

LCD Display 설비 Contents의 Kinect기반 동작제어 기술 구현에 관한 연구

A Study on Implementing Kinect-Based Control for LCD Display Contents

노 정 규*
(Jungkyu Rho)

Abstract - Recently, various kinds of new computer controlled devices have been introduced in a wide range of areas, and convenient user interfaces for controlling the devices are strongly needed. To implement natural user interfaces(NUIs) on top of the devices, new technologies like a touch screen, Wii Remote, wearable interfaces, and Microsoft Kinect were presented. This paper presents a natural and intuitive gesture-based model for controlling contents of LCD display. Microsoft Kinect sensor and its SDK are used to recognize human gestures, and the gestures are interpreted into corresponding commands to be executed. A command dispatch model is also proposed in order to handle the commands more naturally. I expect the proposed interface can be used in various fields, including display contents control.

Key Words : Natural user interface, Gesture recognition, Depth sensor, Kinect, Display contents

1. 서 론

최근 음성인식, 동작인식과 같은 다양한 휴먼 인터페이스(human interface)들의 발달에 따라 키보드나 마우스를 대신하는 새로운 인터페이스 장치들이 등장하여 마케팅, 의료, 비즈니스, 컴퓨터 과학, 엔터테인먼트, 로보틱스 등과 같이 다양한 분야에 응용되고 있으며[1], 동작인식 기술의 사용이 확산됨에 따라 별도의 컨트롤러를 손에 쥐거나 신체에 장착하지 않고 가전제품, LCD 디스플레이 전시물, 영상 광고판 등을 조작할 수 있는 자연스럽게 직관적인 동작인식 인터페이스 기술의 중요성이 점차 높아지고 있다[2]. 본 연구에서는 이러한 추세에 맞추어 사람과 IT 기술 기반의 디스플레이 콘텐츠 간의 자연스러운 교감을 가능케 하는 Kinect기반 동작제어 기술을 제안하고 구현하였다. Depth 센서 기술의 발달에 따라 Kinect를 활용할 수 있는 분야가 다양해지고 있는데[3,4,5,6], Kinect는 실시간으로 depth 정보를 제공함과 동시에 인체의 관절 정보 및 RGB 영상을 제공한다는 특징을 가지고 있다[7].

본 연구에서는 Kinect가 제공하는 depth 및 동작인식 정보를 이용하여 디스플레이 콘텐츠의 매끄러운 제어뿐만 아니라 휴먼 인터페이스가 필요한 다양한 분야에서 활용할 수 있는 동작인식 제어기술을 제안하였다. 2장에서는 관련 연구와 기술 수준을 살펴보고, 3장에서는 연구에서 사용된 디스플레이 콘텐츠 설비와 제어 기술에 대하여 설명한다. 4장에서는 디스플레이 콘텐츠 제어 실험 결과를 기술하고 5장에서 결론을 맺는다.

2. 관련 연구

컨트롤러를 사용한 인간 컴퓨터 상호작용 기술은 가속도 센서, 자이로 센서, 적외선 센서 등을 이용하여 기기의 움직임을 인식하는 기술이 발전함에 따라 꾸준히 연구가 이루어져 왔다. 컨트롤러를 이용한 제어 기술 중 널리 알려진 것으로는 프랑스의 Parrot사에서 개발한 무선 조종 비행기인 AR. Drone이 있는데, 스마트폰을 사용하여 기체의 움직임을 제어하거나 장착된 카메라를 통해 촬영한 영상을 제공받을 수 있다[8]. MIT 미디어랩에서 개발한 Sixth-Sense는 일종의 웨어러블 제스처 인터페이스(wearable gesture interface)로 카메라와 모바일 프로젝터, 그리고 손가락에 부착하는 네 개의 칼라 마커를 사용하여 사진을 찍거나 정보를 검색하는 등 다양한 기능을 수행할 수 있다[9]. Carnegie Mellon 대학의 J. C. Lee는 Wii Remote 컨트롤러를 이용하여 손이나 머리의 움직임을 추적하는 방법을 구현하였고, 이를 이용하여 머리의 움직임에 따라 변하는 가상현실 디스플레이를 제안하였다[10]. 연세대학교의 장수형 등의 연구에서는 Wii Remote 컨트롤러를 이용하여 사용자의 이동 및 시선을 파악하고 그 정보를 바탕으로 원거리의 로봇을 움직여 사용자가 원하는 정보를 수신할 수 있는 시스템을 제안하였다[11].

위의 연구들은 컨트롤러를 이용하기 때문에 상호작용을 할 수 있는 사용자가 제한적이라는 단점을 가지고 있는데, Kinect와 같은 이미지와 depth를 인식할 수 있는 장비를 이용하면 컨트롤러나 마커의 도움 없이 동작을 인식하는 인터페이스를 가능하게 한다. A. Sanna 등의 연구에서는 손, 발을 들거나 몸통을 기울여 AR. Drone을 조종하는 인터페이스를 제안하였다[12]. T. Osunkoya 등의 연구에서는 머리, 어깨, 몸의 중심 등으로부터 손을 뻗은 거리와 방향에 따라 마우스 이벤트를 발생시키고 이를 이용하여 PowerPoint 프

* Corresponding Author : Dept. of Computer Science,
Seokyeong University, Korea

E-mail : jkrho@skuniv.ac.kr

Received : March 24, 2014; Accepted : March 28, 2014

리젠테이션을 제어할 수 있음을 보여주었다[13]. K. T.-M. Tran 등의 연구에서는 양 손의 위치 변화를 이용하여 3차원 이미지를 이동시키거나 확대, 축소, x축, y축 방향 회전을 가능하게 하는 인터페이스를 제안하였다[14].

불특정 사용자가 임의로 접근할 수 있는 전시물이나 광고판과 같은 LCD 디스플레이 설비 콘텐츠의 제어는 아무런 사전 지식이 없는 사용자가 직관적으로 간단하게 사용법을 익히고 조작할 수 있어야 한다. 그런데 기존의 Kinect기반 인터페이스들은 여러 종류의 자세를 각각의 명령에 대응시키기 때문에 사용자가 미리 자세를 알고 익혀야 한다는 단점이 있다. 뿐만 아니라 기존 연구에서는 신체 각 부위의 상대적인 위치를 이용하여 사용자의 의도를 인식하는데, 이는 정적인 자세를 이용한 것으로 사용자의 동적인 움직임을 이용한 것이라고 보기 힘들다. 보다 자연스럽게 직관적인 인터페이스를 제공하기 위해서는 자세뿐만 아니라 동적인 제스처를 인식하여 사용자의 의도를 파악하는 기술이 요구된다. 본 연구에서는 Kinect를 이용하여 사용자의 동작을 인식하고 이를 이용하여 LCD 디스플레이 설비 콘텐츠를 직관적으로 제어하는 방법을 제안한다.

3. Kinect를 이용한 LCD Display 설비 Contents 제어기술의 구성 및 특징

본 장에서는 Kinect를 이용하여 사용자의 움직임을 추적하여 제스처를 인식하고 디스플레이 콘텐츠를 제어하는 기술에 대하여 설명한다. Kinect는 실시간으로 사물의 depth와 RGB 영상을 제공하는데 이를 이용하여 사용자를 감지하고 인체 관절(joint)을 추적하여 위치 정보를 얻을 수 있다. 이를 이용하면 제스처 인식을 위해 사람의 신체부위를 검출하고 자세를 추정하는 수고를 덜 수 있으므로 동작인식 응용 개발이 용이해진다. Depth 센서가 제공하는 정보로부터 사용자를 감지하면 골격(skeleton) 정보를 얻을 수 있는데, 머리부터 발목까지 20개의 관절을 인식할 수 있다. 본 연구에서는 스와이프(swipe) 제스처를 인식하기 위하여 2개 관절 정보(어깨 중심, 왼손)를 사용하였다. 감지된 객체의 관절 정보는 제스처 인식을 위하여 사용되는데, 관절 정보는 특징 추출(feature extraction) 및 분석(analysis)과정을 거쳐서 제스처 인식(gesture recognition) 여부를 결정한다. 제스처 인식이 순조롭게 이루어지면 명령 처리(command dispatch) 단계로 넘어가며, 제스처 인식이 되지 않은 경우 특징 추출 단계로 피드백 되어 제스처 인식 과정을 수행하게 된다. 명령 처리 단계에서는 제스처가 대응되는 명령으로 해석되고 명령을 수행한다. 명령 수행을 위해서는 명령 수행가능 여부와 수행순서를 결정하는 명령 스케줄링을 거쳐 디스플레이 콘텐츠 제어로 전달된다. 그림 1에는 본 논문에서 제안하는 Kinect기반 제스처 제어 프레임워크를 나타내었다.

본 연구에서는 사용자의 의도를 자연스럽게 처리하는 인터페이스를 제공하기 위하여 인식된 제스처를 디스플레이 콘텐츠 제어에 바로 적용하지 않고 명령처리 단계를 거치도록 하였다. 명령처리 단계는 제스처 해석 과정과 명령 스케줄링 과정으로 구성된다. 하나의 제스처에 해당하는 명령이 수행 중일 때 다른 제스처가 인식되어 또 다른 명령이 전달되어 수행되면 사용자의 의도와 다르게 조작될 수 있는데, 이

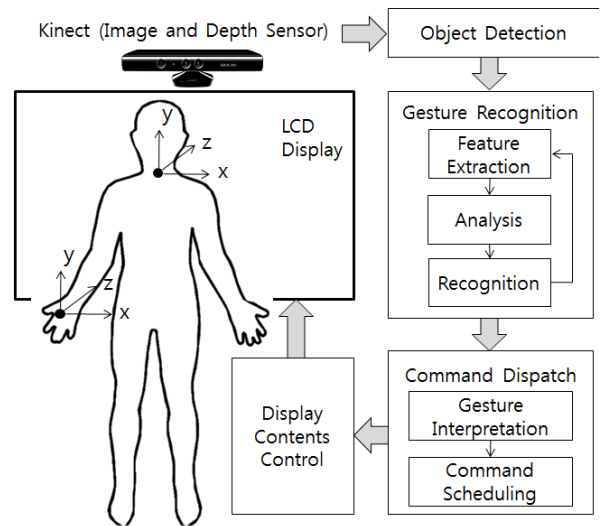


그림 1 디스플레이 콘텐츠의 Kinect기반 제스처 제어 프레임워크

Fig. 1 The framework of Kinect-based gesture control for display contents

를 방지하기 위하여 명령 스케줄링이 필요하다. 그림 2는 하나의 명령이 수행 중인 경우 다른 명령을 수행하지 않도록 하는 스케줄링의 예를 보여준다.

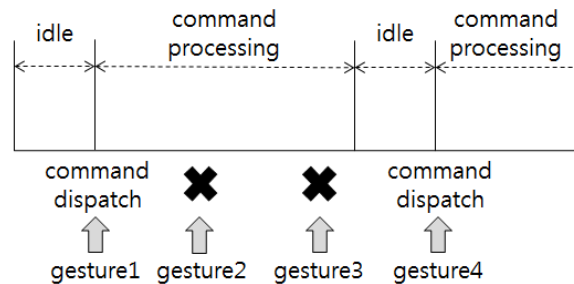


그림 2 명령 스케줄링 예

Fig. 2 An example of command scheduling

명령 처리 단계의 명령 스케줄링 과정을 이용하여 하나의 Kinect 센서로부터 순차적으로 인식된 제스처를 자연스럽게 처리할 수 있으며, 다수의 Kinect 센서로 구성된 환경에서 동시에 제스처가 인식되는 경우에도 매끄러운 처리가 가능하다. 뿐만 아니라 동작인식을 위하여 여러 종류의 센서를 사용하는 경우 제스처 해석 과정을 이용하여 서로 다른 센서를 통해 인식한 서로 다른 제스처에 같은 의미를 부여하고 명령을 수행할 수 있다. 즉, 사용자의 의도는 같지만 다른 제스처를 취하거나 다른 종류의 센서에 의해 제스처가 감지되는 경우를 자연스럽게 처리할 수 있다.

4. Kinect기반 LCD Display 설비 Contents 제어 실험

4.1 Depth 센서를 이용한 제스처 인식 실험

그림 3은 본 연구에서 사용한 디스플레이 콘텐츠 제어 실

험 장비이다. 통상적인 방법대로 Kinect를 디스플레이 상단에 장착하였으며 640x480 해상도와 초당 30 프레임으로 이미지와 depth 정보를 전송하도록 하였다. Kinect의 depth 센서를 통하여 사람이 인식되면 각 관절의 위치가 연속적으로 추적되어 실시간으로 데이터가 얻어진다. 그림 4는 Kinect를 이용하여 객체 감지가 이루어진 후 제스처 인식을 위하여 사용되는 골격 정보를 그림으로 나타낸 것이다.



그림 3 디스플레이 콘텐츠 제어 실험 장비
Fig. 3 Equipments for the experiment of display contents control

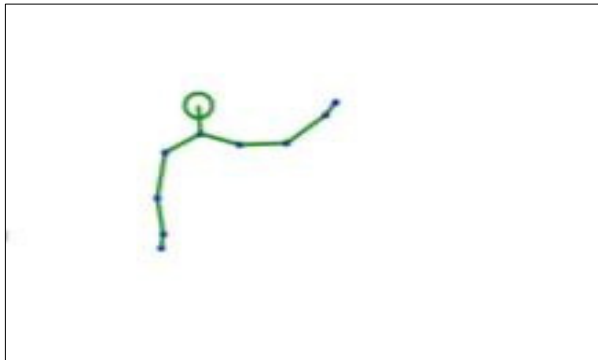


그림 4 Kinect depth sensor를 이용한 관절 위치 감지
Fig. 4 Tracking joint position using Kinect depth sensor

실험에서는 왼손의 네 종류 제스처(왼쪽, 오른쪽, 위쪽, 아래쪽 스와이프)를 인식하도록 하였는데, 디스플레이와 사용자 간의 거리가 100cm, 200cm, 300cm인 경우를 각각 실험하였다. 스와이프는 한쪽 방향으로 손을 미는 동작으로, 손을 뺀고 있는 자세와는 구분된다. 그림 5, 그림 6, 그림 7, 그림 8은 200cm 거리에서 왼손으로 왼쪽, 오른쪽, 위쪽, 아래쪽 스와이프 동작을 반복적으로 하였을 때 왼손의 상대 좌표를 각각 나타낸 것이다.

본 연구에서 스와이프 제스처 인식에 x, y 좌표 값만 사용하였으며 depth 정보는 사용하지 않았다. 실험을 단순화하기 위하여 왼손 제스처만 인식하도록 하였으며 간단하게 양손 제스처 인식으로 확장될 수 있다. 스와이프 동작 패턴은 사용자에게 따라 다르게 나타나는데, 손을 밀고 당기는 속도, 손을 미는 거리, 반복적인 동작 중에 잠시 멈추는 시간, 반

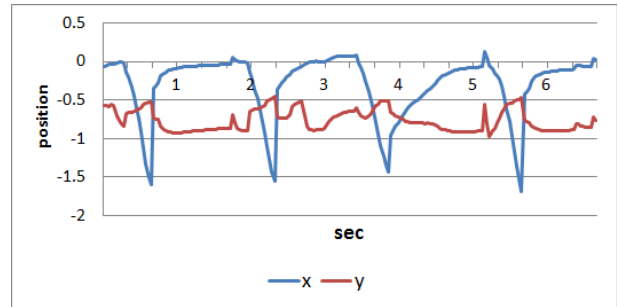


그림 5 반복적인 왼쪽 스와이프 동작에서 손 위치
Fig. 5 Hand position of repeated swipe left gestures

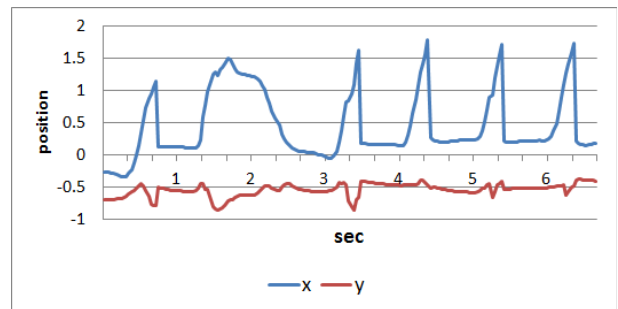


그림 6 반복적인 오른쪽 스와이프 동작에서 손 위치
Fig. 6 Hand position of repeated swipe right gestures

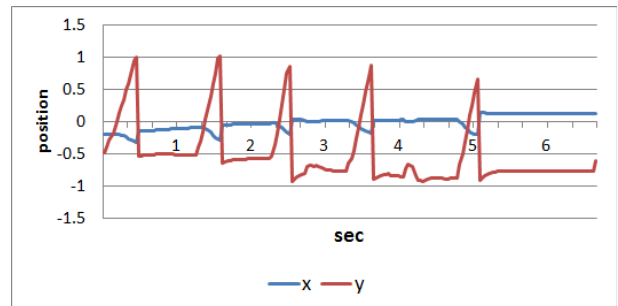


그림 7 반복적인 위쪽 스와이프 동작에서 손 위치
Fig. 7 Hand position of repeated swipe up gestures

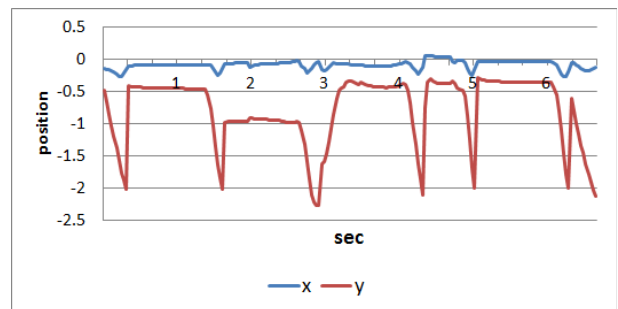


그림 8 반복적인 아래쪽 스와이프 동작에서 손 위치
Fig. 8 Hand position of repeated swipe down gestures

복적인 동작의 주기 등을 고려하여 자연스러운 제스처 인식이 이루어질 수 있도록 제스처 인식 모델의 지속적인 개선이 요구된다.

4.2 LCD Display Contents 제어 실험

본 절에서는 제스처를 이용하여 디스플레이 콘텐츠를 제어하는 실험에 대하여 기술한다. 실험에서는 Google Earth Street View[15]를 제어 대상 디스플레이 콘텐츠로 사용하였으며, 표 1과 같이 각 제스처를 화살표 키 명령으로 대응하였다. Google Earth Street View에서 각 화살표 키 명령은 표 1과 같은 행동으로 대응된다. 5명의 성인 남자를 대상으로 제어 실험을 수행하였으며 비교적 자연스러운 제어가 가능하였으나 디스플레이 콘텐츠를 보다 부드럽게 제어하기 위한 명령 처리 모델의 개선이 필요하다.

표 1 제스처와 명령 간의 대응

Table 1 Mapping between gestures and commands

Gesture (left hand)	Command	Action in Google Earth Street View
Swipe left	Left arrow key	Turn left
Swipe right	Right arrow key	Turn right
Swipe up	Up arrow key	Move forward
Swipe down	Down arrow key	Move backward



그림 9 디스플레이 콘텐츠의 제스처 기반 제어 실험
Fig. 9 Experiment of gesture-based control for display contents

5. 결 론

영상 광고판과 같은 디스플레이 설비 콘텐츠의 제어는 사전 지식이 없는 사용자일지라도 직관적이고 즉각적으로 조작할 수 있어야 하는데, 기존의 Kinect기반 제스처 인식 인터페이스들은 복잡한 제어 방식으로 미리 사용법을 알고 있어야 조작할 수 있다는 단점이 있다. 본 논문에서는 이를 해결하기 위한 자연스러운 제스처 인식 및 디스플레이 콘텐츠 제어가 가능한 기술을 제안하였는데, 동적인 제스처를 인식하여 사용자의 의도를 파악하므로 보다 직관적인 인터페이스가 가능하다. 본 연구에서 제안한 기술은 향후 마우스, 터치스크린과 같은 기존의 포인팅 장치를 대신할 수 있으며, 마케팅, 의료, 비즈니스, 컴퓨터 과학, 엔터테인먼트, 로봇틱스 등과 같은 다양한 분야에서 활용될 수 있을 것으로 기대된다.

Acknowledgement

This research was supported by Seokyeong University in 2013.

References

- [1] I. Oikonomidis, N. Kyriazis, and A.A. Argyros, "Efficient Model-based 3D Tracking of Hand Articulations using Kinect," Proc. of British Machine Vision Conf., pp. 101.1-101.11, 2011.
- [2] D. Wigdor and D. Wixon, Brave NUI World: Designing Natural User Interfaces for Touch and Gesture, Morgan Kaufmann Publishers, 2011.
- [3] J. Sung, C. Ponce, B. Selman, and A. Saxena, "Human Activity Detection from RGBD Images," AAAI 2011 Workshop, pp. 47-55, 2011.
- [4] W. Li, Z. Zhang, and Z. Liu, "Action Recognition Based on A Bag of 3D Points," Computer Vision and Pattern Recognition Workshops(CVPRW), pp. 9-14, 2010.
- [5] R. Munoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas, and A. Carmona-Poyato, "Depth Silhouettes for Gesture Recognition," Pattern Recognition Letters, vol.29, no.3, pp. 319-329, 2008.
- [6] P. Suryanarayan, A. Subramanian, and D. Mandalapu, "Dynamic Hand Pose Recognition using Depth Data," Int'l Conf. on Pattern Recognition, pp. 3105-3108, 2010.
- [7] Kinect for Windows, <http://www.microsoft.com/en-us/kinectforwindows/>
- [8] P. J. Bristeau, F. Callou, D. Vissiere, N. Petit, "The Navigation and Control Technology Inside the AR.Drone Micro UAV," IFAC World Congress, pp. 1477-1484, 2011.
- [9] P. Mistry and P. Maes, "SixthSense: A Wearable Gestural Interface," Proc. of SIGGRAPH Asia 2009 Sketches, Yokohama, Japan. 2009.
- [10] J. C. Lee, "Hacking the Nintendo Wii Remote," Pervasive Computing, IEEE, vol.7, issue 3, pp. 39-45, 2008.
- [11] S.-H. Jang, J.-W. Yoon, and S.-B. Cho, "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.
- [12] A. Sanna, F. Lamberti, G. Paravati, and F. Manuri, "A Kinect-based Natural Interface for Quadrotor Control," Entertainment Computing, vol.4, issue 3, pp. 179-186, 2013.
- [13] T. Osunkoya and J. C. Chern, "Gesture-based Human Computer Interaction using Kinect for Windows Mouse Control and Powerpoint

Presentation," Proc. of Midwest Instruction and Computing Symposium, Wisconsin, USA, 2013.

[14] K. T.-M. Tran and S.-H. Oh, "Hand Gesture Recognition for 3D-Heritage-Tourism using Microsoft Kinect Sensor," Advanced Science and Technology Letters, vol.30, pp. 145-148, 2013.

[15] Google Earth, <http://www.google.com/earth/>

저 자 소 개



노 정 규 (盧 定 奎)

1991년 서울대학교 계산통계학과 이학사,
1993년 서울대학교 전산학과 이학석사,
1999년 동 대학원 이학박사, 1999~2002
년 삼성전자 통신연구소 책임연구원, 현
재 서경대학교 컴퓨터과학과 조교수
관심분야 : 동작인식기술, 소프트웨어설
계, 분산시스템, 데이터처리