

특집논문 (Special Paper)

방송공학회논문지 제31권 제5호, 2026년 9월 (JBE Vol.31, No.5, September 2026)

<https://doi.org/10.5909/JBE.2026.31.5.894>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

센서 재킷 기반 가상 지휘 동작 생성 및 실시간 저작 방법

박 성 실^{a)†}

Virtual Conducting Gesture Generation and Real-Time Authoring Using a Sensor Jacket

Soungsill Park^{a)†}

요 약

비전문가는 3차원 아바타의 지휘 동작을 제작할 때 반복적인 박자 패턴을 유지하면서 셸여림과 양손의 음악적 표현을 함께 수행하는 데 어려움이 있다. 본 논문은 기본 지휘 패턴과 센서 재킷을 결합하여 변화가 필요한 구간만 수정하는 가상 지휘 동작 저작 방법을 제안한다. 2박자, 3박자, 4박자 패턴을 이미지에서 추출 및 정규화하고 SMPL-X 아바타에 역기구학을 적용하여 기본 동작을 생성한다. 또한 6개의 IMU 센서로 박자 변화, 음악적 지시, 셸여림을 감지하여 기본 패턴의 속도, 궤적 크기, 음악적 지시 구간을 수정한다. 비전문가 3명의 실험에서 평균 시간 오차는 전체 동작 수행 시 72.75ms에서 변화 구간 수정 시 7.71ms로 감소하였으며, 실제 음악 실험에서도 102.96ms에서 16.81ms로 감소하였다. 이를 통해 기본 박자 패턴을 유지하면서 필요한 구간만 수정하는 방식이 비전문가의 박자 수행 오차를 줄이는 데 효과적임을 확인하였다.

Abstract

Nonexperts often struggle to accurately represent beat timing, dynamics, and the coordinated use of both hands when authoring conducting gestures for 3D avatars. This paper proposes an adaptive authoring method that combines predefined conducting patterns with a sensor jacket. Two-beat, three-beat, and four-beat patterns are extracted from images and applied to an SMPL-X avatar using inverse kinematics. Six upper-body IMU sensors detect tempo changes, musical instructions, and dynamics, allowing users to modify only the required segments. Experiments with three nonexperts showed that the mean timing error decreased from 72.75ms to 7.71ms, while experiments using actual music showed a decrease from 102.96ms to 16.81ms. These results show that maintaining the basic beat pattern while modifying only the required segments reduces timing errors in nonexpert conducting performance.

Keyword : Conducting Gesture, Gesture Authoring, Sensor Jacket, Adaptive Authoring, Inertial Measurement Unit

a) 숭실대학교 전자정보공학부(Department of Electronic Engineering, Soongsil University)

† Corresponding Author : 박성실(Soungsill Park)

E-mail: pssil1321@gmail.com

Tel: +82-2-828-7194

ORCID: <https://orcid.org/0000-0002-6113-6919>

※ 이 논문은 2025년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원(RS-2018-II180599, SW컴퓨팅산업원천기술개발사업(SW스타랩)의 지원을 받아 수행된 연구임.

· Manuscript received July 20, 2026; Revised August 18, 2026; Accepted August 18, 2026.

I. 서론

3차원 아바타는 가상 공연, 방송, 실감형 콘텐츠 등 다양한 분야에서 활용되고 있으며, 실제 사람의 동작을 가상 캐릭터로 표현하기 위한 동작 제작이 중요하다. 특히 지휘 동작은 반복적인 박자 패턴과 함께 BPM 변화, 썸머림, 큐, 컷오프와 같은 다양한 음악적 표현을 동시에 포함한다. 이러한 동작을 정확하게 제작하기 위해서는 박자와 음악적 표현을 함께 고려해야 하기 때문에 관련 경험이 부족한 비전문가가 전체 지휘 동작을 직접 제작하거나 수행하기에는 어려움이 있다. 기존의 가상 지휘 동작 제작에는 수작업 애니메이션과 모션 캡처가 주로 활용된다. 수작업 방식은 음악의 변화에 따라 동작을 반복적으로 수정해야 하기 때문에 많은 시간과 노력이 필요하고 광학식 모션 캡처는 정밀한 동작을 획득할 수 있으나 다수의 카메라와 마커, 별도의 측정 공간이 필요하다. 하지만 IMU 기반 모션 캡처는 비교적 적은 장비와 공간으로 동작을 획득할 수 있어 착용형 입력 장치로 활용할 수 있다.

그러나 장비가 간편해지더라도 비전문가의 지휘 수행에서 발생하는 문제는 여전히 존재한다. 지휘 경험이 부족한 사용자는 오른손의 반복적인 박자 패턴과 왼손의 큐, 컷오프, 크레센도 등의 음악적 표현을 동시에 수행하는 과정에서 박자 간격과 손끝 궤적이 불안정해질 수 있다. 특히 전체 동작을 센서로 직접 측정하면 이러한 수행 오차가 가상 동작에 그대로 반영될 수 있다. 따라서 본 연구에서는 기본 박자 패턴을 유지하면서 필요한 음악적 표현만 선택적으로 수정하는 저작 방식을 적용하였다. 이는 비전문가의 가상 지휘 동작 제작과 음악 교육 콘텐츠에 활용될 수 있다.

이에 본 연구에서는 IMU 기반 센서 재킷과 미리 생성한 기본 지휘 패턴을 결합한 가상 지휘 동작 생성 및 실시간 저작 방법을 제안한다. 기본 지휘 패턴은 SMPL-X^[1] 아바타에 적용하여 3차원 지휘 동작으로 생성하며, 이후 박자 변화, 썸머림, 음악적 지시가 발생하는 구간만 센서 입력으로 수정한다. 비전문가 3명을 대상으로 전체 동작 수행 방식과 변화 구간 수정 방식을 비교한 결과, 제안 방식에서 박자 시점의 시간 오차가 감소하고 기본 박자 패턴이 보다 안정적으로 유지됨을 확인하였다. 제안 방법은 가상 공연, 음악 교육, 실감형 콘텐츠에서의 지휘 동작 제작과 수정에

활용될 수 있다.

본 논문의 구성은 다음과 같다. II장에서는 모션 캡처 기반 동작 획득, 아바타 동작 저작, 가상 지휘 동작 생성에 관한 관련 연구를 살펴본다. III장에서는 기본 지휘 패턴 생성과 센서 재킷 기반 실시간 저작 방법을 설명하며, IV장에서는 실험 결과를 분석하고 V장에서 결론을 제시한다.

II. 관련 연구

1. 모션 캡처 기반 동작 획득

3차원 아바타 동작은 광학식 모션 캡처와 관성 센서 기반 방식으로 획득할 수 있다. 광학식 모션 캡처는 신체에 부착한 마커를 여러 대의 카메라로 추적하여 정밀한 골격 동작을 얻는다^[2]. 그러나 장비가 고가이고 사전 설치와 별도의 측정 공간이 필요하다. 이러한 한계를 보완하기 위해 IMU 기반 동작 측정 기술이 활용되고 있다. IMU는 가속도와 각속도, 자기장 정보를 이용하여 신체의 움직임과 방향을 측정한다. Crema 등은 IMU 데이터를 이용한 운동 동작 감지 및 분류 방법을 제안하였다^[3]. Jimenez 등은 발에 부착한 IMU로 보행과 경사로 이동을 추정하였으며^[4] Patil 등은 이중 센서를 결합한 자세 추적 플랫폼을 제안하였다^[5]. Qiu 등은 센서 간 시간 및 회전 관계를 실시간으로 보정하였다^[6]. 최근에는 IMU 동작을 3차원 아바타에 실시간으로 적용하는 연구가 이루어지고 있다. Gonzalez-Alonso 등은 무선 IMU로 신체 부위의 방향을 측정하고 시각화하였다^[7]. Yin 등은 저비용 IMU 기반 모션 캡처 시스템을 개발하였으며^[8] Chen과 Tan은 IMU와 다중 카메라를 결합하여 측정 오차와 폐색 문제를 보완하였다^[9].

기존의 IMU 기반 동작 획득 연구는 사용자의 전체 자세를 복원하거나 센서 데이터를 아바타에 직접 적용하는 데 주로 중점을 두었다. 그러나 지휘 동작의 경우 비전문가의 수행 오차가 박자 간격과 손끝 궤적의 불안정성으로 이어질 수 있다. 이에 본 연구에서는 착용이 간편한 센서 재킷을 동작 저작 입력 장치로 활용하여, 전체 동작을 직접 매핑하지 않고 기본 지휘 패턴을 유지하면서 박자 변화, 썸머림, 음악적 지시가 필요한 구간만 선택적으로 수정한다.

2. 아바타 동작 저작

비전문가가 3차원 아바타 동작을 제작하려면 애니메이션과 인체 동작에 관한 전문 지식이 필요하다. 이를 줄이기 위해 자동 및 반자동 아바타 저작 기술이 연구되어 왔다. Fei 등은 시각적 입력을 이용한 증강현실 콘텐츠 저작 시스템을 제안하였다^[10]. Apostolakis와 Daras는 비전문가용 아바타 저작 도구를 개발하였으며^[11] Balasubramanyam 등은 관절 움직임을 각도 공간에 시각화하는 Motion-Sphere를 제안하였다^[12].

자동 저작 방식은 사용자의 직접적인 조작을 줄이고 입력 정보에 따라 콘텐츠나 동작을 구성한다. Wang 등은 사람의 행동이 포함된 증강현실 응용 프로그램 저작 도구를 제안하였고^[13] Anasingaraju 등은 학습 내용에 따라 강사 아바타의 동작을 구성하였다^[14]. 반자동 방식은 사용자가 수행한 동작을 획득한 후 필요한 부분을 수정한다. Chi 등은 반복적인 신체 시연으로 사람의 움직임을 저작하는 방법을 제안하였다^[15]. 최근에는 적은 수의 센서 입력으로 전신 동작을 복원하거나 다른 골격의 아바타에 동작을 적용하는 연구도 수행되고 있다. Reda 등은 최소한의 센서 입력으로부터 전신 동작을 복원하고 서로 다른 골격 구조의 캐릭터에 적용하는 물리 기반 동작 리타게팅 방법을 제안하였다^[16].

기존의 아바타 동작 저작 연구는 사용자의 직접적인 조작을 줄이거나 제한된 입력을 이용하여 동작을 생성하고 수정하는 데 중점을 두었다. 그러나 전체 동작을 직접 시연하는 방식은 반복되는 동작까지 매번 수행해야 하므로 비전문가에게 부담이 될 수 있다. 특히 지휘 동작은 오른손의 반복적인 박자 패턴과 왼손의 음악적 지시가 동시에 나타나기 때문에 기본 박자 동작은 유지하고 변화가 필요한 구간만 선택적으로 수정하는 방식이 필요하다. 본 연구에서는 센서 재킷을 저작 입력 장치로 활용하여 박자 변화, 썸머림, 음악적 지시를 감지하고, 해당 구간의 기본 지휘 동작을 선택적으로 수정한다.

3. 가상 지휘 동작 생성

가상 지휘 동작 생성은 규칙 기반 방식과 음악 기반 학습

방식으로 구분할 수 있다. 규칙 기반 방식은 미리 정의한 박자 패턴과 지휘 규칙에 따라 동작을 생성한다. Nijholt 등은 가상 지휘자를 활용한 음악 학습 시스템을 제안하였으며^[17] 이 방식은 기본 박자 구조를 명확하게 표현할 수 있으나 다양한 지휘 스타일을 반영하기 어렵다.

음악 기반 학습 방식은 음악과 동작의 관계를 학습하여 지휘 동작을 생성한다. Wang 등은 커널 기반 은닉 마르코프 모델로 동적 시퀀스를 생성하였다^[18]. Dansereau 등은 음악 특징으로 지휘봉 움직임을 예측하였으며^[19] Liu 등은 음악과 지휘 동작의 동기화 관계를 자기지도 방식으로 학습하였다^[20].

최근에는 생성 모델을 활용한 연구가 진행되고 있다. Zhao 등은 확산 모델을 이용한 음악 기반 지휘 동작 생성 방법을 제안하였다^[21]. Oh 등은 제한된 데이터 환경을 위한 전이학습 기반 3차원 지휘 동작 생성 방법을 제시하였다^[22]. Severi 등은 지휘자의 머리와 손 움직임을 가상 아바타로 전달하는 원격 지휘 시스템을 제안하였다^[23].

학습 기반 방법은 자연스럽게 다양한 동작을 생성할 수 있으나 기본 박자 구조를 유지하거나 특정 구간을 직접 조절하는 데 한계가 있다. 이에 본 연구에서는 기본 지휘 패턴을 기반으로 필요한 동작 요소를 선택적으로 저작하는 방식을 사용한다.

III. 센서 재킷 기반 가상 지휘 동작 생성 및 실시간 저작 방법

제안 방법은 그림 1과 같이 기본 지휘 동작 생성과 센서 재킷을 이용한 실시간 저작 과정으로 구성된다. 본 연구에서 저작의 대상은 전체 아바타 동작을 새롭게 생성하는 것이 아니라, 미리 생성된 기본 지휘 동작의 박자 속도, 동작 크기, 음악적 지시 구간을 수정하는 것이다. 센서 재킷으로부터 박자 변화가 감지되면 해당 구간의 동작 속도를 조절하고, 썸머림 변화에 따라 손끝 궤적의 크기를 변경한다. 또한 큐, 컷오프, 크레센도, 데크레센도와 같은 음악적 지시가 감지되면 해당 구간에 왼손 동작을 적용한다. 별도의 변화가 없는 구간에서는 기존의 기본 지휘 동작을 유지한다.

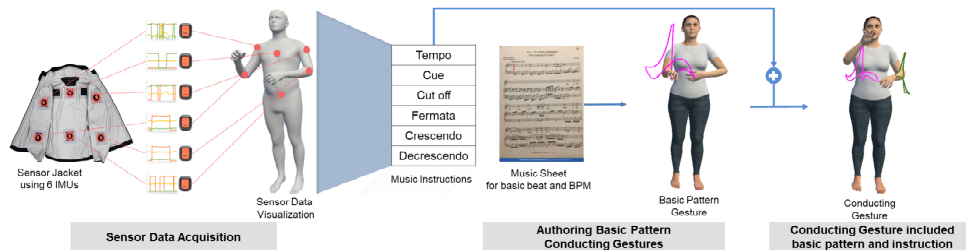


그림 1. 센서 재킷을 이용한 가상 지휘 동작 생성 및 실시간 저작 과정
Fig. 1. Process of virtual conducting gesture generation and real-time authoring using a sensor jacket

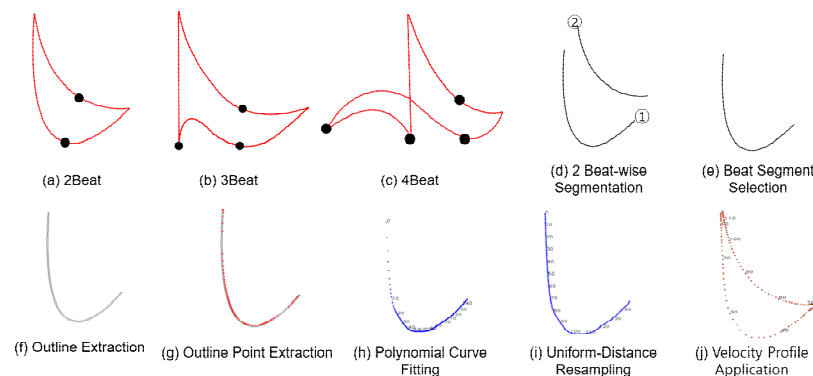


그림 2. 기본 지휘 패턴 이미지로부터 박자별 손끝 궤적을 생성하는 과정
Fig. 2. Procedure for generating beat-wise fingertip trajectories from basic conducting patterns

1. 기본 박자 패턴 생성

지휘 동작은 박자에 따라 서로 다른 궤적을 갖는다. 제안 방법에서는 그림 2의 (a), (b), (c)와 같이 2박자, 3박자, 4박자 지휘 패턴 이미지를 입력으로 사용한다. 각 패턴은 박자 경계점을 기준으로 개별 박자 구간으로 분할하며, 그림 2 (d)는 2박자 패턴을 두 개의 박자 구간으로 분리한 예시이다. 이후 그림 2 (e)와 같이 분할된 박자 구간 중 하나를 선택하여 해당 구간의 손끝 궤적을 처리한다. 선택된 박자 구간에서는 그림 2 (f)와 같이 패턴의 외곽선을 추출하고, 그림 2 (g)와 같이 외곽선을 구성하는 대표 좌표점을 진행 순서에 따라 획득한다.

이미지에서 직접 추출된 좌표점은 곡선의 형태와 픽셀 분포에 따라 점 간 간격이 불균일할 수 있으므로, 그림 2 (h)와 같이 좌표점의 진행 순서를 기준으로 (x)와 (y) 좌표를 각각 다항 곡선으로 적합하여 연속적인 2차원 손끝 궤적을 생성한다. 이후 그림 2 (i)와 같이 다항 곡선으로 적합된

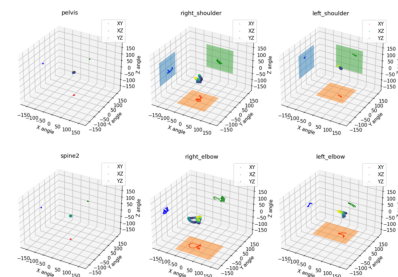


그림 3. 기본 4박자 지휘 패턴의 각도 공간(Angle space) 표현
Fig. 3. Angle-space representation of the basic four-beat pattern

궤적을 따라 인접한 좌표 사이의 누적 이동 거리를 계산하고, 이를 기준으로 전체 궤적을 일정한 거리 간격으로 재샘플링한다. 이를 통해 이미지에서 추출된 좌표의 불균일한 간격을 보정하고, 각 박자 구간에 대해 균일한 공간 간격을 갖는 손끝 궤적을 생성한다. 그림 2 (j)와 같이 실제 지휘 동작에서 나타나는 구간별 속도 변화를 반영하기 위해 속도 프로파일(Velocity Profile)을 적용하여 각 궤적점의 시

간적 진행을 조절한다. 동일한 과정을 각 박자 구간에 적용한 후 순차적으로 연결하여 최종 기본 지휘 패턴을 구성한다.

정규화된 2차원 지휘 패턴을 3차원 아바타에 적용하기 위해 아바타의 전방 방향과 어깨 방향을 기준으로 가상의 지휘 평면(Conducting Plane)을 정의한다. 평면은 기본 지휘 자세에서 명치 부근과 손끝 사이의 거리를 기준으로 아바타 전방에 배치하며, 정규화된 2차원 좌표를 평면의 가로, 세로 범위에 대응시켜 3차원 손끝 목표 위치로 변환한다. 이후 역기구학을 적용하여 손끝이 해당 궤적을 따라 이동하도록 어깨와 팔꿈치의 회전을 계산한다. 또한 섀미림은 동일한 패턴을 유지한 상태에서 지휘 평면의 크기를 확대하거나 축소하여 손끝 이동 범위를 조절하는 방식으로 표현한다.

2. 센서 재킷을 이용한 동작 획득

센서 재킷은 사용자의 체형 차이를 고려하여 S, M, L의 세 가지 크기로 구성하였으며, 사용자는 자신의 신체 크기에 적합한 재킷을 선택하여 착용한다. 상체 지휘 동작을 측정하기 위해 Xsens DOT 기반의 6개 IMU 센서를 사용하였으며, 각 센서는 재킷 내부의 지정된 신체 부위에 고정하였다. 센서 데이터 획득에 앞서 각 IMU의 방향을 동일한 전역 좌표계에 정렬하기 위해 heading 초기화를 수행한다. 초기화가 완료된 센서는 재킷 내부의 지정된 위치에 고정하며, 사용자는 재킷을 착용한 후 센서와 실행 프로그램을 연결한다. 이러한 구조를 통해 센서를 신체에 개별적으로 부착하는 과정을 줄이고, 각 센서가 대응하는 신체 부위의 움직임을 일관되게 측정할 수 있다.

각 IMU 센서에서는 시간에 따른 신체 분절의 orientation을 quaternion 형태로 획득한다. 획득한 quaternion은 센서

의 초기 방향을 기준으로 상대적인 회전 변화를 계산한 후, SMPL-X의 관절 계층 및 좌표계에 대응하는 local rotation으로 변환한다. 이후 지휘 동작의 관절 회전 패턴을 분석하기 위해 Euler angle로 변환하며, 본 연구에서는 양쪽 어깨와 팔꿈치를 주요 분석 관절로 사용한다. 어깨는 3축의 회전 변화를 사용하고, 팔꿈치는 지휘 동작에서 주요하게 나타나는 2축의 회전 변화를 사용한다. 본 연구에서는 센서의 절대 위치나 사용자의 팔 길이를 이용하여 동작을 추정하지 않고, 어깨와 팔꿈치에서 측정되는 회전 변화를 주요 특징으로 사용한다. 사용자 간 체형 차이에 대응하기 위해 사용자는 자신의 체형에 적합한 크기의 재킷을 선택하여 착용한다. 또한 센서 장착에 따른 방향 차이는 측정 전에 수행하는 초기화를 통해 보정한다.

센서 데이터는 아바타의 전체 동작에 직접 적용하지 않고, 기본 지휘 패턴을 수정하기 위한 정보로 사용한다. 이는 센서 측정 오차와 비전문가의 수행 오차가 전체 지휘 동작에 직접 반영되는 것을 줄이기 위한 것이다. 따라서 사용자는 수정이 필요한 구간만 센서 재킷을 이용하여 수행하며, 별도의 변화가 없는 구간에는 미리 생성된 기본 지휘 패턴을 유지한다. 센서 입력에서 음악적 지시가 감지되면 해당 지시에 대응하는 사전 제작된 애니메이션을 선택하여 감지된 시점에 재생한다. 센서 데이터는 동작의 종류와 적용 시점을 결정하는 데 사용하며, 오른손의 기본 박자 패턴은 유지한 상태에서 필요한 구간의 왼팔 동작만 변경한다.

3. 음악적 지시 동작 감지

제안 방법에서는 그림 4와 같이 왼손으로 표현되는 음악적 지시를 큐와 컷오프, 크레센도와 데크레센도의 네 종류로 구분한다. 큐는 음악의 시작이나 특정 연주의 진입 시점



그림 4. 센서 재킷으로 감지한 음악적 지시 동작: (a) 크레센도 (b) 큐 (c) 데크레센도 (d) 컷오프

Fig. 4. Musical instruction gestures detected by the sensor jacket: (a) crescendo (b) cue (c) decrescendo and (d) cutoff

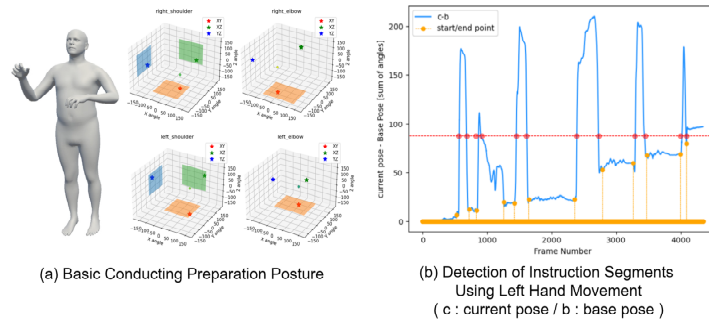


그림 5. 기본 지휘 준비 자세의 각도 공간과 센서 측정 데이터
Fig. 5. Angle-space representation and sensor measurements of the basic conducting ready position

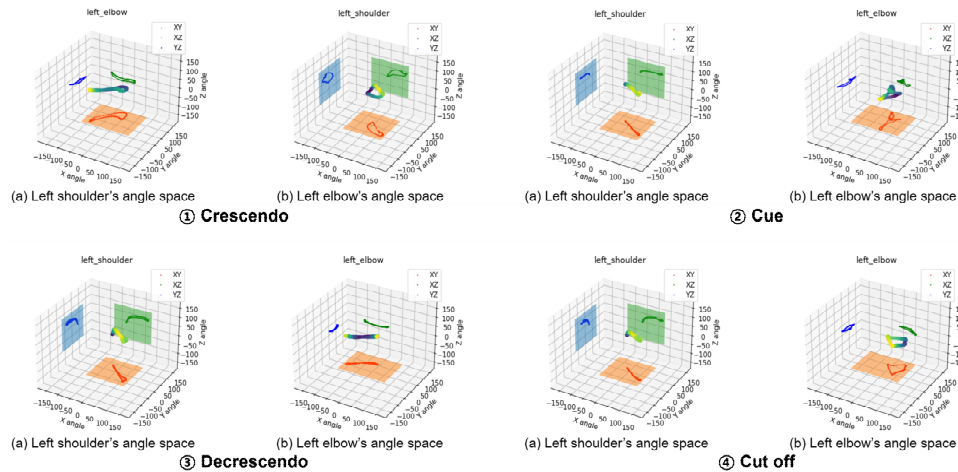


그림 6. 네 가지 음악적 지시 동작의 각도 공간 표현
Fig. 6. Angle-space representations of four musical instruction gestures

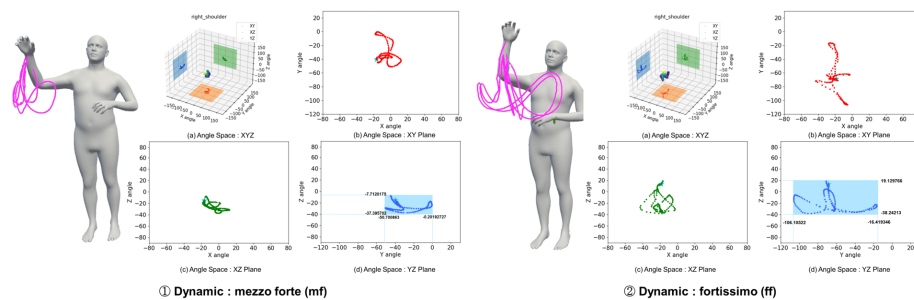


그림 7. 셈여림에 따른 지휘 동작의 각도 공간 비교: (a) 메조포르테 (b) 포르티시모
Fig. 7. Comparison of angle-space representations of conducting gestures according to dynamics: (a) mezzo-forte and (b) fortissimo

을 나타내며 컷오프는 연주의 종료 시점을 의미한다. 크레센도와 데크레센도는 각각 음량이 점차 커지거나 작아지는 변화를 표현한다. 음악적 지시가 없는 경우 왼손은 명치 아

래의 기본 자세에 둔다. 현재 자세와 기본 자세 사이의 어깨와 팔꿈치 각도 차이가 설정한 임계값을 초과하면 지시 동작이 시작된 것으로 판단한다. 이후 각도 차이가 다시 임계

값 이하로 감소하면 동작이 종료된 것으로 판단한다. 그림 5와 같이 감지된 구간의 관절 움직임 분석하여 지시 동작의 종류를 구분한다.

크레센도는 왼손을 아래에서 위로 천천히 올리는 동작이며 데크레센도는 위에서 아래로 내리는 동작이다. 큐는 기본 자세에서 팔꿈치를 펴는 동작으로 표현한다. 컷오프는 손으로 원형 궤적을 그린 뒤 동작이 끝나는 시점을 종료 시점으로 사용한다. 각 동작은 그림 6과 같이 시작 위치와 어깨와 팔꿈치의 회전 패턴이 다르므로 각도 공간에서 구분할 수 있다.

4. 셈여림 표현

셈여림은 지휘 동작의 크기로 표현한다. 작은 동작은 약한 음량을 나타내며 큰 동작은 강한 음량을 나타낸다. 본 연구에서는 오른손의 박자 동작 범위를 계산하여 셈여림을 추정한다. 그림 7과 같이 어깨와 팔꿈치의 각도 변화를 이용하여 지휘 궤적의 범위를 구한 뒤 기준 동작과 비교한다.

실험에서는 피아노(p)와 메조포르테(mf), 포르티시모(ff)의 세 단계로 셈여림을 구분하였다. 메조포르테(mf) 동작을 기준으로 설정한 뒤 관절의 각도 범위가 일정 비율 이상 증가하거나 감소하면 셈여림이 변화한 것으로 판단한다.

감지된 셈여림에 따라 기본 박자 궤적의 크기를 조절하여 아바타 동작에 적용한다. 이를 통해 동일한 박자 패턴을 유지하면서 음악의 강약 변화를 표현할 수 있다.

5. 센서 입력을 이용한 기본 지휘 동작 저작

최종 지휘 동작은 미리 생성된 기본 박자 패턴을 유지하면서 센서 재킷에서 감지된 정보에 따라 필요한 구간을 수정하여 생성한다. 본 연구에서 센서 입력을 통해 저작되는 요소는 박자에 따른 동작 속도, 셈여림에 따른 궤적 크기, 왼손의 음악적 지시 동작이다.

먼저 박자 변화가 감지되면 해당 구간의 기본 궤적 형태는 유지하고 재생 시간을 조절하여 입력된 BPM에 맞게 동작 속도를 변경한다. 셈여림 변화가 감지되면 기본 손끝 궤적의 중심과 형태를 유지하면서 궤적의 이동 범위를 확대하거나 축소하여 동작 크기를 조절한다. 큐, 컷오프, 크레센

도, 데크레센도와 같은 음악적 지시가 감지되면 해당 시간 구간의 왼팔 동작에 애니메이션을 적용하며, 오른손은 기존의 기본 박자 패턴을 유지한다.

따라서 센서 데이터는 전체 지휘 동작을 직접 대체하지 않고, 변화가 필요한 구간의 저작 정보로 사용된다. 별도의 변화가 감지되지 않은 구간에서는 기존의 기본 박자 동작을 그대로 유지한다. 이를 통해 사용자는 전체 지휘 동작을 반복적으로 수행하지 않고도 기본 박자 구조를 유지하면서 필요한 박자 변화와 음악적 표현을 선택적으로 수정할 수 있다.

IV. 실험 및 결과

제안 방법을 검증하기 위해 지휘 경험이 없는 비전문가를 대상으로 목표 BPM에 따른 박자 수행 정확도를 평가하였다. 지휘 경험이 부족한 사용자는 반복적인 박자 동작과 음악적 표현을 동시에 수행하는 과정에서 목표 박자 시점과 실제 수행 시점 사이에 오차가 발생할 수 있다. 이에 본 연구에서는 목표 BPM으로부터 설정된 기준 박자 시점과 센서 재킷으로 측정된 실제 박자 시점 사이의 시간 차이를 박자 수행 오차로 정의하였다. 실험에는 지휘 경험이 없는 비전문가 3명이 참여하였으며, 각 실험자는 자신의 체형에 적합한 크기의 센서 재킷을 선택하여 착용하였다. 이후 전체 지휘 동작을 직접 수행하는 방식과 기본 박자 패턴을 유지하면서 변화가 필요한 구간만 수정하는 방식을 비교하여 제안 방법의 박자 수행 오차 감소 효과를 확인하였다.

1. 기본 박자 패턴 생성 결과

먼저 2박자, 3박자, 4박자에 해당하는 기본 지휘 패턴을 생성하였다. 각 패턴은 이미지에서 추출한 궤적을 정규화한 후 SMPL-X 아바타의 손끝 목표 위치로 적용하였다. 이후 역기구학을 이용하여 어깨와 팔꿈치의 회전값을 계산하고 기본 지휘 동작을 생성하였다. 각 박자 패턴은 박자 수에 따라 서로 다른 수의 궤적 좌표로 구성하였다.

2박자는 1박자당 150개, 3박자는 1박자당 100개, 4박자는 1박자당 75개의 좌표를 사용하였다. 이를 통해 박자 수

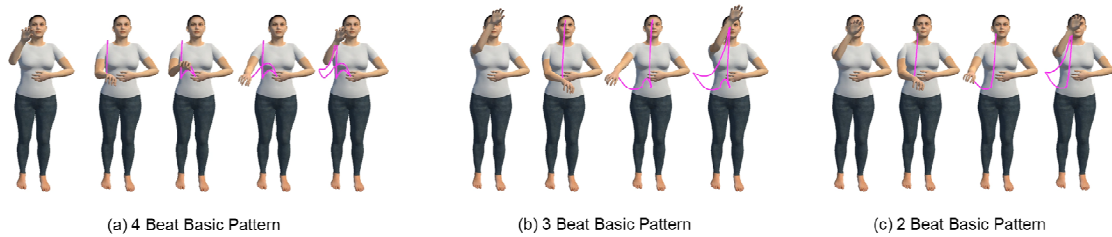


그림 8. 기본 지휘 패턴을 적용한 아바타 동작 생성 결과: (a) 4박자 (b) 3박자 (c) 2박자
Fig. 8. Avatar motion generation results using basic conducting patterns: (a) four-beat, (b) three-beat, and (c) two-beat patterns



그림 9. 기본 박자 패턴에 음악적 표현을 적용한 지휘 동작 생성 결과: (a) 큐 (b) 포르티시모
Fig. 9. Conducting gesture generation results with musical expressions applied to the basic beat pattern: (a) cue and (b) fortissimo

에 따른 궤적 길이를 구성하고, 입력된 BPM에 따라 각 궤적의 재생 속도를 조절하였다. 그림 8은 4박자, 3박자, 2박자에 해당하는 아바타 지휘 동작 생성 결과를 보여준다.

기본 박자 패턴에는 그림 9와 같이 왼손의 음악적 지시 동작과 셈여림 표현을 추가할 수 있다. 큐와 컷오프, 크레센도와 데크레센도는 기본 오른손 박자 동작과 함께 적용하였다. 셈여림은 지휘 궤적의 크기를 조절하는 방식으로 표현하였다. 실험에서는 피아노(p), 메조포르테(mf), 포르티시모(ff)의 세 단계를 사용하였다. 동작의 세기가 커질수록 손의 이동 범위와 지휘 궤적의 크기도 증가하도록 설정하였다.

2. 센서 재킷 기반 동작 저작 결과

센서 재킷을 이용한 실험은 네 가지 상황으로 구성하였다. 첫 번째는 음악의 속도 변화에 따라 박자 동작의 속도를 조절하는 경우이다. 두 번째는 특정 시점에 크레센도와 데크레센도 같은 음악적 지시를 추가하는 경우이다. 세 번째는 박자 동작과 큐 또는 컷오프를 함께 수행하는 경우이다.

네 번째는 지휘 동작의 크기를 변화시켜 셈여림을 조절하는 경우이다. 박자 변화 실험에서는 4박자를 90 BPM으로 수행한 후 중간 구간에서 120 BPM으로 변경하였다. 실험은 전체 박자 패턴을 처음부터 끝까지 직접 수행하는 방식과 박자가 변화하는 구간만 수행하는 방식으로 나누어 진행하였다.

전체 패턴을 직접 수행한 경우에는 각 박자의 수행 시간이 일정하지 않았으며 변화 구간만 수행한 경우에는 기존의 기본 패턴이 유지되므로 박자 간격이 비교적 일정하게 나타났다. 그림 10부터 그림 12는 각 실험자의 박자 변화 감지 결과를 보여준다. 표 1과 같이 실험자 A의 평균 시간 오차는 전체 동작 수행 시 79.63ms에서 변화 구간 수정 시 11.44ms로 감소하였으며, 실험자 B는 69.31ms에서 4.81ms, 실험자 C는 69.30ms에서 6.87ms로 감소하였다. 세 실험자의 평균 시간 오차는 전체 동작 수행 시 72.75ms, 변화 구간 수정 시 7.71ms로 나타났다. 이를 통해 모든 실험자에서 변화 구간만 수정하는 방식이 전체 동작을 직접 수행하는 방식보다 낮은 박자 수행 오차를 보이는 것을 확인하였다.

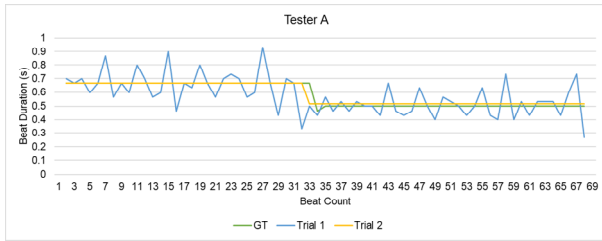


그림 10. 실험자 A의 박자 변화 감지 결과
Fig. 10. Beat change detection results for Tester A

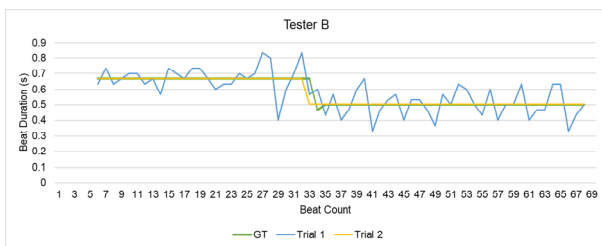


그림 11. 실험자 B의 박자 변화 감지 결과
Fig. 11. Beat change detection results for Tester B

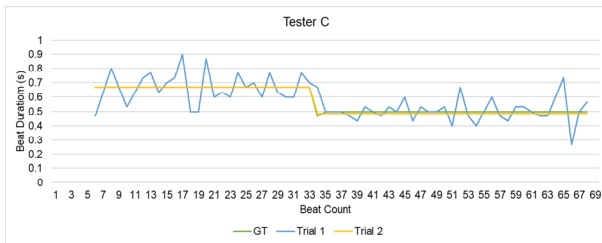


그림 12. 실험자 C의 박자 변화 감지 결과
Fig. 12. Beat change detection results for Tester C

표 1. 박자 변화 실험의 평균 시간 오차 비교
Table 1. Comparison of Mean Timing Errors in the Beat-Change Experiment

Tester	Trial 1	Trial 2
A	79.63	11.44
B	69.31	4.81
C	69.30	6.87
Mean	72.75	7.71

3. 실제 음악 적용 결과

실제 음악을 이용한 실험에서는 모차르트의 곡을 사용하

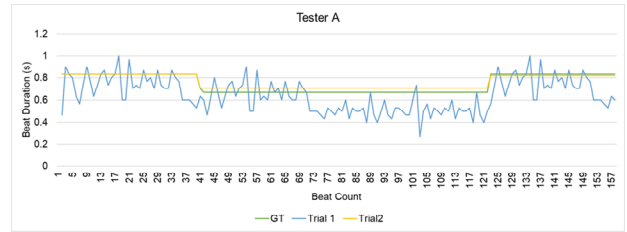


그림 13. 실험자 A의 실제 음악 기반 지휘 동작 저작 결과
Fig. 13. Conducting gesture authoring results using real music for Tester A

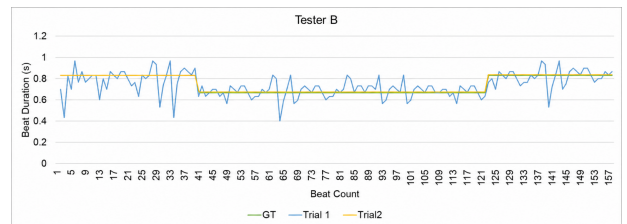


그림 14. 실험자 B의 실제 음악 기반 지휘 동작 저작 결과
Fig. 14. Conducting gesture authoring results using real music for Tester B

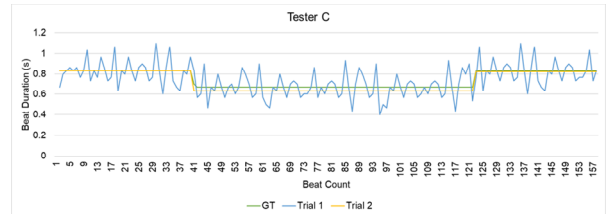


그림 15. 실험자 C의 실제 음악 기반 지휘 동작 저작 결과
Fig. 15. Conducting gesture authoring results using real music for Tester C

표 2. 실제 음악 적용 실험에서의 평균 시간 오차 비교
Table 2. Comparison of Mean Timing Errors in the Real Music Experiment

Tester	Trial 1	Trial 2
A	153.70	23.17
B	63.48	6.81
C	91.71	20.44
Mean	102.96	16.81

였다. 음악의 BPM 정보를 정확하게 반영하기 위해 악보 인식 도구를 이용하여 박자 정보를 획득하였으며, 음악 중간에 BPM이 변화하도록 설정하여 이에 맞추어 지휘 동작

을 수행하였다. 실험은 전체 박자 동작을 음악에 맞추어 지속적으로 수행하는 방식과, 기본 박자 패턴을 유지하면서 BPM이 변화하는 구간만 센서 재킷을 이용하여 수정하는 방식으로 구분하였다. 전체 동작을 직접 수행한 경우 목표 박자 시점과 실제 수행 시점 사이의 오차가 상대적으로 크게 나타난 반면, 변화 구간만 수정한 경우에는 기본 패턴의 시간 구조가 유지되어 보다 안정적인 박자 수행이 가능하였다.

그림 13부터 그림 15는 실험자별 실제 음악 기반 지휘 동작 수행 결과를 보여주며, 표 2는 두 방식에 따른 평균 시간 오차를 비교한 결과이다. 전체 동작을 직접 수행한 경우 실험자 A, B, C의 평균 시간 오차는 각각 153.70ms, 63.48ms, 91.71ms로 나타났으며, 변화 구간만 수정한 경우에는 각각 23.17ms, 6.81ms, 20.44ms로 감소하였다. 세 실험자의 전체 평균 시간 오차는 전체 동작 수행 시 102.96ms에서 변화 구간 수정 시 16.81ms로 감소하였다. 이를 통해 실제 음악을 적용한 경우에도 기본 박자 패턴을 유지하면서 변화 구간만 수정하는 방식에서 비전문가의 박자 수행 오차가 감소하는 것을 확인하였다.

4. 결과 분석

실험 결과 전체 지휘 동작을 직접 수행하는 방식에 비해 변화가 필요한 구간만 수정하는 방식에서 시간 오차가 일관되게 낮게 나타났다. 이는 비전문가가 반복적인 박자 동작과 왼손의 음악적 지시를 동시에 수행하는 부담을 줄이면서, 목표 박자 시점에 보다 안정적으로 동작을 수행할 수 있음을 보여준다. 또한 제안 방법은 기본 박자 패턴을 유지한 상태에서 박자 변화, 썸머링, 음악적 지시가 필요한 구간만 선택적으로 수정하기 때문에 사용자의 수행 오차가 전체 동작에 반영되는 것을 줄일 수 있다.

하지만 본 실험은 지휘 경험이 없는 비전문가 3명을 대상으로 수행하였으며, 박자 시점의 시간 오차를 중심으로 평가하였다. 따라서 손끝 궤적, 썸머링 표현, 양손 협응에 대한 정량적 평가는 제한적이었다. 향후에는 실험자 수를 확대하고 추가적인 평가를 통해 제안 방법의 활용성을 검증할 필요가 있다.

V. 결 론

본 논문에서는 기본 지휘 패턴과 센서 재킷을 결합한 가상 지휘 동작 생성 및 실시간 저작 방법을 제안하였다. 2박자와 3박자, 4박자 패턴을 이미지에서 추출하고 정규화한 뒤 SMPL-X 아바타에 역기구학을 적용하여 기본 지휘 동작을 생성하였다. 또한 상체에 부착된 6개의 IMU 센서를 이용해 박자 변화, 큐, 컷오프, 크레센도, 데크레센도를 감지하고 썸머링 정보를 필요한 구간에 적용하였다.

비전문가 3명을 대상으로 한 실험에서는 전체 박자 동작을 직접 수행하는 방식과 변화 구간만 수정하는 방식을 비교하였다. 실험 결과 변화 구간만 수정한 경우 시간 오차가 감소하였으며 기본 박자 패턴도 보다 안정적으로 유지되었다. 실제 음악을 적용한 실험에서도 동일한 경향을 확인하였다. 이를 통해 기본 박자 패턴을 유지하면서 필요한 구간만 수정하는 방식이 비전문가의 수행 오차가 전체 동작에 반영되는 것을 줄이는 데 효과적임을 확인하였다.

제안 방법은 고가의 광학식 모션 캡처 장비 없이도 기본 박자 패턴을 유지하면서 박자 변화, 음악적 지시, 썸머링을 실시간으로 저작할 수 있다는 장점이 있다. 따라서 가상 지휘 동작 제작 및 비전문가를 위한 동작 저작 인터페이스에 활용될 수 있다. 향후에는 실험자 수를 확대하고 손끝 궤적, 썸머링, 양손 협응에 대한 평가를 추가하며, 다양한 음악적 표현과 지휘 스타일을 적용할 예정이다.

참 고 문 헌 (References)

- [1] G. Pavlakos, V. Choutas, N. Ghorbani, T. Bolkart, A. A. A. Osman, D. Tzionas, and M. J. Black, "Expressive Body Capture: 3D Hands, Face, and Body from a Single Image," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, pp.10975-10985, June 2019.
doi: <https://doi.org/10.1109/CVPR.2019.01123>
- [2] A. Kirk, J. F. O'Brien, and D. A. Forsyth, "Skeletal Parameter Estimation from Optical Motion Capture Data," ACM SIGGRAPH 2004 Sketches, Los Angeles, CA, USA, p.29, August 2004.
doi: <https://doi.org/10.1145/1186223.1186260>
- [3] C. Crema, A. Depari, A. Flammini, E. Sisinni, T. Haslwanter, and S. Salzmann, "IMU-Based Solution for Automatic Detection and Classification of Exercises in the Fitness Scenario," 2017 IEEE Sensors Applications Symposium (SAS), Glassboro, NJ, USA, pp.1-6,

- March 2017.
doi: <https://doi.org/10.1109/SAS.2017.7894068>
- [4] A. R. Jiménez, F. Seco, F. Zampella, J. C. Prieto, and J. Guevara, "PDR with a Foot-Mounted IMU and Ramp Detection," *Sensors*, Vol.11, No.10, pp.9393-9410, October 2011.
doi: <https://doi.org/10.3390/s111009393>
- [5] A. K. Patil, A. Balasubramanyam, J. Y. Ryu, B. Chakravarthi, and Y. H. Chai, "An Open-Source Platform for Human Pose Estimation and Tracking Using a Heterogeneous Multi-Sensor System," *Sensors*, Vol.21, No.7, Article No.2340, April 2021.
doi: <https://doi.org/10.3390/s21072340>
- [6] K. Qiu, T. Qin, J. Pan, S. Liu, and S. Shen, "Real-Time Temporal and Rotational Calibration of Heterogeneous Sensors Using Motion Correlation Analysis," *IEEE Transactions on Robotics*, Vol.37, No.2, pp.587-602, April 2021.
doi: <https://doi.org/10.1109/TRO.2020.3033698>
- [7] J. González-Alonso, D. Oviedo-Pastor, H. J. Aguado, F. J. Díaz-Pernas, D. González-Ortega, and M. Martínez-Zarzuela, "Custom IMU-Based Wearable System for Robust 2.4 GHz Wireless Human Body Parts Orientation Tracking and 3D Movement Visualization on an Avatar," *Sensors*, Vol.21, No.19, Article No.6642, October 2021.
doi: <https://doi.org/10.3390/s21196642>
- [8] M. Yin, J. Li, and T. Wang, "A Low-Cost Inertial Measurement Unit Motion Capture System for Operation Posture Collection and Recognition," *Sensors*, Vol.24, No.2, Article No.686, January 2024.
doi: <https://doi.org/10.3390/s24020686>
- [9] M. Chen and G. Tan, "3D Human Pose Estimation Based on Wearable IMUs and Multiple Camera Views," *Electronics*, Vol.13, No.15, Article No.2926, August 2024.
doi: <https://doi.org/10.3390/electronics13152926>
- [10] G. Fei, X. Li, and R. Fei, "AVATAR: Autonomous Visual Authoring of Tangible Augmented Reality," *Proceedings of the 11th ACM SIGGRAPH International Conference on Virtual-Reality Continuum and Its Applications in Industry*, Singapore, pp.21-24, December 2012.
doi: <https://doi.org/10.1145/2407516.2407521>
- [11] K. C. Apostolakis and P. Daras, "RAAT—The Reverie Avatar Authoring Tool," *2013 18th International Conference on Digital Signal Processing (DSP)*, Santorini, Greece, pp.1-6, July 2013.
doi: <https://doi.org/10.1109/ICDSP.2013.6622788>
- [12] A. Balasubramanyam, A. K. Patil, B. Chakravarthi, J. Y. Ryu, and Y. H. Chai, "Motion-Sphere: Visual Representation of the Subtle Motion of Human Joints," *Applied Sciences*, Vol.10, No.18, Article No.6462, September 2020.
doi: <https://doi.org/10.3390/app10186462>
- [13] T. Wang, X. Qian, F. He, X. Hu, K. Huo, Y. Cao, and K. Ramani, "CAPturAR: An Augmented Reality Tool for Authoring Human-Involved Context-Aware Applications," *Proceedings of the 33rd Annual ACM Symposium on User Interface Software and Technology*, Virtual Event, USA, pp.328-341, October 2020.
doi: <https://doi.org/10.1145/3379337.3415815>
- [14] S. Anasingaraju, M. L. Wu, N. Adamo-Villani, V. Popescu, S. W. Cook, M. Nathan, and M. Alibali, "Digital Learning Activities Delivered by Eloquent Instructor Avatars: Scaling with Problem Instance," *SIGGRAPH Asia 2016 Symposium on Education*, Macau, Article No.5, pp.1-7, December 2016.
doi: <https://doi.org/10.1145/2993352.2993355>
- [15] P. Y. Chi, D. Vogel, M. Dontcheva, W. Li, and B. Hartmann, "Authoring Illustrations of Human Movements by Iterative Physical Demonstration," *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, Tokyo, Japan, pp.809-820, October 2016.
doi: <https://doi.org/10.1145/2984511.2984559>
- [16] D. Reda, J. Won, Y. Ye, M. van de Panne, and A. Winkler, "Physics-Based Motion Retargeting from Sparse Inputs," *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, Vol.6, No.3, Article No.33, pp.1-19, August 2023.
doi: <https://doi.org/10.1145/3606928>
- [17] A. Nijholt, D. Reidsma, R. Ebberts, and M. ter Maat, "The Virtual Conductor: Learning and Teaching about Music, Performing, and Conducting," *2008 Eighth IEEE International Conference on Advanced Learning Technologies*, Santander, Spain, pp.897-899, July 2008.
doi: <https://doi.org/10.1109/ICALT.2008.43>
- [18] T. Wang, N. N. Zheng, Y. Li, Y. Q. Xu, and H. Y. Shum, "Learning Kernel-Based HMMs for Dynamic Sequence Synthesis," *Proceedings of the 10th Pacific Conference on Computer Graphics and Applications*, Beijing, China, pp.87-95, 2002.
doi: <https://doi.org/10.1109/PCCGA.2002.1167842>
- [19] D. G. Dansereau, N. Brock, and J. R. Cooperstock, "Predicting an Orchestral Conductor's Baton Movements Using Machine Learning," *Computer Music Journal*, Vol.37, No.2, pp.28-45, June 2013.
doi: https://doi.org/10.1162/COMJ_a_00173
- [20] F. Liu, D. L. Chen, R. Z. Zhou, et al., "Self-Supervised Music Motion Synchronization Learning for Music-Driven Conducting Motion Generation," *Journal of Computer Science and Technology*, Vol.37, No.3, pp.539-558, June 2022.
doi: <https://doi.org/10.1007/s11390-022-2030-z>
- [21] Z. Zhao, J. Bai, D. Chen, D. Wang, and Y. Pan, "Taming Diffusion Models for Music-Driven Conducting Motion Generation," *Proceedings of the AAAI Symposium Series*, Vol.1, No.1, pp.40-44, October 2023.
doi: <https://doi.org/10.1609/aaais.v1i1.27474>
- [22] J. Oh, J. Jeong, and Y. Chai, "A Transfer Learning Approach for Music-Driven 3D Conducting Motion Generation with Limited Data," *Proceedings of the 30th ACM Symposium on Virtual Reality Software and Technology*, Trier, Germany, Article No.66, pp.1-2, October 2024.
doi: <https://doi.org/10.1145/3641825.3689531>
- [23] L. Severi, M. Sacchetto, A. Bianco, C. Rottondi, G. Abbate, A. Paolillo, and A. Giusti, "Remote Orchestral Conduction via a Virtual Reality System," *2024 IEEE 5th International Symposium on the Internet of Sounds (IS2)*, Erlangen, Germany, pp.1-6, September 2024.
doi: <https://doi.org/10.1109/IS262782.2024.10704182>

저 자 소 개



박 성 실

- 2014년 : 덕성여자대학교 디지털미디어학과 (학사)
- 2016년 : 중앙대학교 첨단영상대학원 영상학과 영상공학-CV/VR (석사)
- 2017년 ~ 2019년 : 한국전자기술연구원 연구원
- 2019년 ~ 2021년 : (주)싸인텔레콤 개발부 과장
- 2026년 : 중앙대학교 첨단영상대학원 영상학과 영상공학-CG/VR (박사)
- 2026년 ~ 현재 : 숭실대학교 전자정보공학부 및 G-LAMP 사업단 박사후연구원
- ORCID : <https://orcid.org/0000-0002-6113-6919>
- 주관심분야 : 의료 멀티모달 학습, 인간 동작 생성, 생성형 모델