환경에서 세 가지 좌표 처리 방식을 비교하였다는 점에 의의가 있다. 이를 통해 제안 방식이 원시 좌표 직접 매핑보다 좌표 흔들림과 제어 중단을 줄일 수 있는 반면, 지연 시간과 프레임 처리 성능에서는 일정한 손실이 발생한다는 것을 확인하였다.

또한 본 연구는 제안 방식의 우수성만을 강조하지 않고, 1€ 필터 방식과 비교하여 정확도 차이가 크지 않으며 반응 속도에서는 오히려 낮은 결과가 나타났음을 함께 제시하였다. 이러한 결과는 디지털 그림자 인형 조종에서 좌표 안정성과 반응 속도가 동시에 최대화되기 어렵고, 시스템 목적에 따라 두 성능 사이의 균형을 설정해야 함을 보여준다.

그러나 본 연구는 몇 가지 한계를 가진다. 첫째, 단일 웹캠과 단일 사용자 환경을 기준으로 실험하였기 때문에 실제 박물관이나 전시 공간의 복잡한 배경, 다중 관람객 및 다양한 카메라 설치 조건을 충분히 반영하지 못하였다. 둘째, 동일 입력 조건을 확보하기 위해 사전 녹화된 영상을 사용하였으므로 관람객이 화면 피드백을 확인하면서 자신의 동작을 실시간으로 수정하는 상호작용 과정은 분석하지 못하였다.

셋째, 본 연구의 제안 방식은 고정된 평활화 강도와 테드존 값을 사용하였다. 따라서 느린 동작과 빠른 동작에 동일한 처리값이 적용되어 빠른 동작에서 반응 지연이 증가할 수 있다. 향후에는 손의 이동 속도와 동작 유형에 따라 필터 강도, 테드존 및 상태 유지 시간을 자동으로 조절하는 적응형 처리 방식을 적용할 필요가 있다.

넷째, 실험에서 사용한 제스처는 손목 이동, 검지 방향, 집기 및 손바닥 개폐의 네 가지로 제한하였다. 실제 그림자극의 세밀한 관절 조종과 다양한 도구 사용을 구현하기 위해서는 두 손 입력, 복합 제스처 및 다중 관절 매핑에 대한 추가 검증이 필요하다.

V. 결 론

본 연구는 손 랜드마크 추적 기반 디지털 그림자극 조종 인터페이스를 구현하고, 원시 좌표 직접 매핑 방식(M0), 1€ 필터 방식(M1), 좌표 평활화와 데드존 및 상태 유지를 적용한 제안 방식(M2)의 성능을 동일 조건에서 비교하였다.

실험 결과, 전체 체크치 인식 정확도는 M0 87.1%, M1 91.5%, M2 93.0%로 나타났다. M2는 M0보다 높은 정확도를 보였으나 M1과의 차이는 통계적으로 유의하지 않았다. 반면 정지 상태 좌표 흔들림은 M2가 M0보다 64.0%, M1보다 29.9% 감소하였으며, 제어 중단 빈도도 M1보다 56.6% 낮게 나타났다. 이를 통해 제안 방식이 디지털 인형의 정지 안정성과 제어 연속성을 개선하는 데 효과가 있음을 확인하였다.

그러나 평균 지연 시간은 M2가 122.6 ms로 M0과 M1보다 길었으며, 평균 FPS도 46.8 FPS로 가장 낮게 나타났다. 따라서 제안 방식은 모든 성능 지표에서 우수한 것이 아니라, 좌표 안정성을 높이는 대신 반응 속도와 처리 성능이 일부 저하되는 특성을 가진다. 세밀한 관절 제어와 정지 안정성이 필요한 장면에는 적합하지만, 빠른 반응이 필요한 동작에는 필터 강도를 낮추거나 적응형 처리 방식을 적용할 필요가 있다.

본 연구는 서로 다른 좌표 처리 방식을 동일한 입력과 하드웨어 조건에서 비교하고, 제안 방식의 우위와 열위를 함께 제시하였다는 데 의의가 있다. 다만 단일 카메라와 제한된 실내 환경에서 실험하였으므로, 향후 실제 전시 공간과 다양한 사용자 조건에서 추가적인 검증이 필요하다.

참고문헌

- [1] UNESCO, Chinese Shadow Puppetry, Representative List of the Intangible Cultural Heritage of Humanity, 2011.
- [2] F. Padovani, "Varying Discourse and Use of Intangible Cultural Heritage in Shaanxi Huaxian Shadow Puppets," *China Perspectives*, No. 132, pp. 39-48, 2023. <https://doi.org/10.4000/chinaperspectives.14601>
- [3] F. Lu, F. Tian, Y. Jiang, X. Cao, W. Luo, G. Li, ... and H. Wang, "ShadowStory: Creative and Collaborative Digital Storytelling Inspired by Cultural Heritage," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Vancouver, Canada, pp. 1919-1928, 2011. <https://doi.org/10.1145/1978942.1979221>
- [4] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 19, No. 7, pp. 677-695, 1997. <https://doi.org/10.1109/34.598226>
- [5] F. Zhang, V. Bazarevsky, A. Vakunov, A. Tkachenka, G.

Sung, C.-L. Chang, and M. Grundmann, "MediaPipe Hands: On-Device Real-Time Hand Tracking," arXiv:2006.10214, 2020. <https://doi.org/10.48550/arXiv.2006.10214>

- [6] P. Xu, "A Real-time Hand Gesture Recognition and Human-Computer Interaction System," arXiv:1704.07296, 2017. <https://doi.org/10.48550/arXiv.1704.07296>
- [7] G. Casiez, N. Roussel, and D. Vogel, "1€ Filter: A Simple Speed-Based Low-Pass Filter for Noisy Input in Interactive Systems," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Austin: TX, pp. 2527-2530, 2012. <https://doi.org/10.1145/2207676.2208639>
- [8] Derivative, Math CHOP, TouchDesigner Documentation, 2022.
- [9] Derivative, Filter CHOP, TouchDesigner Documentation, 2025.
- [10] M. Pu, C. Y. Chong, and M. K. Lim, "Robustness Evaluation in Hand Pose Estimation Models Using Metamorphic Testing," in *Proceedings of the 2023 IEEE/ACM 8th International Workshop on Metamorphic Testing (MET)*, Melbourne, Australia, pp. 31-38, 2023. <https://doi.org/10.1109/MET59151.2023.00012>

조홀(Yi Zhao)



2023년 : 경희대학교 디지털콘텐츠학과 (학사)
2024년 : 경희대학교 디지털콘텐츠학과 (석사)

2024년~현 재: 경희대학교 디지털콘텐츠학과 석사과정
※관심분야 : 인터랙션 디자인, 미디어아트, 디지털 문화유산, NUI, 실감형 콘텐츠 등

김혜경(Hye-Kyung Kim)



1987년 : 고려대학교 심리학과 (문학학사)
1992년 : Pratt Institute 대학원 (M.F.A.-컴퓨터그래픽스)

1993년~1999년: LG소프트
1999년~현 재: 경희대학교 예술디자인대학
디지털콘텐츠학과 교수
※관심분야 : 모션그래픽스, 인터랙션디자인, 미디어 아트 등