

## 손동작 식별 규칙을 이용한 컴퓨터의 프레젠테이션 제어

이규원\*

### Presentation control of a computer using hand motion identification rules

Kyu-Won Lee\*

Department of Electronics and Communications Engineering, Daejeon University, Daejeon, 34520 Korea

#### 요 약

손동작 인식을 통하여 컴퓨터 프레젠테이션을 제어하는 시스템을 제안한다. 제안하는 시스템은 손 동작의 다양한 운동 형태를 인식, 구분함으로써 부가적인 제어용 장치 없이 프레젠테이션을 제어한다. 손동작의 인식을 위하여 얼굴영역 검출과 손영역 검출을 시행한다. 하르분류기(Haar classifier)를 이용하여 얼굴영역을 검출하며, HSV 컬러모델상에서 피부 색상 정보에 따라 손영역을 검출한다. 얼굴 영역은 손동작의 시작과 끝, 동작의 크기 및 방향을 판단하는 기준으로 삼는다. 얼굴 영역으로부터 가로, 세로 중심축을 설정하고 제안하는 모션 식별률에 따라 다양한 손동작을 인식하고 컴퓨터 제어에 이용한다. 약 1200회의 동작 인식 실험에서 97.2%의 인식률을 얻어 제안하는 알고리즘이 유효함을 확인하였다.

#### ABSTRACT

A system that control computer presentations by using the hand motion recognition and identification is proposed. The system recognizes and identifies various types of motion in hand motion, controls the presentation without additional control devices. To recognize hand movements, it performs a face and hand region detection. Facial area is detected using Haar classifier and hand region is extracted according to skin color information on HSV color model. The face area is used to determine the beginning and end of hand gestures, the size and direction of motion. It recognizes various hand gestures and uses them to control computer presentations according to the hand motion identification rules that are proposed and set horizontal and vertical axes from the face area. It is confirmed that 97.2% recognition rate is obtained in about 1200 hand motion recognition experiments and the proposed algorithm is valid in presentation control.

**키워드** : 동작, 프리젠테이션, 하르분류기, 컬러모델

**Keywords** : Motion, Presentation, Haar classifier, Color model

Received 9 July 2018, Revised 1 Aug 2018, Accepted 30 Aug 2018

\*Corresponding Author Kyu-Won Lee(E-mail:kwlee@dju.kr, Tel:+82-42-280-2558)

Department of Electronics and Communications Engineering, Daejeon University, Daejeon, 34520 Korea

Open Access <http://doi.org/10.6109/jkiice.2018.22.9.1172>

print ISSN: 2234-4772 online ISSN: 2288-4165

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/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. 서론

프리젠테이션이란 발표자가 청중들에게 정보, 기획, 안건 등을 제시하고 설명하는 행위이다. 편리한 프리젠테이션을 위하여 다양한 도구 및 기술이 개발되고 있으며[1] 컴퓨터 비전 기술의 발달로 카메라나 레이저 포인터를 활용해 프레젠테이션을 제어하는 시스템도 연구, 개발되고 있다.[2] 또한 깊이 카메라를 이용하여 포인터 추출과 포인터의 깊이 정보와 포인터의 이동 방향을 이용한 이벤트 분석을 통한 제어 방법이 제안되고 있지만 환경적 제약이 많다는 한계가 있다. 이러한 한계로 대부분의 프리젠테이션 제어는 무선주파수 혹은 적외선 신호를 이용하는 전용 제어용 기구를 활용하는 실정이다. 본 논문에서는 제스처 인식의 일환으로 손동작을 인식하여 컴퓨터 프레젠테이션을 제어하는 알고리즘을 제안한다. 즉 손 동작의 다양한 운동 형태를 인식, 구분함으로써 부가적인 제어용 장치 없이 프레젠테이션을 제어하는 시스템이다. 사용자는 이와 관련된 컴퓨터 명령을 미리 설정할 수 있고, 손의 동작 형태에 따라 미리 지정된 작업을 수행할 수 있다. 제안하는 시스템은 손동작의 시작과 끝, 동작의 크기 및 방향을 판단하는 기준으로 이용하기 위하여 얼굴 영역 검출을 시행하며, 얼굴 영역으로부터 세로 중심축을 설정하고 이를 기준으로 다양한 손동작을 인식하고 컴퓨터 제어에 이용한다. 따라서, 발표자는 카메라의 FOV(Field of View)내에서 얼굴과 손이 노출되는 범위 안에서 전후좌우의 움직임에 제한을 받지 않으며 자연스러운 동작 상황에서도 손 동작 인식에 의한 컴퓨터 제어가 가능하다.

제 2장 본문에서는 제안하는 알고리즘의 구조와 흐름을 기술한다. 카메라로부터 실시간 영상을 입력받아 Haar Classifier와 HSV 색상모델 상에서 얼굴과 손을 검출한다. 추출된 영상에서 발표자의 손 무게중심 위치정보와 얼굴 무게중심 정보를 얻어 현재 위치 상태를 확인한다. 이후에 모션 식별 룰을 적용하여 해당 프레젠테이션 명령이 실행된다. 시스템의 전체적인 흐름도는 그림 1과 같다.

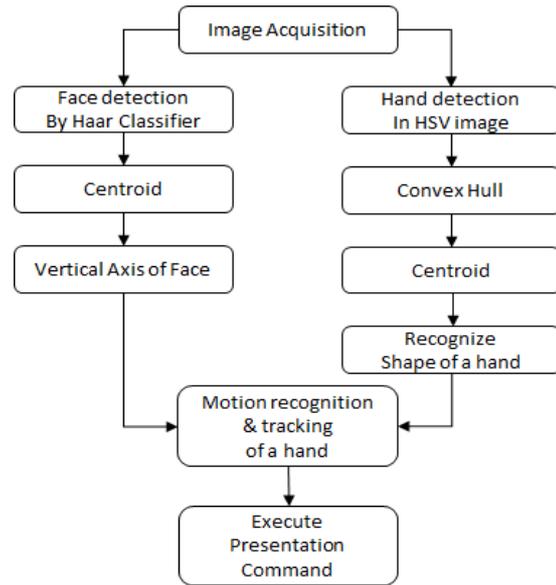


Fig. 1 System Flow

## II. 본론

### 2.1. 얼굴 영역검출

SVM(Support Vector Machines)이나 신경망을 이용한 검출기[3]는 구별 기능이 매우 뛰어난 것으로 알려져 있지만 학습에 많은 시간이 소요된다는 단점을 가지고 있기 때문에 실시간 처리를 요구하는 제안 시스템에서는 입력 영상에서 얼굴 영역을 검출하기 위해 AdaBoost 학습 알고리즘의 Haar-like Feature를 사용하였다.[4]

Haar-like Feature는 Haar Wavelet Transform을 근간으로 하며, 이는 데이터를 분할하거나 압축하는 데 있어서 적은 에러를 가지고 원 데이터를 재구성을 할 수 있도록 하는 손실 압축기법 중의 하나이다. 본 논문에서는 픽셀 자체를 이용하는 것보다 영상이 가지고 있는 특징을 더 잘 대변할 수 있는 Haar-like Feature를 얼굴 영역을 검출하는 인식자로 사용하였다. 이는 학습 알고리즘에서 학습 입력 정보는 줄여주는 반면에 학습 결과의 다양성을 증가시켜주어 더욱 쉬운 분류가 가능하게 한다. 그림 2는 Haar-like Feature의 프로토타입 예를 보여준다.

학습 기반의 물체 추적을 하기 위해서는 대개 각 물체에 대하여 방대한 양의 표본이 있어야 한다. 추가적으로 이 방대한 양의 표본은 각각 레이블을 직접 붙여줘야만

한다. 학습의 과정은 레이블링된 데이터를 학습시키고 이 과정이 끝나게 되면 분류기를 이용하여 레이블링이 되지 않은 시료군을 평가하는 구조로 이루어져 있다.

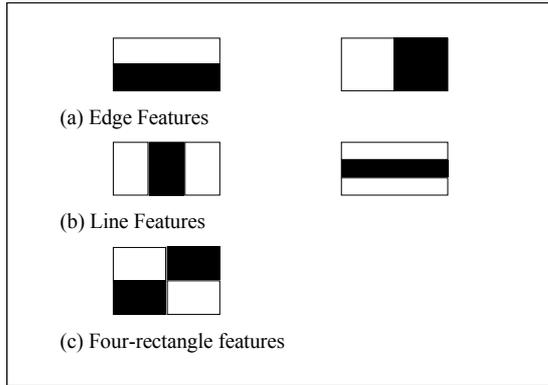


Fig. 2 Prototype of Haar-like feature

이러한 학습을 통한 물체 추적의 대표적인 방법으로는 Adaptive boosting[5]이 있다. 부스팅은 분류 기능이 약한 여러 개의 약 분류기(classifier)를 조합하여 강한 분류기를 만드는 학습(learning)방법이다. 분류 기능이 약한 약 분류기는 간단하지만 연산 속도가 빠르다는 것이 장점이다. 이러한 분류 기능이 약한 약 분류기의 수를 늘리면 분류 기능이 강하고 빠른 연산 속도를 얻을 수 있다. AdaBoost 학습 알고리즘을 이용한 Haar-like Feature로 얼굴 영역을 검출한 결과는 그림 3과 같다.[6]

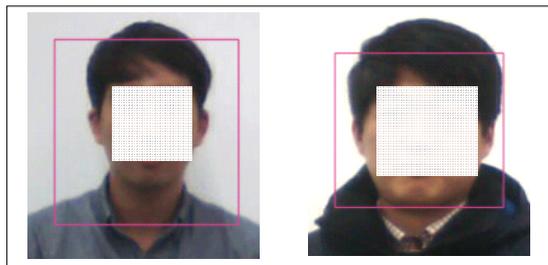


Fig. 3 Results of the face region detection

### 2.2. 손의 경계 및 상태 확인

배경학습과 HSV색상 모델의 Hue 성분 분석에 의하여 영상 내의 손 영역을 검출한 뒤, 손의 상태를 확인하기 위해 손 객체의 외곽 극점의 좌표들의 상대적 위치 및 거리로부터 손의 상태를 파악한다. 우선 검출한 손 영역의 경계를 추출하고, 경계 화소를 대상으로 추적을 수행한다. 외곽선을 이루는 화소간 상대적인 각도 정보

를 이용하여 Convex Hull을 구하고 이를 토대로 손의 상태를 확인한다.

#### 2.2.1. HSV 색상 모델을 통한 손 영역 검출

HSV 색상 모델은 화소의 RGB 성분으로부터 색상(Hue), 채도(Saturation), 명도(Value)의 성분으로 변환된 모델로서, 색상 분석에 있어서 명도나 채도의 변화에 덜 민감한 특징이 있다[4]. 따라서 피부색에 해당하는 적절한 분류기준을 설정하여 환경에 따라 변화가 심한 채도와 명도의 값에 일정한 범위를 부여하여 값의 변화에 따른 영향을 억제하였으며, 피부색에 맞는 색상의 H 값을 기준으로 피부색 화소를 검출하였다[5]. 카메라로부터 추출한 실시간 영상에 대한 정보를 RGB값에서 HSV값으로 변경하기 위해서는 변환함수가 필요하며, 식 (1)과 같이 정의한다.

$$H = \begin{cases} 0 + (G - B) \times 60 / S, & \text{if } V = R \\ 180 + (B - R) \times 60 / S, & \text{if } V = G \\ 240 + (R - G) \times 60 / S, & \text{if } V = B \end{cases}$$

$$\text{if } H < 0 \text{ then } H = H + 360$$

$$S = \begin{cases} (V - \min(R, G, B)) \times 255 / V, & \text{if } (V \neq 0) \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

$$V = \max(R, G, B)$$

배경학습을 통해 전경과 분리한 손영역으로부터 HSV를 이용한 피부색 화소를 추출하면, 보다 정확한 손영역의 검출이 가능하다. 이와 같은 방법으로 손영역을 추출한 영상은 그림 4와 같다.

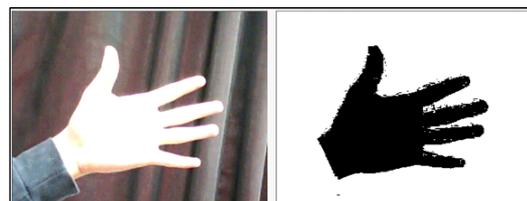


Fig. 4 Hand region Detection using background learning and HSV color model

#### 2.2.2. 엣지 검출

영상에서 상이한 영역을 분리, 검출하는 데 있어서 색상의 값이 현저하게 변화하거나 밝기의 변화가 큰 부분이 특정한 물체의 경계를 나타내는 성질을 이용한다. 따라서 영상 내에서 밝기 변화 값의 차이가 큰 부분을 찾아 객체의 경계를 추출할 수 있다. 엣지 추출은 1차 미

분, 즉 밝기 변화율의 기울기를 검출한다. 기울기가 크다는 것은 밝기 값의 변화가 크게 일어난다는 의미로, 영역의 경계 면이 존재한다는 것을 나타낸다. 이는 엣지가 상대적으로 다른 명암도를 가진 두 영역간의 경계이기 때문이다. 손 영역에 대한 엣지 검출 결과는 그림 5와 같다.

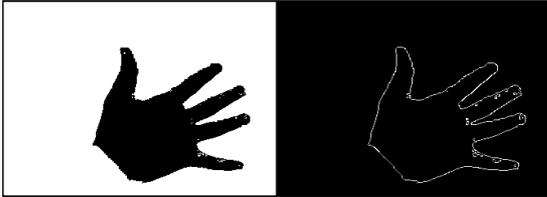


Fig. 5 hand region and edge image

### 2.2.3. 외곽선 검출

검출한 엣지를 이용하여 최외곽 경계들을 좌표값으로 저장하고 처리하기 위해 외곽선 추적을 통하여 외곽 경계를 검출한다. 외곽선은 처음으로 나타나는 임의의 경계에 좌표를 정하고 해당 픽셀의 이웃해 있는 화소의 8방향 이웃점의 존재유무를 검사하는 방식으로 추출할 수 있다.[7] 그림 6은 원본이미지에서 손 영역의 엣지정보로부터 외곽선을 검출한 영상이다.

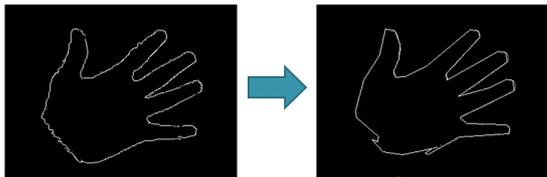


Fig. 6 Hand region and contour extraction

### 2.2.4. Convex Hull

외곽선 추적을 이용하여 윤곽을 검출한 뒤, 윤곽점군 중에 이웃화소와 상대각이 일정 크기 이상이 되는 점들을 연결하여 Convex Hull을 구한다. 외곽선은 개체를 구성하는 점들을 모두 모아놓은 것이고, Convex Hull은 그 중에서도 기울기가 가장 큰 점들을 연결하여 보여준 것이다. 외곽선과 Convex Hull을 이용하면, 개체의 가장 외곽 좌표들과 개체의 가장 안쪽 좌표들을 구할 수 있으며, 손을 기준으로 외곽선과 Convex Hull을 이용하면 손가락들의 끝부분과 손가락의 사이부분을 검출할 수 있게 된다. 그림 7은 외곽선을 이용한 Convex Hull에서 최 외곽점의 연결을 표시한 영상이다. 또한 점군들

내에 분포한 픽셀들의 위치 평균을 이용하여 무게중심을 구할 수 있다. 무게중심 계산식은 (2)와 같다.

$$G_n(x,y) = \left( \frac{\sum_{k=1}^N x_k}{N}, \frac{\sum_{k=1}^N y_k}{N} \right) \quad (2)$$

식(2)에서 N은 검출객체의 화소수,  $G_n$ 은 n번째 프레임의 무게중심 좌표를 의미한다.

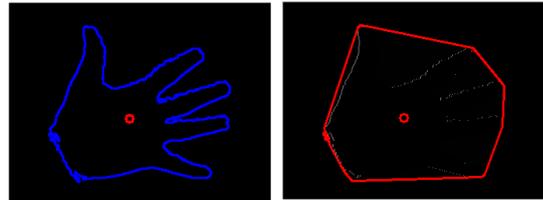


Fig. 7 Contour and Convex Hull

## 2.3. 손의 형태 및 운동 상태 판별

하르 분류기(Haar classifier)와 HSV 색상 모델을 이용하여 입력 영상에서 얼굴과 손 영역을 검출한 뒤, 손 객체의 무게중심점과 Convex Hull로부터 얻어진 다각형의 각 극점과의 연결관계로부터 손의 형태를 인식한다. 이후 얼굴영역의 중심을 기준으로 세로 및 가로 기준축을 설정하고 기준축과 손의 무게중심의 상대적 위치를 참조하여 손 위치 및 운동 상태를 판별한다.

### 2.3.1. 손의 형태 판별

손가락의 끝과 손가락 사이의 좌표를 이용하여 손의 상태를 확인할 수 있다. 하나의 손가락을 기준으로 손가락의 시작점과 끝점을 연결하는 선을 긋는다면 손가락을 펴고 굽히는 상태에 따라 선의 길이가 바뀌게 된다. 선들의 길이로 각각의 손가락의 상태를 확인할 수 있고, 선들의 위치, 길이 및 상대 각도로부터 특정 손가락의 굽힘과 펴에 대한 판단이 가능하다. 이를 위하여 우선적으로 외곽선과 Convex Hull 두 윤곽점들의 공통된 점과 차이가 가장 큰 점을 확인해 보면 손가락의 끝과 사이 이외에도 손목부분에도 좌표들이 확인된다. 이는 손의 중심과 두 개의 점들의 정보(길이, 각도)들을 이용하여 제거할 수 있다.

즉 손가락의 굽힘과 펴에 대한 판단은 다음과 같은 손 형태의 식별 규칙 H-1, H-2의 범주에 포함되며 이를 이용하여 손의 형태를 판별한다.

H-1. 무게중심에서 손가락의 시작점까지의 길이는 중심에서 손가락의 끝점까지의 길이보다 짧다.

H-2. 손의 중심에서 손가락 끝까지 길이는 손가락 시작점에서 끝점의 길이보다 일정 비율 이상 길어질 수 없다.

손의 무게중심점으로부터 손가락의 시점, 종점까지 거리의 상대 정보를 이용하여 손목 이하의 정보를 제거하며, 손가락의 굽힘 혹은 펴 여부를 판별한다. 이에 따라 5종의 명령을 구분하여 실행할 수 있으며 각 명령은 사용자가 사전 정의하여 사용할 수 있도록 한다. 그림 8은 위의 규칙에 따라 손가락의 굽힘과 펴를 판별한 결과이다.

### 2.3.2. 손동작의 운동상태 판별

얼굴 영역으로부터 가로 및 세로 기준축을 설정하고 이를 기준으로 손동작의 상하, 좌우 이동여부를 판단함으로써 컴퓨터 제어에 이용한다. 가로 및 세로 기준축은 다음 식(3)과 같이 정의한다.

$$y = x_c : \text{세로 기준축} \tag{3}$$

$$x = y_c + \frac{1}{2}v : \text{가로 기준축}$$

여기서  $(x_c, y_c)$ 은 무게중심 좌표이며,  $v$ 는 얼굴영역의 세로길이이다.



Fig. 8 Identification of finger status

손의 운동방향에 대한 판단은 다음과 같은 이동방향 식별 규칙 M-1, M-2에 의한다.

M-1. 손의 무게중심점이 기준축  $y$ 를 기준으로 수평방향 이동하는 경우 좌,우 이동으로 판별한다.

M-2. 손의 무게중심점이 기준축  $x$ 를 기준으로 수직방향 이동하는 경우 상, 하 이동으로 판별한다.

식(3)의 기준에 따라 손의 운동상태를 판별한 결과를 그림 9와 그림 10에 나타내었다.

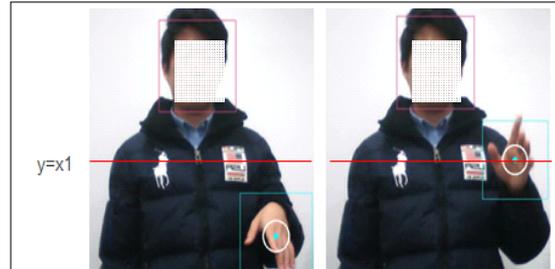


Fig. 9 Vertical movement discrimination of hands using a face region

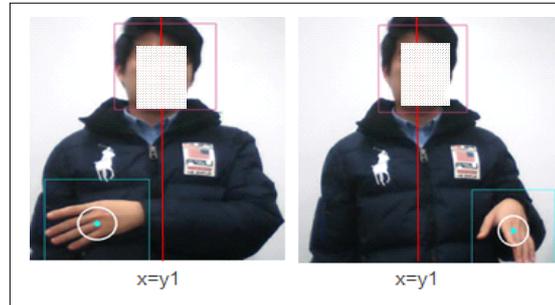


Fig. 10 Horizontal movement discrimination of hands using a face region

### 2.3.3. 손 영역 이동 위치에 따른 모드 전환

손 영역의 위치가 입력되는 영상을 기준으로 다음 그림 11과 같은 방향으로 이동했을 때 모드 전환을 적용하여 다음 슬라이드를 이동하는 기능을 반대로 이전 슬라이드로 이동하도록 만드는 제스처이다.

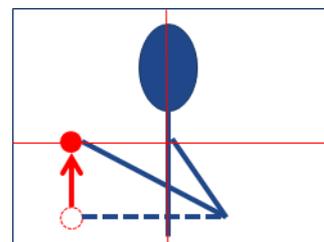


Fig. 11 Movement for switching between forward and backward direction of presentation slides

### III. 실험 및 결과고찰

일반적으로 컴퓨터에 내장 혹은 외장의 형태로 장착되는 카메라의 입력영상으로부터 발표자의 손동작을 인식함으로써 별도의 제어도구 없이 프레젠테이션 제어가 가능한 알고리즘을 구현하였다. 제안하는 시스템은 손동작의 시작과 끝, 동작의 크기 및 방향을 판단하는데 이용하기 위하여 얼굴영역 검출을 시행하며, 얼굴영역으로부터 가로, 세로 기준축을 설정하고 이를 기준으로 다양한 손동작을 인식 및 구분하여 프레젠테이션 제어에 이용한다.

미리 훈련해놓은 객체 인식 특징 파일을 이용하여 해당되는 특정 영역을 검출할 수 있도록 해 주는 하르 분류기(Haar classifier)를 사용하여 얼굴영역을 검출하였다. 또한 손 영역을 검출하기 위한 처리과정으로 HSV 컬러 모델상에서 피부색에 해당하는 Hue 정보의 범위를 설정하여 피부 화소를 검출 하였다. 검출된 피부 화소의 상,하,좌,우 극점의 위치를 근간으로 바운딩 박스로서 얼굴 영역 및 손영역을 정의하였다. 그림 12는 위의 방법으로 검출한 얼굴영역 분할 결과이다.

얼굴영역의 무게중심으로부터 가로, 세로 기준축을 설정한 후 이를 기준으로 한 손의 형태 및 운동 상태 판별 규칙에 따라 사전 설정한 네가지의 운동형태를 판별하는 실험을 행하였다. 손 동작에 따라 다음 슬라이드 이동 동작, 모드를 변경하는 동작, 이전 슬라이드 이동 동작, 그리고 전체화면 동작 총 4가지로 구분하며 그림 13으로 표현하였다.

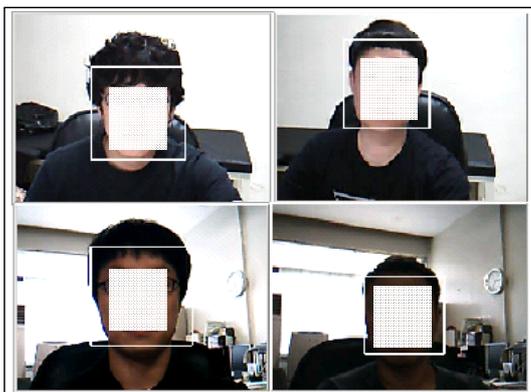


Fig. 12 Results of face region segmentation

전체화면 동작은 우측방향에서 위쪽방향으로 동작을

취하면 전체화면 명령이 실행되고, 손을 우측에서 좌측 방향으로 넘기듯이 손 동작을 취하면 다음 슬라이드 명령이 실행된다. 손을 좌측상단으로 이동하게 되면 모드 전환 명령이 실행되어 다음 슬라이드 동작을 이전 슬라이드 동작으로 전환 할 수 있는 모드 전환이 실행된다.

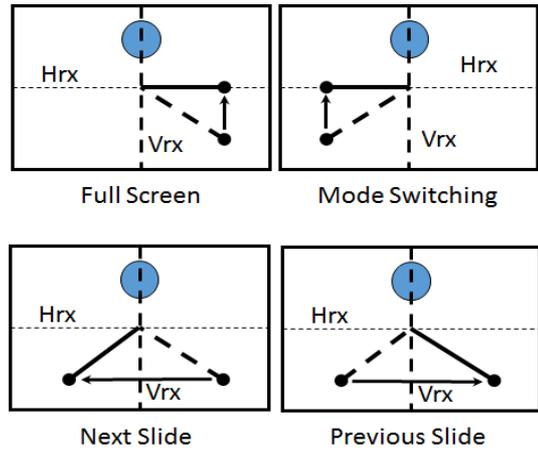


Fig. 13 Presentation control by using motion identification rules

그림 13에서 Hrx는 얼굴영역으로부터 상대적 거리를 고려한 가로 기준축이며 Vrx는 얼굴영역 세로방향 중심을 고려한 세로 기준축이다. 모션 식별 룰에 의한 손동작 인식은 주간 사무실 환경에서 상반신이 노출되는 거리에 카메라를 위치시킨 후 3인에 대하여 그림 13에 제시한 모션 식별 룰을 인식하여 프레젠테이션을 제어하는 방식으로 실시하였다. 세 사람에게 대하여, 손동작 식별 룰 한 동작을 50번씩 두 세트를 반복하여 100회 실시하였고 한 사람당 총 실험 횟수는 400회 반복 실시 하였다. 이에 대한 실험 결과는 표 1에 나타내었다.

Table. 1 Results of presentation control

Motions	Person 1	Person 2	Person 3	Average (%)
Full Screen	96	98	97	97.00
Next/Prev. Slide	97	96	97	96.67
Mode Switch	98	99	97	98.00
Average	97.00	97.67	97.00	97.2

#### IV. 결론

손동작을 인식하여 컴퓨터 프레젠테이션을 제어하는 알고리즘을 제안하였다. 손 동작의 다양한 운동 형태를 인식, 구분함으로써 부가적인 제어용 장치 없이 프레젠테이션을 제어하는 시스템이다. 하르 분류기를 통한 얼굴검출과 HSV 색상모델상에서 피부색 검출을 통한 손영역 추출을 통하여 강건한 손동작 인식을 실현하였다. 본 논문에서 구현한 시스템의 프리젠테이션 제어 실험에서 97.2%의 인식률을 얻어 제안한 알고리즘이 실제 환경에서 효과적으로 사용될 수 있음을 확인하였다. 실험과정에서 발생한 에러는 손동작의 시작 지점과 끝 지점이 가로 혹은 세로 기준축을 양복하는 경우, 모션 식별률에서 정의하지 않은 타원 운동의 경우에 발생하였는데 이러한 에러는 조작자의 손동작 식별률에 대한 숙지 혹은 모션 식별률의 정교화에 의하여 개선될 수 있을 것으로 기대된다.

#### REFERENCES

[ 1 ] B. J. Kim, M. H. Tak, J. Jung and S. H. Park, "A Method of Interaction with AR Object Using Hand Gesture and Arm Movement Recognition in Augmented Reality Advertising," *Asia-pacific Journal of Multimedia Services Convergent with Art, Humanities, and Sociology*, vol.7, no.6, pp. 817-826, June 2017.

[ 2 ] M. S. Chang, S. D. Kwak and S. M. Kang, "Presentation Control System using Gesture Recognition and Sensor," *Journal of Fuzzy Logic and Intelligent Systems*, vol. 21, no. 4, pp. 481-486, Aug. 2011.

[ 3 ] M. H. Huh and H. M. Park "Visualizing SVM Classification in Reduced Dimensions," *Communications for Statistical Applications and Methods*, vol. 16, no. 5, pp. 881-889, Sep. 2009.

[ 4 ] Y. K. Park, H. J. Seo, K. W. Min and J. K. Kim, "A Real-Time Face Detection/Tracking Methodology Using Haar-wavelets and Skin Color," *Journal(B) of Korea Information Processing Society*, vol. 13, no. 3, pp. 283-294, Jun. 2006.

[ 5 ] S. J. Hwang, H. Y. Ko and J. H. Baek, "AdaBoost-based Gesture Recognition Using Time Interval Window Applied Global and Local Feature Vectors with Mono Camera," *Journal of the Korea Institute of Information and Communication Engineering*, vol. 22, no. 3, pp. 471-479, Mar. 2018.

[ 6 ] W. Xu. and E. J. Lee, "A Novel Multi-view Face Detection Method Based on Improved Real Adaboost Algorithm," *Korean Society for Internet Information Transactions on Internet and Information Systems*, vol. 7, no. 11, pp. 2720-2736, Nov. 2013.

[ 7 ] D. Lee, D. Shin and D. Shin, "A Finger Counting Method for Gesture Recognition," *Journal of Korean Society for Internet Information*, vol. 17, no. 2, pp. 29-37, Apr. 2016.



이규원(Kyu-Won Lee)

1986년 연세대학교 전자공학과 졸업(공학사)  
 1988년 연세대학교 대학원 전자공학과(공학 석사)  
 1998년 연세대학교 대학원 전자공학과(공학 박사)  
 1988년 2월 ~ 1989년 8월 (주)LG산전연구소 연구원  
 1989년 9월 ~ 2000년 2월 티티 선임연구원  
 2003년 7월 ~ 2004년 7월 Univ. of Massachusetts 방문연구원  
 2000년 3월 ~ 현재 대전대학교 정보통신공학과 교수  
 ※관심분야 : 영상처리, Robot Vision