

무대 공연을 위한 제스처 인식 기반 동적 프로젝션 맵핑 프레임워크 구현

고유진*, 김태원*, 최유주*[†]

*서울미디어대학원대학교 미디어공학과, [†]교신저자
ssummerr8.8@gmail.com, wingtgnw@naver.com, yjchoi@smit.ac.kr

Implementation of Dynamic Projection Mapping Framework based on Gesture Recognition for Stage Performance

You-Jin Koh*, Tae-Won Kim*, Yoo-Joo Choi*

*Dept. of Newmedia, Seoul Media Institute of Technology

요 약

본 논문에서는 미디어영상을 기반한 무대 공연의 다양한 미디어 효과를 분석하고, 무대 공연을 위한 제스처 기반 동적 프로젝션 맵핑 프레임워크를 설계 구현한다. 이를 위하여, 동적 프로젝션 맵핑 기반 기존 공연에서 공연자의 제스처와 이에 따른 미디어 효과를 분석하고, 동적 프로젝션 맵핑 기술을 효율적으로 구현하기 위하여 모션 히스토리 이미지를 이용한 CNN(Convolutional Neural Network) 기반의 제스처 인식 기술을 구현한다. 또한, 구현된 제스처인식 기술을 기반으로 공연자의 서로 다른 제스처와 미디어 효과를 매칭시킬 수 있는 프레임 워크 구현 내용을 소개한다.

1. 서론

기존의 프로젝션 맵핑 공연은 이미 만들어진 비디오와 공연 안무의 싱크를 맞추어 실제 공연 때는 마치 맵핑 비디오가 사람의 동작에 실시간 반응을 하는 것처럼 보이게 했다. 최근 제스처 인식 기술이 발전하면서 최근에는 공연자의 제스처에 따라 실시간 인터랙션 퍼포먼스가 가능한 동적 프로젝션 맵핑 연출이 관심을 모으고 있다. 대다수의 미디어 아트 퍼포먼스가 라이브 공연인 만큼 공연자의 동작에 따른 효율적인 미디어 이펙트에 대한 처리와 이를 보다 용이하게 조작할 수 있는 프레임워크가 요구되고 있다.

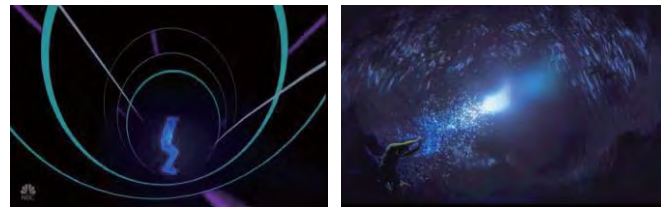
본 논문에서는 대표적인 동적 프로젝션 맵핑 공연 연출과 제스처 인식 연구 현황에 대해 간단히 알아보고 보다 효율적인 제스처 인식 구현을 위해 모션 히스토리 이미지를 이용한 CNN(Convolutional Neural Network)기반의 제스처 인식 실험을 진행하고, 제스처 인식을 기반한 무대 공연에 적합한 동적 프로젝션 맵핑 프레임워크를 설계 구현한다.

2. 프로젝션 맵핑 기반 미디어 아트 퍼포먼스와 제스처 인식기술 현황

2.1 미디어 아트 퍼포먼스와 동적 프로젝션 맵핑

프로젝션 맵핑은 투사하는 대상물의 평면에 가상의 무대 공간을 창조할 수 있기 때문에 빠른 무대 장면 전환을 통한 이야기 전개가 가능하다. (그림 1)은 공연자를 기준으로 가상의 3D 배경을 입혀 공연자가 마치 실제와 같은 효과를 준다. 여기에 제스처 인식 기술이 더해진 동적 프로젝션 맵핑은 공연자가 무대 위에서 자신의 몸짓만으로 메시지 전달을 할 수 있게 한다. (그림 2) 같이 제스처 인식과 인터랙션이 가미된 프로젝션 맵핑 기술은 예술가와 관객이

같은 공연 안에서 함께 즐길 수 있도록 만든다.



(그림 1,2) 프로젝션 맵핑 퍼포먼스 예시 1,2[2-3],

2.2 딥러닝 기반의 제스처 인식

최근 딥러닝을 기반한 제스처 인식 방법이 높은 성능을 보이고 있다. 제스처 및 동작 인식을 위한 딥러닝 접근법은 (그림 4)와 같이 크게 세 가지로 나뉜다. 하나는 학습에 사용하는 신경망의 고도화와 입력데이터의 전처리 그리고 RNN(Recurrent Neural Network)과 같이 연속 데이터(Sequence Data)를 이용한 시간적 방법들(temporal methods) 들이 있다[3].

본 논문에서는 데이터의 사용량을 줄여 성능을 높이는 모션기반 입력 특징값을 활용하는 방법을 택하여 모션 히스토리 이미지를 학습 데이터로 사용하는 CNN 기반의 제스처 인식 방법을 구현하고 이를 사용하였다.

3. 구현 및 실험

3.1 구현 제스처

<표 1>은 직관적 제스처 ‘기모이기’, ‘장풍 쏘기’, ‘발차기’로, 본 논문에서는 해당 제스처들을 인식하고 각 제스처에 해당하는

인터랙션을 구현한다.

<표 1> 구현 제스처 종류



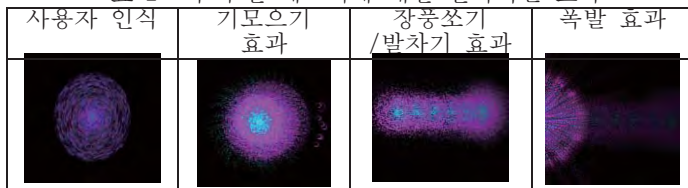
3.2 모션 히스토리 이미지 기반 제스처 인식

본 논문에서는 제스처 인식 실험에 대한 입력데이터로 Motion History Image(MHI)를 사용한다. MHI는 동작이나 제스처의 진행 경로를 이해할 수 있는 정적 이미지 템플릿으로, 여러 프레임으로 나누는 하나의 동작 데이터를 적은 메모리로 활용할 수 있다[4].

3.3 제스처 별 타겟 영상 효과

<표 2>는 제스처와 매칭되는 인터랙션 효과를 나타내고 있다. <표 6>의 ‘만세’ 제스처의 경우, 사용자를 인식하여 프로그램 시작을 판단하는 척도로 사용된다. ‘기모으기’ 제스처는 주변에 발산하던 빛들이 사용자의 중심으로 모여 뭉쳐지는 효과를 준다. ‘장풍 쏘기’는 공연자의 두 손 중심에서 손을 뻗는 방향으로 파티클 덩어리가 발사된다. ‘발차기’는 장풍쏘기와 같은 효과로 연결되고, 장풍쏘기/발차기의 마지막 시점에 불꽃 형태의 파티클이 주변으로 발산하여 폭발효과를 표현한다.

<표 2> 각 구현 제스처에 대한 인터랙션 효과



3.4 제스처와 미디어 효과의 매칭 정의

제스처와 매칭되는 미디어 효과는 텍스트로 구성되는 프로젝션 맵핑 구성 파일에 정의한다. <표 3>은 구성파일에서 제스처와 매칭 효과를 정의하는 규칙과 이에 따라 정의한 형식을 보여준다.

<표 3> 제스처와 미디어 효과 매칭을 위한 구성 파일 규칙

제스처/미디어 효과 정의 규칙	의미
Num_Gesture: 3	타겟 제스처 수
GestureDir: "gesture dir 명"	학습완료 배포데이터디렉토리명
Num_Effect:4	미디어 효과 수
Effect1: "ParticleEffect1"	파티클 효과번호
Effect2: "ParticleEffect2"	파티클 효과번호
Effect3: "ParticleEffect3"	파티클 효과번호
Effect4: "ParticleEffect4"	파티클 효과번호
Performer check 1: 1	Performer가 있다고 판단되는 1의 경우, 1번이펙트 매칭
Gesture2: 2	Gesture2과 2번이펙트 매칭
Gesture3: 3	Gesture3과 3번이펙트 매칭
Gesture3: 4	Gesture4과 4번이펙트 매칭

3.5 실험 방법

총 5명의 참가자가 한 동작마다 10회씩 반복한 후 각 영상을 프레임별로 쪼개어 이미지 데이터를 생성하였다. 4명의 데이터를 학습데이터로 사용하고 1명의 데이터는 테스트 데이터로 사용하였다. 최종

생성된 데이터셋에 대해 CNN을 활용하여 흑백 영상인 모션 히스토리 이미지 기반의 제스처 인식 실험을 진행하였다. <표 4>는 <표 1>의 제스처에 관한 MHI이다.

<표 4> 각 타겟 제스처에 대한 모션 히스토리 이미지



3.6 실험 결과

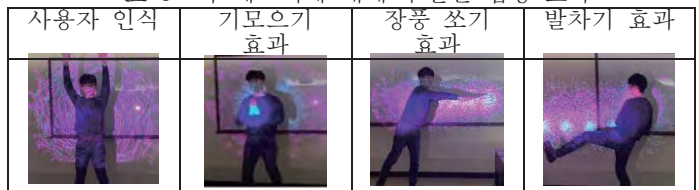
<표 5>는 각 제스처에 대한 인식 실험을 진행한 결과이다.

<표 5> 모션 히스토리 이미지 입력에 따른 제스처 인식률 및 하나의 제스처 인식에 걸리는 시간

Number of filters in Convolution Layer1	Number of filters in Convolution Layer2	Number of nodes in FC layer	Accuracy
32	32	32	0.90
16	16	16	0.95
8	8	8	0.91

<표 6>은 최종적으로 구현된 동적 프로젝션 맵핑 애플리케이션을 사용하여 실제로 제스처를 취했을 때 보여지는 인터랙션 맵핑 효과이다. 결과이다.

<표 6> 각 제스처에 대해 구현된 맵핑 효과



4. 결론

우리는 기존의 동적 프로젝션 맵핑 공연 기술의 유연한 공연 연출을 위해 MHI 기반의 제스처 인식 방법을 통한 동적 프로젝션 맵핑 프레임워크 구현을 시도하였다. 해당 논문에서 제안된 방법은 MHI라는 보다 가벼운 메모리를 통해 제스처를 인식하고 아웃풋을 재현한다. 본 논문에서 구현한 동적 프로젝션 맵핑 프레임워크는 동적 프로젝션에 대해 사용자의 경험을 보다 쉽게 유도할 수 있고 연출하고자 하는 맵핑 연출의 빠른 테스트가 가능할 것으로 예상된다.

사사의 글

이 연구는 “한국연구재단
이공학개인기초연구지원사업
(NRF-2017R1D1A1B03035718)”의 지원을 받아서 수행되었다.

참고문헌

- [1] The Most Amazing Multimedia Act Gets A Simo Cowell Standing Ovation!-America's Got Talent 2018, <https://youtu.be/B3ZlW4-BI4>
- [2] Multimedia Show / video mapping 360 / dome projection / Danel Stryjecki, <https://youtu.be/zrF52zwk284>
- [3] Mayam Asadi-Aghbolaghi, Albert Clapes, Marco Bellantonio. A survey on deep learning based approaches for action and gesture recognition in image sequences. Washington, D.C., USA IEEE 2017
- [4] Motion History Images from Wikipedia, https://en.wikipedia.org/wiki/Motion_History_Images