

퀴즈게임의 체감형 제스처 인터페이스 프로토타입 개발

안정호* 고재필**

요약

우리는 본 논문에서 사용자 제스처 인터페이스 기반 퀴즈게임 콘텐츠를 제안한다. 우리는 기존의 아날로그 방식으로 수행해 오던 퀴즈게임의 요소들을 파악하여 디지털화함으로써 퀴즈 진행자의 역할을 콘텐츠 프로그램이 담당할 수 있도록 하였다. 우리는 키넥트 카메라를 사용하여 깊이영상을 획득하고 깊이영상에서 사용자 분할, 머리 위치 검출 및 추적, 손 검출 등의 전처리 작업과 손들기, 손 상하이동, 주먹 모양, 패스, 주먹 쥐고 당김 등의 명령형 손 제스처 인식기술을 개발하였다. 특히 우리는 사람이 일상생활에서 물리적인 객체를 조작하는 동작으로 인터페이스를 위한 제스처를 정의함으로써 사용자가 이동, 선택, 확인 등의 추상적인 개념을 인터페이스 과정에서 체감할 수 있도록 디자인하였다. 앞서 발표되었던 선행 작업과 비교할 때, 우리는 승리 팀에 대한 카드보상 절차를 추가하여 콘텐츠의 완성도를 높였으며, 손 상하이동 인식과 주먹 모양 인식 알고리즘 등을 개선하여 문제 보기선택의 성능을 크게 향상시켰고, 체계적인 실험을 통해 만족할 만한 인식 성능을 입증하였다. 구현된 콘텐츠는 실시간 테스트에서 만족스러운 제스처 인식 결과를 보였으며 원활한 퀴즈게임 진행이 가능하였다.

A Study on Tangible Gesture Interface Prototype Development of the Quiz Game

Jung-Ho Ahn* Jaepil Ko**

Abstract

This paper introduce a quiz game contents based on gesture interface. We analyzed the off-line quiz games, extracted its presiding components, and digitalized them so that the proposed game contents is able to substitute for the off-line quiz games. We used the Kinect camera to obtain the depth images and performed the preprocessing including vertical human segmentation, head detection and tracking and hand detection, and gesture recognition for hand-up, hand vertical movement, fist shape, pass and fist-and-attraction. Especially, we defined the interface gestures designed as a metaphor for natural gestures in real world so that users are able to feel abstract concept of movement, selection and confirmation tangibly. Compared to our previous work, we added the card compensation process for completeness, improved the vertical hand movement and the fist shape recognition methods for the example selection and presented an organized test to measure the recognition performance. The implemented quiz application program was tested in real time and showed very satisfactory gesture recognition results.

Keywords : Quiz Game Contents, Tangible Gesture Interface, Hand Shape Recognition.

1. 서론

최근 정보통신 기술의 발달로 컴퓨터를 통한 가상공간(Virtual space)을 만들어 현실 세계의 시간과 공간의 한계를 극복하려는 노력이 계속되고 있다. 이러한 노력은 실제와 유사하지만 실

※ 제일저자(First Author) : 안정호
접수일:2012년 04월 23일, 수정일:2012년 05월 17일
완료일:2012년 06월 19일
* 강남대학교 컴퓨터미디어정보공학부
jungho@kangnam.ac.kr
** 금오공과대학교 컴퓨터공학과
nonezero@kumoh.ac.kr
▣ 본 연구는 문화체육관광부 및 한국콘텐츠진흥원의 2010년도 콘텐츠산업기술지원사업의 연구결과로 수행

되었음.

제가 아닌 어떤 특정한 환경이나 상황에서 사용자의 오감을 자극하며 실제와 유사한 공간적, 시간적 체험이 가능한 가상현실(Virtual Reality)에 대한 연구로 이어졌다. 또한 가상현실에 대한 연구는 더 나아가 가상현실 세계와 인간이 살고 있는 현실 세계가 서로 소통하려는 연구로 이어져 증강현실(Augmented Reality)이라는 연구 분야가 탄생하였다.

이러한 증강현실에 대한 연구와 더불어, 크고 작은 특정한 가상공간을 만들어 이와 소통할 수 있는 체감형 인터페이스(Tangible interface) 기술 개발이 최근 큰 주목을 받고 있다. 대표적인 예로 키넥트(Kinect)를 이용한 모션인식 게임을 들 수 있다. 키넥트 게임은 게임을 위한 가상공간을 만들고, 실시간 사용자 모션 인식을 통해 사용자의 의도를 파악하여 이 메시지를 가상공간의 객체들에 전달함으로써 가상공간의 환경을 변화시킨다.

본 논문에서 소개하는 퀴즈 게임 콘텐츠도 퀴즈 게임 진행을 위한 가상공간과 현실 세계의 사용자들이 소통하는 방식을 취한다. 우리는 기존의 아날로그 방식으로 진행되어 오던 퀴즈게임의 요소들을 파악하고 이를 프로그래밍 하여 디지털 콘텐츠화 함으로써 퀴즈 진행을 위한 가상 콘텐츠를 만들었다. 그리고 카메라를 통하여 현실 세계의 사용자들의 제스처(동작)를 인식함으로써 퀴즈 콘텐츠와 소통할 수 있는 제스처 인터페이스를 구현하였다. 제스처는 기존 아날로그 방식의 제스처를 최대한 수용하여 정의하였고, 입력 카메라는 최근 컴퓨터 비전 연구에서 주목받기 시작한 키넥트 카메라를 사용하였다.

제안하는 콘텐츠는 퀴즈 게임과 카드 보상의 두 부분으로 구성되어 있다. 퀴즈 게임은 4 인이 두 팀으로 나누어 참가하며 퀴즈 게임에서 승리한 팀에게는 영어 학습을 위한 영어 단어 카드가 보상으로 주어진다. 제한된 공간에서 한 명의 사용자가 컴퓨터를 이용하여 참여하는 기존의 게임 방식과 달리 제안하는 퀴즈 게임은 사회성(sociality)을 강조하여 팀 별로 수행하도록 기획하였다.

우리는 2절에서 제안하는 퀴즈게임 콘텐츠의 전체적인 개요를 설명하고, 3절에서 퀴즈 문제 풀이를 위한 개선된 제스처 인식 방법론을 소개한다. 4절에서는 카드 보상 단계에서 정의된 제

스처를 인식하기 위한 방법론을 소개한다. 그리고 5절에서 제안하는 퀴즈게임 콘텐츠의 한계점과 실험결과를 기술하고, 6절에서 결론 및 토의 사항을 기술한다.

1.1 관련 연구

우리는 퀴즈 게임 콘텐츠를 위한 제스처 인터페이스에 대한 초기 연구의 결과를 발표한 바 있다[1]. 이는 퀴즈 게임의 디지털화를 위한 콘텐츠 구성과 사용자 인식 및 해당 제스처 인식 알고리즘에 대한 연구였다.

우리는 이후 많은 실시간 테스트를 통해 초기 방법론의 문제점을 파악하고 개선하였다. 또한, 퀴즈문제 풀이 후 승리한 팀의 팀원들에게 주어지는 보상 절차에 대한 콘텐츠와 관련된 체감형 제스처 인터페이스를 추가하였다. 본 논문은 이 개선 사항과 추가된 콘텐츠 및 인식 기술들 등 최종 퀴즈게임 콘텐츠를 기술하였다.

<표 1>은 우리가 제안하는 퀴즈게임 콘텐츠를 위해 개발한 컴퓨터비전 기술을 보여 준다.

<표 1> 개발 세부 기술

개발 기술	세부 기술
사용자 인식	- 사용자 분할 - 사용자 위치 및 신원 파악 - 손들기 제스처 인식
퀴즈 보기 선택 제스처 인식	- 손 이동: 상하 이동 거리 인식 - 손 모양 인식 (주먹)
카드 선택 제스처 인식	- 패스 제스처 인식 - 손 모양 인식 (주먹) - 손 당김: 전후 이동 거리 인식

손 검출 및 손 모양 인식에 대한 연구는 휴먼 컴퓨터 인터페이스(HCI)를 위해 많은 연구가 되어왔다. 손은 보는 각도에 따라 모양, 모션, 텍스처의 변화가 심하기 때문에 어떤 특징을 추출하는가에 따라 인식 성능이 크게 좌우된다. 손 특징 추출 방법으로 손과 관절의 운동학적인 모수(kinematic parameter)를 추정하여 영상으로부터 추출된 2D 특징을 에지기반 3D 손 모델에 대응시켜 손동작을 인식하는 모델기반 접근방식이

한 때 주된 연구 방법이 있다.[13] 하지만 실제 2D 영상에서 이러한 운동학적 모델에 매칭하는 과정의 부정확성으로 2D 손 제스처 영상들의 집합으로 제스처를 모델링한 외형기반 접근 방식 [7]과 하위 레벨 영상의 특징 추출에 기반을 둔 제스처 인식 방식이 있다[12]. 본 논문에서 제시하는 손 특징 추출 방식은 하위 레벨 특징 추출 방식에 속한다.

손 제스처 인식 방법론은 입력 특징들 사이에 수동으로 부호화된 일련의 규칙(rule)을 구성하여 하위 레벨 특징들로 부터 손동작을 인식하는 규칙기반 인식 방법[14]과, HMM(Hidden Markov Model)과 같이 인식 모델을 이용한 기계학습(Machine learning)기반 인식 방법이 있다 [2, 15, 16].

손 검출을 위한 방법론으로 손의 하위레벨 특징을 이용한 접근 방법과 기계학습 방법론을 이용한 접근 방법이 있다. 손 검출을 위해 사용되는 대표적인 하위 레벨 특징으로는 깊이[1], 컬러[2], 블랍, 리지(Ridge), 광흐름(Optical flow) [4] 등이 있으며, 기계학습 방법론으로는 입자 필터(Particle filter) 모델[4], Adaboost 모델[6] 등을 사용한 방법론들이 연구되어 졌다.

또한, 몸의 연속적인 동작을 인식하기 위한 방법론으로 MHI(Motion History Image), MEI (Motion Energy Image)[3, 5] 등이 있으며 여러 동작들이 혼재되어있는 장면에서 주요 제스처를 선택적으로 인식할 수 있는 방법론에 대한 연구 [10]가 있다.

사용자 분할을 위한 대표적인 방법론으로는, 시공간적 정보를 평균 이동 군집화(mean shift clustering)와 GMM(Gaussian Mixture Model)으로 모델링한 후 개선된 GrabCut 알고리즘을 이용하여 정밀한 사람 분할과 얼굴 및 포즈 인식을 수행한 연구[8]와, 연속된 프레임을 시공간적 3D 입체로 구성하여 조건적 랜덤 필드(Conditional random field)의 최적화 해를 이용한 연구[18]가 있다. 제안하는 퀴즈 콘텐츠는 정밀한 사용자 분할이 시스템 성능을 향상에 큰 도움이 되지 않기에, 사용자들 간의 경계를 짓는 방식으로 분할하였다.

제스처를 이용한 휴먼 컴퓨터 인터페이스 시스템에 대한 대표적인 연구로, 7개의 손 모양 인식결과를 조합해 5개의 명령형 제스처를 인식한

Kim et. al[9]과, 5개의 손 모양 인식을 통해 로봇의 동작을 제어한 Malima et. al[11], 음악 플레이어와 농구 게임 콘텐츠를 제작한 Yoon et. al[17], 등이 있다.

2. 퀴즈게임 콘텐츠의 개요

제안한 퀴즈 콘텐츠를 이용하여 교실에서 퀴즈게임을 수행하기 위해, 콘텐츠 UI를 TV 화면에 출력하고 키넥트 카메라는 TV를 설치한 테이블 위에 놓는다. 네 명의 학생들은 카메라와 1.8m ~ 2.3m의 거리에 착석하여, TV 화면에 표시된 지시에 대해 제스처로 반응하며 퀴즈 게임을 하게 된다. 제안하는 퀴즈게임 콘텐츠의 순차적인 흐름은 다음과 같다.

- (1) 출석체크
- (2) 팀 파악
- (3) 퀴즈문제 제시
- (4) 도전자 선택
- (5) 문제 풀이(보기 선택)
- (6) 정답 확인
- (7) 승리 팀 인식
- (8) 보상카드 선택
- (9) 카드정보 전송

퀴즈게임 프로그램이 실행 전 관리자(선생님)는 학생의 아이디와 닉네임, 보상카드 종류를 프로그램에 입력한다. 퀴즈게임을 시작하면 먼저 출석체크를 실시한다. 참여 학생의 닉네임이 화면에 표시되면 해당학생은 손을 들어 자신이 착석한 위치를 프로그램에게 알린다. (그림 1)은 출석체크 시 손들기 제스처의 예를 보여준다.



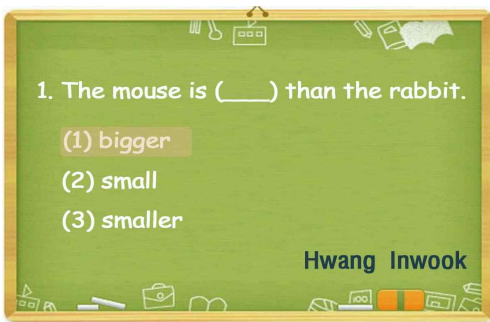
(그림 1) 출석체크를 위한 손들기 제스처

팀을 파악하기 위해 “A팀 손드세요.”라는 메시지가 화면에 표시되면 A팀의 팀원들은 손을 들어 자신이 속한 팀을 프로그램에 알린다. B팀도 동일하다. (그림 2)는 팀 파악을 위해 A팀의 팀원들이 손을 든 모습을 보여 준다.



(그림 2) 팀 인식을 위한 손들기 제스처

출석체크와 팀 파악이 완료되면 퀴즈문제 풀이를 시작한다. (그림 3)과 같이 세 개의 보기와 함께 영어 퀴즈 문제가 화면에 제시되면 정답을 맞히고자 하는 학생은 손을 든다. 이때 먼저 손 든 학생이 도전권을 획득한다.

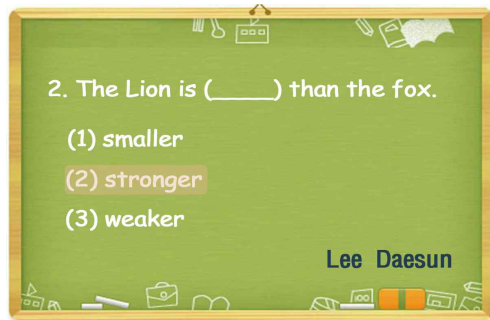


(그림 3) 퀴즈문제 초기화면

도전자는 제스처 인터페이스를 통해 원하는 보기를 선택할 수 있다. 초기 화면은 그림 3과 같이 1번 보기가 활성화되어 있는데, 손을 앞으로 뻗은 상태에서 손가락을 펴고 손바닥이 정면을 향한 상태에서 상하로 팔을 움직이면 다른 보기를 활성화시킬 수 있다. (그림 4)는 보기를 이동시키기 위해 정의된 제스처의 예를 보여 준다. (그림 5)는 보기이동 제스처에 의해 다음 보기로 이동한 콘텐츠 영상을 보여준다.



(그림 4) 보기 이동 제스처

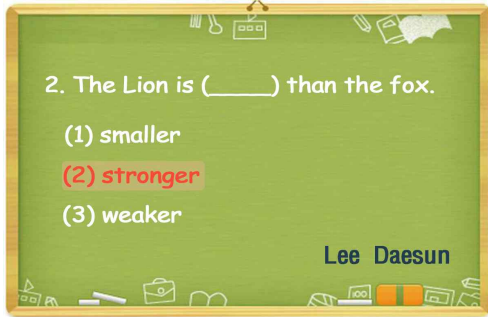


(그림 5) 손 이동 제스처에 의한 보기이동

도전자는 원하는 보기가 활성화되었을 때 주먹을 움켜쥐어 해당 보기를 선택할 수 있다. 다음 (그림 6)은 주먹을 쥔 보기 선택 제스처를 보여 준다. (그림 7)은 보기 선택 제스처에 의해 보기가 선택되었을 때의 모습을 보여 준다.



(그림 6) 보기 선택 제스처

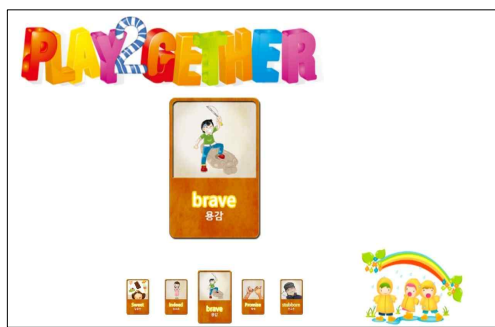


(그림 7) 보기 선택 제스처에 의한 보기 선택

선택한 보기가 정답일 경우, 도전자가 속한 팀의 점수가 100점 가산되며 다음 문제로 이동한다. 오답일 경우, 동일한 문제에 대한 풀이로 돌아가며 먼저 손든 학생에게 도전권이 주어진다. 두 번째 도전자가 정답을 맞힐 경우 해당 팀에 100점이 가산되며, 두 번째 도전자도 오답을 제시한 경우 다른 세 번째 보기가 정답이므로 다음 문제로 이동한다. 오답 시에는 해당 팀의 점수가 50점 감점되어 화면에 표시된다.

준비된 문제에 대한 풀이를 모두 마치면 승리팀이 화면에 표시되고 승리한 팀의 각 팀원들에게 순차적으로 카드 선택 기회가 제공된다. 무승부의 경우 퀴즈 게임이 종료된다.

승리 팀의 팀원들은 제스처 인터페이스를 통해 5장의 카드 중 원하는 카드를 선택한다. 카드 보상 단계에서는 (그림 8)과 같이 TV 화면의 아래쪽에 선택할 수 있는 5장의 카드가 작게 표시되고, 그 중 선택된 한 장의 카드만 확대되어 화면 중앙에 표시된다. 확대된 카드는 현재 선택할 수 있는 카드를 의미한다.



(그림 8) 카드 보상 초기 화면

사용자가 손을 가로 방향으로 짓는 패스 제스처를 취하면 순차적으로 다음 카드가 화면 중앙에 표시된다. (그림 9)는 왼쪽 끝 사람이 패스 제스처 수행하고 있는 모습을 보여준다.



(그림 9) 패스 제스처

원하는 카드로 이동한 후 손을 앞으로 뻗어 카드를 움켜쥐듯이 주먹을 주면 선택되었다는 표시로 해당 카드는 1회 회전한다. 그리고 카드를 움켜쥔 상태로 가슴 방향으로 당기면 카드 선택이 재확인되며, 카드 정보가 사용자 아이디와 함께 서버에 전송된다. 다음 (그림 10)은 카드선택확인 제스처를 보여준다.



(그림 10) 카드선택확인 제스처

손을 저음, 주먹을 쥐, 당김 등의 과정을 통해 카드를 취하는 절차는 일상에서 사용자가 원하는 카드를 찾기 위해 카드를 넘기고, 카드를 쥐고, 카드 뭉치에서 빼내는 선택 과정을 은유(metaphor)적으로 디자인한 것이다. 이러한 인터페이스를 체감형 제스처 인터페이스(Tangible gesture interface)라 부른다.

제안한 퀴즈게임 콘텐츠는 초등학교 4, 5학년 을 대상으로 PC, 스마트폰, IPTV 등 3개의 스크린을 활용한 다중 공간 체감형 영어 교육용 게임 개발의 일부로 진행되었다. PC에서는 SNG (Social Network Game)를 통한 웹기반 교육 엔터테인먼트 프로그램, 스마트폰에서는 현장에서 위치기반 증강현실 체험학습 프로그램, IPTV에서는 교실에서 선생님과 오프라인 학습 후 진행하는 퀴즈게임 프로그램이 유기적으로 연계되어 진행된다. 참여 학생들은 각 매체에서 영어교육 관련 미션을 해결함으로써 보상으로 카드를 획득하게 되는데, 필요한 카드를 모두 모으면 다음 단계(stage)로 이동할 수 있다. 제안한 퀴즈 게임의 마지막에 수행하는 보상 카드 선택 및 전송의 필요성은 이 3-스크린 게임의 맥락에서 이해될 수 있다.

3. 문제 풀이를 위한 제스처인식

퀴즈 풀이 시, 원하는 보기를 선택하는 제스처는 사물을 취할 때 손을 핀 상태에서 팔을 상하로 이동하여 원하는 보기로 이동하고 주먹을 움켜쥐는 과정을 적용하여, 사용자들이 움켜쥐는 손의 체감과 선택이라는 능동적인 개념이 서로 연결될 수 있도록 하였다.

3.1 손 상하이동 인식

손의 이동을 감지하기 위해 우리는 추출된 손 영역의 하단 y 좌표(행)를 큐(Queue)에 저장하고 다음 판단함수 f_M 에 의해 보기 이동 여부를 결정한다.

$$f_M(y_t) = \begin{cases} -1 & d_{t,n_1} < T_M \text{ or } d_{t,n_2} < T_M \\ +1 & d_{t,n_1} > T_M \text{ or } d_{t,n_2} > T_M \\ 0 & \text{else} \end{cases}$$

여기서 -1 , $+1$ 은 각각 아래로 한 칸, 위로 한 칸 이동을 의미하며, $d_{t,n_1} = y_t - y_{t-n_1}$, $d_{t,n_2} = y_t - y_{t-n_2}$ 로 정의한다. 카메라와 사용자의 거리에 의존하는 이동 거리 한계 임계치 T_M 는 실험에서 5로 설정하고, n_1 , n_2 는 실험에서 각각 2와 4로 설정하였다. 사용자마다 손을 상하

로 이동시키는 속도의 차이가 있기 때문에, 인식 시간 오프셋을 의미하는 n_1 , n_2 설정하여 이동 속도에 독립적으로 보기 이동이 가능하도록 개선하였다.

3.2 주먹 모양 인식

우리는 이전 논문[1]에서 손바닥과 주먹, 기타 손 모양을 인식하였다. 제시한 콘텐츠는 손바닥 모양을 한 상태에서 보기를 이동하고, 주먹을 쥐면 해당 보기가 선택되어진다. 하지만 실제 콘텐츠 실행 상황에서 사용자마다 손바닥 모양이 일정하지 않아 인식률이 저하되었다. 따라서 우리는 유사 손바닥(palm-like)과 주먹(fist)의 두 클래스를 설정하고 특징 영역(feature space)에서 이를 구별할 수 있는 주먹 영역을 기술하였다.

주먹의 특징은 엄지 기준 세로길이 비(ratio)와 2사분면의 피부 비율 등 두 가지를 추출하였다. 엄지 기준 세로 비는, (그림 11)과 같이 추출된 손 영역에서 왼쪽 끝점(엄지)을 기준으로 위 부분의 높이 대비 아래 부분의 높이의 비를 의미한다. 2사분면의 픽셀비율은 (그림 11)에서 왼쪽 상단 박스의 크기 대비 검출 픽셀의 수로 정의하였다.



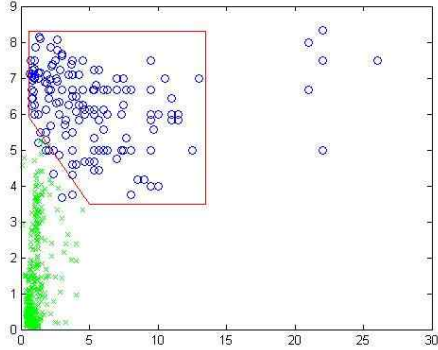
(a) (b)
(그림 11) 손 모양 특징 추출

이를 수식으로 정리하면 다음과 같다.

$$\left(\frac{h_L}{h_H}, \frac{n_S}{n_B} \right)$$

여기서 h_L 과 h_H 는 각각 왼쪽 끝점 기준 아래 부분의 높이와 위 부분의 높이이고, n_S 와 n_B 는 각각 2사분면의 검출된 픽셀 수와 전체 픽셀 수를 의미한다. 우리는 주먹 모양 인식을 위한 학습을 위해 4명의 사람으로부터 손바닥 모양과

주먹 모양의 특징을 추출하였다. 다음 (그림 12)는 추출된 2차원 데이터를 좌표 평면에 표시한 결과이다.



(그림 12) 손 모양 특징 공간

(그림 12)의 o는 주먹 모양, x는 유사 손바닥 모양으로부터 추출된 특징 벡터를 나타낸다. 다양한 손 모양이 존재하기 때문에 우리는 주먹 모양을 기술할 수 있는 영역을 그림의 실선과 같이 정의하였다. 0과 5사이에서 손바닥과 주먹 모양이 분리 가능(separable)하지 않기에, 우리는 주먹이 오인식되는 오류(false negative error)보다 손바닥 모양이 주먹으로 인식되는 오류(false positive error)를 중요시 생각하여 직선 경계를 결정하였다. 또한 직선 밖의 주먹 특징 벡터들은 이상점(Outlier)으로 간주하였다. 주먹 결정 함수(decision function)를 수식으로 정리하면 다음과 같다.

- i) $0.6 < x_1 < 5$ 일 때,
 $0.55x_1 + x_2 > 6.25$ and $x_2 < 8.3$
- ii) $5 \leq x_1 < 13.5$ 일 때,
 $3.5 < x_2 < 8.3$

여기서 x_1 은 엄지 기준 세로길이 비이며, x_2 는 2사분면 피부비율에 10을 곱한 값이다. 또한, 손 모양 인식의 성능향상을 위하여 큐에 위 판단 값을 저장하여 투표방식(voting)에 의해 해당 프레임의 손 모양을 판단한다.

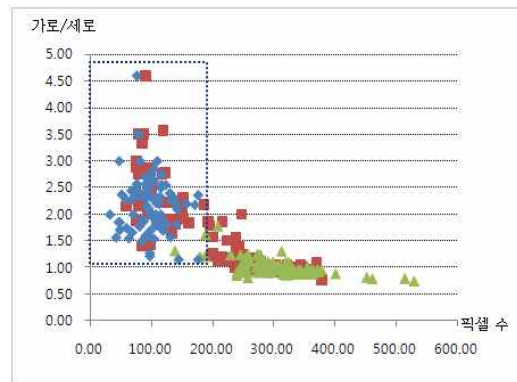
이 판단 함수는 손등과 팔의 각도가 90°에 가

까울 경우 정확도가 매우 높으나 (그림 11)과 같이 손등을 약간 기울일 경우 엄지손가락 또는 손바닥 영역의 깊이 값이 정확히 검출되지 않아 손바닥 모양이 주먹으로 오 인식되는 경우가 있다.



(그림 13) 45° 기울임시 손바닥 모양

이러한 이유로 우리는 검출된 손 영역 픽셀 개수와 박스의 세로 대비 가로의 길이 비율의 특징을 추가하여 기울임 시 주먹 인식 성능을 향상하였다. 다음 (그림 13)은 90°, 45°, 30° 손 기울임 시 손바닥을 주먹으로 인식된 경우의 픽셀수와 가로/세로 비율 특징을 추출한 학습데이터를 보여준다.



(그림 14) 가로/세로비와 픽셀 수 특징추출

우리는 (그림 14)의 왼쪽 점선 박스 영역을 오 인식 영역으로 구분하여 다음과 같이 주먹이 아닌 경우를 판단하였다.

$$n_p < 180, L_w / L_h < 1.05$$

여기서 n_p, L_w, L_h 는 각각 검출된 손 영역의 픽셀 수, 가로 길이, 세로 길이를 의미한다.

4. 카드보상을 위한 제스처인식

승리한 팀의 팀원들에게는 각자 원하는 카드를 선택할 수 있는 기회가 주어진다. 카드 선택 화면은 위 (그림 8)과 같이 나열된 5장의 카드 중 패스 제스처를 이용하여 다른 카드로 이동할 수 있고, 주먹을 꺾으로써 카드를 선택하고, 주먹을 쥔 상태에서 당기는 카드선택확인 제스처를 통해 카드 획득 정보를 서버에 전송하게 된다.

4.1 패스 제스처 인식

패스 제스처는 손이 사람의 중심점을 중심으로 한쪽 측면에서 다른 쪽 측면으로 이동하는 것으로 정의한다. t 번째 프레임에서 검출된 손 영역의 중심점의 x 좌표를 x_t 이라 할 때, 다음 조건을 만족하는 경우 패스 제스처로 인식하였다.

- (a) $x_{t-2} \leq x_{t-1} \leq x_t$
- (b) 적당한 t_0 ($t - T_p < t_0 < t$)에 대하여,
 $x_{t_0-1} \leq C_x \leq x_{t_0}$
- (c) $|x_t - x_{t-d}| > T_d$

여기서 조건 (a)는 손의 x 좌표가 단조 증가함을 의미하고, 조건 (b)는 손이 몸의 중심점을 통과했음을 의미한다. 여기서 C_x 는 카드 선택자의 몸 중심점의 x 좌표이다. 조건 (c)는 손이 일정 거리 이상을 이동함을 의미한다. 손의 이동 속도에 의존하는 상수 d 와 임계치 T_p, T_d 는 실험에서 각각 5, 5, 30으로 설정하였다.

조건 (a)는 오른쪽에서 왼쪽으로 손을 저었을 때를 의미한다. 조건(b), (c)와 다음 (a')을 만족하는 경우, 왼쪽에서 오른쪽으로 손을 짓는 상태를 인식할 수 있다.

- (a') $x_{t-2} \geq x_{t-1} \geq x_t$

이 경우 카드는 역순으로 확대되어 표시된다.

4.2 카드선택 제스처 인식

카드 선택을 위한 주먹 인식 알고리즘은 위 3.2에서 설명한 퀴즈 보기선택 시 사용된 알고리즘과 동일하다. 카드 확대 및 전송은 주먹을 쥔 상태에서 몸 쪽으로 당길 경우 수행된다. 즉, 손 모양이 지속적으로 주먹으로 인식된 상태에서 다음을 만족할 때, 카드 확대 및 전송이 수행된다.

$$d_h^c - d_B < (d_h^p - d_B) / T$$

여기서 d_h^p 와 d_h^c 는 최초 주먹 인식 시 주먹의 깊이 값과 현재 프레임의 주먹 깊이 값을 의미하고, d_B 는 몸의 최대 깊이 값으로 패스제스처 수행 전 카드보상을 위한 메시지 출력 시 사용자별로 설정된다. 참고로, $d_h^c > d_B$ 과 $d_h^p > d_B$ 는 항상 성립한다. 임계치를 결정하는 상수 T 는 실험에서 2로 설정되었다.

5. 실험 결과

5.1 사용자 인터페이스

우리는 실시간 퀴즈 게임 진행을 위해 (그림 15)와 같이 메인 콘텐츠 화면을 구성하였다.

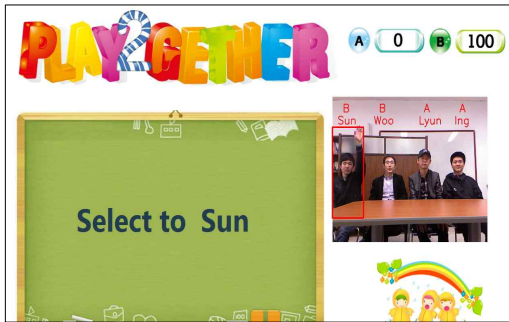


(그림 15) 콘텐츠 화면 구성도

질판은 메인 뷰(Main view)부분으로 게임 진행을 위한 메시지, 퀴즈문제가 표시되고, 오른쪽 상단은 A팀, B팀의 점수가 표시되는 영역이며, START, EXIT은 프로그램의 시작과 종료를 위

한 버튼이다.

퀴즈게임이 진행되면 START, EXIT 버튼은 사라지고, (그림 16)과 같이 키넥트 카메라를 통해 입력된 실시간 영상이 표시된다. 실시간 영상에는 사용자 별로 출석체크를 통해 인식한 아이디와 팀 인식 결과가 표시된다. 또한, (그림 16)에서 보듯이 문제를 맞히기 위해 먼저 손을 사람을 인식하여 해당 영역에 박스를 표시하여 도전자로 선택된 사람을 알려준다.



(그림 16) 퀴즈게임 진행 시 화면 인터페이스

5.2 제스처 인식 성능

제안한 제스처 방법론들의 성능을 평가하기 위해 우리는 다음 <표 2>와 같이 4개의 테스트 데이터 셋을 촬영하였다.

<표 2> 테스트 데이터 셋 구성

데이터 셋	데이터 구성	데이터 개수
손들기 동작	왼손 들기	10
	오른손 들기	10
보기선택 동작	90도 손 기울임	10
	45도 손 기울임	10
패스 동작	보통 빠르게	7
	빠르게	6
	유사 패스동작	7
카드선택 동작	선택(주먹+당김)	10
	비선택(손바닥+당김)	5
	비선택(주먹+당기지 않음)	5

여기서 유사 패스동작은 패스 제스처가 아닌

동작이며, 비선택(손바닥+당김)은 손바닥으로 당기는 동작이고, 비선택(주먹+당기지 않음)은 주먹으로 선택한 후 당김 동작을 하지 않은 경우를 의미한다. 우리는 위 테스트 데이터 셋에 대해 다음 <표 3>과 같은 인식 결과를 얻었다.

<표 3> 인식 결과

데이터 셋	데이터 구성	인식률(%)
손들기 동작	왼손 들기	100
	오른손 들기	100
보기선택 동작	90도 손 기울임	100
	45도 손 기울임	80
패스 동작	보통 빠르게	100
	빠르게	83.4
	유사 패스동작	100
카드선택 동작	선택(주먹+당김)	80
	비선택(손바닥+당김)	100
	비선택(주먹+당기지 않음)	100

손을 많이 기울인 경우 손의 깊이 값이 잘 검출되지 않아 올바른 인식이 되지 않는 경우가 있었으며, 매우 빠르게 패스제스처를 수행하는 경우에는 프레임 입력 속도(Frame rate)가 떨어짐으로 인해 올바른 인식이 되지 않았다. 또한 카드선택 제스처에서는 주먹이 잘 인식 되지 않음으로 인해 오류가 발생하였다.

이전 알고리즘[1]은 보기선택 동작 데이터 셋에 대해 90도 손 기울임에 대해 90%, 45도 손 기울임에 대해 30%의 인식률을 보였다.

실시간 퀴즈게임에서는 제스처인식에 대해 만족스러운 인식 결과를 얻을 수 있었다. 간혹 첫 번째 시도에서 오 인식이 되는 경우가 발생하였는데, 재차 시도에서 인식에 성공함으로써 오 인식으로 인한 불편함을 느끼지 않았다.

6. 결론 및 토의

우리는 본 논문에서 제스처 인식을 이용한 퀴즈 게임 콘텐츠를 제안하였다. 우리는 오프라인 퀴즈 게임의 요소들을 파악하여 이들을 디지털 콘텐츠화 하였으며, 출석체크, 팀 인식, 문제 도

전권 획득, 문제 보기 및 보상 카드의 이동과 선택 과정의 제스처 인터페이스를 사용자가 일상 생활에서 수행하는 제스처와 근접한 방식으로 고안하였다.

즉, 손들기 제스처는 교실에서 출석체크나 팀 확인 시 선생님의 호명에 대해 학생들이 응답으로 취하는 손들기 방식과 동일하게 정의하였고, 손의 상하 이동과 주먹 쥐는 두 과정으로 정의한 보기선택 제스처는 손을 위 아래로 이동할 때 위 또는 아래 보기가 활성화되도록 하여 사용자가 자신의 명령이 즉각적으로 반영되는 것을 느낄 수 있고, 원하는 보기가 활성화되었을 때 손을 힘 있게 움켜쥐는 과정에서 선택을 체감(體感)할 수 있도록 정의하였다.

카드선택 시 적용되는 패스 제스처, 카드선택 제스처, 카드선택 확인 제스처는 일상에서 원하는 카드를 고르기 위해 다음 장으로 넘기고 원하는 카드를 움켜쥐고, 카드를 움켜쥐 상태에서 끌어당김으로써 여러 카드로부터 원하는 카드를 골라내는 동작과 결부시킴으로써, 선택과 이동의 능동적인 추상 개념들을 퀴즈게임 시 체감할 수 있도록 정의하였다.

이전 결과[1]와 비교할 때, 손의 상하이동 인식, 주먹 인식 방법론을 개선하여 제스처 인식 성능을 크게 향상시킬 수 있었다. 또한, 패스 제스처 및 카드선택 제스처, 카드선택확인 제스처의 인식을 통해 승리한 팀의 팀원들이 카드를 획득할 수 있는 콘텐츠를 추가함으로써 퀴즈게임의 완성도를 높일 수 있었다.

제한한 퀴즈게임 콘텐츠는 실시간 테스트에서 안정적으로 동작하였으며 매우 만족스러운 제스처 인식 성능을 보였다.

참 고 문 헌

- [1] 안정호, "제스처인식을 이용한 퀴즈게임 콘텐츠의 사용자 인터페이스에 대한 연구", 디지털콘텐츠학회 논문지, Vol. 13, No. 1, pp91-99. 2012.
- [2] N. D. Binh, E. Shuichi and T. Ejima, "Real-Time Hand Tracking and Gesture Recognition System", GVIP Conference, 2005.
- [3] A. F. Bobick, and J. W. Davis, "The Recognition of Human Movement Using Temporal Templates", IEEE Transactions on Pattern Recognition and Machine Learning, Vol. 23, No. 3, 2001.
- [4] L. Bretzber, I. Laptev and Tony Lindeberg, "Hand Gesture Recognition using Mult-Scale Colour Features, Hierarchical Models and Partial Filtering", IEEE International Conference on Automatic Face and Gesture Recognition, pp.423-428. 2002.
- [5] J.W. Davis, "Hierarchical Motion History Images for Recognizing Human Motion", IEEE workshop on Detection and Recognition of Events in Video, pp.39-46, 2001.
- [6] Y. Fang, K. Wang, J. Cheng and H. Lu, "A real-time hand gesture recognition method", IEEE International Conference on Multimedia and Expo, pp. 995-998, 2007.
- [7] N. Gupta, P. Mittal, S. R. Dutta, S. Chaudhury and S. Banerjee, "Developing a Gesture-Based Interface", IETE Journal of Research, Vol. 48, No. 3, pp. 237-244, 2002.
- [8] A. Hernandez, M. Reyes, S. Escalera, and P. Radeva, "Spatio-Temporal GrabCut Human Segmentation for Face and Pose Recovery", IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, pp. 33-40, 2010.
- [9] H. Kim, G. Albuquerque, S. Havemann and D. W. Fellner, "Tangible 3D: Hand Gesture Interaction for Immersive 3D Modeling", IPT & EGVE Workshop, 2005.
- [10] T. Kirishima, K. Sato and K., Chihara, "Real-Time Gesture Recognition by Learning and Selective Control of Visual Interest Points", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 3, pp.351-364, 2005.
- [11] A. Malima, E. Ozgur and M. Cetin, "A Fast Algorithm for Vision-based Hand Gesture Recognition for Robot Control", IEEE Conference on Signal Processing and Communications Applications, 2006.
- [12] R. J. New, E. HasanBellu and M. Aguilar, "Facilitating User Interaction with Complex Systems via Hand Gesture Recognition", In Proc. of the 41st ACM Southeastern Conference, Savannah, 2003.
- [13] B. Stenger, P. Mendonca and R. Cipolla, "Model-Based 3D Tracking of an Articulated Hand", IEEE Conference on Computer Vision and Pattern Recognition, pp. 310-315, 2001.
- [14] M.-C. Su, "A Fuzzy Rule-Based Approach to Spatio-Temporal Hand Gesture Recognition", IEEE Transactions on Pattern Recognition and Machine Learning, Vol. 23, No. 3, 2001.

actions on Systems, Man and Cybernetics, Appl. and Review, Vol. 30, No. 2, pp. 276-281, 2000.

[15] A. Wilson and A. Bobick, "Parametric Hidden Markov Models for Gesture Recognition", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 9, pp. 884-900, 1999.

[16] H.-S. Yoon, J. Soh, Y. J. Bae and H. S. Yang, "Hand gesture recognition using combined features of location, angle and velocity", Pattern Recognition, Vol. 34, pp. 1491-1501, 2001.

[17] J.-H. Yoon, J.-S. Park and M. Y. Sung, "Vision-based Bare-Hand Gesture Interface for Interactive Augmented Reality Applications", 5th International Conference on Entertainment Computing, 2006.

[18] L. Zhang and Q. Ji, "Segmentation of Video Sequences using Spatio-temporal Conditional Random Fields", IEEE Workshop on Motion and Video Computing, pp. 1-7, 2008.



안 정 호

1998년 : 연세대학교 수학과 (이학석사)
 2001년 : Texas A&M University 통계학과(이학석사)
 2006년 : 연세대학교 컴퓨터과학과 (공학박사)

2006년 : 연세대학교 BK21 지능형모바일서비스를 위한 차세대 단말 소프트웨어 사업단 박사후 연구원
 2007년~현재: 강남대학교 컴퓨터미디어정보공학부 교수
 관심분야 : 패턴인식, 제스처인식, 얼굴인식, 영상처리, 컴퓨터비전



고 재 필

1996년 : 연세대학교 전산학과 (이학학사)
 1998년 : 연세대학교 컴퓨터과학과 (공학석사)
 2004년 : 연세대학교 컴퓨터과학과 (공학박사)

2004년~현재: 금오공과대학교 컴퓨터공학과 교수
 관심분야 : 패턴인식, 제스처인식, 얼굴인식, 영상처리, 컴퓨터비전