

얼굴인식을 활용한 실시간 인터랙티브 프로젝션 매핑+

조인재*, 김도희*, 이주현**, 김경아***, 최유주*,****

*서울미디어대학원대학교 뉴미디어학부 미디어공학전공

**동아방송예술대학 뉴미디어콘텐츠학과

***명지전문대학 컴퓨터공학과

****서울미디어대학원대학교 실감미디어연구소, 교신저자

e-mail : injae1028@gmail.com, galam33@hanmail.net, vincelee21@gmail.com,

kakim@mjc.ac.kr, yjchoi@smit.ac.kr

Real-time Interactive Projection Mapping Using Face Recognition

In-Jae Jo*, Do-Hui Kim*, Joohun Lee**, Kyong-Ah Kim, Yoo-Joo Choi*****

*Dept. of Newmedia, Seoul Media Institute of Technology

**Dept. of Newmedia Content, Dong-Ah Institute of Media and Arts

***Dept. of Computer Science and Engineering, Myongji College

****Immersive Media Lab., Seoul Media Institute of Technology

요 약

본 논문에서는 사각형의 형태를 벗어나 임의의 다각형 평면에 원하는 “카메라 입력 영상”, “비디오 클립”, 혹은 “3차원 그래픽 실시간 렌더링 영상”등을 보다 쉽게 매핑 시킬 수 있는 인터랙티브 프로젝트션 매핑 소프트웨어 시스템을 설계 구현하였다. 제안 시스템은 얼굴 인식 기능을 통하여 사용자 혹은 관객이 프로젝트션 매핑 작품 앞에 등장하였음을 인식하고, 관객의 모습이 미디어 콘텐츠의 일부로 실시간 포함되어 임의의 평면에 매핑하는 기능을 포함하고 있다. 제안 시스템은 프로젝트션 매핑의 초보자가 쉽게 사용할 수 있도록 텍스트 기반의 구성 파일 (Configuration File)에 매핑 평면과 미디어 콘텐츠의 형태 및 내용을 정의해 주도록 하는 구조로 구성하였다. 제안 시스템의 유용성을 확인하기 위하여, 육면체, 원구형, 사각 평면 형태의 실제의 객체에 다양한 형태의 미디어 콘텐츠를 매핑 한 미디어 작품을 제작하였다.

1. 서론

기술의 발전으로 문화(Culture)와 기술(Technology)이 만나 새롭게 형성된 하나의 기술 분야인 문화기술(CT) 분야가 빠르게 성장하고 있으며 그중 증강현실(AR)에 대한 관심이 많아지고 있다.

증강현실의 한 분야인 프로젝트션 매핑은 특정 공간이나 오브젝트에 그래픽 이미지나 영상을 투영하는 방식으로 새로운 공간 및 오브젝트를 만드는 기술을 말한다[1].

증강현실(AR)은 특정기기를 신체에 착용하고 증강되어진 정보를 받아들여 관객에게 보여주지만 프로젝트션 매핑은 다수의 관람자가 함께 관람할 수 있고 신체에 착용하는 것이 없어 행동이 자유롭기 때문에 작품 감상에 몰입도가 높아 광고, 전시, 및 공연과 미디어 아트 분야에서 많이 사용되어지고 있는 기법이다. 과거의 프로젝트션 매핑은 정적인 콘텐츠로 다양한 표면으로 이루어진 특정 오브젝트나 건물 외벽에 사용되었다. 기술의 발달됨에 따라 과

거의 프로젝트션 매핑 기법에 인터랙티브적인 요소가 합쳐진 콘텐츠가 제작되어지고 있다[2-5].

이러한 인터랙티브 프로젝트션 매핑을 제작하기 위해서는 다양한 하드웨어와 소프트웨어들이 필요하다. 기본적으로 필요한 장비는 컴퓨터, 프로젝터, 프로젝트션 매핑 소프트웨어, 영상 전환을 위한 스위처 등 작품제작자의 인터랙티브적인 요소를 담을 수 있는 다양한 기기 등이 필요하다.[2]

본 논문과 관련한 이전 연구에서는 초보자들이 프로젝트션 매핑을 정교하게 제작하고자 할 때 많은 학습시간을 필요로 한다는 점을 고려하여 보다 쉽고 정교하게 제작할 수 있고 다양한 연출을 할 수 있도록 애니메이션 기능이 포함된 프로젝트션 매핑 프레임워크를 제안하였다.

본 논문에서는 제안된 프로젝트션 매핑 프레임워크를 확장하여 얼굴인식 기능을 통하여 관객의 출연을 인지하고, 프로젝트션 되는 영상 콘텐츠가 실시간으로 바뀔 수 있는 확장된 프로젝트션 매핑 프레임워크를 제안하고자 한다. 제안 프레임워크의 유용성을 검증하기 위하여, 제안 프레임워크를 이용하여 관객이 전시물 앞에 머물렀을 때, 관객의 얼굴을 촬영하고, 관객의 얼굴이 프로젝트션 매핑

+ 본 연구는 한국연구재단 이공학개인지초연구지원사업 기본연구지원사업(NRF-2015R1D1A1A01059304)과 2017년도 서울미디어대학원대학교 엑스프로그램 지원을 받아 수행된 연구임(SMIT-X2.0-2017-012).

되는 전시물의 일부로 흡수되어 모자이크 영상을 실시간으로 생성하는 미디어 콘텐츠를 제작하였다.

2. 관련연구

최근 인터랙티브적인 요소를 미디어 아트에 적용하기 위해 다양한 기기들이 등장하고 있어 더 이상 작품이 고정되어 있는 것이 아니라 게임적인 요소를 적용함으로써 관객과 상호작용을 통해 작품을 재미있고 다양하게 감상할 수 있도록 유도하고 있다. 이에 미디어 아트 분야에서 다양하고 독창적인 시도들이 이루어지고 있다[6-8].

미디어 아트에 인터랙티브적인 요소를 적용하는 방식에는 관객의 이미지, 소리, 터치, 움직임 등 다양한 방법들이 있으며 관객의 행동을 입력하기 위한 마이크, 카메라, Kinect, Leap-motion 등 다양한 기기들이 인터랙티브한 작품을 만드는데 사용되어지고 있으며 인간의 감각과 비슷한 센서에 대해 표 1에서 보여주고 있다[9].

논문에서 제안된 프로젝션 매핑 프레임워크에서의 인터랙티브적인 요소는 관람객의 얼굴을 인식하여 콘텐츠와의 상호작용을 하기 위한 기기로 이미지 센서, 즉 카메라를 사용하였다.

표1. 인간의 감각과 센서의 종류

인간의 감각		센서의 종류
시각	빛	광센서
	이미지	이미지센서
청각	소리	음향센서
촉각	온도	온도센서
	습도	압력센서
오감이 아닌 센서		중력센서
		자가센서

3. 제안방법

3.1 사각형과 비사각형 형태 매핑 모드 분리

컴퓨터 그래픽에서 기본적으로 3차원 메쉬 모델을 표현할 때, 사각형 메쉬 또는 삼각형 메쉬로 나타내어 사용한다. 프로젝션 매핑 소프트웨어에서 다양한 형태의 대상체에 대한 매핑을 지원하기 위하여 보다 세밀한 메쉬 구조의 2차원 혹은 3차원 모델을 구성하고 이를 실제의 공간에 놓인 오브제에 프로젝션한다. 2차원 비디오 클립을 실제의 오브제에 프로젝션하기 위해서는 일반적으로 사각형 2차원 모델을 구성하고, 여기에 비디오 영상을 텍스처 매핑(texture mapping) 하여 표현한다. 이때, 2차원 사각형 모델은 그림 1에서 제시하는 것과 같이 다양한 세분화 레벨의 모델로 표현할 수 있다. 본 연구에서는 매핑된 텍스처의 왜곡을 최소화 하기 위하여 그림 1(d)와 같은 좌우상하 대칭 구조의 세분화 된 메쉬 구조를 사용하였다.

사각형태의 오브제에 매핑을 하는 경우에도 프로젝션되는 각도에 따라 발생할 수 있는 영상의 왜곡을 최소화 하기 위하여 그림 1(d)와 같은 세분화 된 메쉬 구조가 유리하다[10]. 이 경우, 사각형의 외부 꼭지점 (그림 1에서

V_0, V_1, V_2, V_3) 만을 조정하면 내부 구성 메쉬의 정점(vertex) 들이 네 개의 외부 꼭지점들의 위치에 따라 자동으로 보간(interpolation) 되는 기능이 요구된다. 그러나 대부분의 프로젝션 매핑 소프트웨어 들에서는 2차원 모델이 세분화된 메쉬구조로 구성된 이후에는 외부 정점을 이동함에 따라 내부의 정점이 자동으로 보간되는 기능이 지원되고 있지 않아, 내부의 구성 정점들을 하나하나 수동으로 위치 조정해주어야 하는 불편함이 있었다.

비사각형 형태의 오브제에 매핑을 하는 경우에는 그림 1(d) 형태의 사각형 모델을 우선 매핑 시키고자 하는 오브제의 위치로 옮겨 중점을 일치 시키고, 사각형 모델의 각 정점들을 세세히 이동하여 비사각형 형태의 오브제에 일치하도록 조정하여야 한다. 이때 정점 이동에 따른 이웃 정점에 대한 자동 보간 처리가 비활성화 되어야 한다. 이와 같이 프로젝션 매핑 대상 형태가 사각형인지 비사각형인지에 따라 내부 정점의 자동 보간 처리를 활성화 혹은 비활성화 하도록 조정함으로써 보다 편리하게 오브제의 형태에 프로젝션 되는 콘텐츠의 형태를 맞출 수 있도록 입력 모듈을 구성하였다.

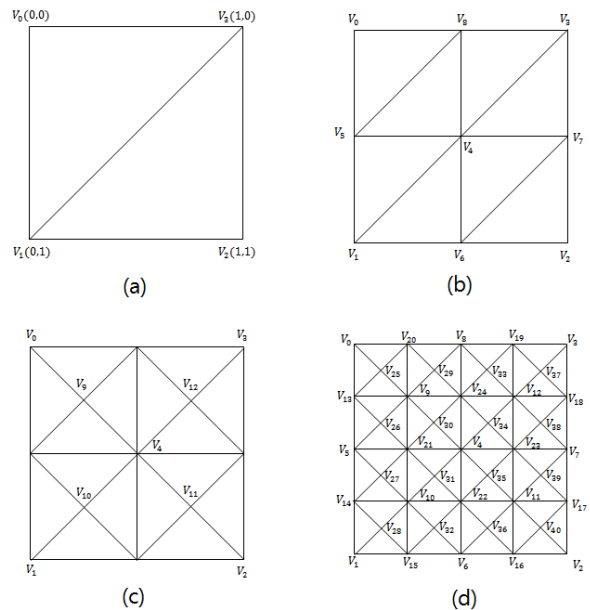


그림 1. 다양한 메쉬의 세분화 레벨. (a) Two_Triangles 가장 기본적인 메쉬 방법, (b) Eight_Triangles, (c) Sixteen_Triangles, (d) Forty_Triangles 제안된 방법

3.2 실시간 인터랙티브 프로젝션 매핑 설계

제안 시스템에서는 관람객이 전시물의 관람영역에 입장하였음을 인지하기 위하여 얼굴인식 기능을 사용하였다. 얼굴인식은 OpenFrameworks에서 제공하는 ofxCVHaarFinder 기능을 사용하였다[11]. 이때, 전시물 앞에 설치된 카메라로부터 초당 30 프레임의 속도로 카메라 영상을 입력받게 되고, 매 프레임 영상에 대하여 얼굴인식

을 수행하여 해당 프레임에 얼굴이 감지되는지를 검사한다. 연속된 프레임에서 30프레임 이상 얼굴이 감지되면 관람객이 전시물 관람 영역에 입장한 모드로 세팅하고 프로젝션된 콘텐츠의 변형을 수행하기 시작한다. 반대로 연속 30 프레임 이상에서 얼굴 감지가 연속적으로 이루어지지 않으면 관람객이 전시물 관람 영역을 벗어난 것으로 인지하고, 이에 따른 콘텐츠의 변형을 수행한다. 30프레임 이상으로 얼굴감지를 수행하여야 얼굴인지가 수행된 것으로 판단한 이유는 노이즈에 따른 얼굴감지 오류가 발생할 수 있기 때문이다.

실시간으로 관람객의 출현과 퇴장에 따라 프로젝션되는 그래픽 디스플레이의 내용을 수정하기 위하여, 매 프레임 영상의 렌더링 결과를 FBO(FrameBuffer Object) 에 저장하고, 이를 오브제에 매핑되는 2차원 모델의 텍스처로 사용하였다. FBO를 사용함으로써 임의의 오브제의 표면에 실시간으로 수정되어 렌더링되는 인터랙티브 콘텐츠를 프로젝션 시킬 수 있다.

본 논문에서 제안하는 얼굴인식을 활용한 인터랙티브 프로젝션 매핑 기술을 이용하여 관람객의 얼굴로 생성되는 실시간 모자이크 영상 프로젝션 매핑 콘텐츠를 제작하였다. 그림 2는 제안하고 있는 시스템에 대한 클래스 다이어그램을 보여주고 있다..

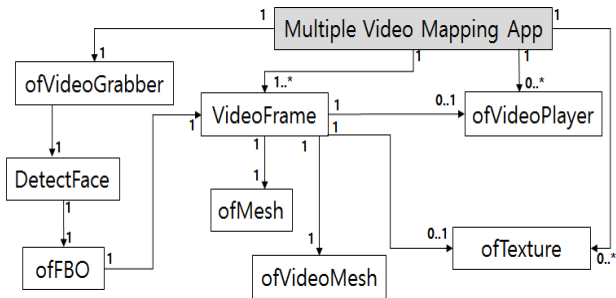


그림 2. 제안된 시스템 클래스 다이어그램

표 2. 프로젝션 매핑을 위한 구성파일 파라미터

입력 파라미터	정의
numVideo	사용되는 비디오 파일의 수
Video#	동영상 #번째 동영상의 파일이름
numImage	사용되는 이미지 파일의 수
Image#	이미지 #번째 이미지의 파일이름
numFrame	Frame 수
Frame#_v%_0	Frame크기 제어 부분
Frame#_v%_1	
Frame#_v%_2	
Frame#_ivideo	비디오, 웹캠영상, FBO 영상 선택
Frame#_iimage	사용되는 이미지 파일의 번호
Frame#_irefineMesh	Mesh 세분화 On/Off

제안 프레임워크는 프로젝션 매핑의 경험이 없는 초보자도 쉽게 프로젝션 매핑을 수행할 수 있도록 하기 위

하여 텍스트 형태의 구성 파일 (Configuration File)에 프로젝션 매핑의 내용을 입력하고, 이를 기반으로 프로젝션 화면이 로딩될 수 있도록 하였다. 프로젝션 화면이 로딩된 후에는 마우스를 이용하여 매핑되는 2차원 모델의 위치와 형태를 수정할 수 있도록 제작되었다. 표 2는 구성파일의 파라미터의 이름과 의미를 보여 주고 있다. 표 3는 구성파일의 예로서 1개의 사각 평면에 카메라 영상에 반응하는 FBO 영상을 프로젝션 매핑 시키는 구성파일의 예를 보여 주고 있다. 구성 파일은 “파라미터: 파라미터값” 으로 구성되어 있다. 여기서 physicsScreen은 관람객의 얼굴을 인식하여 저장된 이미지들이 FBO(Framebuffer Object)의 영상으로 만들어지도록 하는 기능을 사용하겠다는 의미이다. 이렇게 만들어진 텍스처에 사진은 멀리서 보면 하나의 큰 이미지이나 가까이에서 바라보면 관람객의 이미지가 랜덤하게 들어가 있게 되는 모자이크 영상으로 구성되었다. 또한 irefineMesh 파라미터를 통해 사각형 객체인지 비사각형 객체인지를 정의하고 이에 따라 내부 정점의 자동 보간 기능을 활성화시킬 것인지 비활성화 시킬 것인지가 결정되도록 구성되었다.

표3. 구성파일의 예

```
%YAML:1.0
numVideos: 1
Video1: "1.mov"
numImages: 1
Image1: "landscapeF.png"
physicsScreen: 1
thumbnailDir: "C:WW\ImagesWWtest_images"
mainImageFile: "C:WW\ImagesWWGogh_room_2.jpg"
outputImageFile: "C:WW\ImagesWWoutput.png"
fadeInSec: 3
fadeOutSec: 2
numFrames: 1
Frame1_irefineMesh: 0
Frame1_v0_0: 41
Frame1_v0_1: 343
Frame1_v0_2: 0
Frame1_v1_0: 52
Frame1_v1_1: 702
Frame1_v1_2: 0
Frame1_v2_0: 555
Frame1_v2_1: 582
Frame1_v2_2: 0
Frame1_v3_0: 508
Frame1_v3_1: 286
Frame1_v3_2: 0
Frame1_ivideo: -2
Frame1_iimage: 0
```

본 논문에서 제안하고 있는 프로젝션 매핑 프레임워크에서는 프로젝션 되는 미디어 콘텐츠의 포맷으로 2차원 스틸 이미지, 동영상, 카메라영상, 인터랙티브한 3D 그래픽 FBO 영상 등을 사용할 수 있다.

4. 실험결과

제안된 프로젝션 매핑 프레임워크는 Windows 10 환경

에서 OpenFrameworks로 구현되었다. 그림 3은 프로젝션 매핑이 이루어지기 전의 오브제의 설치 장면이다. 오브제들은 백색의 불투명 재질로 제작되었다. 그리고 그림 4는 각각의 오브젝트에 이미지와 영상을 투영한 결과 영상을 보여 주고 있다. 그림 5는 파티클 형태로 구성되어 애니메이션 되는 모자이크 FBO 결과 영상을 보여 주고 있다. 그림 5의 왼쪽 영상은 관람객의 출현 전의 영상이고, 오른쪽 영상은 관람객 출현 후 얼굴이 촬영되어 모자이크 영상안에 관람객의 얼굴이 추가 된후 모자이크 파티클이 애니메이션되고 있는 영상을 보여 주고 있다.



그림 3. 프로젝션 매핑 전의 오브제

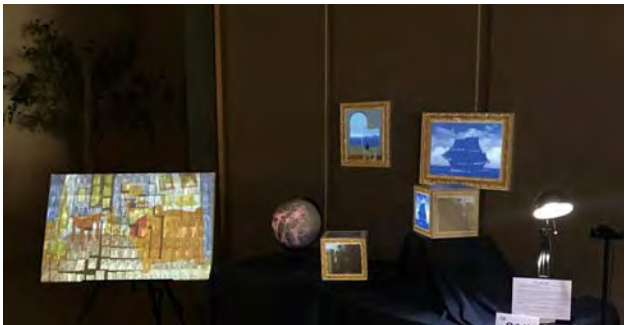


그림 4. 제안한 방법으로 구현된 시스템 구동화면



그림 5. 파티클 형태로 구성되어 애니메이션 되는 모자이크 영상

구현된 콘텐츠는 평상시에는 테이블 위에 있는 각각의 오브젝트와 벽에 있는 가상의 액자 그리고 캔버스에 한대의 빔 프로젝트로 10개의 면에 이미지와 영상이 투영되고 있으며, 관람객의 얼굴이 카메라를 통해 인식 되었을 경우 캔버스의 있는 이미지가 일정시간동안 인터랙티브하게 움직여 찰칵 소리와 함께 관람객 얼굴을 촬영되고 이미지로 저장된다. 캔버스의 이미지는 멀리서 바라보았을 때 하나의 이미지이지만 가까이에서 작품을 바라보면 저장된 이미지가 조각으로 배치되어 각각의 이미지로 보이게 된다.

5. 결론

기존의 프로젝션 매핑 콘텐츠는 관람객이 가만히 바라보는 콘텐츠였다면 논문에서 제안하는 방법의 프로젝션 매핑을 감상한 관람객들의 다수는 감상을 하였지만, 자신의 얼굴이 인식되는 것을 알게 된 순간 움직이는 그림에 흥미를 가지게 되는 결과를 볼 수 있었다. 또한 오브젝트에 이미지 또는 영상을 매핑시 조절할 때 필요한 키보드 조작은 현장에서 불편한 점이 있었다. 향후 연구로 제안된 시스템에서 더 나아가 다양한 기기를 활용한 인터랙티브 프로젝션 매핑 환경을 구축할 수 있는 프레임워크에 대해 연구 및 구현하고자 한다.

참고문헌

- [1] O. Bimber and R. Raskar, "Spatial Augmented Reality: Merging Real and Virtual Worlds", A. K. Peters, 2005.
- [2] T. P. V. Dumont, 온보란, 남양희, "Projection-mapping for Flexible Surface of Object", 한국 HCI 학술대회 논문집, pp. 155-157, 2014.
- [3] 이재운, 김연진, 김동호, "유연한 형태를 갖는 동적 객체 대상의 실시간 프로젝션 매핑", 한국 HCI 학술대회 논문집, pp. 187-190, 2014.
- [4] J.Y. Cho, Y. S. Moon, J. Y. Kim, "A Study on Real-Time Projection Mapping for Rotation Object", In Proceeding of ICCO, pp. 217-218, 2015.
- [5] 이청운, 박상운, "손 위치 트래킹 기반의 프로젝션 매핑 시스템 및 응용", 컴퓨터그래픽학회논문지 22(4), pp. 1-9, 2016.
- [6] Beam Art Group, Interactive projection mapping. <http://vimeo.com/beamart>.
- [7] 송민지, 박진완, "프로젝션 매핑 콘텐츠의 카메라 움직임 유형에 관한 연구", 한국콘텐츠학회논문지 제14권 제8호, pp. 1-12, 2014
- [8] 최봉화, "공연 예술을 위한 마스킹 프로젝션 매핑 시스템 개발", 학위논문(석사), 숭실대학교, 2013
- [9] 조옥희, 이원형, "인터랙티브 미디어 아트에 적용 가능한 제스처 인식 시스템 구현 연구", 디지털디자인학연구, 12(2), pp. 123-130, 2012. 4.
- [10] In-Jae Jo, Joo-hun Lee, Yoo-Joo Choi, "Simple Method of Video Mapping of Multiple Targets", Advances in Computer Science and Ubiquitous Computing, Lecture Notes in Electrical Engineering, Vol. 421, pp. 665-673, 2016. 11.
- [11] openFrameworks. <http://openframeworks.cc/documentation/ofxCvOpenCv/ofxCvHaarFinder/>