

# 제스처인식을 이용한 퀴즈게임 콘텐츠의 사용자 인터페이스에 대한 연구

안정호\*

## 요약

우리는 본 논문에서 아날로그 영역의 퀴즈 게임을 디지털화시키는 작업을 소개한다. 우리는 퀴즈 진행, 퀴즈 참가자 파악, 문제 제시, 먼저 손든 참가자 인식, 정오답 판단, 점수 합산, 승리팀 판단 등 기존의 퀴즈 게임이 아날로그 방식으로 수행해온 작업을 디지털화시키는 작업을 수행하였다. 이를 자동화하기 위해 최근 주목받기 시작한 키넥트 카메라를 이용하여 깊이 영상을 입력받아, 사용자들의 위치를 파악하고 사용자 위주로 정의된 제스처를 인식하는 알고리즘을 고안하였다. 영상의 깊이 값의 분포를 분석하여 퀴즈 참가자들의 상체를 검출하고 사용자들의 분할하였고 손 영역을 검출하였다. 또한 손바닥, 주먹, 기타 손 모양을 인식하기 위한 특징 추출 및 판단 함수를 고안하여 사용자가 퀴즈 보기를 선택할 수 있게 하였다. 구현된 퀴즈 응용 프로그램은 실시간 테스트에서 매우 만족스러운 제스처 인식 결과를 보였으며 원활한 게임 진행이 가능하였다.

## A Study on User Interface for Quiz Game Contents using Gesture Recognition

Jung-Ho Ahn\*

## Abstract

In this paper we introduce a quiz application program that digitizes the analogue quiz game. We digitize the quiz components such as quiz proceeding, participants recognition, problem presentation, volunteer recognition who raises his hand first, answer judgement, score addition, winner decision, etc, which are manually performed in the normal quiz game. For automation, we obtained the depth images from the kinect camera which comes into the spotlight recently, so that we located the quiz participants and recognized the user-friendly defined gestures. Analyzing the depth distribution, we detected and segmented the upper body parts and located the hands' areas. Also, we extracted hand features and designed the decision function that classified the hand pose into palm, fist or else, so that a participant can select the example that he wants among presented examples. The implemented quiz application program was tested in real time and showed very satisfactory gesture recognition results.

Keywords : Quiz Game Contents, User Interface, Gesture Recognition, Kinect Camera

## 1. 서론

최근 디지털 콘텐츠의 사용자 인터페이스는 다양한 플랫폼에서 사용자가 사용하기 편리한 방향으로 급속한 발전을 이루어 나가고 있다. 스마트폰의 출현이후 버튼방식의 핸드폰 인터페이스는 사라지고 터치 사용자 인터페이스(TUI)로 전환되었다. 터치 방식도 꾸준히 발전하여 이제 화면 전환, 콘텐츠 선택, 확대/축소, 어플리케이션

※ 제일저자(First Author) : 안정호  
접수일:2012년 03월 01일, 수정일:2012년 03월 15일  
완료일:2012년 03월 24일  
\* 강남대학교 컴퓨터미디어정보공학부  
[jungho@kangnam.ac.kr](mailto:jungho@kangnam.ac.kr)  
■ 이 논문은 2010년도 강남대학교 교내연구비 지원에 의한 것임.

선 실행 등의 기능들은 다양한 손가락 터치 방식에 의해 조작 가능하게 되었다. 태블릿 PC는 기존 PC의 마우스와 키보드에 의한 전통적인 사용자 인터페이스 방식을 손가락 터치와 가상 키보드를 이용한 방식으로 바꾸어 놓았다.

이러한 하드웨어적인 측면에서의 진화와 더불어, 앞으로는 소프트웨어적인 측면에서도 디지털 콘텐츠의 사용자 인터페이스 진화가 광범위하게 이루어질 것으로 기대된다. 이는 기존의 디지털 콘텐츠의 인터페이스 발전 뿐 아니라, 아날로그 콘텐츠의 디지털화로 이어질 것이다. 컴퓨터를 이용하지 않고 수행되어 온 아날로그 콘텐츠들을 디지털화하는 데 많은 어려움이 있다. 예를 들어, 본 논문의 연구에서 수행했던 퀴즈 게임을 위한 사용자 인터페이스 개발은 아날로그 방식으로 수행되어 오는 퀴즈 콘텐츠의 디지털화를 위한 시도이다.

퀴즈 게임이란 참가자들이 준비된 퀴즈 문제를 풀어나가는 과정으로 진행자와 참가자들의 소통을 통한 문제 제시와 풀이, 점수합산, 승자에 대한 보상 등 여러 가지 아날로그적인 활동들이 결부되어 있는 콘텐츠이다. 우리는 이러한 활동들을 사람이 아닌 컴퓨터 응용 프로그램이 퀴즈 참가자들과 소통할 수 있게 함으로써 가능하도록 하는 것에 초점을 맞추었다. 최대한 사용자 위주의 제스처를 정의함으로써 프로그램과 사용자가 최대한 자연스러운 소통이 가능하도록 고안하였다.

### 1.1 퀴즈 게임 개요

우리는 다수의 참가자가 팀을 이루어 경쟁할 수 있는 영어 교육용 퀴즈 게임을 위한 응용 프로그램 (Quiz Application Program, 이하 QAP)을 개발하였다. 초등학교 4, 5학년생을 대상으로 4 또는 5명의 참여 학생이 두 팀으로 나누어 팀 대항 퀴즈게임을 수행한다. 팀원 중 한명이 퀴즈 문제를 맞히면 팀의 점수는 100점이 가산되고 틀릴 경우 50점이 감점된다. 게임의 프로시저는 다음과 같다.

QAP는 교실의 TV 스크린에서 보여 지고 TV위에는 키넥트(Kinect) 카메라가 설치되어 퀴즈 참가자들의 모습을 입력받아 QAP에 전달한다. 퀴즈풀이 전 선생님은 참여 학생의 이름과 3지선다형 퀴즈 문제들을 QAP에 입력하고 학생

들은 두 팀으로 나뉘어 TV 앞 의자에 착석한다. 이 때 팀원들이 모여 앉을 필요는 없다.

퀴즈 응용프로그램이 시작되면 학생들의 출석체크가 먼저 이루어진다. 참여 학생들의 이름이 스크린에 표시되면 해당학생은 손을 든다. 이때 응용 프로그램은 학생의 위치를 파악한다. QAP는 출석체크 후, 팀 구성원을 알기 위해 “A팀 손드세요”라는 메시지를 화면 출석하여 A팀의 구성원들을 파악한다. 마찬가지로 B팀의 구성원도 파악한다.

팀 파악후 QAP는 선생님이 입력한 문제를 순차적으로 제시한다. 문제가 제시된 후 정답을 맞히고자 하는 학생은 손을 들어 의사를 표시한다. 이때 먼저 손든 학생에게 기회가 주어진다. QAP는 먼저 손든 학생의 이름을 표시하고 해당 학생은 손 제스처를 통해 문제의 보기중 하나를 선택한다. 보기 선택을 위한 손 제스처는 손바닥을 화면으로 향한 상태에서 위 아래로 이동하여 보기를 이동할 수 있고, 주먹을 쥐면 해당 보기가 선택되는 것으로 정의하였다. 이때 맞추면 100점이 가산되고 오답을 선택한 경우 다른 학생에게 기회가 주어진다. 즉 오답인 경우 다른 중 먼저 손든 학생에게 기회를 제공한다. 한 문제에 대해 오답이 두 번 나올 경우, 나머지 보기가 정답이므로 다음 문제로 넘어간다. 입력된 문제가 모두 수행되었으면 팀 별 총점이 계산되어 TV 스크린에 표시되고 퀴즈 게임이 종료된다.

### 1.2 관련 연구

손 검출 및 모양 인식에 대한 연구는 휴먼 컴퓨터 인터페이스(HCI)를 위해 많은 연구가 되어 왔다.

Kim et. al[6]은 7개의 손 모양인식을 통한 인터페이스 시스템을 고안하였다. 손 모양 인식을 통해 Grab, Rotate, Scale, Point, Pause 등 5개의 정의된 제스처를 인식하였다. 이들은 손가락에 흰색 골무를 착용하여 손가락 인식률을 높일 수 있었다. Yoon et. al.[10]은 제스처 인터페이스 기반 음악 플레이어와 농구 게임을 제작하였다. 이들은 손 모양 인식 대신 움직임 궤적을 분석하여 가상 오브젝트 접근과 메뉴 접근을 구분하였다.

Binh et. al.은 피부색 모델을 이용한 손 검출 후 개선된 칼만 필터를 이용하여 손을 추적하였

으며 2-DHMM을 이용하여 제스처를 인식하였다. 이들의 실험 환경은 사용자의 손을 카메라 정면에 위치시켜야 하는 제약 사항이 있다. Bretzber et. al.[3]은 IUV 컬러와 블랍(Blob), 리지(Ridge) 특징을 이용하여 손 모양을 추출하였으며 입자 필터(particle filtering)와 계층적 샘플링 방법을 이용하여 손 추적을 실시하였다. Fang et. al.[5]은 Adaboost 방법을 적용한 손 검출과 광흐름(Optical Flow)과 컬러 정보를 이용한 손 추적 방법론을 사용하였다. 또한 범위-공간(scale-space) 특징 검출 방법을 고안하여 손바닥과 손가락을 구별하였는데 실험 환경이 다소 제한적이다. Malima et. al.[8]은 5개의 손 모양 인식을 통해 로봇의 동적을 제어하였는데, 이들은 손 중심점을 기준으로 원형 영역의 연결성분 분석을 통해 손 모양을 인식하였다. Yoon et. al.[9]은 피부색과 모션 정보를 이용하여 손 검출을 수행하였으며 위치, 각도, 속도 특징 정보를 이용한 HMM 분류기에 의해 손동작을 인식하였다.

몸 전체를 이용한 제스처 인식 방법으로 MHI(Motion History Image)와 MEI(Motion Energy Image)에 대해 추출된 Hu 모멘트를 Mahalanobis 거리를 이용하여 몸동작을 인식한 사례가 있다[2][4]. 또한 미지 정의되어 있지 않은 제스처에서 선택적으로 주요 제스처를 실시간으로 인식하는 방법론에 대한 연구[7]도 대표적인 관련 연구라 할 수 있다.

## 2. 사용자 분할

### 2.1 키넥트를 이용한 깊이영상 생성

마이크로소프트의 XBOX 360이라는 게임기를 위해 개발되어진 키넥트(Kinect)는 두 개의 카메라 렌즈와 한 개의 적외선 센서를 포함하고 있다. 하나의 렌즈는 일반 RGB 영상 획득을 위해 사용되며 다른 렌즈는 적외선 센서의 값을 측정하여 깊이 값을 측정하는데 사용된다. 키넥트 SDK는 카메라로부터 대략 0.9m~3.5m 떨어진 물체의 깊이(거리) 값을 측정한다. 0.9m 보다 가까이 있는 물체의 깊이 값은 0으로 주어지며 0.9m 거리의 물체의 깊이 값은 800, 최댓값은 3900값 정도이다. 최대 측정 가능 거리를 넘어가

는 멀리 있는 물체의 깊이 값은 0으로 주어진다. 우리는 깊이 값이 800일 때 255로, 3900일 때 0이 되도록 선형 보간(interpolation)하고 깊이 값이 0일때는 0값을 갖도록 깊이 영상(depth map)을 생성하였다. 우리는 깊이 영상만을 사용하여 제스처 인식을 수행하였다.

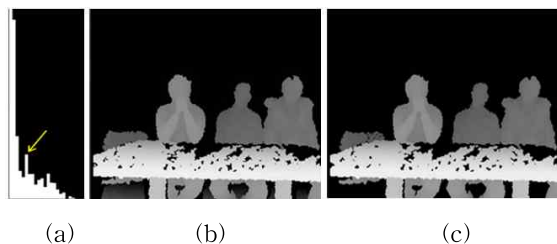
### 2.2 상체 검출 및 사용자 수 인식

우리는 우선 주어진 깊이 영상에서 불필요한 배경영역을 제거하고 앉아 있는 사용자의 상체 부분을 검출하고 사용자의 수를 파악한다. 검출된 상체 영역에서 사용자의 위치를 파악하고 손 제스처 인식을 수행하였다.

배경을 제거하기 위해 매 프레임 전 영상에 대한 깊이 값에 대한 프레임 히스토그램  $H = \{h_i\}$ 을 구한다. 카메라와 앉아 있는 학생 사이에 물체가 없다는 가정 하에서 배경은 뒷부분에서 나타나기 때문에 작은 값을 가지는 깊이 값을 적절히 제거해야 한다. 우리는 이를 위해 깊이 값이 급격히 변화하는 시점부터 전경영역이라고 판단하여 다음 의 깊이 값  $d_{min}$  이상인 영역을 전경 영역으로 분류하였다.

$$d_{min} = \operatorname{argmin}_i (h_i / h_{i-1} > T_F)$$

우리는 실험에서 히스토그램의 빈(bin) 사이즈는 4로 설정하였고, 임계치  $T_F$ 는 2.5로 설정하였다. 다음 (그림 1)의 (a)는 프레임 히스토그램(빈 사이즈=4), (b)는 배경 제거 전, (c)는 배경 제거 후의 모습을 보여 준다.



(a) (b) (c)  
(그림 1) 프레임 히스토그램 및 배경 제거 결과. (a) 프레임 깊이 히스토그램, (b) 배경제거 전 깊이영상, (c) 배경제거 후 깊이영상.

상체 검출은 화면 중앙에 퀴즈 응시자들이 있

다는 가정 하에, 다음과 같이 응시자 수를 파악하고 이 정보를 이용하여 상체를 검출하였다.

화면의 기준 행  $r_0 (=100)$ 을 기준으로 아래 방향으로 50행의 연결 성분의 개수를 파악한다. 이 중 가장 많은 연결 성분의 개수를 사람 수로 설정한다. 이는 퀴즈 풀이 환경에서 앉아 있는 사람의 상체가 약 100번째 라인을 기준으로 분포함에 근거한다.



(그림 2) 상체 검출 결과

상체의 행 영역  $[r_{\min}^B, r_{\max}^B]$ 는 다음과 같이 결정한다.

$$r_{\min}^B = \min_{r_0 < r} (0 < N_r^c)$$

$$r_{\max}^B = \max_{r < r_0} (1 < N_r^c, N_{r+1}^c < 9) - T_M^B$$

여기서  $N_r^c$ 은 행  $r$ 의 연결 성분의 개수이고  $T_M^B$ 는 상체 검출 마진으로 실험에서는 10으로 설정하였다. 다음 그림은 손들기 동작에서 검출된 상체 영역을 보여 준다.

### 2.3 인체 영역 분할 및 머리 영역 검출

상체 검출 후 각 인체 영역을 수직 분할하고 머리 위치를 파악한다. 인체 영역 수직 분할점은 검출된 상체 깊이 영상의 수직 히스토그램  $K = \{k_i\}$ 의 넓은 영역의 국소점(local minimum)으로 설정하였다. 즉, 수직 분할점  $i$ 는 다음과 같이 구한다.

$$\operatorname{argmin}\{k_i | k_i < k_i < k_{i+1} < T_S^M, i - T_S < i < i + T_S\}$$

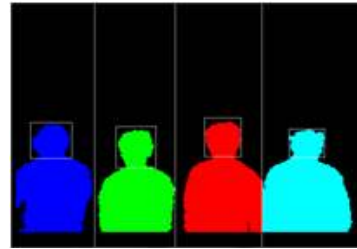
여기서  $T_S^M$ 은 국소점의 한계 최댓값이며,  $T_S$ 는 분할점의 검출지점이 조밀하지 않도록 하는

정규 조건(regularity condition)으로 실험에서는  $T_S^M = 30$ ,  $T_S = 50$ 으로 설정하였다.

머리 상단점은 퀴즈풀이 초기 상태에서 수직 분할 점 사이의 최대 빈 값을 가지는 지점으로 설정하였고 이후 그 지점을 추적(tracking)하였다.  $t$ 번째 프레임의  $i$ 번째 사용자의 머리 상단점  $p_i^t$  추적 방정식은 다음과 같다.

$$p_i^t = \operatorname{argmin}_j (|k_j^t - k_{p_i^{t-1}}| + C|j - p_i^{t-1}|)$$

여기서  $k_j^t$ 는  $t$ 번째 프레임의  $j$  열의 히스토그램 값이고  $k_{p_i^{t-1}}$ 는 이전 머리 상단점 위치  $p_i^{t-1}$ 의 열 히스토그램 값이다.  $C$ 는 히스토그램 값 차이  $|k_j^t - k_{p_i^{t-1}}|$ 와 위치의 차  $|j - p_i^{t-1}|$ 의 상댓값을 정의하는 가중치(weight)로 실험에서는 3로 설정하였다.



(그림 3) 분할된 인체 영역

머리 박스는 머리 상단점으로부터 아래방향 깊이 값이 현격히 감소하는 점으로 설정하여 검출하였다. 아래 (그림 3)은 4명의 사용자에게 대해 분할된 인체 영역과 머리 검출 결과를 보여 준다.

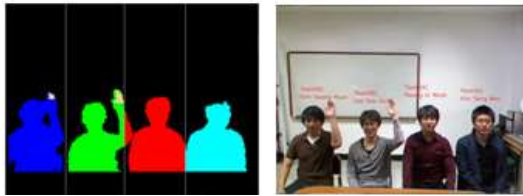
## 3. 손 동작 인식

### 3.1 사용자 ID와 손들기 제스처 인식

퀴즈 게임의 사용자 ID 인식은 응용프로그램 QAP의 출석 체크 모듈에 의해 이루어진다. 먼저 사용자의 이름이 TV 스크린에 표시되면 해당 사용자는 손을 든다. 이때 QAP는 이 사용자의 위치를 파악한다.

손들기 제스처는 두 수직 분할점 사이에서 머

리 상단 부분 위로 깊이 값이 다수 검출되면 해당 사람이 손을 든 것으로 인식하였다. 다음 (그림 4)는 손들기 제스처의 인식 결과를 보여준다. 그림 4-(a)에서 머리 위로 올라간 손 영역을 색깔로 표시하였다. (그림 4)-(b)는 출석 체크와 팀 구분시 인식된 성명과 팀명을 학생들 머리 위에 빨간색 텍스트로 표시한 결과를 보여 준다.



(a) 인체 분할 결과 (b) 출석체크 결과  
(그림 4) 손들기 제스처 인식

### 3.2 보기 선택 제스처 인식

퀴즈 문제가 제시된 후 문제의 정답을 맞히고자 하는 학생은 손을 들어 의사를 표현한다. 여러 명의 학생이 손을 드는 경우 먼저 손든 학생이 도전자로 선택된다.

도전자가 된 학생은 손을 앞으로 내밀어 손 제스처를 통해 보기를 선택한다. TV 화면에 손바닥이 보이도록 하여 손을 위 아래로 움직이면 해당 보기가 활성화 된다. 보기 선택 초기에는 1번 보기가 활성화되어 있다. 보기를 이동하다가 주먹을 쥐면 현재 활성화된 보기가 선택된다. 우리는 이 프로시저를 구현하기 위해 인체로부터 앞으로 내민 손 영역을 분할하고, 분할된 영역이 손바닥 모양인지 주먹 모양인지 판단한다.

#### 3.2.1 손 영역 분할

도전자가 선택되면 도전자 영역의 깊이 값들을 히스토그램에 저장한 후, 손이 어느 깊이 값을 취하는지 알아낸다.

도전자가 선택되고 사용자의 동작이 멈춘 순간 도전자의 수직 분할 영역 안의 인체 깊이 값의 최대치  $d^B$ 를 구한다. 동작이 멈추었음은 차영상의 깊이 일정 임계치를 넘지 않음으로 판단한다. 이후 이  $d^B$  이상의 깊이 값이 검출되면 이 깊이 값에 해당하는 픽셀들을 후보 손 영역이라 판단한다.

후보 손 영역은 손 이외에 팔의 일부분을 포함하는 경우가 많다. 팔 부분을 제외시키기 위해 우리는 후보 손 영역 중 봉우리 부분을 검출하였다.

봉우리 검출은 최대 깊이값  $d_{max}$ 부터 작은 방향으로 샘플링된 위치에서의 깊이 히스토그램 값이 감소하는 지점  $d_H$ 을 손 깊이값의 최솟값으로 설정하였다. 즉

$$h_i - h_{i-2} < T_H < 0$$

여기서  $i = d_{max}, d_{max} - 2, \dots, d^B$ 로 샘플링 단위를 2로 설정하였다.  $T_H$ 는 노이즈에 강인한 검출을 위해 0이 아닌 음수 값을 설정하였다. 실험에서  $T_H$ 는 -2로 설정하였다.



(a) (b)  
(그림 5) 깊이 영상에서 검출된 손 영역.  
(a) 인체영역 깊이 히스토그램, (b) 손 영역 검출 결과

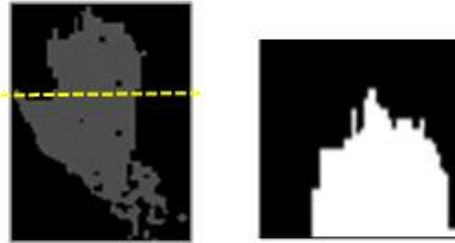
위 (그림 5)의 (a)는 도전자 인체 영역의 깊이 값 히스토그램이고 (b)는 검출된 손 영역을 보여준다. 그림 (a)의 빨간색 동그라미는 손 영역의 깊이값 범위를 표시한 것이다.

#### 3.2.2 손 모양 인식

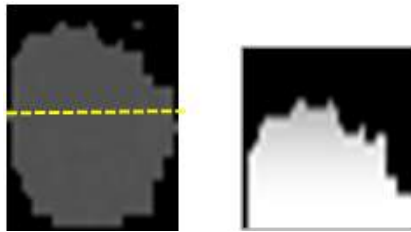
우리는 손 모양 인식을 위해 손바닥(P, palm), 주먹(F, fist), 기타(E, else)의 세 클래스(class)를 구성하였다. 검출된 손 영역을 포함하는 외곽박스(bounding box)의 깊이 영상을 다음 판단 기준에 의해 이 클래스들 중 하나로 분류하였다.

분류를 위해 외곽박스에서 0이 아닌 깊이값을 255로 치환하여 이진 영상을 만든 후, 255값을 가지는 왼쪽 끝점 위치의 y 좌표를 기준으로 위 부분의 열 히스토그램(column histogram)을 분

석하였다. 다음 (그림 6)과 (그림 7)은 획득된 손바닥과 주먹 모양의 외곽 박스와 열 히스토그램을 보여 준다. 외곽박스의 노란색 점선은 왼쪽 끝점으로부터 그은 수평선이다. 이 수평선 위 부분 값에 대해 열 히스토그램을 구한 것이다.



(그림 6) 손바닥 외곽 박스와 히스토그램



(그림 7) 주먹 외곽박스와 히스토그램

히스토그램의 왼쪽 빈 값의 존재 여부는 엄지 손가락이 펴졌는지 여부에 따라 나뉜다. 우리는 이 점을 착안하여 다음과 같이 판단 함수 (decision function)를 고안하였다.

$$f(H) = \begin{cases} P & \text{if } n_0 > T_0 \text{ and } h_{n_0} \leq h_{n_0+1} \leq h_{n_0+2} \\ F & \text{if } n_0 < T_0 \text{ and } h_{n_0} \leq h_{n_0+1} \leq h_{n_0+2} \\ E & \text{otherwise} \end{cases}$$

여기서  $P$ 는 손바닥,  $F$ 는 주먹,  $E$ 는 기타 손 모양을 의미한다.  $H = \{h_i\}_{i=1}^{n_H}$ 는 평활화된 열 히스토그램으로 다음과 같이 구한다.

$$\begin{aligned} h_1 &\leftarrow (h_1 + h_2)/2 \\ h_i &\leftarrow (h_{i-1} + h_i + h_{i+1})/2, \quad i = 1, \dots, n_H \\ h_{n_H} &\leftarrow (h_{n_H-1} + h_{n_H})/2 \end{aligned}$$

$n_0$ 는 히스토그램의 0아닌 값이 처음으로 나타나는 인덱스로 다음과 같이 구한다.

$$n_0 = \operatorname{argmax}_i (h_1 \cdots h_i = 0) + 1$$

$n_0$ 에 대한 임계치  $T_0$ 는 실험에서 5로 사용하였으며 식  $h_{n_0} \leq h_{n_0+1} \leq h_{n_0+2}$ 은  $n_0$ 부터 단조 증가함을 의미한다.

## 4. 실험 결과

### 4.1 사용자 인터페이스 구성

우리는 실시간 퀴즈 게임 진행을 위해 (그림 8)과 같이 메인 콘텐츠 화면을 구성하였다. 칠판은 메인 뷰(main view)부분으로 키넥트 카메라를 통한 실시간 영상, 게임의 지시사항, 점수 등이 표시된다. START, EXIT 버튼은 프로그램의 시작과 종료 버튼이며, 게임을 시작하게 되면 빨간 점선으로 부분 안에 실시간 촬영된 사용자의 모습이 보이게 된다.



(그림 8) 콘텐츠 화면 구성도

### 4.2 실시간 퀴즈 게임

퀴즈 게임이 시작되면 (그림 9)와 같이 실시간 영상이 우측 화면에 출력되고 START, EXIT 버튼이 비활성화 된다. 칠판에는 출석체크를 위한 메시지가 보인다. 이때 프로그램 QAP는 사용자의 상체를 검출, 사용자 수 파악, 인체 수직 분할, 머리 박스 검출 등의 작업을 수행하게 된다.



(그림 9) 퀴즈 게임 시작

출석체크가 시작되면 칠판에 미리 입력된 이름이 순서대로 하나씩 표시된다. 다음 (그림 10)은 표시된 이름의 예를 보여준다.



(그림 10) 출석 체크용 이름

화면의 이름에 해당하는 사람은 손을 들면 QAP는 이를 인식해 인체 수직 분할선을 기준으로 해당 사람 객체에 이름 부여한다. 다음 (그림 11)은 TV위에 설치된 키넥트 카메라에 의해 촬영된 모습이다. 해당되는 한 사람이 손을 들고 있다.



(그림 11) 출석 확인을 위한 손들기 과정

참여 학생에 대해 출석체크가 완료되면, 팀 분할을 위한 메시지가 (그림 12)와 같이 차례로 화

면의 칠판 영역에 표시된다.



(그림 12) 팀 구분을 위한 메시지

이때 해당 팀원들은 (그림 13)과 같이 손을 들어 프로그램에 팀 구성을 알려 주게 된다.



(그림 13) 팀 구분을 위한 손들기 과정

팀 구성이 완료되면 “첫 번째 문제를 시작합니다.”와 “정답을 알면 손을 드세요.”라는 메시지가 뜨고 미리 입력된 3지선다형 문제가 제시되고 문제를 풀고자 하는 학생들은 손을 들어 의사를 표시한다. 이때 먼저 손을 든 학생이 도전자로 선택되고 도전자의 이름이 (그림 14)와 같이 칠판 영역에 표시된다.



(그림 14) 도전자 선택

도전자가 선택되면 보기 1번이 디폴트로 활성화된다. 선택된 도전자는 (그림 15)와 같이 손바닥 모양을 한 상태에서 팔을 상하로 움직여 다른 보기를 활성화시킬 수 있다. 원하는 보기가 활성화되었을 때 주먹을 쥐어 보기를 선택한다.

(그림 15)의 상단 문제 이미지는 2번 보기가 활성화된 상태를 보여주고, 하단 문제 이미지는 2번 보기를 선택했을 때의 모습을 보여준다.



(그림 15) 보기 선택 제스처

모든 문제의 풀이가 끝나면 (그림 16)과 같이 팀별 점수가 합산되고 이긴 팀이 표시된다.



(그림 16) 팀별 점수 표시 화면

우리는 보기 선택 제스처의 손모양 인식률을 측정하기 위해, 실시간으로 100회의 테스트를 수행하였다. 이때 인식률은 79%를 보였고, 인식이 실패된 경우 재차 시도하여 100% 원하는 보기를 선택할 수 있었다.

## 5. 결론 및 토의

우리는 실내에서 제스처 인터페이스를 통해 팀별 퀴즈 게임을 진행할 수 있는 응용 프로그램을 개발하였다. 최근 주목받기 시작한 키넥트 카메라를 이용하여 선형 보간법에 의해 깊이 영상을 추출하였다. 이 깊이 영상의 분포를 분석하

여 배경제거, 사용자 영역 분할, 손 검출 알고리즘을 개발하였다. 또한 효율적인 손바닥과 주먹 모양 인식 방법론을 제시하여 사용자가 손 제스처를 이용하여 콘텐츠를 제어할 수 있게 하였다. 이를 게임 흐름에 맞게 디자인된 사용자 인터페이스 화면들에 결합시켜 실시간 응용프로그램을 완성하였다.

제시한 알고리즘의 제스처 인식 성능은 실시간 테스트에서 만족스러웠으나 앞으로 손 검출 및 인식 알고리즘을 좀 더 보완하려 한다. 제시한 손 검출 알고리즘은 손을 높이 드는 경우 소매 부분이 같이 검출되는 약점이 있어 TV를 높이 설치할 경우 손 인식 결과가 좋지 못하다. 따라서 우리는 검출된 손 영역에서 적절한 손 중심 위치를 찾고 이를 기준으로 부분 매칭 방식을 이용한 손 모양 인식 방법론을 고려하고 있다.

## 참 고 문 헌

- [1] N. D. Binh, E. Shuichi and T. Ejima, "Real-Time Hand Tracking and Gesture Recognition System", GVIP Conference, 2005.
- [2] 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.
- [3] L. Bretzber, I. Laptev and Tony Lindeberg, "Hand Gesture Recognition using Mult-Scale Colour Features, Hierarchical Models and Partical Filtering", IEEE International Conference on Automatic Face and Gesture Recognition, pp.423-428. 2002.
- [4] 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.
- [5] 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.
- [6] H. Kim, G. Albuquerque, S. Havemann and D. W. Fellner, "Tangible 3D: Hand Gesture Interaction for Immersive 3D Modeling", IPT & EGVE Workshop, 2005.
- [7] 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.

[8] 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.

[9] 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.

[10] 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.



### 안정호

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, 얼굴인식, 컴퓨터비전, 로봇비전, 영상처리 등