

# 바디 제스처 인식을 위한 기초적 신체 모델 인코딩과 선택적 / 비동시적 입력을 갖는 병렬 상태 기계

## Primitive Body Model Encoding and Selective / Asynchronous Input-Parallel State Machine for Body Gesture Recognition

김 주 창<sup>1</sup>, 박 정 우<sup>2</sup>, 김 우 현<sup>3</sup>, 이 원 형<sup>4</sup>, 정 명 진<sup>†</sup>

Kim Juchang<sup>1</sup>, Park Jeong-Woo<sup>2</sup>, Kim Woo-Hyun<sup>3</sup>, Lee Won-Hyong<sup>4</sup>,  
Chung Myung-Jin<sup>†</sup>

**Abstract** Body gesture Recognition has been one of the interested research field for Human-Robot Interaction(HRI). Most of the conventional body gesture recognition algorithms used Hidden Markov Model(HMM) for modeling gestures which have spatio-temporal variabilities. However, HMM-based algorithms have difficulties excluding meaningless gestures. Besides, it is necessary for conventional body gesture recognition algorithms to perform gesture segmentation first, then sends the extracted gesture to the HMM for gesture recognition. This separated system causes time delay between two continuing gestures to be recognized, and it makes the system inappropriate for continuous gesture recognition. To overcome these two limitations, this paper suggests primitive body model encoding, which performs spatio/temporal quantization of motions from human body model and encodes them into predefined primitive codes for each link of a body model, and Selective/Asynchronous Input-Parallel State machine(SAI-PSM) for multiple-simultaneous gesture recognition. The experimental results showed that the proposed gesture recognition system using primitive body model encoding and SAI-PSM can exclude meaningless gestures well from the continuous body model data, while performing multiple-simultaneous gesture recognition without losing recognition rates compared to the previous HMM-based work.

**Keywords:** Body Gesture Recognition, Primitive Body Model Encoding, SAI-PSM, Multiple-Simultaneous Gesture Recognition

### 1. 서 론

본 연구는 사람-로봇 상호작용(Human-Robot Interaction;

HRI)을 위한 바디 제스처 인식을 위한 알고리즘을 제안한다. 특히, 최근 활발히 연구되고 있는 깊이 센서 기반 몸의 자세 추정 연구를 바탕으로 얻어진 바디 제스처 인식을 다루고 있다<sup>1-3)</sup>. 기존의 바디 제스처 인식 알고리즘은 대부분 은닉 마르코프 모델(Hidden Markov Model, 이하 HMM)을 바탕으로 바디 제스처 인식을 수행한다<sup>4-7)</sup>.

Yang<sup>7)</sup>의 연구에서는 제스처가 아닌 동작을 가비지 제스처(garbage gesture)로 정의하고 이를 검출하기 위한 garbage gesture model을 HMM을 이용해 설계하여 강인한 바디 제스처 인식을 구현하였다. 하지만 사람의 동작은 헤아릴 수 없이 다양하기 때문에, 의미 없는 제스처를 유한 개의 HMM으로 모델링 하여 의미 있는 제스처로부터 구별

Received : Jul. 13. 2012; Reviewed : Aug. 6. 2012; Accepted : Dec. 4. 2012

※ This research was funded by the MKE(The Ministry of Knowledge Economy), Korea, under the Human Resources Development Program for Convergence Robot Specialists support program supervised by the NIPA(National IT Industry Promotion Agency) (NIPA-2012-H1502-12-1002)

<sup>†</sup> Corresponding author: Electrical Engineering, KAIST, Guseong-dong, Yuseong-gu, Daejeon, Korea (mjchung@ee.kaist.ac.kr)

<sup>1</sup> Electrical Engineering, KAIST(kjc0419@kaist.ac.kr)

<sup>2</sup> Electrical Engineering, KAIST(pjw@rr.kaist.ac.kr)

<sup>3</sup> Electrical Engineering, KAIST(ishsrain@rr.kaist.ac.kr)

<sup>4</sup> Electrical Engineering, KAIST(leestation@rr.kaist.ac.kr)

해 내는 데에는 한계가 있으며, 이는 인식하고자 하는 제스처의 종류가 많아질수록 더 어려워진다. 또, HMM을 이용한 제스처 인식은 제스처의 끝점을 인식한 후 거꾸로 거슬러 올라가 시작점을 찾는 제스처 분할 과정을 거친 후 HMM을 통해 제스처 인식을 수행하기 때문에, 지연 없는 연속적 제스처 인식이 어렵다는 문제점도 존재한다.

Kim<sup>[4]</sup>은 기존의 HMM을 이용한 제스처 인식에서, 제스처 분할 이후 제스처 인식을 수행할 때 생기는 지연으로 인해 연속적인 제스처 인식이 불가능하다는 한계점을 극복하기 위해 forward spotting accumulative HMM을 이용한 연속적인 바디 제스처 인식을 제안하였다. 하지만 여전히 HMM을 통한 제스처 모델링을 하기 때문에, 의미 있는 제스처와 의미 없는 제스처간의 강인한 구분이 어렵다.

Peng<sup>[5,6]</sup>은 다중 카메라를 통한 view-invariant 바디 제스처 인식과 동시에, HMM을 이용한 의미 없는 제스처의 구별을 이전보다 강화하기 위해 일반적인 가비지 제스처 모델(general garbage gesture model)과 특정 가비지 제스처 모델(specific garbage gesture model)을 동시에 학습시키고, 이를 제스처 HMM 모델과 병렬로 연결하여 보다 강인한 제스처 인식을 제안하였다. 의미 없는 제스처 모델링을 구체적으로 정의하고 학습시킴을 통해 HMM을 이용한 의미 없는 제스처의 구별 성능을 향상시켰다는 장점이 있지만, 제스처 분할 이후 제스처 인식을 수행하면서 발생하는 지연으로 인한 문제점을 여전히 가지고 있다.

본 논문에서는 선행 연구들이 가지는 HMM 기반 제스처 인식 알고리즘이 갖는 문제점들을 해결한 새로운 제스처 인식 알고리즘을 제안한다. 제안된 알고리즘은 깊이 센서 데이터로부터 추정된 신체 모델의 동작을 시공간적으로 양자화 및 코드화하는 기초적 신체 모델 인코딩(Primitive Body Model Encoding)과, 이로부터 생성된 코드를 바탕으로 동시다발적 제스처 인식을 수행하기 위한 선택적/비동시적 입력을 갖는 병렬 상태 기계(Selective / Asynchronous Input-Parallel State Machine, SAI-PSM)로 이루어져 있다. 실험을 통해 기초적 신체 모델 인코딩과 SAI-PSM을 이용한 제스처 인식이 기존의 HMM 기반 알고리즘의 인식률을 유지한 채 의미 없는 제스처의 제외에 강인하며 제스처간의 지연 없이 연속적인 제스처 인식을 수행할 수 있음을 확인하였다.

## 2. 기초적 신체 모델 인코딩

### 2.1 깊이 영상으로부터의 신체 모델 추정

바디 제스처를 인식하기 위해서는 센서로부터 사람 몸의 자세를 추정해야 하는데, 깊이 센서를 이용한 깊이 영상 기반 신체 모델 추정 기술이 현재까지 연구된 것들 중 가장 좋은 성능을 나타낸다. Shotton<sup>[3]</sup>의 연구에서는 단일 깊이 영상으로부터 빠르고 정확하게 신체 관절의 3차원 좌표를 추정하는 알고리즘을 제안하였다. Shotton<sup>[3]</sup>은 기존의 영상 내 신체 모델 추정 알고리즘이 겪는 트레이닝 데이터의 부족을 극복하기 위해 모션 캡처를 통한 방대한 양의 데이터베이스를 구축하고, 이를 바탕으로 깊이 영상에서 신체 부위를 분류한 후 각 신체 관절의 3차원 좌표를 추정하는 알고리즘을 개발하였다. 본 연구에서는 Shotton<sup>[3]</sup>의 연구 내용을 바탕으로, 깊이 영상으로부터 총 15개의 관절 좌표를 추정하고, 이를 통해 12개의 링크를 가지는 스펙트럼 신체 모델을 생성한다.

### 2.2 기초적 신체 모델 인코딩

기초적 신체 모델 인코딩(primitive body model encoding)은 영상으로부터 추정된 신체 모델의 연속적인 시퀀스를 시공간적으로 양자화 하는 과정이다. 신체 관절의 좌표로부터 얻어지는 관절과 관절 사이의 링크(link)의 3차원 각도를 계산하고, 각 링크별로 각도의 변화를 관찰하여 링크의 동작을 시공간적으로 양자화 하게 된다. 각 링크의 양자화 된 동작은 하나의 코드(code)로 표현하게 되며, 이 코드를 기초 코드(primitive code)라고 정의한다. 즉, 기초적 신체 모델 인코딩은 신체 모델을 각 링크에 해당하는 기초 코드로 변환하는 과정이다.

### 2.3 기초 코드의 정의

기초 코드는 각 링크의 동작을 양자화 한 정보를 담고 있으며, 기초 동작 코드(Primitive Motion Code; PMC)와 기초 자세 코드(Primitive Pose Code; PPC)로 구성된다. 두 개의 서브 코드는 서로 다른 정보를 양자화 하기 위해 각각 다른 알고리즘을 통해 인코딩 된다. 기초 동작 코드는 링크의 회전 방향 및 회전 각도를 시공간적으로 양자화 한 정보를 가지며, 기초 자세 코드는 회전의 끝점에 해당하는 링크의 자세를 공간적으로 양자화 한 정보를 가진다.

기초 동작 코드는 각 링크의 회전 방향을 나타내는 코드 워드  $\alpha$ 와, 회전 각도를 나타내는 코드 워드  $\beta$ 로 이루어져 있다. 회전 방향은 세 개의 궤도에 대해 각각 두 개의

방향인 총 6개의 방향으로 정의된다. 회전 방향으로의 회전 각도의 양자화 간격은 20°로 결정하였는데, 양자화 간격이 너무 크거나 작아지면 이를 코드 화했을 때 코드의 개수가 너무 적거나 많아지게 된다는 점을 고려하며 실험적으로 결정하였다. 결과적으로, 각 링크의 회전은 총 55개의 기초 동작 코드를 통해 표현되게 된다(Table 1).

각 궤도에 해당하는 회전을 검출하기 위해서 먼저 각 궤도상의 회전을 가질 수 있는 링크의 각도 범위를 정의하고, 각 궤도에 해당하는 범위에 링크가 존재할 때 그 궤도에 해당하는 회전 여부 및 회전 각도를 검출한다. 링크가 각 궤도에 해당하는 범위에 존재할 때, 링크의 3차원 벡터를 궤도 위로 투사하여 궤도상의 벡터로 변환한 후, 궤도 위에서의 각도 변화를 관찰하여 회전 여부 및 회전 각도를 검출한다. Fig. 1은 수직 궤도에 대한 회전 검출 과정을 예시로 보여준다.

정의된 기초 동작 코드는 링크의 회전 방향 및 각도에 대한 정보만을 담고 있기 때문에 회전의 시작과 끝을 알 수 있는 절대적인 정보가 필요하다. 기초 자세 코드는 회전의 끝점의 정보, 즉 회전의 절대적인 위치 정보를 양자화한 정보를 담게 된다. 링크의 자세를 구 좌표계를 이용해  $\varphi$ 와  $\theta$ 로 나타내게 되면, 기초 자세 코드는  $\varphi$ 를 양자화 한 코드 워드  $\gamma$ 와  $\theta$ 를 양자화 한 코드 워드  $\delta$ 로 이루어진다. 이 때 구 좌표계의 원점은 각 링크의 시작 관절이 된다. 각

Table 1. 55 Primitive Motion Codes

Primitive Motion Code = ( $\alpha$ , $\beta$ )				
Trajectory & Direction			Angle	
Trajectory	$\alpha$	Direction	$\beta$	Angle
	0	Static	0	0°
Vertical Trajectory	1	Clockwise	1	20°
			2	40°
			3	60°
Horizontal Trajectory	3	Clockwise	4	80°
			5	100°
			6	120°
Circular Trajectory	5	Clockwise	7	140°
			8	160°
			9	180°

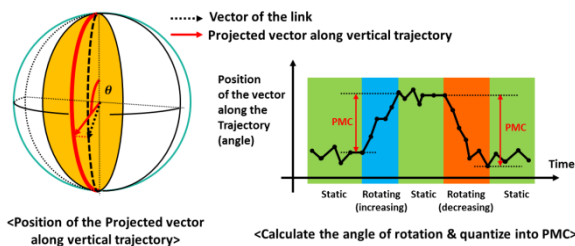


Fig. 1. Detection of rotation along the vertical trajectory

코드 워드는  $\varphi$ 와  $\theta$ 를 45°간격으로 분할한 양자화 레벨에 대해 정의되며, 이 양자화 간격도 기초 동작 코드의 경우와 마찬가지로 실험적으로 결정하였다. 즉, 각 회전의 끝점에 해당하는 자세는 26개의 기초 자세 코드를 통해 표현되게 된다(Table 2).

Table 2. 26 Primitive Pose Codes

Primitive Pose Code = ( $\gamma$ , $\delta$ )			
$\gamma$	$\phi$	$\delta$	$\Theta$
0	0°	0	
1	45°	1	0°
		2	45°
		3	90°
2	90°	4	135°
		5	180°
		6	225°
3	135°	7	270°
		8	315°
		4	180°

### 3. 선택적/비동시적 입력을 갖는 병렬 상태 기계

#### 3.1 기초 코드와 유한 상태 기계

기초 코드로부터 연속적인 바디 제스처를 인식하기 위해서는 기초 코드를 이용해 제스처를 모델링하고, 실시간으로 들어오는 신체 모델의 기초 코드로부터 제스처가 발생했는지를 검출해야 한다. 기초 코드의 순차적 발생을 제스처로 정의하고, 이를 검출할 수 있는 제스처 인식 모델을 제안한다. 기초 코드의 순차적 발생을 검출하기 위한 모델로 유한 상태 기계(Finite State Machine; FSM)의 개념을 이용한다. 유한 상태 기계는 시스템이 처할 수 있는 모든 상황을 유한개의 상태로 정의하고, 주어지는 입력에 따라 다른 상태로 전이를 하거나 출력 또는 액션이 일어나게 하는 장치를 나타내는 모델이다. Mitra<sup>[8]</sup>에 의하면, “제스처는 유한 상태 기계를 이용해 시공간적 상태의 시퀀스로 모델링 될 수 있다.”라고 하였다. 그러므로 유한 상태 기계는 기초 코드의 순차적 발생을 검출하는데 있어서 가장 직관적이고 적합한 알고리즘이라고 할 수 있다.

유한 상태 기계를 이용해 기초 코드의 순차적 발생을 검출하기 위한 시스템을 모델링 하기 위해, 유한 상태 기계

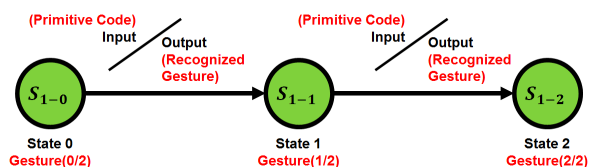


Fig. 2. Gesture recognition using finite state machine

의 상태는 현재 인식하고자 하는 제스처의 인식 과정의 단계를 나타내고 입력과 출력은 각각 기초 코드와 인식된 제스처로 정의한다. 이 같은 정의를 통해, 유한 상태 기계에 기초 코드가 순차적으로 입력되었을 때 상태가 전이됨에 따라 제스처의 인식 과정의 단계가 올라가고, 최종 상태에 도달하게 되면서 출력을 통해 제스처가 인식되게 된다(Fig. 2).

### 3.2 선택적/비동시적 입력을 갖는 병렬 상태 기계

일반적인 유한 상태 기계와 기초 코드를 이용한 제스처 인식 과정에서 발생하는 몇 가지 한계점을 극복하기 위해, 본 연구에서 동시다발적인 제스처 인식을 위한 선택적/비동시적 입력을 갖는 병렬 상태 기계(Selective / Asynchronous Input-Parallel State Machine, 이하 SAI-PSM)을 제안한다.

SAI-PSM은 유한 상태 기계에 세 가지 특성을 추가하여 정의된다. 첫 번째로, SAI-PSM은 여러 유한 상태 기계가 동시에 존재하는 병렬 상태 기계(Parallel State Machine; PSM)의 구조이다. 이 같은 병렬적인 구조를 통해 하나의 제스처 당 하나의 유한 상태 기계를 설계하여 동시다발적인 제스처 인식을 수행한다. 병렬 상태 기계 내에 총  $n$ 개의 유한 상태 기계가 존재한다고 가정했을 때,  $i$ 번째 유한 상태 기계의 현재 상태를  $S_i(t)$ 라 하면 SAI-PSM의 상태는 식 (1)와 같이 정의된다.

$$S(t) = \{S_1(t), S_2(t), \dots, S_n(t)\} \quad (1)$$

두 번째로, SAI-PSM은 약간의 시간차를 두고 입력되는 다중 입력을 처리 할 수 있다.  $n$ 개의 다중 입력을 통한 상태 전이는, 처음 입력이 들어온 순간부터  $n$ 개가 모두 입력되는데 까지 걸리는 시간이 일정 시간 이하일 때에 성공하게 된다. 이를 통해, 서로 다른 신체 부위로부터 약간의 시간차를 두고 입력되는 동작에 대한 제스처 인식을 수행할 수 있게 된다.

마지막으로, SAI-PSM 내의 각 유한 상태 기계는 처리할 수 있는 고유 입력 집합을 가진다. 즉, 각 유한 상태 기계는 인식하고자 하는 제스처에 해당하는 신체 부위의 기초 코드만을 선택적으로 입력 받는다. 이와 같은 선택적인 입력 처리를 통해 SAI-PSM을 이용한 제스처 모델링을 보다 효율적으로 수행할 수 있으며, 제스처에 해당하는 신체 부위의 동작만을 입력 받으므로 강인한 제스처 인식을 수행할 수 있다.

### 3.3 SAI-PSM을 이용한 바디 제스처의 모델링

SAI-PSM과 기초 코드를 통한 동시다발적 제스처 인식을 수행하기 위해 바디 제스처를 SAI-PSM을 이용해 모델링 해야 하는데, 이 과정은 SAI-PSM의 상태를 정의하는 것과 입력에 따른 상태 전이를 정의하는 것으로 나뉜다. SAI-PSM의 상태는 인식하고자 하는 제스처의 특징적인 자세나 동작을 기준으로 정의되어야 한다. 아래의 가정 (1)과 (2)로부터 각 제스처에 해당하는 SAI-PSM의 상태를 정의하게 된다.

가정 (1) : 인식하고자 하는 제스처의 설명은, 제스처가 가지는 의미와 더불어 그 제스처를 수행하기 위해 어떤 신체부위를 어떻게 움직여야 하는지에 대한 정보를 포함한다.  
 가정 (2) : SAI-PSM의 상태는 제스처의 설명으로부터 사용자에 의해 별 다른 규칙 없이 직관적으로 정의될 수 있다.

SAI-PSM의 상태가 정의되고 나면, 제스처 인식을 위한 입력과 상태 전이를 정의해야 한다. Table 3은 SAI-PSM의 상태 전이에 해당하는 입력을 정하는 규칙을 정의한다. 각 상태 전이에 해당하는 입력은 반복적인 제스처의 수행을 통한 강화 학습(reinforcement learning)을 통해 학습된다.

Table 3. Definition of State Transitions in SAI-PSM

State Transition	Description	Primitive Codes
Forward State Transition	Every state transition except static & backward state transition	Set of primitive codes of the motion for the next state
Static State Transition	State transition of staying on the current state	Set of primitive codes which can be seen between the motion for the current state and the motion for the next state
Backward State Transition	State transition that return to the base state	Every primitive codes except primitive codes for forward state transition and backward state transition

## 4. 실험 결과 및 분석

실험을 위해 엔터테인먼트 로봇과의 인터랙션을 위한 15개의 바디 제스처를 정의하였다. 각각의 바디 제스처는 팔, 다리, 목 등 신체 모델의 모든 링크의 다양한 움직임을 반영함과 동시에 일반적으로 통용되는 동작 위주로 정의함을 통해, 본 실험에 사용된 제스처들이 다른 연구에 사용된

제스처들에 비해 일반성을 잃지 않도록 하였다. 본 논문은 제안된 알고리즘의 성능을 총 세 가지의 실험을 통해 검증하였다.

### 4.1 의미 없는 행동들의 제외

제안된 알고리즘이 기존의 HMM을 바탕으로 한 가비지 제스처 모델링 없이도 의미 없는 행동들을 효율적으로 제외할 수 있음을 검증하고자 한다. 제안된 알고리즘은 제스처 분할 과정 없이 실시간으로 입력되는 기초 코드를 바탕으로 제스처 인식을 수행하기 때문에, 기존의 HMM 기반 확률적 제스처 인식과는 다르게 제스처에 해당하는 기초 코드가 순서대로 정확히 입력되어야 제스처가 인식된다. 그러므로 제스처가 아닌 다른 행동들로부터 기초 코드들이 SAI-PSM에 입력되더라도, 학습된 제스처에 해당하는 순서대로 정확히 입력되지 않으면 출력을 만들어 내지 못해 제스처가 인식되지 않는다. 이를 보여주기 위해, 카메라를 향해 팔과 다리를 움직일 수 있는 모든 방법을 동원하여 다채롭게 몸동작을 수행했을 때 이러한 움직임은 잘 제외하는지를 정성적으로 확인하였다.

Fig. 3은 5000 프레임(약 500초)동안 카메라를 향해 팔과 다리를 끊임없이 다채롭게 움직였을 때 SAI-PSM 내 모든 유한 상태 기계들의 상태 및 출력을 나타낸다. 의미 없는 행동들이 발생시킨 기초 코드에 의해 다수의 유한 상태 기계가 상태 전이를 했지만 대부분이 출력을 발생시키는 데에는 실패하고 다시 처음 상태로 돌아가게 된다. 실험 도중 “Kick”과 “Hello” 제스처가 한 번씩 인식되었는데, 이는 자유롭게 몸을 움직이는 과정에서 각 제스처에 해당하는 동작을 무의식적으로 수행하여서 발생된 출력이다. 이 실험을 통해, 제안된 시스템이 정의된 제스처 이외의 행동들을 의미 없는 행동으로 잘 제외할 수 있는 특성을 가짐을 확인하였다.

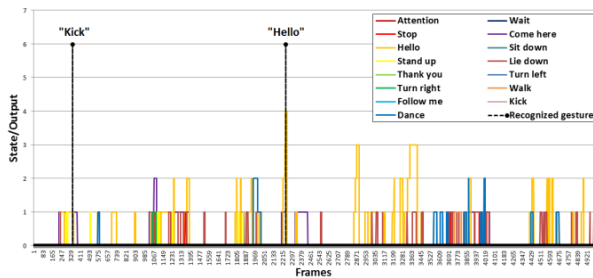


Fig. 3. State and output of each finite state machine in SAI-PSM while performing meaningless gestures

### 4.2 동시다발적 제스처 인식

제안된 시스템이 여러 제스처를 동시다발적으로 인식할 수 있음을 검증하기 위해 서로 다른 두 제스처를 동시다발적으로 수행하고 이에 따른 병렬 상태 기계의 상태 전이 및 출력을 분석했다. Fig. 4는 오른 팔을 이용한 “Hello” 제스처와 오른쪽 다리를 이용한 “Kick” 제스처에 해당하는 유한 상태 기계의 상태와 출력을 시간에 따라 관찰한 그래프이다. 17-71 프레임 구간과 109-125 프레임 구간에서는 각 제스처를 단독으로 수행하였으며, 172-220 프레임 구간과 255-300 프레임 구간에서는 두 제스처를 동시에 수행하였다. 결과로부터 각각의 유한 상태 기계가 병렬적으로 제스처 인식을 수행함을 통해 “Hello”와 “Kick” 제스처가 동시에 인식 되는 것을 볼 수 있다. 이 결과를 통해 실험에 사용된 두 제스처 이외에 다른 모든 제스처들의 동시다발적 인식이 가능함을 알 수 있다. 즉, 제안된 시스템은 어떠한 두 개 또는 다수의 제스처들을 동시다발적으로 인식할 수 있으며, 이는 곧 제스처와 제스처 사이에 지연 없이 연속적으로 인식할 수 있는 시스템임을 검증하였다.

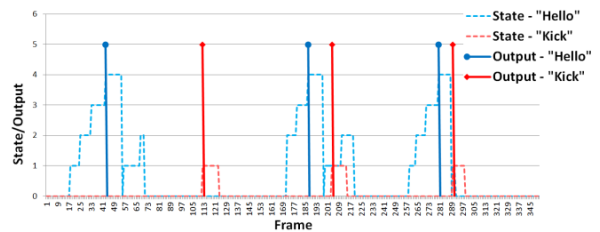


Fig. 4. State and output of each finite state machine while performing both “Hello” and “Kick” gestures

### 4.3 시공간적 변화를 가지는 제스처의 인식을 검증

같은 제스처라도 제스처를 수행하는 사람이 바뀌거나 제스처를 여러 번 수행할 때 그 형태는 시공간적 변화를 가진다. 기초 코드와 SAI-PSM을 이용한 제스처 인식이 시공간적 변화를 가지는 제스처를 잘 모델링 하는 지를 검증하기 위해, 여러 명의 사람이 정의된 제스처를 수행하였을 때의 인식을 확인하여 제스처의 시공간적 변화에 강인한지를 검증한다. Table 4는 10명의 실험자에게 제스처에 대한 설명을 제시하고, 각 제스처에 대해 10번의 제스처를 수행하고 성공 횟수를 체크하여 인식을 계산한 결과이다. 각 제스처에 대한 10명의 평균 인식률은 최소 70%에서 최대 99%였으며, 종합적인 인식률은 87.7%였다. “Come here”, “Sit down”, “Stand up”, “Thank you” 제스처의 경우에는 평균 인식률이 70% 대로 다른 제스처에 비해 낮게 나타났

Table 4. Gesture Recognition Rate from Various Subjects

Participant Gesture	1	2	3	4	5	6	7	8	9	10	Sum	Recogniti on Rate
Attention	10	8	10	2	6	10	10	10	10	9	85	85%
Wait	10	8	10	3	6	10	10	10	9	8	84	84%
Stop	9	10	10	10	10	10	10	10	10	10	99	99%
Come here	7	7	5	9	8	3	8	8	8	7	70	70%
Hello	9	10	10	10	9	10	10	8	10	8	94	94%
Sit down	8	6	6	9	9	7	10	4	5	9	73	73%
Stand up	5	8	6	5	10	8	6	10	10	10	78	78%
Lie down	9	9	10	10	5	10	10	9	9	10	91	91%
Thank you	9	10	5	10	8	4	7	5	8	9	75	75%
Turn left	9	10	7	8	10	8	10	10	10	7	89	89%
Turn right	10	10	10	10	8	10	8	10	10	8	94	94%
Walk	10	10	10	10	10	10	10	10	8	10	98	98%
Follow me	10	9	8	10	9	10	10	9	10	10	95	95%
Kick	9	10	9	9	10	10	10	10	10	10	97	97%
Dance	7	10	9	10	9	10	8	10	10	10	93	93%
Sum	131	135	125	125	127	130	137	133	137	135		Overall
Recognition Rate	87%	90%	83%	83%	85%	87%	91%	89%	91%	90%		87.7%

는데, 이는 깊이 센서로부터 자세를 추정하는 과정<sup>[3]</sup>에서 팔에 해당하는 링크가 카메라의 정면을 향할 때 발생하는 오차에 의한 인식률의 하락이 그 원인이다. 이처럼 자세 추정 기술 측면에서 발생하는 오차를 감안한다면, 제안된 제스처 인식 시스템이 시공간적 변화를 갖는 제스처 인식에 있어서 다른 HMM 기반 제스처 인식 시스템에 견줄 만 한 성능을 보인다는 것을 알 수 있다.

#### 4.4 선행 연구와의 비교 및 분석

Table 5는 앞선 세 실험 결과들을 바탕으로, 제안된 시스템을 이전의 HMM 기반 제스처 인식 시스템들과 종합적으로 비교하여 보여준다. 제안된 시스템은 다양한 사람으로부터 여러 제스처 인식을 연속적으로 수행하였음에도 불구하고 의미 없는 행동을 잘 제외하는 유일한 연구라는 것을 알 수 있다. 실험에 사용된 제스처의 개수도 가장 많았으며, 인식률 계산을 위해 수행된 제스처의 횟수는 1500회로 다른 연구들에 비해 월등히 많았다. 비록 제안된 알고리즘의 인식률이 불연속적인 방법들, 즉 제스처의 분할 과정을 거친 후에 인식을 수행하는 방법들<sup>[4,5,7]</sup>에 비해서는 인식률이 떨어지지만, 연속적인 동작들의 흐름 속에서 제스처를 찾아내어 인식한다는 점에서는 Peng<sup>[6]</sup>의 연구보다 더

Table 5. Overall Comparison with Other Works

	Proposed Algorithm	Yang(2007)	Kim(2007)	Peng(2008)	Peng(2011)
Excluding Non-Gestures	O	Δ	Δ	O	O
Continuous Gesture Recognition	O	X	O	X	X
Number of Gestures (Total Number of Input Gestures)	15 (1500)	10 (500)	8 (480)	11 (330)	11 (539)
Tested by Several Subjects (Number of Subjects)	O (10 subjects)	X	X	O (10 subjects)	O (6-10 subjects)
Gesture Recognition Rates: Recognition After Segmentation	-	97.4%	95.42%	94.59%	-
Gesture Recognition Rates: Spotting from Continuous Data	87.7%	-	-	-	80.15%

좋은 성능을 나타냈다.

## 5. 결론

본 연구에서는 기존의 HMM을 이용한 바디 제스처 인식에서 발생하는 문제점을 해결하고자 기초적 신체 모델 인코딩과 SAI-PSM을 통한 동시다발적 바디 제스처 인식 알고리즘을 제안했다. 신체 모델 각 링크의 동작을 시공간적으로 양자화 및 코드화 하는 기초적 신체 모델 인코딩을 이용해 의미 없는 제스처를 강인하게 제외하였으며, SAI-PSM을 통해 동시다발적인 제스처 인식을 수행하였다. 실험을 통해 본 시스템이 기존의 HMM 기반 제스처 인식에서 벗어난 새로운 형태의 제스처 인식 시스템이며, 성능 면에서도 뛰어난 것을 검증하였다. 추후에, 제스처 인식을 위한 SAI-PSM의 설계과정에서 사용자의 개입 없는 학습 방법에 관한 연구가 필요하다.

## 참고문헌

- [1] V. Ganapathi, C. Plagemann, D. Koller, and S. Thrun, "Real Time Motion Capture Using a Single Time-Of-Flight Camera," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June, 2010.
- [2] T. Gonzalez-Sanchez and D. Puig, "Real-Time Body Gesture Recognition Using Depth Camera," Electronics Letters, vol. 47, no. 12, pp. 697-698, June, 2011.
- [3] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, and M. Finocchio, "Real-Time Human Pose Recognition in Parts from Single Depth Images," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, June, 2011.
- [4] D.H. Kim, J. Song, and D. Kim, "Simultaneous Gesture Segmentation and Recognition Based on Forward Spotting Accumulative HMMs," *Pattern Recognition*, vol. 40, no. 11, pp. 3012-3026, November, 2007.
- [5] B. Peng, G. Qian, and S. Rajko, "View-Invariant Full-Body Gesture Recognition from Video," Proceedings of the International Conference on

- Pattern Recognition, December, 2008.
- [6] B. Peng and G. Quan, "Online Gesture Spotting from Visual Hull Data," *Pattern Analysis and Machine Intelligence*, vol. 33, no. 6, pp. 1175-1188, June, 2011.
  - [7] H.D. Yang, A.Y. Park, and S.W. Lee, "Gesture Spotting and Recognition for Human-Robot Interaction," *Robotics*, vol. 23, no. 2, p. 256-270, April, 2007.
  - [8] S. Mitra and T. Acharya, "Gesture Recognition: A Survey," *Systems, man, and Cybernetics*, vol. 37, no. 3, pp. 311-324, May, 2007.



**김 주 창**

2010 KAIST 전기및전자공학  
과(공학사)  
2012 KAIST 전기및전자공학  
과(공학석사)  
2012~현재 KAIST 전기및전자  
공학과 박사과정

관심분야 : Body Gesture Recognition, Deep Learning



**박 정 우**

2005 경북대학교 전자전기공  
학부(공학사)  
2007 KAIST 전기및전자공학  
과(공학석사)  
2007~현재 KAIST 전기및전자  
공학과 박사과정



**김 우 현**

2007 KAIST 전기및전자공학  
과(공학사)  
2009 KAIST 전기및전자공학  
과(공학석사)  
2009~현재 KAIST 전기및전자  
공학과 박사과정



**이 원 형**

2008 KAIST 전기및전자공학  
과(공학사)  
2010 KAIST 전기및전자공학  
과(공학석사)  
2010~현재 KAIST 전기및전자  
공학과 박사과정



**정 명 진**

1973 서울대학교 공과대학전  
기공학과(공학사)  
1977 미시간대학교 전기공학  
과(공학석사)  
1983 미시간대학교 제어공학  
과(공학박사)

1976 국방과학연구소 연구원  
1981~1983 미시간대학교 CRIM 연구 조교  
1983~현재 한국과학기술원 전기및전자공학과 교수  
관심분야 : Service Robot, Human-Robot Interaction