

3차원 인체 포즈 인식을 이용한 상호작용 게임 콘텐츠 개발

Developing Interactive Game Contents using 3D Human Pose Recognition

최윤지, 박재완, 송대현, 이철우
전남대학교 전자컴퓨터공학부

Yoon-Ji Choi(yoonji@image.chonnam.ac.kr), Jae-Wan Park(cyanlip@image.chonnam.ac.kr),
Dae-Hyeon Song(min2man@nate.com), Chil-Woo Lee(leecw@image.chonnam.ac.kr)

요약

일반적으로 비전기반 3차원 인체 포즈 인식 기술은 HCI(Human-Computer Interaction)에서 인간의 제스처를 전달하기 위한 방법으로 사용된다. 특수한 환경에서 단순한 2차원 움직임 포즈만 인식할 수 있는 2차원 포즈모델 기반 인식 방법에 비해 3차원 관절을 묘사한 포즈모델은 관절각에 대한 정보와 신체 부위의 모양정보를 선행지식으로 사용할 수 있어서 좀 더 일반적인 환경에서 복잡한 3차원 포즈도 인식할 수 있다는 장점이 있다. 이 논문은 인체의 3차원 관절 정보를 이용한 포즈 인식 기술을 인터페이스로 활용한 상호작용 게임 콘텐츠 개발에 관해 기술한다. 제안된 시스템에서 사용되는 포즈는 인체 관절 중 14개 관절의 3차원 위치정보를 이용해서 구성된 포즈 템플릿과 현재 사용자의 포즈를 비교해 인식된다. 이 방법을 이용하여 제작된 시스템은 사용자가 부가적인 장치의 사용 없이 사용자의 몸동작만으로 자연스럽게 게임 콘텐츠를 조작할 수 있도록 해준다. 제안된 3차원 인식 기술을 게임 콘텐츠에 적용하여 성능을 평가한다. 향후 다양한 환경에서 더욱 강건하게 포즈를 인식할 수 있는 연구를 수행할 계획이다.

■ 중심어 : | HCI | 포즈 인식 | 상호작용 인터페이스 |

Abstract

Normally vision-based 3D human pose recognition technology is used to method for convey human gesture in HCI(Human-Computer Interaction). 2D pose model based recognition method recognizes simple 2D human pose in particular environment. On the other hand, 3D pose model which describes 3D human body skeletal structure can recognize more complex 3D pose than 2D pose model in because it can use joint angle and shape information of body part. In this paper, we describe a development of interactive game contents using pose recognition interface that using 3D human body joint information. Our system was proposed for the purpose that users can control the game contents with body motion without any additional equipment. Poses are recognized comparing current input pose and predefined pose template which is consist of 14 human body joint 3D information. We implement the game contents with the our pose recognition system and make sure about the efficiency of our proposed system. In the future, we will improve the system that can be recognized poses in various environments robustly.

■ keyword : | HCI | Pose Recognition | Interactive Interface |

* 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음.

(NIPA-2011-C1090-1111-0008)

접수번호 : #110901-003

접수일자 : 2011년 09월 01일

심사완료일 : 2011년 10월 14일

교신저자 : 이철우, e-mail : leecw@jnu.ac.kr

I. 서론

제스처 인터페이스기술(Gesture Interface Technology)은 최근 미국 매사추세츠공과대학(MIT)이 발표한 '2011년 10대 유망기술'[1] 중 하나로 선정되었다. 현재 제스처 인터페이스 기술은 가상현실, 스마트 감시 시스템, 항공 교통 모니터링 시스템, 게임, 스포츠와 같은 엔터테인먼트 등 여러 분야에서 응용되고 있다. 더욱이 최첨단의 제작기술과 표현기술의 발전은 디지털 콘텐츠 시장을 급속도로 성장시켜 사람과 컴퓨터간의 효과적인 상호작용을 위한 더욱 편리하고 자연스러운 사용자 인터페이스(User Interface)를 요구하고 있다. 따라서 입력 장치를 사용하지 않고 컴퓨터와 상호작용하기 위한 인간-컴퓨터 상호작용 기술의 필요성은 점차 증대되고 있으며 특히 제스처 인식 기술과 멀티모달 인터페이스에 대한 연구의 필요성이 함께 높아지고 있다.

자연스러운 사용자 인터페이스의 목표는 인간 몸동작만으로 컴퓨터가 인간의 의도를 인지하는 것이다. 인간의 포즈는 다수의 3차원 관절의 움직임으로 구성된다. 그렇기 때문에 컴퓨터가 인간의 동작을 분석하고 인식하는 것은 매우 어렵고 복잡한 일이다. 이 같은 문제를 해결하기 위한 방법으로 움직임 감지가 가능한 센서나 기타 장치를 사용자의 신체 일부에 부착하여 사용자의 움직임 정보를 획득하는 센서 기반의 제스처 인터페이스 방법이 사용되어 왔다. 이 방법은 비교적 고가의 장비를 사용하기 때문에 사용자의 정확한 움직임 정보를 획득할 수 있다는 장점을 갖고 있으나 장치를 몸에 부착해야하기 때문에 자연스러운 움직임을 방해 받는다는 단점이 있다. 따라서 부가적인 장치의 사용으로 인한 불편을 해소하기 위해 컴퓨터 비전 기술만을 이용한 시각 기반 제스처 인식 기술 연구가 활발히 진행되고 있다. 시각 기반의 제스처 인식 기술이란 카메라를 통해 입력된 영상을 이용해 인간의 행동을 분석하는 기술로써 카메라 영상으로부터 제스처를 인식하기 위해 하나 혹은 그 이상의 카메라를 이용하여 제스처 영상을 획득한 다음 획득된 영상으로부터 사용자의 움직임을 추적하고 인식한다. 일반적으로 제스처 인식 과정은 입력 영상으로부터 관심영역인 신체의 일부를 분

할(Segmentation)하는 과정, 인체의 움직임 정보에서부터 의미 있는 정보를 추출하는 특징추출(Feature Extraction) 과정, 추출된 특징을 분석해 사용자의 움직임 인식(Recognition)하는 과정을 거친다. 이 방법은 부가적인 장비 사용 없이 몸동작만을 사용하기 때문에 사용자의 직관적인 움직임을 통해서 다양한 콘텐츠와 상호작용할 수 있는 인터페이스로 기대된다[2].

본 논문은 이러한 시각 기술 기반으로 한 3차원 포즈 인식 시스템의 개발과 이것을 인터페이스로 활용하여 사용자의 직관적인 움직임으로 콘텐츠와 상호작용 이 가능한 콘텐츠 개발에 관해 기술한다. 본 논문의 구성은 다음과 같다. II장에서는 시각 기술 기반의 인체 포즈 인식기술에 대하여 살펴본다. III장에서는 상호작용 게임 콘텐츠 개발을 위한 3차원 인체 포즈 인식 기술에 대해 기술한다. IV장에서는 다중 사용자 상호작용 게임 콘텐츠 제작 기술에 대해 기술한다. V장에서는 실험결과를 살펴보고 마지막으로 VI장에서는 결론을 기술하고 논문은 맺는다.

II. 시각기반 인체 포즈 인식 기술

시각 기반의 인체 포즈 인식 연구 분야는 관점에 따라 다양하게 나누어 질 수 있다. 본 논문에서는 시각 기반의 인체 포즈 인식 방법을 2D 방법과 3D 방법으로 나누어 살펴보았다. 2D 기반의 포즈 인식 방법은 인체 부분별 검출을 통한 포즈 추정 방법과 포즈의 모양이나 특징 추출을 통한 포즈 추정 방법으로 나누어진다. 그리고 3D기반의 포즈 인식 방법은 카메라의 수에 따라 단일시점기반 방법(single-view)과 다시점기반 방법(multi-view)로 나눌 수 있다.

2.1 2D 기반 포즈 인식 방법

2.1.1 인체 부분별 검출을 통한 포즈 인식

인체 부분별 검출을 통한 포즈 인식 방법은 신체 각 부분에 대한 단서를 이용해 영상에서 팔, 다리, 머리 등의 위치를 찾고 이를 이용해 포즈를 인식하는 방법을 말한다. 확률적인 방법과 contour 분석을 통해 인체의

각 부분들의 위치를 찾아낸 후 각 부분간의 관계를 고려해 포즈를 인식해야 하기 때문에 각 부분간의 위치 관계에 대한 사전 정보가 필요하다[3]. Haritaoglu et al.[4]은 단일 실루엣 영상을 분석하여 6개의 신체 부위를 검출하였으며 실루엣 영상의 수직, 수평 히스토그램과 데이터베이스 히스토그램과의 비교를 통해 유사도가 가장 높은 포즈로 결정하는 방법을 제안했다. 이러한 방법은 관절의 관계나 자유도를 크게 고려하지 않기 때문에 비교적 간단한 방법이라는 장점이 있지만 실루엣에 나타나지 않은 신체 부위와 관절들을 표현하기 어려운 단점이 있다.

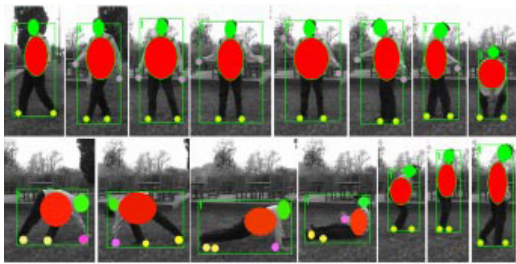


그림 1. 신체 부분별 검출을 통한 포즈인식

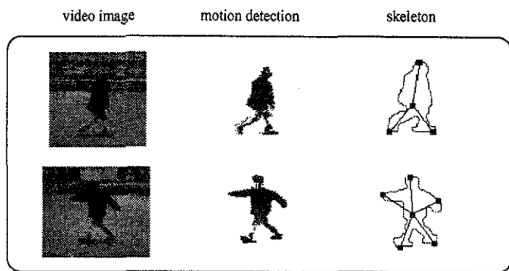


그림 2. Star Skeleton 방법

2.1.2 특징 추출을 통한 포즈 인식

특징 추출을 통한 포즈 인식 방법은 각 관절 사이의 관계나 인체 부분간의 관계를 고려하지 않고 포즈의 에지, 윤곽선, 점의 위치 등과 같은 정보를 이용해 포즈를 인식하는 방법이다. 따라서 특징을 검출하기 전에 신체 영역을 배경으로부터 정확히 분리하는 것이 중요하다. H. Fujiyoshi et al.[5]은 포즈에 대한 실루엣 영상에 별 모양의 Star Skeleton을 생성하여 포즈를 표현하였다.

먼저 포즈의 중심점과 외곽선 위의 점들과의 거리를 구한다. 그 후 중심과 가장 먼 점들을 extremity로 결정하고 중심점과 extremity를 연결하는 벡터를 Star Skeleton을 생성한다. 이 때 생성된 벡터를 포즈 인식을 위한 특징으로 사용한다. Baumbaerg와 Hogg[6]는 실루엣 영상의 에지 중 두드러진 점들을 eigenshape 모델을 이용해 인식하는 방법을 제안하였다. 영상의 특징을 기반으로 한 포즈 인식 방법은 실시간 인식 시스템에 적합하지만 카메라의 위치에 따라 인식률의 변화가 크다는 단점이 있다.

2.2 3D 기반 포즈 인식 방법

3D 기반의 포즈 인식 방법은 카메라로부터 획득한 입력 영상에서 3차원 관절 정보를 추정하여 인체의 포즈를 인식하는 방법이다. 제약조건이 많은 2D 기반의 포즈 인식 방법에 비해 3D 기반의 포즈 인식 방법은 좀더 일반적인 환경에서 복잡한 3차원 움직임 포즈도 인식할 수 있고 관절의 자유도를 대부분 그대로 구현하고 있기 때문에 더 나은 포즈를 추정할 수 있다[7].

2.2.1 단일 시점 기반 방법(Single-view)

단일 시점 기반 방법은 하나의 카메라를 이용해 인체의 3차원 포즈를 추정하는 방법이다. 단일 시점 기반 방법에는 다 시점에서 촬영한 인체의 데이터베이스와 단일 영상에서 추출한 특징을 비교해 인체의 관절을 찾고 인체의 3D 포즈를 추정하는 방법이 있다[8]. 그리고 영상에 투영된 포즈와 실루엣의 외곽선 사이의 관계를 통해 회전각을 결정하고 이 회전각의 변화로 인체의 포즈를 추정하는 방법[9], 가상공간에서 3D 인체를 모델링하고 가능한 모든 포즈에 대한 실루엣 영상을 데이터베이스화하여 입력된 영상과 데이터베이스를 모두 비교해 포즈를 인식하는 방법[10] 등이 있다. 이와 같은 방법들은 단일 카메라를 사용하여 시점에 의존적이 때문에 높은 인식률을 얻기 위해서는 많은 수의 샘플 데이터가 필요하다.

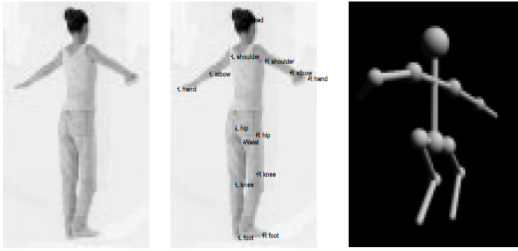


그림 3. 단일 시점 기반 방법(Single-view)

2.2.2 다 시점 기반 방법(Single-view)

다 시점 기반의 3D 포즈 인식 방법은 시점에 의존적인 2D 기반의 포즈인식 방법과 단일 시점 기반의 3D 포즈 인식 방법의 단점을 보완해 준다. 또한 다수의 카메라를 사용하기 때문에 회전, 가려짐 등의 문제에 대해서 비교적 자유롭다. D.M. Gavril et al[11]은 4대의 카메라를 설치한 후 카메라 파라미터 추정을 통해 3차원 정보를 얻어, 이 정보를 이용해 3차원 인체 모델을 구성한 후 각 카메라 영상에서 얻은 이차원 영상과 모델과 chamfer 매칭을 통해 3차원 포즈를 검출하였다. Chohen et al[12]은 4대의 카메라를 통해 얻어진 실루엣 영상으로 3차원 visual hull을 생성하여 3차원 포즈를 인식하였다. 박재완[13]은 2개의 적외선 카메라를 사용하여 상반신의 정면과 측면에 대한 적외선 영상을 획득하고, 획득된 포즈들의 연속으로 상반신 제스처를 인식하였다.

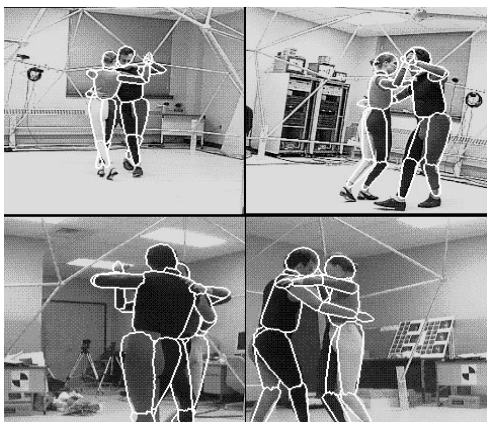


그림 4. 다시점 기반 방법(Multi-view)

3차원 인체 포즈 인식 방법의 사용은 신체 부위의 모양정보(shape)와 관절 사이의 한계각도 정보를 선행지식으로 사용할 수 있는 장점이 있으며 이러한 정보를 이용하여 신체 부위별 겹침 현상에 대한 문제점을 해결할 수 있다.

III. 3차원 인체 포즈 인식 시스템 구현

3.1 관절각의 추정

인체의 3차원 포즈를 추정하기 위해서는 입력된 영상에서 배경을 제거하고 인체의 형상을 정확히 분할해야 한다. 하지만 영상에서 인체를 분할하는 것은 매우 복잡하고 어려운 과정이다. 본 논문에서는 포즈의 특징을 정의하고 인식하는 것에 중점을 두고 있기 때문에 인체 형상을 분할하는 문제를 간단히 하기 위해 Microsoft사의 Kinect[14] 카메라와 NUI Skeleton API[15]를 사용해 인체 영역을 분할하고 인체의 관절각 정보를 획득하였다. NUI Skeleton API는 Kinect 센서 앞에 있는 사람을 인식해 신체의 주요 20개 부위에 해당하는 관절 [표 1] 좌표와 회전 정보를 제공해 준다. 인체의 각 관절은 (x, y, z)의 3차원 좌표 값을 갖는다.

표 1. Kinect에서 제공하는 20개 신체 관절

번호	관절 이름	번호	관절이름
1	HEAD	11	SPINE
2	SHOULDER_CENTER	12	HIP_CENTER
3	SHOULDER_RIGHT	13	HIP_RIGHT
4	SHOULDER_LEFT	14	HIP_LEFT
5	ELBOW_RIGHT	15	KNEE_RIGHT
6	ELBOW_LEFT	16	KNEE_LEFT
7	WRIST_RIGHT	17	ANKLE_RIGHT
8	WRIST_LEFT	18	ANKLE_LEFT
9	HAND_RIGHT	19	FOOT_RIGHT
10	HAND_LEFT	20	FOOT_LEFT

NUI Skeleton API에서는 다음과 같은 과정을 통해 인체의 3차원 관절 정보를 얻는다[그림 5]. 먼저 Kinect의 적외선 카메라로부터 영상이 들어오면 깊이 영상이

생성되고 깊이 정보를 이용해 사람과 배경을 분리한다. 배경과 분리된 인체 영역은 학습된 데이터와 깊이 영상의 특징에 의해, 관절의 관심 영역에 위치한다고 추정되는 신체의 각 부분으로 나누어 레이블링 된다. 레이블링 된 신체의 각 부분은 다시 3차원 공간으로 재투영된다. 각각의 신체 부분의 공간적인 분포에 따라 신뢰 가중치를 부여한 후 각 관절의 3차원 위치가 결정된다[16].

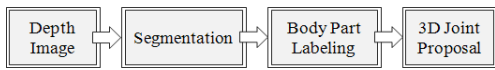


그림 5. 관절 정보 획득 과정

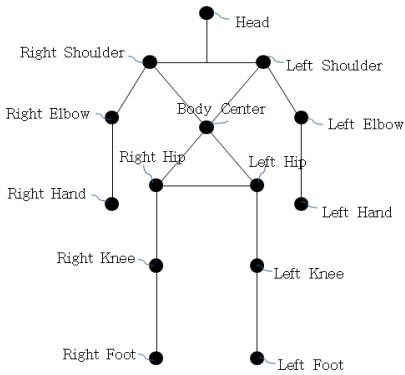


그림 6. 실험에 사용한 14개의 신체 관절

3.2 특징 추출

본 논문에서는 총 16개의 포즈를 정의하였으며, 각 포즈별로 개별적인 포즈 템플릿을 구성해 포즈를 인식하였다. 각 포즈 템플릿은 14개의 관절에 대한 관절 벡터의 모음으로 정의할 수 있다. 각 포즈에 대한 특징을 추출하기 위해 NUI Skeleton API에서 제공하는 20개의 인체 관절 중 왼쪽 손목, 오른쪽 손목, 왼쪽 발꿈치, 오른쪽 발꿈치, 어깨 중심점 그리고 엉덩이 중심점을 제외하고 14개 관절(머리, 오른쪽 어깨, 왼쪽 어깨, 오른쪽 팔꿈치, 왼쪽 팔꿈치, 오른손, 왼손, 몸 중심, 오른쪽 엉덩이, 왼쪽 엉덩이, 오른쪽 무릎, 왼쪽 무릎, 오른발, 왼발)을 사용했다. 제외시킨 6개의 관절은 포즈가 달라져도 큰 변화를 갖지 않거나 손, 발과 같은 관절과 매우

가까워 해당 관절 값의 사용 유무가 포즈를 정의함에 있어 큰 의미를 갖지 않는 관절이다. 획득한 14개 신체 관절[그림 6]의 3차원 좌표(x,y,z)를 얻은 후 14개 신체 관절의 3차원 위치 좌표 정보를 바탕으로 각 포즈에 대한 개별적인 포즈 템플릿을 구성하였다. 각 포즈 템플릿은 14개 관절에 대한 관절 벡터로 구성된다. 즉, 하나의 포즈 템플릿은 14개의 벡터 특징을 갖는다.

$$P_i = [V_1, V_2, V_3, \dots, V_{14}] \quad (1)$$

식(1)의 $P_i(i=1 \text{ to } 16)$ 는 각 포즈에 해당하는 포즈 템플릿을 의미하고 V_i 는 관절 벡터를 의미한다.

3.3 포즈 인식 방법

앞 절에서 설명한 것과 같이 16가지 유형의 포즈를 인식하기 위해서 각 포즈에 해당하는 템플릿을 구성해야 한다. 제안하는 포즈 시스템은 2명의 사용자에게 대한 포즈를 각각 인식할 수 있게 구성하였다. 이 시스템은 사용자가 들어오면 캘리브레이션 및 배경 제거 후 각 사용자에게 대한 3차원 좌표 정보를 획득한다. 사용자가 정의된 16가지 포즈 중 임의의 포즈를 취하면 사용자의 포즈와 가장 가까운 포즈 템플릿을 계산해서 구한다. 포즈 번호를 구하기 위한 방법은 아래 식(2)로 표현한다.

$$PoseNum = \operatorname{argmin}(|P_i - X|^2) \quad (2)$$

식(1)에서 P_i 는 미리 정의된 포즈 템플릿을 의미하고 X 는 현재 사용자의 신체 관절 정보를 의미한다. 포즈는 사용자에게 따라 약간의 오차가 발생할 가능성이 있기 때문에 템플릿을 구성하는 관절 관계의 특징에 대해 5%의 오차 범위를 두고 포즈를 인식했다. [그림 7]은 포즈 인식의 개략적인 흐름을 보여준다.

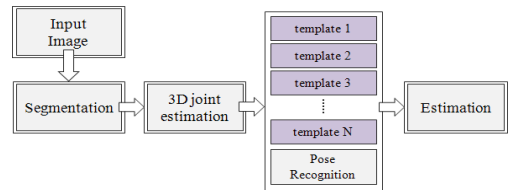


그림 7. 포즈 인식

IV. 다중사용자 포즈 인식 콘텐츠

4.1 콘텐츠의 개요

본 논문에서 제안하는 다중 사용자 포즈 인식 콘텐츠는 3차원 깊이 카메라를 이용한 포즈 인식과 모바일을 인터페이스로 활용한 포즈 인식 게임으로 화면에서 다가오는 벽을 통과하는 게임이다. 사용자는 다가오는 벽에 뚫린 구멍과 똑같은 포즈를 취해야 벽을 통과해 점수를 획득하며 벽은 16개의 포즈 중에 임의로 선택된다. 두 명의 사용자가 함께 참여할 수 있으며 협업모드와 대결모드를 선택한다. 그리고 각 사용자는 모바일의 진동으로 벽과의 충돌 여부를 알게 되며, 모바일을 사용해 공격과 방어 아이템을 사용한다.

시스템은 [그림 8]과 같이 콘텐츠의 시나리오 진행에 따른 영상부분, 사용자의 포즈를 인식하는 인식 부분과 모바일 인터페이스를 위한 통신부분으로 구성된다.



그림 8. 시스템 개요

4.2 포즈 정의



그림 9. 포즈 정의

본 콘텐츠는 [그림 9]와 같이 총 16개의 포즈를 정의한다. 포즈는 8개의 정면포즈(Pose1, 2, 3, 5, 6, 7, 10, 11)와 8개의 측면포즈(Pose 4, 8, 9, 12, 13, 14, 15, 16)로 구성된다. 정면포즈의 경우 포즈 템플릿을 정의할 때 깊이 값이 포즈 변화에 큰 영향을 미치지 않기 때문에 x와 y좌표 값만을 사용하였다. 정면포즈와 달리 측면포즈의 경우 각 관절의 깊이 정보가 포즈 변화의 특징이 되기 때문에 3차원 좌표 값을 모두 사용해 포즈 템플릿을 정의하였다.



그림 10. 모바일 대기 화면

4.3 모바일 인터페이스

게임의 중요한 인터페이스 요소는 사용자의 몸동작이다. 이와 더불어 게임에 실감나는 효과를 더하기 위해 모바일 인터페이스를 추가로 구성하였다. 모바일은 도입부에서 사용자의 준비 상태를 전송받기 위해, 게임 도중 아이템을 사용하여 상대방을 공격하거나 방어하기 위해, 그리고 게임의 결과와 게임진행시 찍힌 사진을 전송 받기 위한 목적으로 사용된다. 또한 모바일의 햅틱 기능을 이용해 사용자가 게임 도중 벽에 부딪혔을 때 진동과 효과음으로 실감나는 효과를 주었다.

모바일은 안드로이드 플랫폼을 기반으로 하는 갤럭시S를 사용하였으며 모바일과 PC 사이의 데이터 전송은 Wi-Fi 무선 통신을 통해 이루어진다. 통신은 TCP/IP 소켓 통신을 사용하며, PC는 서버로 동작하고 사용자가 가진 두 개의 모바일은 클라이언트로 동작하게 된다. 콘텐츠가 실행되면 모바일은 무선공유기를 통해 PC에 접근한 뒤 데이터를 주고받는다. PC와 모바일의 원활한 데이터 통신을 위해 별도의 프로세스에서 통신이 이루어지기 때문에 통신 부분은 콘텐츠 외에 별도의 인터페이스로 처리된다.

4.4 콘텐츠 구현

전체 시나리오는 도입부분, 게임 진행부분, 평가부분으로 구성된다. 본 콘텐츠는 협업모드와 대결모드로 구성되어 있어 사용자가 게임모드를 선택할 수 있다. 게임은 사용자의 선택에 따라 진행되며 평가부에서는 게임 결과를 평가 받게 된다. 본 콘텐츠의 전체 시나리오는 [그림 11]과 같다.

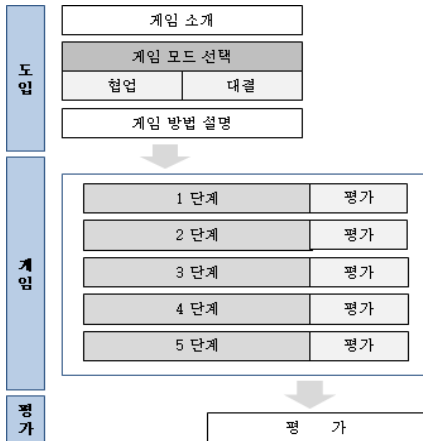


그림 11. 콘텐츠 전체 시나리오

4.4.1 도입부

도입부에서는 간단한 게임 방법 소개와 함께 게임 모드 선택에 따른 게임 방법을 설명한다. 먼저 콘텐츠가 실행되고 두 명의 사용자가 인식되면 디스플레이에서는 도입부의 애니메이션이 종료되고 게임 모드 선택화면으로 넘어간다. 게임 모드는 ‘협업 모드’와 ‘대결 모드’로 나뉜다. 이 중 한 가지가 선택되면 게임 방법 설명과 아이템 설명이 차례로 나온다. 마지막으로 두 명의 사용자가 모두 가지고 있는 모바일에서 ‘Ready’ 버튼을 누르고 각 사용자에게 대한 캘리브레이션을 마친 후 게임이 시작된다. [그림 12]은 콘텐츠 도입부 화면을 보여준다.



그림 12. 콘텐츠 도입부 화면

4.4.2 게임 진행 부분

게임은 사용자의 몸동작과 모바일을 사용해 진행된다. 게임이 시작되면 임의로 선택된 벽이 다가오고 사용자는 그에 맞는 포즈를 취해야 한다. 사용자의 포즈와 벽에 뚫린 구멍이 일치하면 벽을 통과 할 수 있고 그렇지 않으면 벽을 통과할 수 없다. 총 5개의 벽이 지나가고 난 후 통과한 벽의 개수로 결과를 평가 받게 된다. 게임은 ‘협업 모드’와 ‘대결 모드’로 나뉜다.

‘협업 모드’에서는 한 개의 큰 벽이 다가오고 두 사용자가 서로 협동해서 다가오는 벽에 뚫린 구멍과 같은 모양의 포즈를 만들어야 한다. 총 5개의 벽이 지나가고 난 후에 결과를 평가 받게 된다.

‘대결 모드’에서는 두 사용자에게 서로 다른 모양의 구멍이 뚫린 벽이 다가오고, 총 5개의 벽이 지나간 후 각 사용자가 통과한 벽의 개수를 비교해 승패가 결정된다. ‘대결 모드’에서는 모바일에서 아이템을 사용해 상대방을 공격하거나 상대방의 공격에 대한 방어를 할 수 있다. 아이템은 [표 2]와 같이 4가지 종류가 있으며 한 스테이지에서 한 번만 사용할 수 있다. 아이템을 사용하고 나면 사용한 아이템은 모바일에서 비 활성화된다. [그림 13]은 아이템을 사용할 때 모바일 화면을 보여준다.

표 2. 아이템 설명

아이템	설명
Hard	상대방 벽의 난이도 높임
Change	사용자 1과 사용자 2의 벽을 바꿈
Speed up	벽이 다가오는 속도 올림
Speed down	벽이 다가오는 속도 줄임

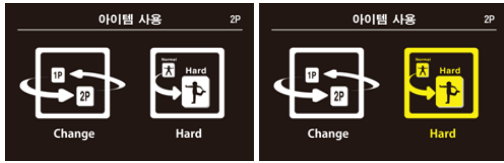
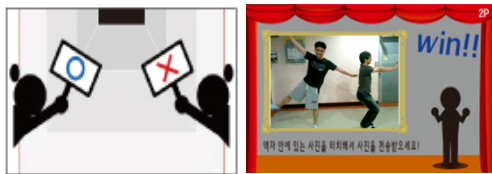


그림 13. 아이템 사용화면

4.4.3 평가부

사용자는 하나의 벽이 지나가는 단계마다 그 결과를 평가 받는다. 적절한 포즈를 취했을 경우에는 [그림 14] (a)의 왼쪽 부분과 같이 'O'꽃말을 든 캐릭터가 나오고 다른 포즈를 취했을 경우에는 'X'꽃말을 든 캐릭터가 나온다.

협업 모드의 경우에는 성공한 횟수에 따라 등급으로 최종 평가를 받게 되고 대결 모드의 경우에는 두 명 사용자의 성공 횟수를 비교해 승패를 가림으로써 사용자가 성취감을 느낄 수 있고 콘텐츠에 몰입할 수 있도록 했다. 또 마지막 평가 화면에서는 평가 결과와 함께 [그림 14]의 (b)와 같이 게임 중 찍힌 사진을 화면에 보여 주고 사용자가 가진 모바일로 사진을 전송받을 수 있는 기능을 넣었다. 사용자는 모바일 상에서 '전송' 버튼을 터치해 전면 디스플레이에 보이는 사진을 모바일로 전송받을 수 있다.



(a) (b)
그림 14. 평가부 화면



그림 15. 콘텐츠 실행 모습

V. 실험결과

[그림 9]에서 정의한 16개의 포즈에 대한 인식률은 [표 3]과 같다. 실험 결과 16개의 포즈 중 14개의 포즈가 90%이상의 인식률을 보였다. 하지만 포즈를 취할 때 정면 혹은 측면에서 몸이 틀어진 경우 다른 포즈로 인식하거나 인식하지 못하는 오류가 있었다. 5번, 9번, 10번과 13번 포즈와 같이 신체 부분이 멀리 떨어져 있는 경우에는 높은 인식률을 보였다. 하지만 8번 12번과 14번 포즈의 경우에는 비교적 인식률이 낮았다. 좌표들 사이의 상대적인 값의 특징을 사용해 포즈 템플릿을 구성하였기 때문에 몸이 틀어지게 되면 좌표 값이 구성된 템플릿의 오차 범위를 벗어나 인식하지 못하거나 다른 포즈로 잘못 인식되었다.

표 3. 포즈 인식 실험 결과

포즈번호	인식률	포즈번호	인식률
POSE 1	95%	POSE 9	98%
POSE 2	95%	POSE 10	98%
POSE 3	95%	POSE 11	93%
POSE 4	93%	POSE 12	92%
POSE 5	99%	POSE 13	99%
POSE 6	92%	POSE 14	89%
POSE 7	92%	POSE 15	92%
POSE 8	87%	POSE 16	95%

VI. 결론

본 논문에서는 인체의 3차원 포즈 인식을 위한 방법으로 3차원 관절 정보를 이용한 포즈 인식 시스템과 이것을 인터페이스로 활용한 상호작용 게임콘텐츠 개발에 대해 기술하였다. 인체에서 특징이 되는 14개 관절의 3차원 좌표 값을 이용해 각 포즈에 대한 포즈 템플릿을 구성하고, 현재 포즈와 포즈 템플릿의 거리를 비교해 포즈를 인식했다. 그리고 측면포즈 8개와 정면포즈 8개로 총 16개의 포즈를 정의해 제안하는 게임 콘텐츠에 적용해 보았다. 평균 94%의 높은 인식률을 보였고 측면 포즈에 비해 정면 포즈의 인식률이 높았다.

3차원 좌표 값만을 이용해 포즈를 인식했기 때문에 비교적 간단하지만 높은 인식률로 정의한 포즈를 인식할 수 있었다. 하지만 3차원 좌표 값만을 이용했기 때문에 자세가 심하게 틀어 졌을 경우에는 인식하지 못하는 문제점을 갖는다. 또한 포즈가 추가될 때마다 템플릿을 구성해 번거로운 단점이 있다. 또한 게임을 하는 도중 다른 사용자가 개입하게 되면 추적 중인 정보를 상실해 게임을 중단해야하는 문제점이 있었다.

향후에는 카메라의 시점에서 사용자의 자세가 틀어진 경우에도 강건하고 효율적으로 포즈를 인식하는 방안에 대해 연구를 수행하고자 한다.

참 고 문 헌

- [1] MIT, "10 Emerging technologies 2011," Technology Review, 2011(5).
- [2] 홍동표, 우운택, "제스처기반 사용자 인터페이스에 대한 연구 동향", Telecommunications Review, 제18권, 제3호, 2008.
- [3] 전성국, 홍광진, 정기철, "회전무관 3D Star Skeleton 특징 추출", 정보과학회논문지, 제36권, 제7호, pp.836-850, 2009.
- [4] I. Haritaoglu, D. Harwood, and L. S. Davis, "Ghost: A human body part labeling system using silhouettes," Proc. of the 14th International Conference on Pattern Recognition, pp.77-82, 1998.
- [5] H. Fujiyoshi, A. J. Lipton, and T. Kanade, "Real-Time Human motion analysis by image skeletonization," IEICE Transactions on Information Systems, pp.113-120, 2004.
- [6] A. Baumberg and D. Hogg, "An adaptive eigenshape model," Proc. of British Machine Vision Conference, Vol.1, pp.87-96, 1995.
- [7] 오치민, 3차원 인체포즈 인식을 위한 그림구조 모델과 파티클필터 추적, 전남대학교 석사학위논문, 2009.
- [8] G. Mori and J. Malik, "Estimating Human Body Configurations using Shape Context Matching," Proc. of the 7th European Conference on Computer Vision Copenhagen(LNCS), Vol.2352, pp.150-180, 2002.
- [9] Y. Kameda, M. Minoh, and K. Ikeda, "Three dimensional pose estimation of an articulated object from its silhouette image," Proc. of the Asian Conference on Computer Vision, pp.612-615, 1993.
- [10] B. Boulay, F. Brenmond, and M. Thonnat "Applying 3d human model in a posture recognition system," Pattern Recognition Letter, Vol.27, pp.1788-1796, 2006.
- [11] D. M. Gavrila and L.S. Davis, "3-D model-based tracking of humans in action: a multi-view approach," Proc. of IEEE Computer Vision and Pattern Recognition, 2006.
- [12] I. Cohen and H. Li, "Posture and gesture Recognition using 3D body Shapes Decomposition," Proc. of IEEE Computer Society Conference on Computer vision and Pattern Recognition, 2005.
- [13] 박재완, 오치민, 이철우, "가상 놀이 공간 인터페이스를 위한 HMM기반 상반신 제스처 인식", 한국콘텐츠학회논문지, 제10권, 제8호, pp.11-17, 2010.
- [14] Microsoft Corp. Redmond WA. Kinect for Xbox360
- [15] <http://research.microsoft.com/en-us/um/redmond/projects/kinectsdk/>
- [16] J. Shotton, A. Fitzgibbon, M. Cook, M. Finocchio, R. Moore, A. Kipman, and A. Blake, "Real-Time Human Pose recognition in Parts from single depth Images," CVPR, 2011.

저 자 소 개

최 윤 지(Yoon-Ji Choi)

정회원



- 2010년 : 전남대학교 전자컴퓨터 공학과(학사)
- 2010년 ~ 현재 : 전남대학교 전자컴퓨터 공학과(공학석사)

<관심분야> : HCI, 포즈인식

박 재 완(Jae-Wan Park)

정회원



- 2007년 : 호남대학교 정보통신 공학과(학사)
- 2009년 : 전남대학교 전자컴퓨터 공학과(공학석사)
- 2009년 3월 ~ 현재 : 전남대학교 전자컴퓨터 공학 박사 과정

<관심분야> : HCI, 테이블탑 디스플레이

송 대 현(Dae-Hyeon Song)

정회원



- 2007년 : 전주대학교 멀티미디어 공학과(학사)
- 2007년 ~ 2009년 : 전남대학교 전자컴퓨터 공학과(공학석사)
- 2009년 3월 ~ 현재 : 전남대학교 전자컴퓨터 공학 박사 과정

<관심분야> : 전통문화 콘텐츠, 멀티미디어, 증강현실

이 칠 우(Chil-Woo Lee)

정회원



- 1992년 : 동경대학교 대학원 전자공학과(공학박사)
- 1992년 4월 ~ 1995년 12월 : 일본 이미지정보과학연구소 수석연구원, 오사카대학 기초공학부 협력연구원 및 리즈메이칸대학

정보학부 특별초빙강사

- 1996년 1월 ~ 현재 : 전남대학교 전자컴퓨터공학부

교수

- 2002년 1월 ~ 2003년 2월 : 미국 NC A&T State University 방문교수
- 2006년 3월 ~ 2008년 2월 : 정보통신부 자체평가위원
- 2008년 10월 ~ 2009년 8월 : KAIST 및 오사카대학 초빙교수
- 2007년 1월 ~ 2008년 8월 : 전남대학교 산학연구부 처장 겸 산학협력부단장
- 2006년 3월 ~ 현재 : 전남대학교 문화콘텐츠기술연구소 소장
- 2009년 3월 ~ 현재 : 전남대학교 차세대휴대폰인터페이스연구센터(ITRC) 센터장

<관심분야> : 컴퓨터 비전, 지능형 휴먼 인터페이스, 디지털 콘텐츠, 컴퓨터그래픽스