

# 최대 공통 부열을 이용한 비전 기반의 양팔 제스처 인식

정회원 최철민\*, 안정호\*\*, 변혜란\*

## Vision-Based Two-Arm Gesture Recognition by Using Longest Common Subsequence

Cheolmin Choi\*, Jung-Ho Ahn\*\*, Hyeran Byun\* *Regular Members*

요 약

본 논문은 비전에 기반한 사람의 양팔 제스처의 모델링과 인식에 관한 연구이다. 우리는 양팔 제스처 인식을 위한 특징점의 추출에서부터 제스처의 분류에 이르는 전체적 틀을 제안하였다. 먼저, 양팔 제스처의 모델링을 위해 색채 기반의 양손 추적 방법을 제안하였고, 추출된 양손의 궤적 정보를 효과적으로 선택하게 하는 제스처 구(Phrase) 분석법을 제시하였다. 선택된 특징 점들의 시퀀스(sequence) 들로 이루어진 훈련 데이터들의 최대 공통 부열(Longest Common Subsequence) 정보를 이용하여 제스처를 모델링하고 이에 따른 유사도 척도를 제안하였다. 제안된 방법론을 공항 등에서 이용하는 항공기 유도 수신호에 적용하였고, 실험을 통해 제안된 방법론의 효율성과 인식성능을 보였다.

**Key Words** : Gesture recognition, Gesture phase analysis, Longest common subsequence

ABSTRACT

In this paper, we present a framework for vision-based two-arm gesture recognition. To capture the motion information of the hands, we perform color-based tracking algorithm using adaptive kernel for each frame. And a feature selection algorithm is performed to classify the motion information into four different phrases. By using gesture phrase information, we build a gesture model which consists of a probability of the symbols and a symbol sequence which is learned from the longest common subsequence. Finally, we present a similarity measurement for two-arm gesture recognition by using the proposed gesture models. In the experimental results, we show the efficiency of the proposed feature selection method, and the simplicity and the robustness of the recognition algorithm.

### I. 서 론

제스처는 신체 일부분 즉, 손가락, 손, 팔, 머리, 그리고 얼굴의 의미 있는 움직임 또는 전신의 움

직임이다. 이러한 제스처를 인지하고 이해하는 것은 인간-기계 상호작용(human-computer interaction) 시스템, 인간-로봇 상호작용(human-robot interaction) 시스템, 그리고 보안 감시 체계에서 가장 활발한 연

※ 본 연구는 한국과학재단 특정기초연구(R01-2005-000-10898-0) 지원 및 한국학술진흥재단(KRF-2005-204-D00041) 지원으로 수행되었습니다.

\* 연세대학교 컴퓨터 과학과 컴퓨터 비전 및 패턴인식 연구실 (wxman, hrbyun@yonsei.ac.kr),

\*\* 강남대학교 컴퓨터 미디어 공학부 (jungho@kangnam.ac.kr)

논문번호 : KICS2007-11-501, 접수일자 : 2007년 11월 9일, 최종논문접수일자 : 2008년 5월 2일

구 분야 중 하나이다. 하나의 제스처는 손의 모양, 손바닥이나 손가락의 방향을 포함하는 손, 움직임, 의미라는 3가지 요소로 부호화될 수 있다. 그러므로 제스처를 인지하기 위해서 우리는 먼저 제스처를 구성하는 인체 부분의 움직임과 구성과 위치에 대한 감지가 필요하다. 이러한 감지는 자기장(magnetic field) 추적기, 데이터-장갑(data-gloves), 그리고 데이터-슈트(data-suit)와 같이 사람이 직접 착용하는 감지 장치를 통한 직접 감지와 카메라와 컴퓨터 비전 기술을 이용하는 간접 감지 방법이 있다. 직접 감지 방식과 달리 비전에 기반한 접근 방법은 환경에 대한 적응력이 있으며, 불편하지도 않고 보다 자연스러운 접근이 가능하며, 움직임 정보의 추출, 형태 정보 추출, 그리고 시공간 정보의 추출과 같은 다양한 지각(perception) 처리가 가능하다. 이러한 이점들은 컴퓨터 비전이 가려짐(occlusion)이라는 본질적 문제점에도 불구하고 인간-컴퓨터 상호작용 시스템, 인간-로봇 상호작용 시스템, 감시 체계 등의 주요한 센서로써 널리 쓰이게 하였다. 최근, 많은 연구자들이 사람의 손, 팔 제스처의 모델링 방법과 인식에 대해서 제안하였는데, 이는 손과 팔의 제스처가 인간의 가장 자연스럽고 일반적으로 쓰이는 제스처 중 하나임을 방증하는 것이기도 하다. 이러한 접근 중에는 은닉 마르코프 모델(hidden Markov models), 유한-상태 머신(finite-state machine)<sup>6, 11</sup>, 인공 신경망<sup>15, 9, 10</sup>, 그리고 통계적 필터링에 기반한 접근 방법<sup>14, 7</sup> 등이 있다.

본 연구에서는 양팔 제스처 인식을 위한 전체적 방법론을 제안한다. 수화와 같은 손 제스처와 달리 양팔 제스처의 의미는 팔의 움직임 경로 즉, 이전 위치에 대한 운동 정보에 크게 의존한다. 이러한 경로 의존성에 따라 우리는 첫째, 양손 제스처의 모델링에 중점을 두었다. 그 다음으로 모델들 간의 유사도를 측정하여 제스처를 분류하는 척도를 제안하였다. 그림 1은 제안된 방법론의 흐름도이다. 사람 얼굴에 대한 피부색 모델을 이용하여 사람의 얼굴을 찾고, 인체의 기하학적 모델을 적용하여 손의 위치를 추정한 다음, 평균 이동(mean shift)에 기반한 추적법<sup>13</sup>으로 양손의 궤적을 추출한다. 추출된 궤적들은 제스처 단계(gesture phase)<sup>18</sup> 분석을 통해 제스처 열(string)로 만들어지고, 제안된 평가 방법을 통해 특정 제스처로 분류된다.

## II. 실시간 손 추적

본 연구에서 우리는 양손의 운동 정보 추출을 위해 평균 이동 추적 알고리즘에 기반한 색채 기반 추적법을 제안한다. 제안된 방법론은 손의 검출을 위해 배경 모델링이나 템플릿을 필요로 하지 않고 서로 다른 조명을 가진 한 사람의 4개의 얼굴영상으로부터 획득된 평균 영상만을 이용한다.

### 2.1 평균 이동 추적법

평균 이동 추적법은 지역적 최상 후보를 찾기 위해 확률 밀도의 그래디언트(gradient)를 찾는 간단하고 강인한 방법이다<sup>14</sup>. 이 방법론은 다음과 같은 세 단계로 구분할 수 있다.

- 1) 목표 객체와 후보 객체의 표현 : 목표 객체와 후보 객체의 probability density function (PDF) 모델은 아래와 같이 정의 될 수 있다<sup>13</sup>. 수식 (1)에서  $\hat{q}_u$ 는 목표 객체,  $\hat{p}_u$ 는 후보 객체를 나타낸다.  $k$ 는 공간적 가중을 위한 프로파일 커널(kernel)이며,  $h$ 는 목표 객체 모델의 대역폭(bandwidth)에 비례한 커널 대역폭,  $C_q$ 와  $C_p$ 는 정규화 요소들이며,  $\delta$ 는 크로네커 델타 함수(kronecker delta function),  $u$ 는 히스토그램 빈의 색인이다.

$$\hat{q}_u = C_q \sum_{i=1}^n k(\|x_i\|^2) \delta[b(x_i) - u], \quad (1)$$

$$\hat{p}_u = C_p \sum_{i=1}^{n_h} k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \delta[b(x_i) - u].$$

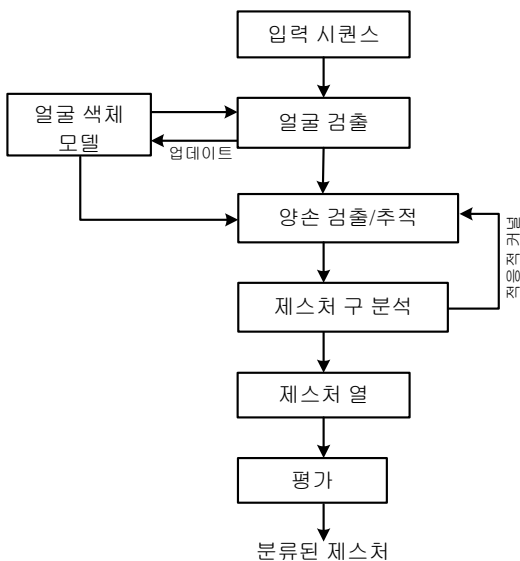


그림 1. 제안된 방법의 흐름도

2) 유사도 측정 : Bhattacharyya 계수를 이용하여 두 히스토그램의 유사도를 수식 (2)와 같이 측정할 수 있다. 여기서 테일러(Taylor) 전개를 이용하면 수식 (2)는 수식 (3)과 같이 정리될 수 있다. 수식 (2)에서  $m$ 은 히스토그램 빈의 수,  $n_h$ 는 공간 커널 속에 포함되는 화소의 수이며,  $y_0$ 는 이전 객체의 위치를 나타낸다.

$$\hat{\rho} = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (2)$$

$$\hat{\rho} \approx \frac{1}{2} \sum_{u=1}^m \sqrt{\hat{p}_u(y_0) \hat{q}_u} + \frac{C_p}{2} \sum_{i=1}^{n_h} w_i k \left( \left\| \frac{y - x_i}{h} \right\|^2 \right) \quad (3)$$

여기서

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \quad (4)$$

3) 평균이동 알고리즘을 이용한 최적화 : 객체의 추적 문제는 유사도 평면상에서의 최대화 문제로 생각될 수 있고, 이 최대화 문제는 평균 이동 방법에 의해 효율적으로 계산될 수 있다. 이전 위치에 기반한 새로운 위치 좌표는 수식 (5)과 같이 계산된다. 여기서,  $g(x) = -k'(x)$ 이다.

$$\hat{y}_1 = \frac{\sum_{i=1}^{n_h} x_i w_i g \left( \left\| \frac{\hat{y}_0 - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^{n_h} w_i g \left( \left\| \frac{\hat{y}_0 - x_i}{h} \right\|^2 \right)} \quad (5)$$

### 2.2 양손의 검출과 추적

사람의 피부색은 인체 검출과 추적 분야에서 가장 널리 이용되는 특징들 중의 하나이다. 얼굴 피부색의 PDF 모델은 평균 얼굴 영상에 공간 커널 (spatial kernel)을 가중시킨 정규화된 빨강, 녹색의 색채 히스토그램을 통해 획득되었다. 히스토그램의 양자화 특성으로 인해 한 사람의 평균 영상만으로도 영상속의 관심 영역에서 사람의 얼굴을 찾을 수 있다 [2]. 그림 2는 얼굴 검출과 이로 인한 양손 검출의 결과 영상이다. 얼굴의 중심점은 평균 이동 추적으로 구할 수 있고, 경계치와 모폴로지 연산을 통

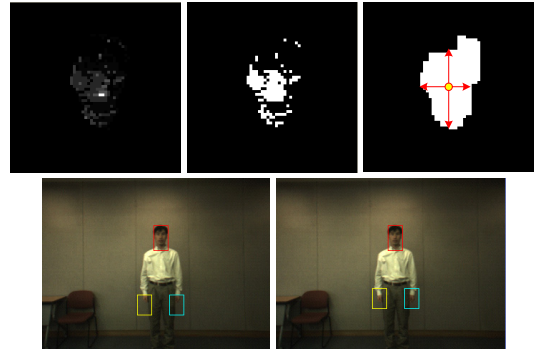


그림 2. 얼굴의 크기와 양손의 검출. 가중치 영상식 (4)에 경계값과 모폴로지를 적용하여 얼굴 크기를 추정한다 (왼쪽). 추정된 얼굴 정보를 이용하여 양손의 초기 탐색영역(아랫줄 왼쪽)을 정하고 양손위치를 검출한다(아랫줄 오른쪽).

해 얼굴을 크기 정보를 추정할 수 있다. 만약 제스처가 카메라 뷰 (view)의 정면에서 수행된다면, 검출된 얼굴 정보와 인체의 기하학적 위치관계를 이용하면, 얼굴의 PDF모델을 이용해서 양손을 검출할 수 있다. 여기서 얼굴 크기 정보는 2가지 중요한 역할을 담당한다. 첫째, 검출된 얼굴의 크기 정보는 새로운 피부색 PDF모델을 형성한다. 이는 조명 변화에 대한 적응력을 갖게 해 준다. 둘째, 양손 추적을 위한 후보 영역의 크기를 추정할 수 있게 한다.

#### 2.2.1 커널(kernel) 위치 예측

양손 제스처의 동작은 때때로 혹은 사람에 따라 그 움직임의 크기가 다양하다. 즉, 이러한 움직임의 다양성으로 손의 위치가 현재 프레임의 커널을 벗어났다면, 추적은 실패하게 된다. 만약 제스처 움직임이 이전 시간에 대해 급격한 변화를 하지 않는다면, 이전 움직임 정보를 이용하여, 커널의 위치를 예측함으로써 이런 단점을 보완할 수 있다.  $t$  시간에서의 예측된 위치,  $P_t^p$ 는 수식 (6)과 같이 계산되어진다. 여기서, 홀드 단계 (Hold Phase; 0)의 의미는 이전 위치와 비교할 때 움직임이 정체 되어 있다는 것으로 자세한 설명은 III장에서 기술하고 있다.

$$P_t^p = P_{t-1} + M_{t-1,t-2} \quad (6)$$

여기서

$$M_{t-1,t-2} = \begin{cases} 0 & \text{where } P_{t-1} \in O \\ P_{t-1} - P_{t-2} & \text{otherwise,} \end{cases} \quad (7)$$

2.2.2 방향 가중 커널

비전 기반 접근법의 잘 알려진 문제점인 가려짐 후의 궤적 복구를 위해 본 연구에서는 방향-가중된 커널을 제안하였다. 방향 가중 커널은 손 움직임의 일관성 가정 하에 이전 움직임 정보를 이용해 커널의 가중치를 적응적으로 변화시키는 창이다. 방향 가중치를 정하기 위해 먼저 움직임 벡터  $M_{t-1, t-d}$  를 구하고, 이 움직임 벡터를 이용하면 t시간에 후보 영역에 속한 픽셀에 대한 n개의 가중 값들을 아래와 같이 정할 수 있다.

$$D_i^t = \{w_1, w_2, \dots, w_n\} \tag{8}$$

여기서,  $0 \leq w_j \leq 1, j=1, \dots, n$  그리고  $D_i^t$ 는 t시간에서 i번째 픽셀의 방향 가중 값으로 커널의 중심으로 부터의 방향 벡터와 움직임 벡터사이의 각도 차이에 반비례 하는 값이다. 식 (1), (3), (4) 그리고 (8)에 따라서, Bhattacharyya coefficient는 다음과 같이 유도되고,

$$w_{i,D} = \sum_{u=1}^m D_i \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{p_{u,D}(y_j)}}, \tag{9}$$

여기서  $\hat{p}_{u,D}$ 는 양손이 서로 가려졌을 때 후보영역에 대한 PDF 표현이며,  $D_i$ 는 후보 영역 픽셀 i의 방향 가중 값이다.

III. 특징 벡터 선택

본 장에서는 간단하고 효율적인 특징점 선택 방법을 제안한다. 양손의 추적 궤적들은 손의 궤적 분석을 통해 식 (10)과 같이 4가지의 제스처 단계(phase)로 구분될 수 있다.

$$P_t \in \begin{cases} H \text{ where } D(P_t, P_{t-1}) < D_T \\ D \text{ where } D(P_t, P_{t-1}) \geq D_T \\ \quad \text{and } \theta(P_t, P_{t-1}) > \theta \\ O \text{ where } D(P_t^r, P_t^l) \geq D_T \\ N \text{ where } P_t \notin H, D, O. \end{cases} \tag{10}$$

여기서,  $D(P_t, P_{t-1})$ 와  $\theta(P_t, P_{t-1})$ 는 각각  $P_t$ 와  $P_{t-1}$  사이의 거리와 각도이며,  $P_t^r, P_t^l$ 는 각각 t 시간에서의 오른손과 왼손의 위치이다. 그리고  $D_T, \theta_T$ 는 각각 거리와 각도에 대한 임계값이다. 집합 H의 의미는 손의 위치가 이전 프레임에 대해 정체하고 있다는 것이며, O는 현 프레임에서 양손이 서로 가

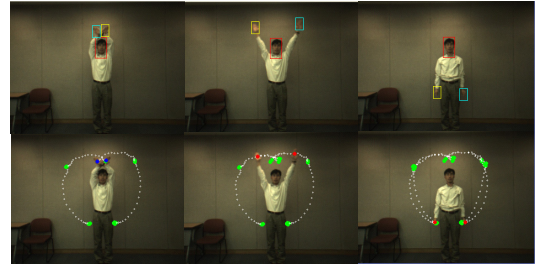


그림 3. 양손 추적과 궤적 구분. IK\_stop 시퀀스의 64, 75, 108번째 영상들(좌로부터). 위쪽은 양손의 추적결과 영상, 아래쪽은 실시간 궤적 분류 결과 영상 (녹색점은 H, D, O에 해당하는 궤적, 파란 점은 현 영상에서 손 위치가 H, D, O에 속할 때, 빨간 점은 현 영상에서의 손이 N에 속할 때, 흰색 점은 누적된 N 궤적).

깝게 위치하고 있음을 의미하며, D는 손이 이전 프레임에 비해 동적으로 움직였음을 의미한다. 식 (10)에 의해 임의의 입력 좌표는 제안된 4가지의 제스처 단계로 분류할 수 있는데, 우리는 제스처 모델을 위한 특징 벡터로 H, D 그리고 O에 해당하는 궤적들만 선택하여, 4차원 특징 벡터(2차원 양손 좌표)를 구성하였다. 그림 3은 양손 추적과 궤적 분류 결과영상이다.

IV. 제스처 모델링

제스처 모델링을 위해 먼저, 우리는 선택된 특징 벡터들을 표상 (symbol) 공간으로 사상 (mapping) 한다. 벡터 양자화 (vector quantization)를 통해 양자화 된 특징 벡터들의 시퀀스는 표상들의 시퀀스로 표현될 수 있다. 제스처의 시간적 분할 (segmentation)이 특징 벡터 선택 과정을 통해 이루어 졌다면 이런 양자화를 통해서는 제스처의 공간적 분할이 이루어진다고 할 수 있다.

4.1 벡터 양자화

선택된 특징 벡터들 간의 스케일 불변 (scale invariant) 를 위해 정규화 단계가 요구되는데, 본 연구에서는 얼굴의 중심좌표와 양손의 초기 검출 위치를 이용하여 신체 중심점을 추정하고, 추정된 신체 중심과 얼굴의 수직 좌표간의 거리를 정규화 상수로 이용하여 선택된 벡터에 정규화를 수행하였다. 정규화된 특징 벡터들은 벡터 양자화 (Vector Quantization) 을 통해 다음과 같은 표상들의 시퀀스 (sequence) 로 표현된다.

$$S = (s_1, s_2, \dots, s_M), \tag{11}$$

여기서,  $s_i, i=1,2,\dots,M-1$ 은 시퀀스의  $i$ 번째 표상이며,  $s_i \neq s_{i+1}$  이다

#### 4.2. 제안된 제스처 모델

훈련 데이터 set의 표상 시퀀스들로부터 각 제스처의 모델을 만들기 위해 우리는 최대 공통 부열과 표상들의 확률 값을 이용한다. 먼저 훈련 데이터 set들의 표상 시퀀스로부터 최대 공통 부열은 동적 계획법(dynamic programming)으로 구할 수 있고, 각 제스처에 대해 표상들의 빈도수로부터 확률 모델을 구할 수 있다. 식 (12)은 제스처 최대 공통 부열 모델이며, 식 (13)은 제스처 확률 모델을 나타낸다.

$$G^i = (g_1^i, g_2^i, \dots, g_{n_i}^i), \quad (12)$$

$$H^i = \{h_1^i, h_2^i, \dots, h_{K_i}^i\}, \quad (13)$$

여기서  $g_l^i, l=1,2,\dots,n_i, i=1,2,\dots,N$ 는  $i$ 번째 최대 공통 부열 모델의  $l$ 번째 표상이며,  $n_i$ 는  $i$ 번째 최대 공통 부열 모델의 길이이다. 또한  $K$ 는 표상들의 총 수 (즉, 코드북의 크기)이며,  $h_p^i, p=1,2,\dots,K$ 는 제스처  $i$ 번째 제스처의  $p$ 번째 표상의 히스토그램 값이다. 여기서 최대 공통 부열의 길이는 각 제스처마다 다를 수 있다. 이러한 길이의 차를 반영하기 위해 최대 공통 부열에 포함되는 표상들에 대한 가중치를 적용하는데,  $G_i$ 의 표상에 대한 가중치를 식 (14)과 같이 정의할 수 있다.

$$w^j = \frac{N_c}{n_i}, \quad (14)$$

여기서  $N_c, N_c \geq n_{\max}$ , 그리고  $n_{\max}$ 는 모든 최대 공통 부열 모델의 최대 길이이다.

### V. 제스처 인식

본 장에서는 제스처 인식을 위한 유사도 척도(similarity measure)를 제안한다. 제안된 척도를 이용하면 임의의 입력 표상 시퀀스에 대해 각 제스처 모델과의 유사도를 평가함으로써 제스처를 인식할 수 있다. 임의의 제스처 표상 시퀀스를  $X=(x_1, x_2, \dots, x_{|X|})$  다음과 같이 정의하자.

$$X = (x_1, x_2, \dots, x_{|X|}), \quad (15)$$

여기서  $|X|$ 는  $X$ 의 길이이다.  $LCS_{X, G^i}$ 를  $X$ 와  $G^i$ 사이의 최대 공통 부열에 속하는 표상들의 집합이라

고 하면,  $X$ 와  $G^i$ 사이의 유사도  $S(X, G^i)$ 는 식 (16)와 같이 정의된다. 측정된 유사도 값들 중에서 최대의 유사도를 가지는 제스처 모델을 선택함으로써  $X$ 에 대한 레이블  $L(X)$ 는 식 (17)과 같이 제스처 인식이 수행된다.

$$S(X, G^i) = \sum_{x_j \in LCS_{XG}} w_i + \sum_{x_j \notin LCS_{XG}} p_{x_j}^i, \quad (16)$$

$$i = 1, \dots, N$$

$$L(X) = \arg \max S(X, G^i). \quad (17)$$

### VI. 실험 결과

본 논문에서 제안한 방법은 Windows XP에서 Visual C++ 6.0을 이용하여 구현하였으며, 펜티엄-IV의 CPU 3.0 GHz와 2GB RAM의 하드웨어에서 실험하였다. 실험데이터는 지상에서 항공기를 유도하는데 쓰이는 양팔 제스처 (aircraft marshalling gesture)<sup>[12]</sup>를 대상으로 10명의 사람들에게 8가지 양손 제스처를 수행하게 하였다. 총 423개의 제스처 영상을 만들었고, 그 중 205개의 영상은 훈련 영상으로 사용하였고, 나머지는 테스트용으로 사용하였다. 그림 4는 본 연구에서 쓰인 양팔 제스처를 보여준다. 벡터 양자화를 위한 코드워드의 수는 20개를 사용하였고, 각각의 코드워드는 1부터 20까지의 숫자를 부여하였다. 표 1은 각 제스처의 최대 공통 부열 모델을 보여준다. Fire의 경우 최대 공통 부열의 길이가 6으로 가장 길었고, S등은 3으로 가장 짧았다. 모든 최대 공통 부열 모델의 첫 번째와 마지막 표상이 13으로 동일하다. 이는 본 연구에

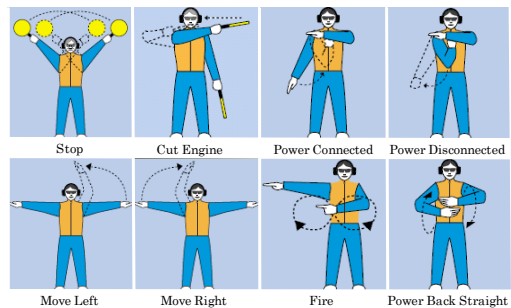


그림 4. 항공기 유도를 위한 양팔 제스처. Stop(S), Cut Engine(CE), Power Connected(PC), Power Disconnected(PDC), Move Left(ML), Move Right(MR), Fire(F), Power Back Straight(PBS).

쓰인 제스처의 시작과 끝이 동일하게 수행되었음을 의미한다. 제안된 특징 벡터 선택방법의 효율성 (총 제적수/선택된 제적수)은 표 2와 같다. 특징 벡터 선택방법의 효율성의 경우 전체 평균이 317%이며, 특히 ML 과 MR의 경우 600%가 넘는 효율성으로 보여, 전체 특징 벡터수의 약 16%에 해당하는 수의 특징 벡터만으로도 제스처 인식이 가능함을 나타냈다. 표 3은 은닉 마르코프 모델을 이용한 인식 결과와 제안된 방법론 간의 인식율을 보여준다. 실험에 이용된 은닉 마르코프 모델은 은닉 상태 (hidden state) 수 2, 상태 (state) 수 20개를 가지는 left-right 모델을 이용하였다. 인식률의 경우 평균 약 98%의 높은 인식률을 보여, 실험에 쓰인 은닉 마르코프 모델의 결과보다 우수하였다. 오분류의 경우 PDC의 5%가 PC로 나머지 5%의 실험 데이터가 PBS로 오분류(misclassification) 되었다. 오분류의 원인은 PDC, PC, PBS 간의 시공간적 유사성에 기인한 것으로 판단된다. 특히 PDC 와 PC의 경우 공간적 구성은 동일하고 시간적 구성이 반대의 경우이다. 또한 제스처를 수행하는 사람간의 변화 (Variation) 도 영향을 미쳤을 것으로 판단된다.

### VII. 결 론

본 논문에서 우리는 최대 공통 부열을 이용하여 양팔 제스처를 모델링 하고 인식하는 방법론에 대해 기술하였다. 먼저 양손의 검출과 추적을 위해 적응적 커널을 갖는 평균 이동 추적법을 제안하였고, 추출된 양손 제적을 효과적이며 효율적으로 분류하는 제스처 단계 분석법을 통해 특징 벡터를 선택하는 방법을 제안하였다. 선택된 특징 벡터들은 표상화 과정을 거쳐 제안된 최대 공통 부열과 확률 값

표 1. 최대 공통 부열 모델

제스처	최대 공통 부열 모델	가중치 값 ( $N_c = 12$ )
CE	13, 8, 20, 13	3
F	13, 18, 14, 11, 11, 13	2
ML	13, 9, 5, 13	3
MR	13, 9, 3, 13	3
PBS	13, 19, 13	4
PC	13, 12, 10, 13	3
PDC	13, 10, 12, 13	3
S	13, 2, 13	4

표 2. 제안된 방법론의 효율성

제스처	총 제적 수 (평균)	선택된 제적수 (평균)	효율성 (%)
CE	120	29	308
F	165	55	300
ML	132	21	629
MR	137	22	623
PBS	114	50	228
PC	101	46	220
PDC	105	56	188
S	123	39	315
합계	1007	318	317

표 3. 인식 결과

Gesture	HMMs	Proposed Method
S	100 %	100 %
CE	100 %	100 %
PC	65 %	100 %
PDC	45 %	90 %
ML	100 %	100 %
MR	100 %	100 %
F	100 %	100 %
PBS	100 %	100 %
Overall	88.75 %	98.75 %

에 기반한 유사도 평가과정을 거쳐 각 제스처로 인식된다. 제안된 특징 벡터 선택방법의 효율성과 제안된 인식기의 높은 인식률을 통해 제안된 방법이 항공기 유도를 위한 양팔 제스처에 적용될 수 있음을 보였다. 다만, 색채 기반 추적법의 조명에 대한 민감성은 향후 연구를 통해 해결되어야 할 것이며, 제안된 제스처 모델의 경우 단일 제스처에만 적용 가능한데, 연속적인 복수 제스처에 대한 일반화 측면에서의 연구를 통해 해결할 것이다.

### 참 고 문 헌

- [1] M. J. Black, and A. D. Jepson, "A probabilistic framework for matching temporal trajectories: Condensation-based recognition of gestures and expressions," *Proc. of the European Conference on Computer Vision*, Vol.1, pp 909-924, 2000.
- [2] C. Choi, J.-H. Ahn, S. Lee, and H. Byun, "Disparityweighted histogram-based object tracking for mobile robot systems," *Lecture Notes in Computer Science*, Vol.4282, pp.584-593, 2006.



[3] D. Comaniciu, and P. Meer, "Kernel-based object tracking," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.25, no. 5, pp.564-577, 2003.

[4] D. Comaniciu, and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, voi. 24, no. 5, pp.603-619, 2002.

[5] F. Dadgostar, A. Sarrafzadeh, C. Fan, L. D. Silva, and C. Messom, "Modeling and recognition of gesture signals in 2D space: A comparison of NN and SVM approaches," *Proc. of the IEEE Int'l Conf. Tools with Artificial Intelligence*, pp.701-704, 2006.

[6] P. Hong, M. Turk, and T. S. Huang, "Gesture modeling and recognition using finite state machine," *Proc. of the IEEE Int'l Conf. Automatic Face Gesture Recognition*, pp.410-415, 2000.

[7] M. Isard, and A. Blake, "CONDENSATION-Conditional density propagation for visual tracking," *Int'l Journal of Computer Vision*, Vol.1, pp.5-28, 1998.

[8] D. McNeil, "*Hand and Mind, What gesture reveal about thought*," University of Chicago Press, 1992.

[9] Q. Munib, M. Habeeb, B. Takruri, and H. A. Al-Malik, "American sign language(ASL) recognition based on Hough transform and neural networks," *Expert Systems with Applications*, Vol.32, pp.24-37, 2007.

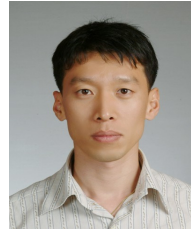
[10] C. Oz, and M. Leu, "Linguistic properties based on American sign language isolated word recognition with artificial neural networks using a sensory glove and motion tracker," *Neurocomputing*, Vol.70, pp.2891-2901, 2007.

[11] M. Yeasin, and S. Chaudhuri, "Visual understanding of dynamic hand gesture," *Pattern Recognition*, Vol.33, pp.1805-1817, 2000.

[12] <http://www.caa.co.uk/docs/33/CAP637.PDF>

**최 철 민 (Cheolmin Choi)**

정회원



1995년 2월 공군사관학교 산업공학과 졸업  
 2000년 2월 연세대학교 전기전자공학과 석사  
 2005년 3월~현재 연세대학교 컴퓨터과학과 박사과정

<관심분야> 원격 탐사, 인공 지능, 패턴인식

**안 정 호 (Jung-Ho Ahn)**

정회원



1996년 2월 연세대학교 수학과 졸업  
 1998년 2월 연세대학교 수학과 석사  
 2001년 12월 Texas A&M Univ. Statistics M.S.  
 2006년 2월 연세대학교 컴퓨터과학과 박사

2007년 2월~현재 강남대학교 컴퓨터미디어공학부 전임강사

<관심분야> 패턴인식, 인공 지능

**변 혜 란 (Hyeran Byun)**

정회원



1980년 2월 연세대학교 수학과 졸업  
 1983년 2월 연세대학교 수학과 석사  
 1987년 6월 Univ. of Illinois at Chicago Computer Science, M.S.

1993년 12월 Purdue Univ. Computer Science, Ph. D.

2004년 3월~현재 연세대학교 컴퓨터과학과 정교수

<관심분야> 패턴인식, 컴퓨터 비전 시스템