모션 캡쳐를 이용한 기타 리듬게임

박동규* · 정정수**

Guitar Rhythm Game Using Motion Capture

DongGyu Park* · JeongSu Jeong**

이 논문은 2011-2012년도 창원대학교 연구비를 지원받았음

요 약

키넥트는 이용자의 신체를 이용하여 게임과 엔터테인먼트를 경험할 수 있는 주변기기로 2010년 초 마이크로소프트사에서 발표하여 공개하였다. 본 논문은 키넥트에서 제공하는 세 가지 센서를 이용한 동작 인식 기능을 이용한 기타리듬 게임의 개발과 관련 기술에 대하여 다룬다. 리듬 게임은 게임의 여러 장르 중에서 매우 단순하고 학습기간이 짧으며, 신체의 활발한 활동성과 리듬성을 이용하기 때문에 피씨, 콘솔기기, 스마트폰 등 다양한 기기에서도 널리 활용되고 있는 장르이다. 본 논문에서 구현한 리듬게임은 화면구성과 게임화면을 DirectX 11버전에서 구현하였으며, 키넥트를 이용하여 게이머의 손동작 인식을 수행하기 위하여 OpenNI API를 사용하였으며 신체 움직임을 표현하기 위하여 OpenGL 라이브러리를 사용하였다.

ABSTRACT

Microsoft released motion sensing device named Kinnect on early 2010, which is available for developing Xbox 360 game and entertainment software. Also many third party libraries and applications are developed for Kinnect. In this paper, we studied and developed guitar rhythm game on Kinnect using three sensor units on it. Rhythm game is the most popular game genre for many game devices including PC, console device, and smart phone. The main reasons for their popularity depends on their simplicity, short learning time and physical activity with rhythms. We developed the game screen and layout on DirectX 11, also we used OpenNI library for recognize physical activity of gamer's body and fingers, and OpenGL for body gestures on the game.

키워드

키넥트, 리듬게임, 음성처리, 이미지처리, 모션 캡쳐

Key word

Kinnect, Rythm Game, Speech Processing, Image Processing, Motion Capture

* 정회원 : 창원대학교 정보통신공학과 부교수(교신저자, dongupak@gmail.com)
** 정회원 : 창원대학교 정보통신공학과 정보시각화 연구실 연구원

접수일자 : 2013. 02. 05
심사완료일자 : 2013. 03. 28

Open Access http://dx.doi.org/10.6109/jkiice.2013.17.5.1106

ⓒThis is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(http://creativecommons.org/li-censes/by-nc/3.0/) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
Copyright Ⓒ The Korea Institute of Information and Communication Engineering.

Ⅰ. 서 론

게임기술은 최근 각광받고 있는 여러 소프트웨어 분야 중에서 가장 인기 있는 카테고리이며 정보통신 산업에서도 매우 중요한 시장을 차지하고 있다. 시장조사기관인 DFC intelligence에 의하면 2012년 670억불에 달하는 콘솔게임시장은 2017년에는 820억불에 달하는 거대한 시장으로 성장할 것으로 예상하고 있다[1-3].

2010년 마이크로소프트사는 별도의 컨트롤러 없이 인체의 움직임을 감지하는 기기인 키넥트라는 게임 주변기기를 개발하여 발표하였다. 키넥트는 RGB 카메라와 IR 카메라 등의 센서와 멀티어레이 마이크로폰을 이용한 동작인식과 거리인식기능이 탐재되어 있으며 이를 통한 3D객체의 깊이인식까지도 가능한 매우 강력한 기능의 디바이스로 각광받고 있다[2-4].

게임의 다양한 장르 중에서도 특히 리듬게임류의 활동성 게임은 음원과 이 음원의 리듬에 따른 점수를 획득하는 구조로 인하여 몰입도가 매우 높은 것이 특징이다. 또한 활동성을 수반하고 있는 특성으로 인하여 손과 발의 협응 훈련등 시지각장애인은 위한 기능성 게임으로서의 장점도 가지고 있다[5]. 반면에 게임의 특성상 몰입도를 높이기 위해서는 그 리듬게임을 위한 고가의 전용장비가 요구된다는 단점이 있다. 이러한 특성으로 인하여 리듬게임류의 활동성 게임은 오락실과 같이 게임을 위한 전용장비가 확보되지 않은 곳에서는 이용하기가 불가능하다.

본 논문에서는 마이크로소프트사의 키넥트를 이용하여 비교적 저가의 PC 환경에서도 가상의 음원연주 기능을 구현하여 게임의 몰입도를 높이도록 하였다. 본 논문에서는 키넥트의 다양한 센스를 이용하여 동작인식 기능을 구현하였으며, 이 기능과 음원동기화 기법을 통하여 개발한 리듬게임 개발기술에 관하여 다룬다.

Ⅱ. 키넥트 게임 디바이스와 게임 동향

2.1. 키넥트 게임 디바이스

그림 1에 나타난 바와 같이 키넥트는 원래 게임 콘솔인 XBox 360상에서 이용할 수 있는 게임을 위해서 출시되었다.

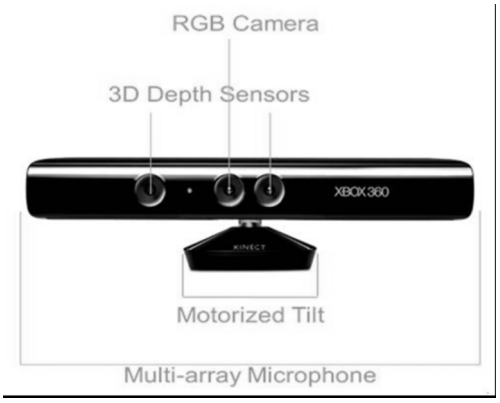

그림 1. 키넥트의 센서와 구조.
Fig. 1 Kinnect sensor and its layout.

기존의 닌텐도나 소니에서 선보였던 손이나 신체에 센서 컨트롤러를 들거나 부착하는 방식이 아니라 사람의 동작을 별도의 외부 가속도 센서 없이 직접 인식하는 혁신적인 방식을 채택하고 있다. 키넥트는 적외선 센서, RGB 카메라, 깊이 센서의 세 가지 센서를 통해서 사람의 동작을 비교적 정확하게 인식하는 것이 가능하다. 또한 마이크 모듈을 통하여 음성을 인식하는 기능까지 갖추고 있다.

키넥트는 부가적으로는 사람을 잘 인식하기 위해서 받침대와 목사이의 관절에 해당하는 모터 피봇(motorized pivot)이 설치되어 있어 동작자를 능동적으로 찾아서 시선을 조정하기도 하는 다양한 종류의 인식기능을 제공해 준다. 이에 따라 다양한 키넥트 기반 게임 소프트웨어들이 출현하고 있으며 피시상에서도 활용할 수 있는 API들도 개발되고 있다[1,4,6].

2.2. 리듬게임 동향

리듬게임은 많은 사람들이 즐겨하는 게임 장르로써 음악성과 활동성을 결합한 게임으로 국내에서 큰 인기를 얻고 있다.

리듬게임 중에서 미국과 유럽지역에서 인기를 얻고 있는 게임으로는 기타 히어로와 같은 게임이 있는데 이 게임은 2005년 하모닉스사에서 출시하여 콘솔, 닌텐도 DS, 윈도, 맥용뿐만 아니라 스마트폰용으로도 출시되어 큰 인기를 얻고 있으며, 여러 종류의 시리즈물로 지속적으로 출시되고 있다. 뿐만 아니라 록밴드와 같은 다양한 리듬게임도 지속적으로 개발되고 있다[7-10].

또한 리듬게임은 음악성과 활동성을 통한 게임 장르라는 특성으로 인하여 중풍환자의 치료를 위한 물리치료에도 활용되고 있으며 건강상의 목적으로 활용하려는 연구가 이루어지고 있다[2,10].

Ⅲ. 시스템 구성과 구현방식

3.1. 시스템 구성

그림 2는 본 논문의 리듬게임의 시스템 구성도이다. 본 논문에서 구현한 시스템은 그림 2와 같이 크게 키넥트 이미지 처리, 피씨 게임 응용프로그램, 스피치 처리부로 나누어져 있다.

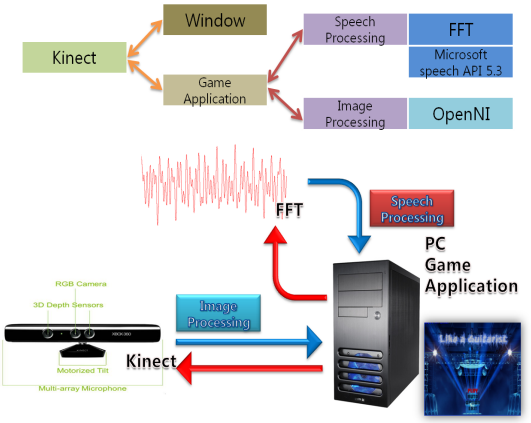

그림 2. 본 논문의 시스템 구성도
Fig. 2 System architecture

키넥트는 삼차원 깊이 센서를 이용하여 사람과 사물의 깊이를 인식한다. 이때 감지한 깊이 데이터는 피씨의 OpenNI(Open Natural Interaction) API가 받아 영상처리를 하게 된다.

OpenNI는 PrimeSence에 의해 개발되고 있는 오픈 소스 프레임워크이다. OpenNI는 전신 모션 트래킹을 지원하는데, Primesense와 ASUS WAVI Xtion 센서 등의 센서를 지원할 뿐만 아니라 윈도우, 리눅스, 맥 운영체제 등의 다양한 운영체제에서도 사용할 수 있다.

또한 OpenNI는 키넥트뿐만 아니라 여러 종류의 오디오나 비주얼 디바이스 센서에서도 작동할 수 있는 C++와 C# 언어로 작성된 표준 API를 제공한다. OpenNI의 주요 API는 음성과 음성 명령 인식, 손 제스처, 신체 움직임 트래킹 등이 있다.

본 시스템은 1) OpenNI API를 통하여 가장 먼저 사람의 골격을 인식하고 팔, 다리, 몸 등을 구분하고, 2) DirectX를 통해 스프라이트, 애니메이션 등 게임엔진을 제작하여 이를 3) OpenNI와 연동을 시키는 단계를 거치게 된다.

또한 이러한 기능을 이용하여 기타연주를 할 수 있는 게임을 개발하였다. 본 논문에서 구현한 게임을 실행하면 최초 사용자는 칼리브레이션을 한 후, 아무런 컨트롤러 없이 게임을 진행하게 된다. 그리고 한 음원의 게임을 클릭하게 되면 시스템은 사용자의 두 손을 추적하여 가상의 기타를 디스플레이에 그리며, 사용자는 이것을 보고 가상의 환경에서 손쉽게 기타 게임을 즐길 수가 있다.

3.2. 시스템 구현 기술

3.2.1. OpenNI 인터페이스

키넥트는 저가의 깊이 카메라 센서를 이용하여 실시간 RGB 영상, 깊이 값, 관절의 추적정보를 얻을 수 있다. 이전의 연구에서는 다양한 순차적 데이터 학습방법등을 이용하여 팔과 신체의 움직임을 분석하고 제스처를 인식하였다[6]. 본 연구에서 키넥트를 사용하기 위한 인터페이스로는 OpenNI를 이용한 모션 캡쳐 기능을 이용하였다. 키넥트는 3가지 센서를 이용하여 사람의 움직임을 인식하게 된다. 이를 통해 손과 발, 무릎, 머리의 위치를 알 수 있다.

키넥트를 이용하여 모션 캡쳐를 하기 위해서는 몇 가지 절차가 필요하다. 우선 칼리브레이션 과정이 필요하다. 그림 3에서는 키넥트의 카메라에서 촬영한 영상을 이용하여 사람의 골격을 추출하기 위한 사전 칼리브레이션 과정을 보여준다. 칼리브레이션은 신체의 전체가 감지될 수 있는 위치에서 키넥트가 트래킹 가능한 동작을 취하여 인식한다. 이 과정은 움직이는 물체를 라벨링을 하고 그 라벨링 된 물체가 특정한 자세를 취하게 되면 그 때 칼리브레이션을 시작하며, 이 과정을 통해 사람의 골격을 추출하게 된다. 이렇게 추출된 골격은 움직임에 따라 같이 움직이게 되며 손, 발, 머리 등의 위치를 추출하게 된다.

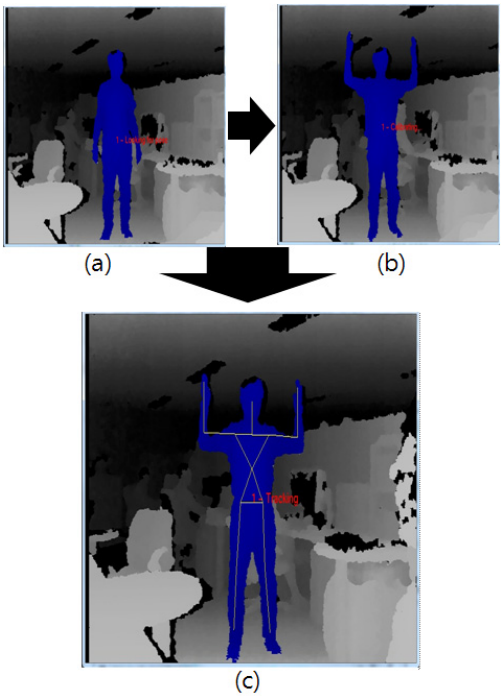

그림 3. (a) 신체골격에 대한 레이블링, (b) 칼리브레이션, (c) 스켈리톤 추출과정
Fig. 3 (a) Labeling on skeleton, (b) calibration step, (c) extracting skeleton

3.2.2. 영상 처리

기타를 연주할 사용자의 캐릭터를 위한 골격 정보를 추출한 다음 과정은 기타 연주를 위한 손의 위치를 추적(Hand tracking)하는 과정이다. 여기서 손 위치추적을 하는 방법으로 히스토그램을 통하여 얻은 사람의 깊이 정보와, 외곽선 끝 단을 이용해서 손이라는 것을 찾을 수 있다. 우선 이진화와 레이블링 작업을 통해서 골격인식과 손의 위치가 추적된 선행 작업들에 의해서 가능하다. 본 논문에서는 그림 4와 같이 컨벡스 헐(Convex Hull), 컨벡시티 디펙트(Convexity Defects)를 이용해서 손과 손가락 모양을 찾아낸다.

컨벡시티 디펙트는 컨벡스 헐과 컨투어 사이의 이탈부(defect)를 찾는 과정이다. 이 방법은 사용자 손을 흑백 실루엣으로 표현할 경우 나타나는 튀어나온 부분(손가락)과 그 골을 추적하는 과정을 통해서 골의 좌표 목록을 얻을 수 있다.

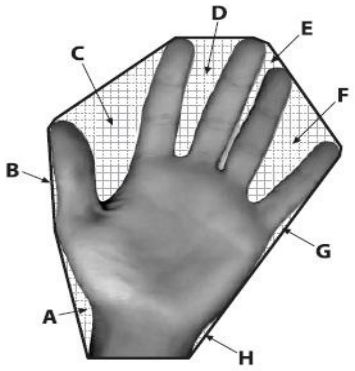

그림 4. 컨벡스 헐 추출과정과 컨벡시티 디펙트
Fig. 4 Convex hull extracting steps and convexity defects

이 과정에서 손가락 추적을 할 때 왼손과 오른손을 따로 트래킹한다. 오른 손잡이의 경우 일반적으로 왼손은 기타의 목 부분을 잡으며, 기타에서 나는 소리의 화음을 제어한다. 또한 오른손은 여러 기타 줄을 한 번에 연주하는 스트러밍(Strumming)이라는 주법을 활용하기위해 트랙킹 한다.

3.2.3. 음성 처리

본 논문에서는 게임에서 제공하는 연주 음원과 연주자가 노래를 직접 불러서 얻어지는 음원을 합치기 위해 사용하는 라이브러리로 마이크로소프트 스피치(Microsoft Speech) API 5.3을 이용하였다. 일차적으로 마이크로소프트사에서 제공하는 스피치 라이브러리(Speech Library)를 이용하여 더 빠르고 질 좋은 음원파일을 얻어내는데, Microsoft에서는 5.1버전부터는 윈도에 기본적으로 내장하여 제공된다. 구조는 그림 5와 같다.

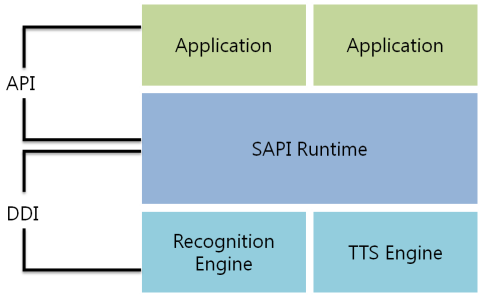

그림 5. 마이크로소프트사의 스피치 API 5.3 아키텍처
Fig. 5 Speech API 5.3 architecture on Microsoft

스피치 API는 음성인식 엔진과 TTS (Text-To-Speech) 엔진을 통한 텍스트 변환 기능을 제공하고 있다. 이러한 엔진기반위에서 SAPI(Speech API) 런타임기를 통한 응용프로그램 개발이 이루어지는 구조를 가지고 있다.

키넥트에서 제공하는 마이크를 이용하여 연주자가 직접 부른 노래에는 주변 환경으로 인하여 여러 가지 잡음이 들어가게 된다. 이러한 주변 소음을 제거하고 더욱 깔끔한 음원을 구하기 위해서 백색잡음의 제거 기능이 필요한데 이를 위하여 직접 구현한 필터를 이용한 잡음 주파수 영역에서의 잡음 제거를 추가하였다.

Ⅳ. 게임 시스템 진행

본 장에서는 실제 게임이 진행되는 순서와 실행 이미지를 이용하여 게임 시나리오를 설명하도록 한다.

그림 6. 구현된 게임의 메인 화면으로 Play, Auto, Option, Exit 메뉴로 구성되어 있다
Fig. 6 Main menu on the game. The menu has Play, Auto, Option and Exit

게임을 실행하게 되면, 위 그림 6과 같은 게임의 메인 화면이 나타나게 된다. 메인 화면에는 Play, Auto, Option, Exit 메뉴가 있으며 Play는 게임진행, Auto는 음원추출을 통한 자동 게임진행, Option은 사운드조절을 비롯한 부가적인 게임 설정작업을 할 수 있다.

그림 7. 음악선택 장면-사용자는 원하는 장르와 곡을 선택할 수 있다
Fig. 7 Music selection scene-users can choose genre and title of the music

그림 7은 음악 선택화면이다. 이 화면에서 사용자는 자기 자신이 원하는 장르의 음악을 고를 수 있고, 이 음악을 플레이 할 수가 있다.

그림 8. 게임 플레이 화면-화면의 좌측은 게임진행 화면이며, 우측은 실제 게이머가 게임을 진행하는 모습이다
Fig. 8 Game play scene-Game scene is on the left and player's game playing scene is on the right

음악 선택 장면에서 음악을 선택하게 되면 위의 그림 8과 같은 게임화면이 나오게 된다. 왼쪽에는 DirectX 기반의 게임 화면이 나오게 되며 음악에 따라 각각의 노드들이 다르게 떨어지는 방식의 리듬 게임이다. 게임 구조는 2D로 제작되었지만, 3D의 느낌으로 노드들이 떨어지게 된다. 그리고 오른쪽 화면은 OpenGL로 구현한 화면으로 키넥트의 카메라에서 비추는 사용자 영상을 보여 준다. 그리고 손 위치를 추적하여 가상의 기타가 나오며, 가상의 기타는 왼손, 즉 코드 잡는 부분이 손가락으로가 아닌 손위치를 변경해주면서 각각의 노드들을 맞추어야 한다. 게임 중에 몰입도를 증가시키기 위한 이벤트성 문구가 나오게 되는데 이것에 맞춰서 율동을 하게 되면 보다 많은 점수를 획득할 수 있게 된다. 율동은 헤드뱅잉과 박수가 있다. 그림 8의 상단은 박수 율동 장면이며, 하단은 해드뱅잉 율동 장면이다. 이와 같은 기법을 통해 이용자는 기타 솔로 연주처럼 다양한 행동과 연주를 하면서 사용자의 흥미를 높일 수 있다.

메인 화면에서 Auto를 선택하게 되면, 기본적으로 제공되지 않는 음원에 대하여 게임을 할 수 있게 FFT를 이용한 음성처리를 이용하여 주파수 데이터를 얻게 되고 이를 저장하고 이를 통하여 게임에 필요한 데이터를 생성하게 된다. 이는 이 작업과 동시에 저장되게 되며, 다음에도 다시 사용할 수 있게 된다. 이러한 방법을 통하여 게임에서 기본적으로 제공되는 음원뿐만 아니라, 사용자가 원하는 노래, 좋아하는 노래를 이용하여 게임을 즐길 수 있게 된다.

Ⅴ. 결론 및 향후 연구 방향

본 연구에서는 PC 기반의 리듬 게임에 음원 추출기능과 합성기능, 그리고 활동성을 위하여 키넥트를 연동한 리듬 게임에 대하여 연구하고 이를 구현하였다. 기존의 PC용 리듬 게임과 다르게 키넥트를 PC와 연동시켜 다양한 영상처리를 통한 활동성 게임을 구현하였다. 그리고 사용자가 원하는 음악으로 게임을 할 수 있도록 음악 파일만 있으면, 음원을 추출하여 그 음악으로 게임을 할 수 있도록 개발하였다. 사용자는 아무런 컨트롤러 없이 오른쪽 키넥트 화면을 보면서 게임을 조작할 수 있으며, 게임 중의 가상의 기타를 이용하여 왼쪽에 나오는 게임 에 맞춰 흥겨운 리듬 게임을 즐길 수 있다. 그리고 다양한 이벤트가 있어서 흥미와 몰입도를 높이고 있다. 이러한 활동성 게임은 저가의 PC 환경에서 활용할 수 있는 장점으로 인하여 시지각 장애인의 시지각 및 협응훈련을 위한 기능성 게임으로 이용될 수 있다.

향후 연구과제로 네트워크 기능을 통한 다자간 협연기능과 점수 기록 및 순위 연동 등의 기능구현과 소셜 미디어를 활용한 공유 기능 등에 대한 연구가 필요하다.

뿐만 아니라 손가락의 위치인식에 대한 정확도 개선이 요구되면 몰입도를 더욱 증가시키기 위한 게임 효과 및 게이머의 피드백에 대한 연구도 필요할 것이다.

감사의 글

본 연구는 2011-2012년도 창원대학교 연구비에 의하여 연구되었음

참고문헌

[1] 윤기선, “키넥트를 이용한 인터랙티브 멀티미디어 작품 제작 연구 : 멀티미디어음악작품 <Kinect Illusion>을 중심으로”, 동국대학교 영상대학원 석사학위 논문, 2012, 2.

[2] M. Siddiqui and G. Medioni, ″Human pose estimation from a single view point, real-time range sensor″, Workshop on Computer Vision for Computer Games at Conference on Computer Vision and Pattern Recognition, pp. 1297-1298, 2010.

[3] 한혜민, “초음파 및 키넥트 센서 융합기반의 실내 환경 모델링”, 고려대학교 대학원 석사학위 논문, 2012, 2.

[4] 이주원, 오경수, “키넥트를 활용한 요가 학습 콘텐츠”, 한국 HCI학회 학술대회, No. 1, pp. 221-223, 2012.

[5] 권정민, 김민영, “기능성 게임을 이용한 지적 장애인 직무훈련 : 직업 기능 교과서의 게임화를 위한 콘텐츠 연구”, 한국컴퓨터게임학회논문지, Vol. 25, No. 4, pp 35-46, 2012.
[6] 조선영, 변혜란, 이희경, 차지훈, “키넥트 센서 기반 슈팅 게임을 위한 팔 제스처 인식”, 한국정보과학회 논문집, Vol. 39, No. 10, pp 796-805, 2012.
[7] Forbes, 2012 년 7월 18일, http://www.forbes.com/sites/johngaudiosi/2012/07/18/new-reports-forecasts-global-video-game-industry-will-reach-82-billion-by-2017/
[8] 워싱턴 타임즈, Healing with Guitar Hero, 2008년 9월 24일, http://www.washingtontimes.com/news/2008/sep/24/you-dont-have-to-be-a-quick-fingered-15-year-old-t/
[9] Unnithan, V.B., Houser, W., Fernhall B., "Evaluation of the energy cost of playing a dance simulation video game in overweight and non-overweight children and adolescents". International Journal of Sports Medicine. Vol. 27, No. 10, pp. 804-809, 2006, Oct.
[10] 강민아, 강진호, 조은선, “키넥트를 이용한 소셜 네트워크 서비스”, 한국정보과학회 학술발표논문집, Vol. 39, No. 1A, pp 361-363, 2012.

저자소개

박동규(DongGyu Park)

1993년 부산대학교 전자계산학과(이학사)
1996년 부산대학교 전자계산학과(이학석사)
1999년 부산대학교 전자계산학과(이학박사)
2000년～2002년 영산대학교 멀티미디어 공학과 전임강사
2002년～현재 창원대학교 정보통신공학과 부교수
※관심분야 : 모바일 콘텐츠, 모바일 게임, 컴퓨터 그래픽스, 보완대체 의사소통도구(AAC)

정정수(JeongSu Jeong)

2008년～현재 창원대학교 정보통신공학과 정보시각화 연구실 연구원
※관심분야 : 모바일 콘텐츠, 모바일 게임, 키넥트 컨트롤러와 게임 콘텐츠