

멀티 제스처 인식을 이용한 실감 미디어 표출 기법

*양지희 **김영애 ***박구만

* **서울과학기술대학교 나노IT디자인융합대학원 방송통신융합프로그램, **서울과학기술대학교 일반대학원 미디어IT공학과

*jihee8929@seoultech.ac.kr **young5636@seoultech.ac.kr ***gmpark@seoultech.ac.kr

A Method of Immersive Media Display using Multi-Gesture Recognition

*Yang, Ji-hee **Kim, Young-ae ***Park, Goo-man

* **Graduate School of Nano IT Design Fusion, Seoul National University of Science & Technology

요약

최근 다양한 실감 미디어 기술과 표출 방법이 제시되고 있다. 특히 개인 방송에서 제작자 및 사용자에게 편리성을 제공하고 시청자의 실감성 체험을 극대화할 필요가 있다. 이에 따라 본 논문에서는 멀티 제스처가 필요한 실감 미디어를 멀티 디스플레이, 시점 선택, 3차원 객체 복원으로 정의하였다. 또한 사용자 및 제작자의 편리성을 위해 필요한 제스처에 대한 종류를 분석하였으며, 사용자의 다양한 제스처를 인식하고 실감 콘텐츠를 표출하도록 시스템을 구성하였다.

1. 서론

최근 다양한 실감 미디어 기술과 표출 방법이 제시되고 있다. 특히 자유시점 영상은 자유롭게 공간을 탐색할 수 있도록 임의의 시점에서 영상을 생성하며 사용자와 상호작용이 가능한 인터페이스 기술을 적용하여 다양한 실감 영상 표출이 가능하다[1]. 기존에 자유시점 영상 표출을 위해 반구형으로 배치된 카메라 배열을 제안하였으며, 이를 통해 실감 영상을 표출이 가능한 8개의 파라미터를 정의하였다[2]. 또한, 실감 영상 표출 중 하나로 디지털 사이니지나 미디어 폴, 비디오 월과 같은 대형 디스플레이를 통해 실감 콘텐츠 및 다양한 정보를 사용자에게 제공하였으며, 사용자와의 양방향 정보 제공 서비스를 위한 실감 미디어 기술이 연구되었다[3, 4].

본 논문에서는 개인 방송 내 표출 할 수 있는 실감 미디어를 3가지로 정의한다. 먼저 멀티 디스플레이는 사용자에게 사실감과 현장감을 제공하기 위해 시야 내에서 광각을 표출할 수 있다. 다시점 영상 획득 환경에서 시점 선택은 사용자가 원하는 시점의 영상을 자유자재로 선택하며, 3차원 객체 복원을 통해 실감 미디어 표출이 가능하다. 이는 1인 시청만 가능한 HMD 기술 보다 여러 사용자들과 공유 시청이 가능하며, 디지털 사이니지나 미디어 폴 등 다양한 실감 미디어 서비스에 응용될 수 있다.

개인방송의 제작 편리성과 실감 체험 극대화하기 위해 NUI 기술이 필요로 한다. 따라서 본 논문에서는 실감 미디어를 제어하기 위한 다양한 제스처에 대한 종류를 분석하고 이에 따른 실감 미디어 시스템을 구현하였다. 2장에서는 본 시스템에 대한 전체적인 구성 및 시나리오, 실험 결과에 대하여 설명한다. 3장에서는 결론 및 향후 연구 계획에 대하여 논한다.

2. 본론

그림 1은 본 논문에서 제안하는 시스템의 전체 구성도이다. 사용자의 제스처에 따라 다양한 실감 영상 표현이 가능하도록 디스플레이에 센서와 사용자 인터페이스를 설치하며, 반구형 구조로 배치된 카메라 배열에서 다시점 영상을 획득한다. 카메라를 통해 획득한 영상 정보에서 인식된 사용자의 제스처에 따라 크게 멀티 디스플레이 제어와 자유시점 영상 제어, 3차원 객체 제어로 나누어 시스템의 모든 동작과 수행이 이루어진다.

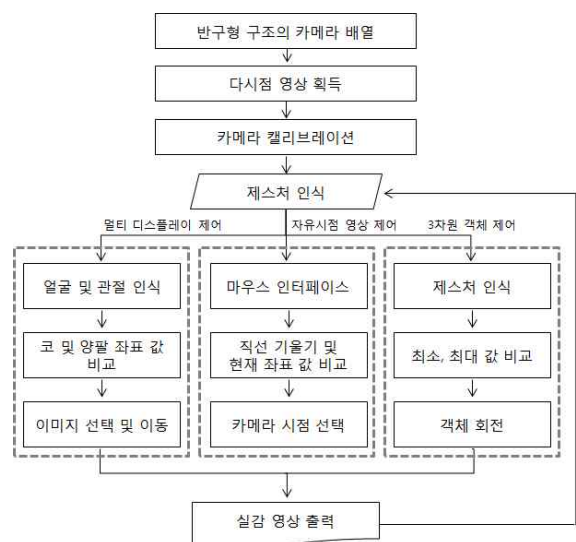


그림 1. 시스템 전체적인 구성도

2-1. 멀티 디스플레이 제어

멀티 디스플레이 제어는 3대의 모니터로 구현한 멀티 디스플레이에 파노라마 이미지를 출력하고 골격과 얼굴 정보에 따라 제어한다. 먼저 키넥트로 한 명의 사용자를 인식하고 신체에서 48개의 관절 정보를 획득하며, 골격의 x, y 좌표를 이용하여 제어에 사용할 정보를 선택한다. 양손을 몸과 평행하게 두는 차렷 자세와 차렷 자세에서 오른손을 머리 위로 올려 제어에 사용할 정보를 전환한다.

골격 정보에 의한 제어에서 파노라마 이미지 5장을 포함하는 1차원 배열을 생성하고, 추적한 양팔의 x, y 좌표 정보를 이용해 이미지를 선택한다. 어깨와 오른손 혹은 왼손을 수평이 되도록 동작을 하여 배열의 인자를 변경하여 이미지 선택이 되며 선택된 이미지를 출력한다.

얼굴 정보에 의한 제어 부분은 사용자 코의 x, y 좌표 정보를 추적하여 동작 패턴에 따라 선택된 이미지 콘텐츠 내에서 시야각을 상, 하, 좌, 우로 이동하며 출력한다. 그림 2와 같이 사용자가 정면을 바라보고 있을 때의 코의 x, y 좌표 값을 $nose(x, y)$ 으로 두고, 인식된 얼굴을 중심으로 정사각형의 관심 영역을 4분할하여 상, 하, 좌, 우를 판단한다. $nose(x, y)$ 와 정사각형의 꼭짓점을 지나는 대각선 2개를 그려 관심 영역을 4분할하며, 각 분할 영역마다 x와 y 값의 범위로 설정된 조건문에 부합할 경우 해당 방향으로 이미지를 이동시키며 출력한다 [4]. 그림 3은 키넥트 제스처 인식을 이용하여 멀티 디스플레이를 제어하며, 골격과 얼굴 동작에 따라 실감 콘텐츠 제어 및 출력되는 결과를 확인할 수 있다.

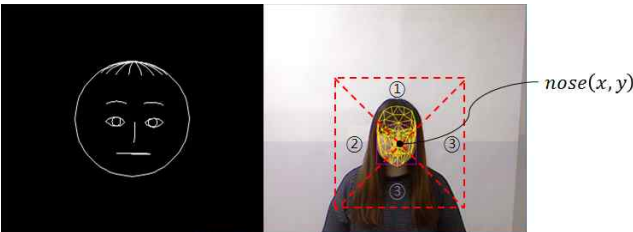


그림 2. 얼굴 정보에 의한 제어 부분



그림 3. 멀티 디스플레이 제어 실험 결과 화면

2-2. 자유시점 영상 제어

자유시점 영상 제어는 마우스의 x, y 좌표 정보에 따라 반구형 구조의 카메라들을 선택하며, 카메라 시점에 따른 영상을 출력한다. 객체가 역동적인 콘텐츠에서 시점이 변화하여도 지속적인 동작을 유지하도록 반구형 구조에서 다시점 동영상을 획득하며 프레임을 나누어 n차원 이미지 배열을 생성한다.

그림 4와 같이 마우스 입력의 초기 값을 $mouse(x, y_0)$ 로 두고, 이를 기준으로 정사각형 영역을 분할하여 상, 하, 좌, 우를 판단한다. $mouse(x_0, y_0)$ 을 중심으로 정사각형 형태의 영역을 지정하고, $mouse(x_0, y_0)$ 와 정사각형의 좌측 상단 점 $pt1(x, y)$ 을 지나는 직선과 우측 상단 점 $pt2(x, y)$ 을 지나는 직선을 그려 영역을 4분할한다. 두 직선에 대한 기울기 $pt1. rad = \frac{pt1.y - mouse.y_0}{pt1.x - mouse.x_0}$ 와 $pt2. Grad = \frac{pt2.y - mouse.y_0}{pt2.x - mouse.x_0}$ 을 각각 -1과 1로 두고, $mouse(x_0, y_0)$ 와 현재 마우스 좌표 값 $mouse(x_t, y_t)$ 을 지나는 직선을 그려 기울기 $Grad = \frac{mouse.y_t - mouse.y_0}{mouse.x_t - mouse.x_0}$ 와 비교하며 방향을 설정한다.

```

if ( Grad > 1 ) &&(mouse.y_t > mouse.y_0)
    CurrentX --;
else if ( Grad > 1 ) &&(mouse.y_t < mouse.y_0)
    CurrentX ++;
else if ( Grad < -1 ) &&(mouse.y_t > mouse.y_0)
    CurrentX --;
else if ( Grad < -1 ) &&(mouse.y_t < mouse.y_0)
    CurrentX ++;
    식 (1)
    
```

```

if ( Grad < 1 ) &&(mouse.x_t > mouse.x_0)
    CurrentY ++;
else if ( Grad < 1 ) &&(mouse.x_t < mouse.x_0)
    CurrentY --;
else if ( Grad < -1 ) &&(mouse.x_t < mouse.x_0)
    CurrentY --;
else if ( Grad < -1 ) &&(mouse.x_t > mouse.x_0)
    CurrentY ++;
    식 (2)
    
```

식 1, 식 2는 반구형 구조로 배치된 카메라에서 각각 위도, 경도를 제어하는 식으로 n차원 배열에서 인자 CurrentX와 CurrentY를 변경하여 이미지를 출력함으로써 자유시점 영상을 표출한다.



그림 4. 마우스 인터페이스에 의한 자유시점 제어부분

2-3. 3차원 객체 제어

사용자의 손이 움직이는 방향을 따라 객체 회전이 가능하도록 다 시점 영상에서 카메라 캘리브레이션을 통해 파라미터를 추정하고 3차원으로 객체를 복원한다. 립모션 컨트롤러를 통해 획득하는 손바닥 중심의 x 좌표 추적을 통해 객체 회전 방향을 제어한다.

먼저 립모션 컨트롤러로 한 명의 사용자에 대한 오른손을 인식하고, 손바닥 중심의 위치에 대한 벡터 값을 획득한다. 그 중 x 좌표 정보만 추출하여 기준 좌표 값 X_{std} 로 설정하고 경로를 추적한다. 제스처는 립모션 컨트롤러의 x축을 따라 X_{std} 을 기준으로 좌 또는 우로 왕복 이동하며 좌측으로 이동할 때에 손바닥 중심의 위치 X_{palm} 에 대한 최소 x 좌표 값 X_{min} 과 우측으로 이동할 때에 최대 x 좌표 값 X_{max} 을 구한다. 손을 좌로 이동 후 반대로 이동할 경우 X_{min} 값 보다 X_{max} 값이 X_{std} 와 비슷해지며, 이와 반대로 우로 이동 후 좌로 이동할 경우 X_{min} 값이 X_{std} 와 비슷해지므로 이를 X_{std} 와 비교하여 이동 방향을 설정한다.

```

if  $X_{palm} < X_{std}$ 
    if  $(X_{palm} < X_{min})$ ,
         $X_{min} = X_{palm}$ ;
    else
         $X_{min} = X_{min}$ ;

else if  $(X_{palm} > X_{std})$ 
    if  $(X_{palm} > X_{max})$ ,
         $X_{max} = X_{palm}$ ;
    else
         $X_{max} = X_{max}$ 
    식 (3)
    
```

식 3은 사용자 손바닥의 중심 좌표가 좌측으로 이동할 경우와 우측으로 이동할 경우를 나누어 X_{min} 과 X_{max} 을 구하는 식을 말한다.

```

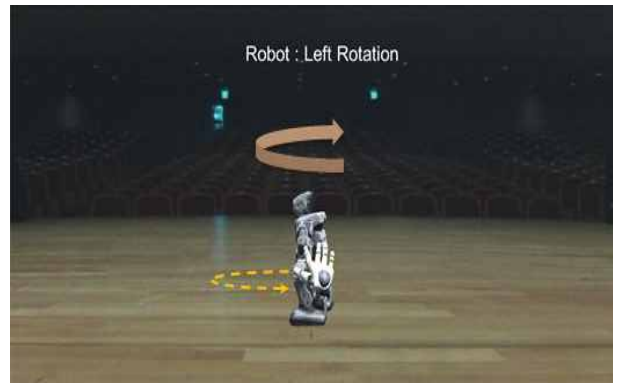
if  $(X_{max} < X_{std} + \alpha) \ \&\& \ (X_{max} < X_{std} - \alpha)$ 
    Object.Rotate(0, 0, 100);

else if  $(X_{min} < X_{std} + \alpha) \ \&\& \ (X_{min} < X_{std} - \alpha)$ 
    Object.Rotate(0, 0, -100);
    식 (4)
    
```

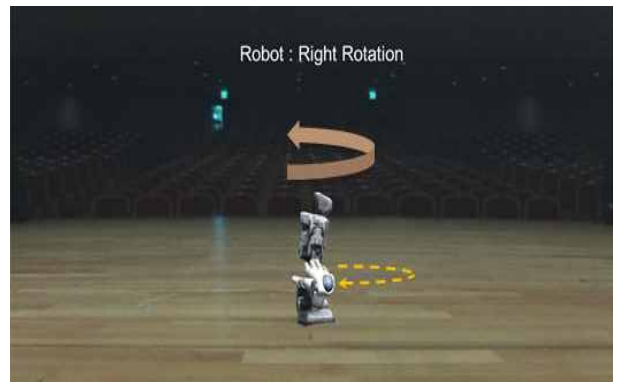
식 4는 제스처가 움직이는 방향에 따라 추정된 X_{min} 과 X_{max} 을 이용하여 객체 회전 방향을 설정하는 식으로 기준 좌표 값에 대해 사용자가 임의로 설정한 범위에 위치할 경우 객체가 회전한다.



(a)



(b)



(c)

그림 5. (a) 제스처에 의한 객체 제어, (b) 객체 좌측 회전 제어, (c) 객체 우측 회전 제어

개발환경은 Visual studio 2010에서 C언어를 기반으로 OpenCV, OpenGL 라이브러리를 사용하여 구현하였으며, Kinect와 LeapMotion을 사용하여 멀티 제스처 인식과 이에 따른 실감 미디어를 구현한 시스템 성능 실험을 진행하였다. 실험을 통해 얼굴, 관절, 손가락, 마우스 동작 등에 따라 실감 콘텐츠가 제어 및 출력되는 결과를 확인할 수 있었다.

3. 결론

본 논문에서는 개인 방송 내에서 표출이 가능한 실감 미디어를 3가지로 정의하였으며, 제작 편리성과 실감 체험 극대화를 위해 필요한 다양한 제스처 종류를 분석하였다. 이에 디스플레이에 설치된 센서를

통해 사용자의 관절, 얼굴, 손가락 동작 등을 인식하고 다양한 실감 미디어 표출 방법을 선택하여 디지털 실감콘텐츠를 제어 및 제공하도록 구성하였다. 이는 사용자와 디스플레이 사이의 상호작용이 가능할 뿐만 아니라 몰입감과 현장감을 극대화 하는 콘텐츠를 제공받을 수 있다.

향후 연구로는 본 논문에서 스마트 기기 및 관련 영상 장비를 통해 VR, AR, '360도 영상' 등 다양한 실감 미디어를 구성할 것이다. 또한 제스처 이외에도 콘텐츠 제어가 가능한 사용자의 인터페이스를 추가하여 양방향 실감 미디어의 성능을 높이기 위한 연구를 진행할 것이다.

감사의 글

본 연구는 미래창조과학부 및 정보통신기술연구진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [B0717-16-0012, 제작 편리성과 실감 시청 체험 극대화를 위한 개인방송 제작 기술 개발]

참고문헌

- [1] 호요성, “FTV 3차원 비디오 부호화 표준화 동향”, 한국정보통신 기술협회, 2015
- [2] Jihee Yang, Jihye Jeon, Gooman Park, " Defining Parameter Set for Interactive Free Viewpoint Video in Hemispherical Camera Placement Framework", EEECS 2016, 2016
- [3] 정효택, 윤기승, 손욱호, “고해상도를 지원하는 디지털 사이니지 기술 및 산업동향”, 차세대 콘텐츠기술 특집, Electronics and Telecommunications Trends, 제29권, 제1호, pp.72-82, 2014
- [4] 김영애, 박구만, “키넥트의 관절 추적과 얼굴 인식을 이용한 멀티 디스플레이 제어 기법”, 한국통신학회 추계종합학술발표회 논문집, pp.63-64, 2015