

---

# 몰입형 가상현실의 착용식 사용자 인터페이스를 위한 Mixture-of-Experts 기반 제스처 인식

## Gesture Recognition based on Mixture-of-Experts for Wearable User Interface of Immersive Virtual Reality

윤종원, Jong-Won Yoon\*, 민준기, Jun-Ki Min\*\*, 조성배, Sung-Bae Cho\*\*\*

---

**요약** 최근 가상현실에 대한 관심이 높아짐에 따라 다양한 서비스와 어플리케이션이 개발되고 있으며, 이와 더불어 몰입형 상호작용을 위한 제스처 기반 사용자 인터페이스가 활발히 연구되고 있다. 본 논문에서는 가상현실 환경 내 사용자 움직임을 효과적으로 반영하고 몰입감을 높이기 위해 제스처 인식 기반 착용식 사용자 인터페이스를 제안한다. 제안하는 인터페이스는 적외선 LED가 부착되어 있는 헬멧과 적외선 수신기를 이용하여 사용자의 머리 움직임을 인식하며, 양 손에 착용한 데이터 글로브로부터 사용자의 손 제스처를 인식한다. 또한, 헤드 마운트 디스플레이 장치(HMD)를 이용하여 직접 사용자의 시점 변화와 가상환경의 시점 변화를 일치시킨다. 손 제스처의 경우 다수의 관절로 이루어져 있는 손의 특성상 다양한 동작이 가능하며 손 크기나 손동작이 사람마다 모두 다르기 때문에 다양한 사용자들을 대상으로 할 때, 일반적인 모델로는 정확한 인식이 어렵다. 본 논문에서는 다양한 사용자를 대상으로 정확히 손 제스처를 인식하기 위해 Mixture-of-Experts 기반 인식 방법을 적용하였다. 제안하는 인터페이스의 유용성을 평가하기 위해 가상 오케스트라 지휘 환경을 구현하여 인터페이스의 동작 성능을 분석하고 사용성 평가를 수행하였다. 그 결과, 사용자들이 쉽고 직관적으로 사용할 수 있으며 흥미를 유발함을 확인하였다.

**Abstract** As virtual reality has become an issue of providing immersive services, in the area of virtual reality, it has been actively investigated to develop user interfaces for immersive interaction. In this paper, we propose a gesture recognition based immersive user interface by using an IR LED embedded helmet and data gloves in order to reflect the user's movements to the virtual reality environments effectively. The system recognizes the user's head movements by using the IR LED embedded helmet and IR signal transmitter, and the hand gestures with the data gathered from data gloves. In case of hand gestures recognition, it is difficult to recognize accurately with the general recognition model because there are various hand gestures since human hands consist of many articulations and users have different hand sizes and hand movements. In this paper, we applied the Mixture-of-Experts based gesture recognition for various hand gestures of multiple users accurately. The movement of the user's head is used to change the perspective in the virtual environment matching to the movement in the real world, and the gesture of the user's hand can be used as inputs in the virtual environment. A head mounted display (HMD) can be used with the proposed system to make the user absorbed in the virtual environment. In order to evaluate the usefulness of the proposed interface, we developed an interface for the virtual orchestra environment. The experiment verified that the user can use the system easily and intuitively with being entertained.

**핵심어:** *wearable interface, gesture recognition, virtual reality, mixture-of-experts*

---

본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음 (NIPA-2010-(C1090-1021-0008))

\*주저자 : 연세대학교 컴퓨터과학과 석사과정; e-mail: jwyoons@slab.yonsei.ac.kr

\*\*공동저자 : 연세대학교 컴퓨터과학과 박사; e-mail: loomlike@slab.yonsei.ac.kr

\*\*\*교신저자 : 연세대학교 컴퓨터과학과 교수; e-mail: sbcho@slab.yonsei.ac.kr

■ 접수일 : 2010년 6월 28일 / 심사일 : 2010년 7월 16일 / 게재확정일 : 2011년 3월 10일

## 1. 서론

최근 가상현실 시스템이 빠르게 발전함과 더불어 사용자로 하여금 가상현실 내에 실제로 존재하는 것처럼 몰입하도록 하는 인터페이스가 활발히 연구되고 있다[1,2]. 이처럼 인간의 감각을 활용하는 인터페이스는 보다 효율적이며 자연스러운 사람과 가상현실 환경과의 상호작용을 가능케 한다. 이에 착안하여 사용자가 가상환경과 더 실감나게 상호작용 할 수 있도록 하는 다양한 인터페이스가 출시되고 있다.

닌텐도의 Wii 제어기의 경우 사용자의 다양한 행동을 입력으로 이용함으로써 사용자의 몰입도와 흥미를 이끌어내는 연구가 진행되어왔으며[3], 데이터 글로브는 손의 제스처를 인식하여 제조업, 헬스케어, 수화 인식 등과 같은 다양한 분야에서 활용되고 있다[4]. 또한 가상현실 게임 분야에서도 사용자의 몰입도를 향상시키기 위해 새로운 인터페이스들이 제안되었다[5,6]. 그러나 현재까지 대부분의 시스템은 제한된 인터페이스만을 이용하여 사용자가 가상현실 시스템과 상호작용할 수 있도록 하고 있으며, 사용자가 경험할 수 있는 몰입감에도 한계가 있다.

따라서 몰입감을 증진시키기 위해서는 확장된 멀티모달 인터페이스를 사용하여 시스템을 제작하는 것이 필요하다. 착용식 멀티모달 인터페이스를 통해 가상현실 기반 시스템에서 향상된 몰입감을 제공하기 위해서는 다음과 같은 사항이 고려되어야 한다[7,8].

- 사용자와 시스템 간에 자연스럽게 직관적인 상호작용
- 사용자가 제공받는 감각 정보의 질 향상

제시된 사항을 고려하여 본 논문에서는 적외선 헬멧, 데이터 글로브, 헤드 마운트 디스플레이 장치(HMD)를 이용한 착용식 멀티모달 인터페이스 기반 가상현실 시스템을 개발한다. 시스템은 적외선 헬멧과 데이터 글로브를 이용한 제스처기반 인터페이스를 사용하여 사용자와 시스템간의 보다 자연스러운 상호작용을 제공하며, HMD를 통해 향상된 감각정보를 제공한다. 특히 제스처 기반 인터페이스를 구현함에 있어서 중요시되는 제스처의 정확한 인식을 위해 적외선 신호기반 머리 움직임 인식 방법과 Mixture-of-Expert 기반 손 제스처 인식 방법을 사용하였다.

## 2. 배경 및 관련 연구

### 2.1 몰입형 인터페이스 기반 가상현실 시스템

몰입형 인터페이스는 사용자가 가상세계에 완전히 몰입할 수 있도록 사용되고 있으며, 사용자는 컴퓨터를 이용하여 만들어진 가상환경에 HMD, 데이터 헤드셋 등의 특수한 인터페이스들을 사용해 완전히 가상환경 속에 몰입함으로써 그 속에서 정의된 세계를 경험하고 상호 대화식으로 정보를 주고받는다. 몰입형 인터페이스 기반 가상현실 시스템은 완벽한 가상환경을 체험케 해줄 수 있다는 장점이 있으며, 그림 1과 같이 사용자가

다양한 인터페이스들을 이용하여 가상현실 시스템을 제어하는 연구들이 진행되고 있다.

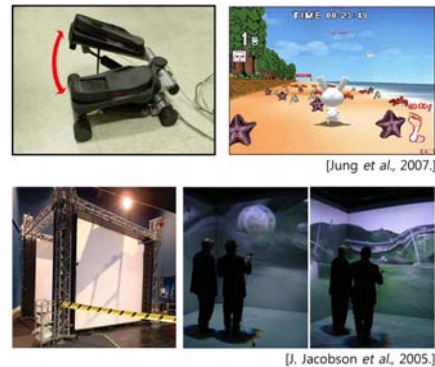


그림 1. 몰입형 가상현실 시스템 관련 연구

Aoki 등은 HMD와 머리추적기를 이용하여 가상현실기반 우주선 조종사들의 비행전 항법교육을 위한 시스템을 제안하였다[9]. 우주선 조종사들은 항법교육을 가상현실을 이용하여 구현된 실제와 유사한 우주선 환경 속에서 수행한다. 또한 Jung 등은 게임을 위한 몰입형 인터페이스인 *Well-Tep*을 제안하였다[5]. 사용자는 인터페이스 위에서 걸어감으로써 가상 캐릭터를 제어한다. Jacobson 등은 시각화 큐브를 이용하여 CaveUT 시스템을 구현하였다[6]. 제안된 시스템은 몰입형 가상현실 인터페이스가 제공되는 정육면체 공간 안에서 전신 행동을 이용하여 1인칭 슈팅게임 *Unreal Tournament 2004*를 즐길 수 있도록 설계되었다.

본 논문에서는 사용자와 시스템간의 자연스러운 상호작용을 제공함과 동시에 사용자에게 향상된 감각정보를 제공하여 사용자로 하여금 몰입감을 느끼게 하는 착용식 멀티모달 인터페이스를 제안한다.

### 2.2 손 제스처 인터페이스

제스처를 이용한 인터페이스는 기존의 키보드나 마우스 등에 비해 직관적인 사용이 가능하며 사용성을 높일 수 있다는 장점이 있다[10]. 그 중 손 제스처는 사용자가 손쉽게 다양한 형태를 취할 수 있어 자연스러운 상호작용이 가능하며 몰입형 인터페이스의 한 요소로 각광받고 있다.

손 제스처 인식은 시각 기반 인식과 센서 기반 인식의 두가지 방법이 존재한다. 센서 기반 인식은 시각 기반 인식에 비해 데이터의 노이즈가 적어 비교적 정확하며 빠른 측정이 가능하다는 장점이 있다. 본 논문에서는 이에 착안하여 데이터 글로브를 이용하여 손 제스처의 데이터를 수집, 인식하는 모델을 설계하였다.

데이터 글로브 기반 손 제스처 인식은 다양한 분야에 접목되어 연구되고 있다. Oz와 Leu는 미국수화의 단어인식을 위해 *Cyberglove*와 *인공신경망*을 이용하였으며[11], Kamel 등은

온라인 서명 확인을 위해 데이터 글로브와 특이값 분해 방법을 사용하였다[12]. 또한 Fels와 Hinton은 데이터 글로브를 이용한 상호작용을 위해 Radial basis function과 인공신경망을 이용하였다[13].

데이터 글로브 기반 손 제스처 인식을 상호작용 수단으로 활용하기 위해서는 다양한 손 제스처의 정확한 인식이 필요하다. 하지만 기존의 인식기로는 다수의 사용자로부터 입력되는 다양한 제스처의 정확한 인식이 힘들다는 한계가 존재한다. 이를 해결하기 위해 본 논문에서는 복잡한 인식에 유리한 다수의 전문가 인식기를 활용하는 Mixture-of-Expert 모델을 사용하여 손 제스처를 인식하는 방법을 적용하였다.

### 3. 몰입형 사용자 인터페이스

사용자가 가상현실 시스템을 사용함에 있어서 보다 향상된 몰입감을 제공하기 위해 본 논문에서는 데이터 글로브와 머리 추적기, HMD로 구성된 착용식 멀티모달 인터페이스를 제안한다. 그림 2는 인터페이스의 구조를 보여준다. 인터페이스는 앞서 제시된 자연스러운 상호작용과 높은 수준의 감각정보 전달이라는 두 가지 목표를 고려하여 설계되었다.

사용자와 시스템간의 자연스러운 상호작용을 제공하기 위해서는 사용자로 하여금 시스템을 제어하기 위해 인위적으로 데이터를 입력한다는 느낌을 갖지 않도록 하는 것이 중요하다. 이를 위해 적외선 LED를 사용한 머리 추적기와 데이터 글로브가 사용되었다. 사용자는 머리 추적기를 이용하여 머리를 움직이므로써 가상환경 내에서 직접 움직이며 시점을 이동할 수 있으며, 데이터 글로브를 통해 손 제스처를 가상환경의 입력으로 활용한다. 사용자의 움직임은 직접 시스템으로 반영되기 때문에 사용자와 시스템간의 부자연스러움을 감소시킬 수 있다. 또한 사용자에게 향상된 감각 정보를 제공하기 위해 HMD 기반 영상출력 방법을 사용한다. 기존의 모니터를 통한 영상 출력은 개방된 공간에서 사용자에게 감각정보를 제공하기 때문에 사용자로 하여금 감각정보의 습득에 대한 의식적인 능동성을 요구하며, 외부의 정보와 섞여 감각정보의 질이 하락할 우려가 있다. HMD는 외부의 정보 차단을 통해 이러한 문제점을 극복하고 사용자에게 향상된 시각정보를 제공해 줄 수 있다.

제안하는 시스템에서는 머리 움직임 인식 모듈과 손 제스처 인식 모듈을 이용하여 사용자의 머리와 손의 움직임을 인식하고 가상환경에 반영한다. 이후 가상현실 영상은 HMD를 통해 사용자에게 제공된다.

#### 3.1 머리 움직임 인식

가상환경에서 사용자에게 몰입감을 제공하기 위해서는 가상환경에서의 사용자 시점이 실제 세계에서의 시점과 일치되어야 한다[14]. 이를 위해 본 논문에서는 적외선 신호 기반 사용자 머리추적 방법을 제안한다.

고가의 기존 머리추적 시스템과는 달리 제안하는 인터페이스에서는 그림 3과 같이 적외선 LED와 저가의 시스템을 위해 적외선 수신기로 Wii 제어기[15]를 사용하여 사용자의 머리 움직임을 추적한다. Wii 제어기는 사용자의 머리 위에 위치하고 있으며, 포함하고 있는 적외선 센서를 이용하여 적외선 신호의 절대 위치를 인식, PC로 전송한다. 사용자는 4개의 적외선 신호기가 장착된 헬멧을 착용함으로써 Wii 제어기와 신호를 주고 받는다.

Wii 제어기로부터 전송받은 센서 데이터는 머리 움직임 인식 모듈로 전달되어 사용자의 위치 및 시선 방향, 머리 움직임을 추정한다. 사용자는 머리를 상하, 혹은 좌우로 회전시킴으로써 가상 작업공간을 탐사할 수 있다. 본 논문에서는 헬멧의 앞뒤에 부착된 두 개의 적외선 LED의 절대위치좌표를 바탕으로 사용자 머리의 상하 회전각도( $\theta_v$ )와 좌우 회전각도( $\theta_h$ )를 계산하는 움직임 모형을 설계하였다. 그림 4는 가능한 사용자 머리의 움직임을 보여준다.

우선 적외선 신호의 초기위치  $I = \{(x_1, y_1), (x_2, y_2)\}$ 을 이용하여, 신호간의 초기 거리  $d$ 와 초기 좌우 절대 각도  $\theta$ 를 수식 (1)과 (2)를 이용하여 계산한다.

사용자의 머리가 움직일 경우, 새로운 적외선 신호의 위치값  $I'$ 을 측정하고, 새로운 거리  $d'$ 과 새로운 좌우 절대 각도  $\theta'$ 을 계산한다. 이 때 상하 회전각도  $\theta_v$ 와 좌우 회전각도  $\theta_h$ 를 수식 (3)을 통해 계산할 수 있다.

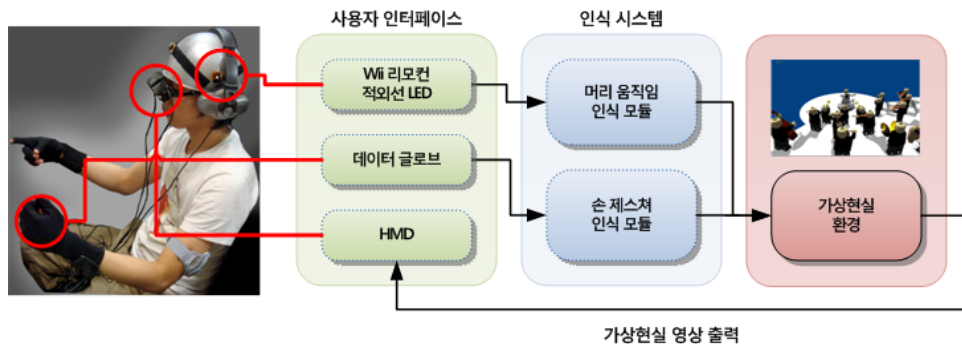


그림 2. 제안하는 착용식 사용자 인터페이스 구조

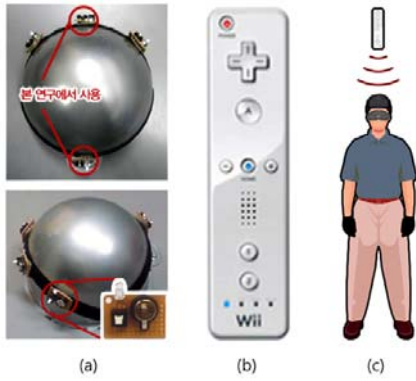


그림 3. 적외선 LED 부착 헬멧(a)과 Wii 리모컨(b), 및 사용 예(c)

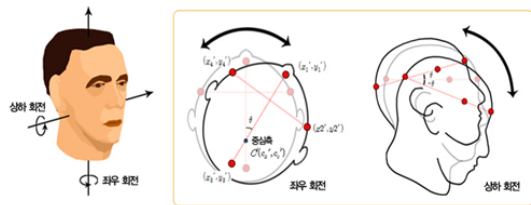


그림 4. 사용자 머리의 움직임

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

$$\theta = \tan^{-1} \left( \frac{x_1 - x_2}{y_1 - y_2} \right) \quad (2)$$

$$\theta_v = \cos^{-1} \frac{d'}{d}, \quad \theta_h = \theta' - \theta \quad (3)$$

측정된  $\theta_v$ 와  $\theta_h$ 를 이용하여 사용자의 머리 움직임을 가상환경 내의 시점 움직임에 반영할 수 있다. 사용자는 머리를 움직임으로써 사용자가 보고 싶은 방향으로 고개를 돌리면 되기 때문에 보다 직관적으로 시점을 전환할 수 있다.

### 3.2 손 제스처 인식



그림 5. 손 제스처 인식을 위한 데이터 글로브

손 제스처 인식 모듈에서는 데이터 글로브를 이용하여 사용자 손 제스처를 인식하고 가상환경에 반영한다. 데이터 글로브를 이용하여 실시간으로 수집되는 양 손의 센서 정보는 입력으로 사용되며 손 제스처 인식 모듈은 입력으로부터 제스처를 인식

한다. 그림 5는 인터페이스를 위해 사용한 데이터 글로브를 보여준다.

사람의 손은 많은 관절로 이루어져 다양한 형태의 제스처를 취할 수 있으며, 인식 모듈은 복잡하고 다양한 손 제스처를 최대한 정확하게 인식해야한다. 본 논문에서는 이를 위해 다수의 전문가를 활용하는 Mixture-of-Experts[16] 구조기반 제스처 인식기법을 제안한다.

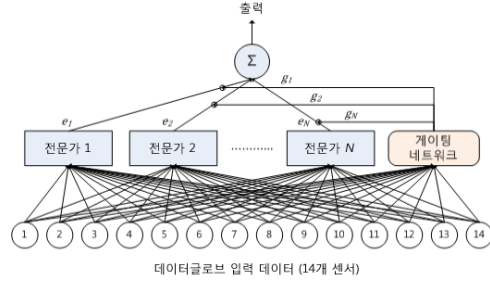


그림 6. Mixture-of-Experts 구조

Mixture-of-Experts 구조는 각기 다른 영역에 작용하는 전문가 네트워크들의 결과를 게이트 네트워크를 통해 결합하여 사용하는 구조이다. 그림 6은 모듈에서 사용하는 Mixture-of-Experts 구조를 보여준다.

Mixture-of-Experts 모델 내에서 모든 전문가들의 의견은 게이팅 네트워크를 이용하여 통합된다. 게이팅 네트워크는 입력 데이터에 기반하여 모든 전문가들의 결과의 반영 정도를 결정한다. 입력  $x$ 에 대한 게이팅 네트워크의 출력  $G(x)$ 는  $N$ 개의 전문가들의 가중치들을 출력하며 수식 (4)와 같이 정의된다.

$$G(x) = \{g_1(x), g_2(x), g_3(x), \dots, g_N(x)\} \quad (4)$$

이 때 입력  $x$ 에 대한  $N$ 개 전문가의 출력  $E(x) = \{e_1(x), e_2(x), \dots, e_N(x)\}$ 를 통합한 최종 출력벡터  $O(x)$ 는 다음과 같이 표현된다.

$$O(x) = G(x)E(x)^T = \sum_{i=1}^N g_i(x)e_i(x) \quad (5)$$

이후 제시된 과정을 통해 통합된 의견들 중에서 가장 높은 값을 가지는 제스처를 인식 결과로 선택한다.

본 논문에서 사용된 전문가들과 게이팅 네트워크는 다층 퍼셉트론(Multi-Layered Perceptron, MLP)을 이용하여 구현되었다. 게이팅 네트워크의 학습을 위한 특정 입력  $x$ 에 대한 기대 출력  $y_g$ 는 각 전문가의 정확도를 계산하여 사용한다. 전문가의 정확도를 계산하기에 앞서, 입력  $x$ 에 대한  $i$ 번째 전문가의 오차량  $Err_i(x)$ 를 다음과 같이 정의한다.

$$Err_i(x) = \sum_t |e_i^t(x) - y^t| \quad (6)$$

$y$ 는 입력  $x$ 에 대응하는 정답 벡터이며  $e_i^t(x)$ 와  $y^t$ 는 각각  $e_i(x)$ 와  $y$ 의  $t$ 번째 값을 의미한다. 이후 다음과 같이 모든 전문가들의 최대 오차량과 최소 오차량을 이용하여 오차량을 정규화한 뒤 전문가의 정확도를 계산한다.

$$Acc_i(x) = 1 - \frac{Err_i(x) - \min_k Err_k(x)}{\max_k Err_k(x) - \min_k Err_k(x)} \quad (7)$$

최종적으로  $x$ 에 대한 게이팅 네트워크의 기대 출력 벡터  $y_g$ 는 수식 (8)과 같이 정의된다.

$$y_g^t = Acc_i(x) \quad (t = 1, 2, \dots, N) \quad (8)$$

게이팅 네트워크는 모든 학습 데이터와 그에 따른 기대 출력 벡터를 이용하여 오류 역전파 알고리즘을 통해 학습된다. 제안하는 Mixture-of-Experts 모델은 초기의 학습된 상태 이외에도 새롭게 들어오는 입력을 활용하여 현재 사용자의 손 모양에 적용한다.

### 3.3 HMD 기반 영상 출력



그림 7. 영상 출력을 위한 HMD

가상환경과의 몰입적인 상호작용을 위하여 사용자는 그림 7의 HMD를 통해 가상현실 환경의 영상을 제공받는다. HMD는 사용자의 머리에 착용하여 사용하는 디스플레이 장치로, 사용자의 눈앞에 영상을 출력해준다. 가상환경의 영상은 사용자의 시점에 따라 변화되며 사용자는 변화가 적용된 가상환경을 HMD를 통해 접할 수 있다.

### 4. 구현 및 평가

제안하는 시스템에서 사용하는 인터페이스는 머리 추적기와 데이터 글로브, 그리고 HMD로 구성된다. 머리 추적기는 자체제작한 적외선 LED 부착 헬멧과 Wii 제어를 사용하였으며, 데이터 글로브는 오른손 14개, 왼손 5개 센서가 부착된 5DT사의 데이터 글로브를 사용하였다. 또한 HMD는 60Hz에서 1920 × 1080 해상도를 제공하는 Deocom사의 DG-450HP HMD를 사용하였다.

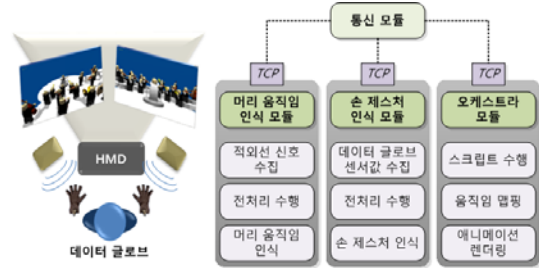


그림 8. 가상 오케스트라 지휘 환경 구성도

제안하는 시스템의 실제 가상현실 환경에의 적용 및 평가를 위해 그림 8과 같이 몰입형 사용자 인터페이스를 이용한 사용자의 시점 변화와 손 제스처가 유용하게 사용될 수 있는 가상 오케스트라 지휘 환경을 개발하였다.



그림 9. 정의된 8가지 지휘 손 제스처

구현을 위해 3D Max, Direct 3D와 FMOD Ex Sound Library가 사용되었으며 사용자는 Lua 기반 스크립트로 짜여진 시나리오에 따라 그림 10과 같이 나타나는 제스처를 타이밍을 맞추어 입력하면서 플레이하게 된다.

사용자는 시점 변화와 손 제스처 두 가지의 행동이 가능하며, 손 제스처의 경우 오른손은 그림 9와 같은 8가지의 미리 정의된 지휘 행동을 인식하며 왼손은 연속적인 센서값을 활용하여 조명 밝기와 음량 크기를 조절한다.

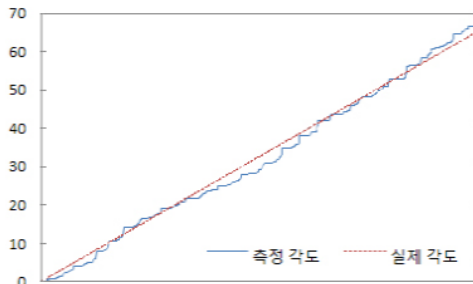
### 4.1 시스템 성능 평가

제안하는 방법의 사용자 행동 인식 성능 평가를 위해 머리 움직임 인식 모듈과 손 제스처 인식 모듈 모두에 대하여 인식 성능을 분석하였다.

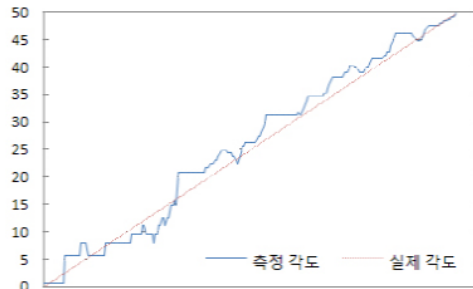
우선 머리 움직임 인식 모듈의 인식 성능을 분석하기 위해 사용자 머리의 좌우 회전각도와 상하 회전각도를 서서히 증가시키면서 실제 사용자 머리의 각도와 인식 모듈을 통해 측정된 각도를 비교하였다.

표 1. 사용자 머리 회전각도 측정의 오차각도

	평균 오류 각도(°)	최대 오류 각도(°)
좌우 회전	1,408	3,967
상하 회전	1,786	5,377



(a) 좌우 회전 각도



(b) 상하 회전 각도

그림 10. 머리 움직임 인식 성능 분석

그림 11(a), (b)는 각각 좌우 회전각도와 상하 회전각도를 증가시키면서 인식된 각도와 실제 각도를 비교한 그래프를 나타낸다. 제안하는 머리 움직임 인식방법은 실제 각도에서 크게 벗어나지 않는 범위 안에서 사용자 머리의 회전각도를 측정하는 것을 확인할 수 있다. 표 1은 좌우 회전각도 인식과 상하 회전각도 인식에 있어서 평균 오차 각도와 최대 오차 각도를 보여준다.

이후 손 제스처 인식 모듈의 인식 성능 분석을 위해, 5명의 사용자로 하여금 25개의 정의된 동작들을 수행하도록 한 뒤에 인식 성능을 분석하였다. 분석을 위해 각 동작 당 1000개의 데이터를 획득하였으며, 기존에 널리 사용되는 인식방법인 Support Vector Machine(SVM), Naive Bayes(NB), 인공신경망(Artificial Neural Network, ANN)과 제안하는 방법(ME)의 성능을 비교하였다.

표 2. 손 제스처 인식 성능 비교 (단위: %, 각 동작별로 8번 실험 후 평균)

인식기	평균 정확도	표준편차	오차
SVM	85,93	0,0515	±0,0026
NB	74,23	0,0051	±0,0025
ANN	85,49	0,0412	±0,0020
<b>ME</b>	<b>91,05</b>	<b>0,0434</b>	<b>±0,0022</b>

표 2와 같이 제안하는 Mixture-of-Experts 기반 손 제스처 인식방법은 기존의 다른 인식기를 사용하여 제스처를 인식하는 방법에 비해 최대 12% 가량의 성능 차이를 보여주었으며 모든 비교 대상 인식기들에 비해 우수한 성능을 보임을 확인할 수 있었다.

#### 4.2 사용성 평가

성능 평가와 더불어 제안하는 시스템이 실제로 사용자를 대상으로 유용하게 사용될 수 있음을 확인하기 위해 사용성 평가를 수행하였다. 사용자는 제안하는 인터페이스를 직접 착용하고 구현된 가상 오케스트라 지휘 시스템을 사용해본 뒤 시스템의 사용성을 평가한다.

실험에는 19세~30세의 남녀 대학생 및 대학원생 48명이 참여하였으며, 참가자 중 약 50% 이상의 피험자들은 비디오 게임을 즐겨하지 않는 것으로 확인되었다. 사용성 평가를 위해 피험자들을 대상으로 SUS질의[17]가 제공되었다. 질의는 10개 문항으로 구성되어 있으며 각 문항당 점수는 “강한 부정”을 나타내는 1점부터 “강한 긍정”을 나타내는 5점까지 5단계로 나누어서 평가되었다. 표 3은 사용성 평가에 사용된 SUS 질의를 보여준다.



(a)



(b)

그림 11. 가상 오케스트라 지휘 화면(a) 및 실제 사용 모습(b)

표 3. SUS질의

1	이 시스템을 자주 사용할 것이다.
2	이 시스템이 불필요하게 복잡하다고 생각한다.
3	이 시스템이 사용하기 쉽다고 생각한다.
4	이 시스템을 사용하기 위해서는 전문가의 도움이 필요하다고 생각한다.
5	이 시스템의 다양한 기능이 조직적으로 잘 결합되어 있다고 생각한다.
6	이 시스템이 불안정한 것 같다고 생각한다.
7	많은 사람들이 이 시스템의 사용법을 매우 빠르게 익힐 것이라고 생각한다.
8	이 시스템이 큰 방해가 된다고 생각한다.
9	이 시스템을 사용하는 데에 자부심을 느낀다.
10	이 시스템을 계속 사용하기 위해 많은 것들을 배워야 한다고 생각한다.

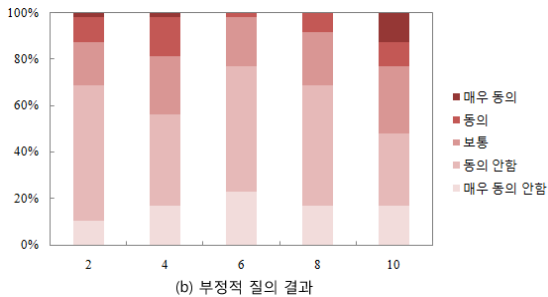
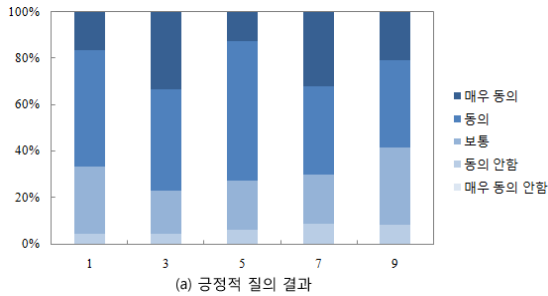


그림 12. SUS질의 결과

사용성 평가 결과, 평균 SUS점수는 100점 만점에 68.5점을 얻었다. 특히 피험자의 77%가 제안하는 인터페이스가 사용하기 쉽다고 응답하였으며 56%가 전문가의 도움 없이도 사용할 수 있다고 응답하여 제안하는 인터페이스가 직관적이며 사용자에게 자연스러운 상호작용 수단을 제공함을 확인하였다. 그림 12는 사용성 평가 결과를 보여주는 것으로, 그림 12(a)는 제안하는 인터페이스에 대한 긍정적인 면에 관한 홀수번 질의 결과며, 그림 12(b)는 부정적인 면에 관한 짝수번 질의 결과이다. 긍정적 질의 경우 과반수의 사용자가 동의하는 결과를 보여주었으며, 부정적 질의 경우 과반수의 사용자가 동의하지 않아 사용자들이 대체로 시스템에 만족한다는 결론을 내릴 수 있다. SUS 질의 외에도 추가적으로 시스템의 사용성을 평가하기 위

해 표 4와 같이 3개의 추가 질의를 제공하였다. 각 문항당 점수는 1점부터 5점까지 나누어서 평가되었다.

표 4. 추가 질의

1	다른 사용자의 플레이 장면이 도움이 되었다.
2	장비의 착용이 불편하다고 생각한다.
3	이 시스템이 재미있다고 생각한다.

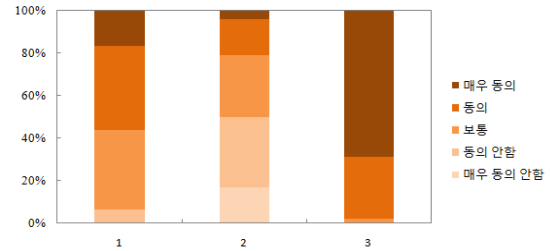


그림 13. 추가 질의 결과

추가 질의 결과, 다수의 사용자가 다른 사람의 플레이 장면을 보고 쉽게 사용법을 터득할 수 있다고 판단하였으며, 착용식 장비를 사용함에 있어서도 큰 불편을 느끼지 않는다고 응답하였다. 또한 대다수의 사용자가 제안하는 시스템에 대해서 재미있다고 판단하였으며 이를 통해 시스템이 가상환경에 대한 몰입감을 향상시킴으로써 흥미를 유발할 수 있음을 확인하였다. 그림 13은 추가 질의 결과를 보여준다.

## 5. 결론

본 논문에서는 가상현실 기반 시스템 사용자를 위한 착용식 멀티모달 인터페이스 시스템을 제안하였다. 시스템은 적외선 LED와 Wii 제어기 기반의 머리추적기, 데이터 글로브, 그리고 HMD로 구성되며, 사용자의 다양한 움직임을 인식하여 가상환경에 반영함으로써 향상된 몰입도를 제공해 줄 수 있다. 제안하는 시스템은 가상 오케스트라 지휘 환경에 적용 되어 구현되었으며, 사용자 움직임 인식 성능 평가 및 실제 사용자들을 대상으로 한 사용성 평가를 수행하여 제안하는 시스템의 유용성을 검증하였다.

향후에는 가속도 센서나 EMG 센서 등의 추가 장비를 이용하여 더욱 다양한 사용자 움직임을 인식하고 이를 가상현실 환경에 반영하는 연구를 진행할 계획이다. 또한 보다 많은 피험자를 대상으로 한 사용성 평가를 진행하여, 가상현실 시스템에 대한 경험 유무와 같이 다양한 각도에서 평가 결과 분석이 이루어져야 할 것이다.

## 참고문헌

- [1] Azuma, R., Y. Baillet, R., Behringer, R., Feiner, S., Julier, S. and MacIntyre, B. Recent advances in augmented reality. *IEEE Computer Graphics and Applications*, vol. 21, no. 6, pp. 34-47, 2001.
- [2] Johnsen, K. and Lok, B. An evaluation of immersive displays for virtual human experiences. In *Proc. of the IEEE Virtual Reality Conf.*, pp. 133-136, 2008.
- [3] Guo, C. and Sharlin, E. Exploring the use of tangible user interface for human-robot interaction: A comparative study. In *Proc. of CHI 2008*, pp. 121-130, 2008.
- [4] Dipietro, L., Sabatini, A.M. and Dario, P. A survey of glove-based systems and their applications. *IEEE Trans. Systems, Man, and Cybernetics-Part C: Applications and Reviews*, vol. 38, pp. 461-482, 2008.
- [5] Jung, D.-G., Jang, C.-S., Oh, J.-H., Kang, J.-G., Oh, I.-G. and Jung, K.-C. An immersive game using a new interface: The Well-Tep. *Lecture Notes in Computer Science*, vol. 4469, pp. 63-68, 2007.
- [6] Jacobson, J., Renard, M. C., Lugin, J.-L. and Cavazza, M. The CaveUT system: Immersive entertainment based on a game engine. In *Proc. of the 2005 ACM SIGCHI Int. Conf. on Advances in Computer Entertainment Technology*, pp. 184-187, 2005.
- [7] Kim, HS., T. Di Giacomo, Egges, A., Lyard, E., Garchery, S. and Magnenat-Thalmann, N. Believable virtual environment: Sensory and perceptual believability. In *Workshop on Believability in Virtual Environment*, 2004.
- [8] Figueiredo, M., Bohm, K. and Teixeira, J. Advanced interaction techniques in virtual environments. *Computers & Graphics*, vol. 17, no. 6, pp. 655-661, 1993.
- [9] Aoki, H., Oman, C. M., Buckland, D. A. and Natapoff, A. Desktop-VR system for preflight 3D navigation training. *Acta Astronautica*, vol. 63, no. 21, pp. 841-847, 2008.
- [10] O'Hagan, R. G., Zelinsky, A. and Rougeatux, S. Visual gesture interfaces for virtual environments. *Interacting with Computers*, vol. 14, no. 3, pp. 231-250, 2002.
- [11] Oz, C. and Leu, M. Linguistic properties based on American sign language isolated word recognition with artificial neural networks using a sensory glove and motion tracker. *Neurocomputing*, vol. 70, pp. 2891-2901, 2007.
- [12] Kamel, N., Sayeed, S. and Ellis, G. Glove-based

approach to online signature verification. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30, pp. 1109-1113, 2008.

- [13] Fels, S. and Hinton, G. Glove-TalkII: A neural network interface which maps gestures to parallel formant speech synthesizer controls. *IEEE Trans. on Neural Networks*, vol. 9, pp. 205-212, 1998.
- [14] Wacksuth, S., Wrede, S. and Handeide, M. Coordinating interactive vision behaviors for cognitive assistance. *Computer Vision and Image Understanding*, vol. 108, no. 1-2, pp. 135-149, 2007.
- [15] Lee, J. Hacking the Nintendo Wii remote. *IEEE Pervasive Computing*, vol. 7, no. 3, pp. 39-45, 2008.
- [16] Jacobs, R., Jordan, M., Nowlan, S. and Hinton, G. Adaptive mixture of local experts. *Neural Computation*, vol. 3, no. 4, pp. 79-87, 1991.
- [17] Brooke, J. SUS: A quick and dirty usability scale. *Usability Evaluation in Industry*, 1996.
- [18] Jang, S.-H., Yoon, J.-W. and Cho, S.-B. User interfaces for visual telepresence in human-robot interaction using Wii controller. *Journal of the HCI Society of Korea*, vol. 3, no. 1, pp. 27-32, 2008.



**윤 종 원**

2005년 3월 ~ 2009년 8월 연세대학교 컴퓨터과학과 졸업(학사). 2009년 9월 ~ 현재 연세대학교 컴퓨터과학과 석사과정. 관심분야는 HCI, 패턴인식 등임.



**민 준 기**

1999년 3월 ~ 2004년 2월 연세대학교 기계전자공학부 졸업(학사). 2004년 3월~2006년 2월 연세대학교 컴퓨터과학과 졸업(석사). 2006년 3월~2011년 2월 연세대학교 컴퓨터과학과 졸업(박사). 관심분야는 상황인지, 패턴인식 등임.



**조 성 배**

1988년 연세대학교 전산학과 졸업(학사). 1990년 한국과학기술원 전산학과 졸업(석사). 1993년 한국과학기술원 전산학과 졸업(박사). 1993년~1995년 일본 ATR 인간정보통신연구소 객원 연구원. 1998년 호주 Univ. of New South Wales 초청연구원. 1995년~현재 연세대학교 컴퓨터과학과 정교수. 관심분야는 신경망, 패턴인식, 지능정보처리 등임.