

논문 2011-48SP-4-4

영상기반의 안정적 수신호 인식을 위한 손동작 패턴 설계 방법

(Hand Motion Design for Performance Enhancement of Vision Based
Hand Signal Recognizer)

손수원*, 배정훈**, 양철중*, 왕한*, 고한석***

(Suwon Shon, Joungoon Beh, Cheoljong Yang, Han Wang, and Hanseok Ko)

요약

본 논문에서는 수신호 인식기에 쓰이기 위한 분별성 있는 손동작을 만드는 방법을 제안한다. 기존의 수화DB에서 손의 움직임을 분석하여 기본 동작이 되는 4가지의 모션 프리미티브를 선정하였으며, 선정된 모션 프리미티브를 조합하여 구별성 있는 '기본 손동작 집합'을 제작하였다. 제안하는 '기본 손동작 집합'의 구별성을 증명하기 위하여 '기본 손동작 집합' 인식기를 만들고 인식결과를 확인하였다. 사용된 인식기는 hidden Markov model (HMM) 을 기반으로 제작되었다. 기본 손동작 인식 task에 대한 성능평가 결과 99.01%로써 각 모델 간에 높은 구별성을 보이는 것을 확인할 수 있었다.

Abstract

This paper proposes a language set of hand motions for enhancing the performance of vision-based hand signal recognizer. Based on the statistical analysis of the angular tendency of hand movements in sign language and the hand motions in practical use, we construct four motion primitives as building blocks for basic hand motions. By combining these motion primitives, we design a discernable 'fundamental hand motion set' toward increasing the hand signal recognition. To demonstrate the validity of proposed designing method, we develop a 'fundamental hand motion set' recognizer based on hidden Markov model (HMM). The recognition system showed 99.01% recognition rate on the proposed language set. This result validates that the proposed language set enhances discernability among the hand motions such that the performance of hand signal recognizer is improved.

Keywords : HMM, hand motion, discernability

I. 서론

컴퓨터와 사람간의 상호작용을 위해 다양한 입력 방법들이 소개되고 개발되어왔다.^[1~4] 그 중 손동작(hand gesture) 기반의 상호작용에 대한 연구가 활발하게 진

행되고 있다. 손동작 기반의 컴퓨터와 사람간의 상호작용을 위한 명령어로서 손쉽게 배우기 적당한 양이면서도 다양한 명령을 내릴 수 있는 적지 않은 개수의 HCI(Human-computer interaction) 명령어 체계가 필요하다. 또한 단순한 고립 손동작 인식이 아닌, 연속 손동작 인식하에서도 높은 인식률을 가질 수 있도록 구별성 있게 설계 되어야 할 것이다.

현재까지 발표된 손동작 인식기는 그 동작의 복잡함에 따라 종류가 나뉘어 질 수 있다. 손가락의 모양과 위치를 특징으로 삼아 그 의미를 인식하게 하는 손모양(static hand posture) 인식기로부터 손모양과 동시에 손의 동작을 인식하는 자동 수화인식기에 이른다. 손모

* 학생회원, *** 정회원, 고려대학교 전기전자전파공학과 (Dept. of Electrical Engineering, Korea University)

** 정회원, University of Maryland Institute for Advanced Computer Studies

※ Authors gratefully acknowledge that this research was support in part by SMG(WR080951) and in part by ONR (N000141010068)

접수일자: 2010년11월24일, 수정완료일: 2011년4월2일

양을 기반한 인식기는 다양한 명령어를 생성하는데 한계가 있다. 손모양은 십여 개의 관절을 통하여 나타나는데, 이 관절이 움직일 수 있는 각도가 제한되어 있고 손가락들 또한 매우 근접하게 붙어있기 때문에 나타낼 수 있는 손모양이 제한적이기 때문이다.^[5~6] 수화 인식의 경우는 반대로 너무 복잡하여 손과 팔 전체를 3D 트래킹 하거나 팔이 occlusion 하지 않아야 하는 전제조건이 필요하다.^[7~8] 또한 수화에 존재하는 모든 단어를 인식하기에 한계가 있으며 명령자가 수화를 익히는데 적지 않은 시간이 필요할 것이다. 여기서 우리가 HCI를 위한 수신호 인식기를 명령 체계의 복잡도 면에서 고려하였을 때, 손모양 인식기와 자동 수화 인식기 사이의 중간 정도의 복잡도가 적합하다. 이는 손쉽게 배우면서도 다양한 명령을 내릴 수 있어야 하기 때문이다. 결과적으로 우리가 고려하는 수신호 인식기는 손모양은 고려하지 않으며, 손이 공간상에서 그리는 궤적을 인식하게 된다.

이러한 손동작 인식관련 연구 분야는 고립 수신호 인식^[9~10], 연속 수신호 인식^[11~12] 등이 있다. 또한 다양한 입력 장치와 카메라, 로봇을 이용한 연구도 진행되어져 왔다.^[1, 4, 13~15] 하지만 열거된 논문에서의 인식기들은 모두 기존에 존재하는 문자나 기호, 혹은 자신들이 임의로 정의한 손동작 집합(gesture set)을 이용하였고 그 종류의 가지 수도 제한적이었다. 즉 미리 정해진 손동작을 인식하되 인식률을 최대로 하는 것이 주된 목적이었다. 수신호 인식기의 성능을 인식률이라고 하였을 때, 본 논문의 목적은 그 성능을 보장할 수 있는 수신호 동작을 설계한다는 점에서 기존 논문과 차별화 된다. 덧붙여, 연속 수신호 인식에 있어 중요한 문제인 비손동작(non-gesture) 부분을 구분하는 문제^[11]를 없애기 위해 모든 수신호의 끝점을 시작점으로 돌아오게 하여 비손동작 부분이 생성되지 않게 설계하는 방법을 제안하였다.

본 논문에서는 Hidden Markov Model (HMM)을 기반으로 인식기를 만들었다. HMM은 시변적인 신호를 통계적으로 모델링 할 수 있어서 음성인식^[16]을 비롯하여 HCI, 및 정보과학 등 다양한 분야에 적용되어져 왔다^[17~19]. 특히 우리는 음성인식의 가변어휘방식^[20]에 착안하여 손동작을 만들었다. 이 방식을 이용하였을 때의 장점은 적은 수의 기본 손동작 패턴만을 가지고 어떠한 동작이든 인식에 사용되는 수신호 모델을 생성해 낼 수 있다는 것이다.

기본 손동작 패턴을 디자인하기 위해서 수화 DB와 기존에 존재하는 수신호 시스템에 대한 분석을 통해 가장 많이 쓰이는 움직임을 알아내었고 모션 프리미티브(motion primitive)라 이름 붙인다. 이렇게 디자인된 기본 손동작은 HMM으로 모델링 된다.

본 논문의 나머지 부분은 다음과 같이 구성된다. II장에서는 본 논문의 실험에 사용된 인식기의 특징 추출 방법과 HMM에 대한 소개가 있다. III장에서는 모션 프리미티브의 정의에 관련된 수화 DB의 분석이 기술되며, IV장에서는 이를 바탕으로 한 손동작 설계 대해서 설명한다. 인식률을 바탕으로 한 성능평가가 V장에서 이루어지며, 마지막으로 VI장에서 결론을 맺는다.

II. HMM 기반 인식기

제안한 수신호 동작은 한쪽 손만으로 이루어진다. 그렇게 정한 이유는 대부분의 의미 있는 손동작들이 한 손으로 이루어지고, 두 손이 같이 움직인다 하더라도 인간의 행동 패턴상, 한 손은 다른 한 손에 mirroring되거나 똑같은 동작을 두 손으로 하기 때문이다^[21]. 따라서 본 논문의 실험에서 사용된 인식기는 특징벡터는 한쪽 손의 이용하여 동작을 취했을 때 시/공간에 나타나는 궤적의 각도와 속도로 구성된다. 특징을 얻는 방법은 웹캠을 이용해 영상을 입력 받은 후, 손이 이루는 픽셀들의 중점을 통해 각 프레임에 따른 손의 좌표들을 얻는다. 그리고 그림 1과 같이 각 좌표의 움직임 벡터들의 각도와 크기를 계산하여 특징을 얻는다.

그림 1에서 A점과 B점이 순서에 따라 관측되면 각이 두 점 사이에서 얻어지는 각도 θ 와 길이 l 을 이용하여 특징벡터 $\mathbf{o} = [\theta \ l]^T$ 나타내며 θ 와 l 을 구하는 식은 다음과 같다.

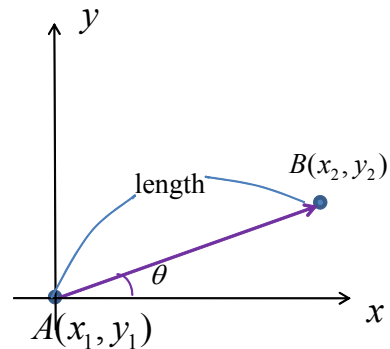


그림 1. 두 좌표 사이의 각도와 크기
Fig 1. Angle and length between the two poin.

$$\theta = \arctan\left(\frac{y_2 - y_1}{x_2 - x_1}\right) \quad (1)$$

$$l = \sqrt{(y_2 - y_1)^2 + (x_2 - x_1)^2} \quad (2)$$

각 특징을 시간 순으로 정렬하여 다음 식 (3)과 같은 M 시간 동안의 관측열(observation sequence)을 얻는다.

$$O = [o_1 o_2 \dots o_M] \quad (3)$$

이러한 특징을 이용하여 우리는 HMM을 통해 각 손 동작에 해당 하는 모델을 얻을 수 있다. HMM의 정의는 다음과 같은 요소들로 구성되어있다.

▶ N : the number of states in the model

$$State S = \{s_1, s_2, \dots, s_N\} \quad (4)$$

▶ M : the number of distinct observation symbols per state

$$Symbols V = \{v_1, v_2, \dots, v_M\} \quad (5)$$

▶ $A = \{a_{ij}\}$: the state transition probability distribution

$$a_{ij} = P[q_{t+1} = s_j | q_t = i], 1 \leq i, j \leq N \quad (6)$$

▶ $B = \{b_j(k)\}$: the observation symbol probability distribution in state j

$$b_j(k) = P[o_t = v_k | q_t = j], 1 \leq k \leq M \quad (7)$$

▶ $\pi = \{\pi_i\}$: the initial state distribution

$$\pi_i = P[q_1 = i], 1 \leq i \leq N \quad (8)$$

위의 파라미터를 이용하여 HMM을 λ 로 놓고, 간단히 $\lambda = \{A, B, \pi\}$ 로 나타낼 수 있고 λ 는 Baum-Welch algorithm^[22]을 통해 추정된다. 즉 추출된 특징 벡터들은 그것의 각도와 길이에 기반을 두어 HMM의 각 state에 속하여 training 될 것이다.

III. 모션 프리미티브

음성인식에서 단어를 인식하기에 앞서 단어를 구성하

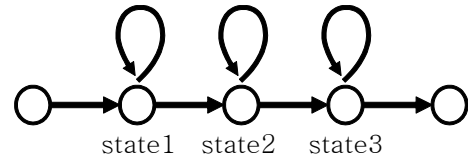


그림 2. HMM 모델

Fig 2. HMM Model.

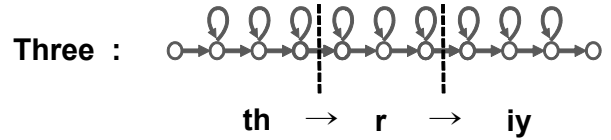


그림 3. 음소의 조합을 이용한 음성인식의 예

Fig 3. Example of speech recognition using phoneme combination.

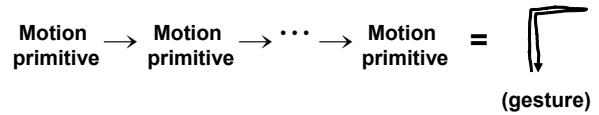


그림 4. 모션 프리미티브의 조합을 통한 손동작 설계의 예

Fig 4. Example of design hand motion with motion primitive combination.

는 음소(phoneme)의 조합을 이용하여 단어를 인식한다. 거꾸로 보면 음성에서 서로 식별 가능한 최소의 단위인 음소의 조합을 통해 단어가 만들어진다고 볼 수 있다. Simple left to right 구조의 HMM을 가정한다면 음성 인식의 각각의 음소는 다음 그림 2와 같은 HMM 모델로 표현될 수 있다.

음성 ‘Three’는 세 가지 음소 th, r, iy로 이루어진다. 세 음소의 HMM을 이어서 그림 3과 같이 음성 ‘Three’에 해당하는 HMM을 만들 수 있고 HMM 기반의 인식기 상에서 확률계산을 통해 인식을 할 수 있다.

손동작 인식분야에서는 높은 인식률을 가지는 고급 손동작 인식기들이 이전 논문에서 제안되었고, 그림 2와 같은 HMM 구조를 통해 수신호 인식기를 구성한 시도가 있어왔다.^[9-10, 23] 본 논문에서도 이러한 음소-단어 HMM 구성 관계를 응용하였다. 수신호에서 음성의 음소에 해당하며 또한 분별성이 큰 모션 프리미티브를 발견하고 그것을 통해 그림 4처럼 모션 프리미티브의 조합을 통하여 더욱 체계적이고 분별성이 큰 수신호를 만들 수 있을 것이다.

수신호의 기본이 되는 모션 프리미티브를 정의하기 위해 사람의 손이 수화나 수신호를 행할 때의 움직임

경로를 분석하였다. Boston university 에서 공개한 미국 수화 (American Sign Language) DB^[24]와 기존의 수신호 시스템들이 분석 대상이다.

1. 미국 수화

미국 수화에 대한 분석은 언어학 연구 목적으로 Boston university 에서 녹화한 공개된 데이터베이스 “RWTH-BOSTON-104”를 이용하였다. “RWTH-BOSTON-104”는 총 201개의 미국 수화 문장으로 구성되어 있고 제공된 양손의 좌표를 이용하여 분석하였다. 시간에 따른 각각의 손의 좌표의 점과 점 사이의 직선이 이루는 각도를 구하여 다음 그림 5와 같은 통계를 얻었다. 이를 통해 0°, 270°, 90°, 180° 를 이루는 점과 점들이 많은 것을 알 수 있었다.

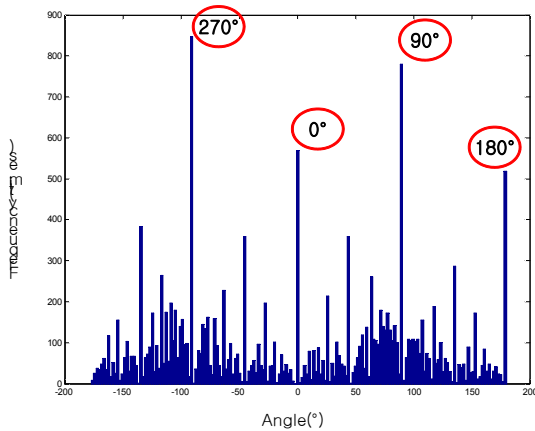


그림 5. “RWTH-BOSTON-104” 데이터베이스의 손 좌표의 각도 통계 결과

Fig 5. Statistical Result of angle in hand position coordinate of “RWTH-BOSTON-104” database

2. 기존의 수신호 시스템

기존에 존재하는 수신호 시스템은 각종 스포츠경기의 공식적으로 정해진 심판의 hand signal과 OSHA (Occupational Safety and Health Administration, U.S. Department of Labor) 에서 지정한 Crane operator 의 수신호 시스템을 이용하였다.^[25] 주로 먼 거리에서 많은 사람들이 어떤 각도에서 보더라도 간단하면서도 명료한 의미를 전달해야 되기 때문에 단순한 동작들로 이루어져 있다. 그 중 특히 구분이 가장 잘되는 수평방향과 수직방향의 신호들이 많이 있다는 것을 알 수 있다. 아래 그림 6 의 각각 축구^[26], 농구^[27], 야구^[28], 하키^[29], crane

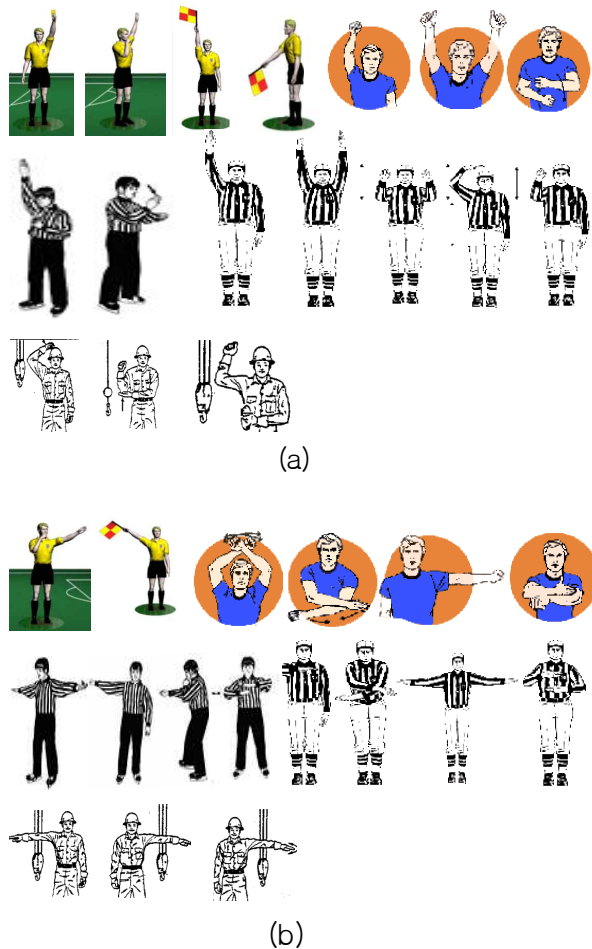


그림 6. 기존 수신호 시스템에서 찾을 수 있는 수신호의 예 (a) 수직방향 (b) 수평방향

Fig 6. Example of hand signal that can be found in existing hand signal system (a) horizontal direction (b) vertical direction

operator의 수신호 시스템에서 찾을 수 있는 수신호 중 (a)는 수직방향을 (b)는 수평방향의 수신호를 나타내고 있다.

또한 위와 같이 공식적인 수신호 시스템은 아닌 지역에 따라 조금씩 차이가 나지만 자전거 주행 시 수신호 등과 같이 다양한 분야의 수신호 시스템에서도 수평 수직방향의 수신호는 중요한 의미를 전달하는 기본적인 의미를 담고 있는 것을 확인할 수 있었다.

3. 네 방향 모션 프리미티브

이러한 분석 결과를 종합하여 손동작의 움직임의 기본이 되는 각도는 수평, 수직 방향의 네 방향 성분으로 나눌 수 있다. 그래서 다음 표 1과 같이 수평, 수직 성분 4개를 모션 프리미티브로 정의하였다.

표 1. 선정된 네 방향 모션 프리미티브
Table 1. Selected four direction motion primitive.

Symbol	Up	Right	Left	Down
Gesture	↑	→	←	↓

IV. 기본 손동작 집합 생성

첫 번째로, 인식이 인식할 손동작 집합을 모션 프리미티브의 조합으로 만든다. 예를 들어 밑의 그림 7에서 왼쪽과 같은 손동작을 만들고 싶다면 모션 프리미티브

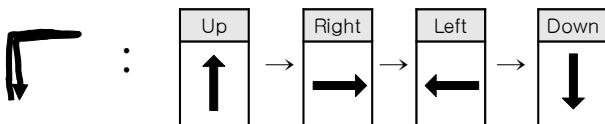


그림 7. 4개의 모션 프리미티브를 이용한 손동작 UR 조합의 예
Fig 7. Example of UR Hand motion combination using 4 motion primitives.

표 2. 모션 프리미티브를 이용하여 만든 12가지 '기본 손동작 집합'
Table 2. 12 class 'fundamental hand motion set' using motion primitives.

Symbol	U	R	L	D
Gesture	↑↓	←→	→←	↓↑
Symbol	UL	UR	DL	DR
Gesture	↙	↘	↖	↗
Symbol	LU	RU	LD	RD
Gesture	↘	↙	↗	↖

를 오른쪽과 같이 4개의 조합을 통하여 만들 수 있으며 손동작의 이름을 UR 이라 하고 이를 수식으로 나타내면 식 (9)와 같다.

$$UR = [UP, RIGHT, LEFT, DOWN] \quad (9)$$

이러한 방식으로 모션 프리미티브를 2개 또는 4개를 조합하여 다음 표 2와 같은 '손동작 집합'을 만들었고, 이것의 인식을 '기본 손동작 집합' 인식기라 한다. '손동작 집합'의 1 부터 4 까지는 모션 프리미티브 2개의 조합으로, 5번부터 12번까지는 모션 프리미티브 4개의 조합으로 만들었다.

이것이 제안하는 기본 손동작으로써 12가지 손동작들의 끝점은 모두 시작점으로 돌아오게 되어있다. 이러한 구조는 연속 손동작 인식에서 중요한 문제점인 동작과 동작 사이의 비손동작 구간부분을 없앨 수 있다.

V. 실험

실험에 쓰인 기본 손동작 집합에 대한 데이터베이스는 28명이 12개의 모델을 3번씩 촬영하여 모았다. '기본 손동작 집합'의 인식률을 최대로 하는 state와 mixture를 찾고 최종적으로 인식률을 평가한다. 실험에 쓰인 인식률을 얻는 방법은 'Leave-one-out cross-validation' 방법을 적용하였다. 27명의 데이터베이스를 훈련용으로 이용하고 남은 1명의 데이터베이스를 인식용으로 쓰는 방법으로 인식용 데이터베이스를 바꿔주며 총 28번의 실험을 통해 인식률을 내어 평균을 얻었다. 즉 실험의 공정성을 위해 다수의 실험을 통해 성능을 검증하였다.

1. 손동작 인식 task를 위한 최적의 mixture, state 평가

각 기본 손동작 집합과 baseline 손동작 집합 데이터베이스의 최적의 mixture와 state를 찾는 실험이다. 첫 번째로 Mixture를 2로 고정하고 state 수의 변화 따른 인식률의 변화를 보았다. 그림 8에서 기본 손동작 집합의 경우 state가 9일 때 98.02% 최대의 인식률을 보였다.

인식률이 최대로 나온 state로 고정시킨 후 mixture를 변화시켜가며 실험해 보았다. 그 결과 그림 9와 같이 기본 손동작 집합은 mixture가 4일 때 99.01%로 최대의 인식률을 보였다.

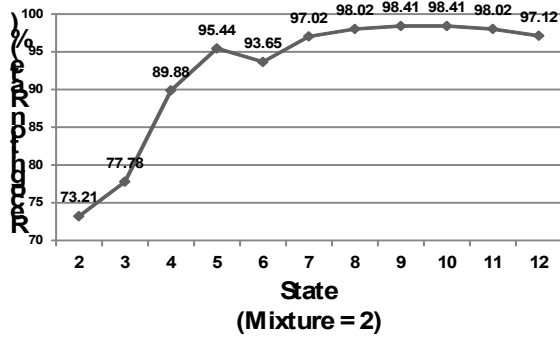


그림 8. '기본 손동작 집합'의 Mixture가 2일 때, state 변화에 따른 인식률

Fig 8. Recognition rate corresponding state variation when mixture is 2. in 'fundamental hand motion set'.

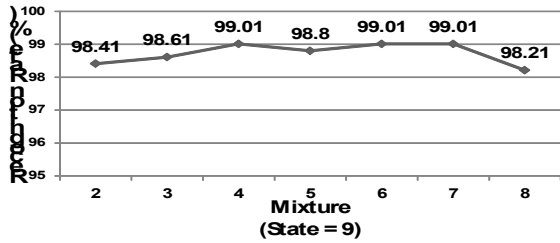


그림 9. '기본 손동작 집합'의 state가 9일 때, Mixture 변화에 따른 인식률

Fig 9. Recognition rate corresponding mixture variation when state is 9 in 'fundamental hand motion set'.

2. 손동작 인식 task에 대한 인식기 성능 평가

모션 프리미티브의 조합으로 만든 '기본 손동작 집합' 인식기의 인식률을 구하여 성능을 평가한다. 실험환경은 위에서 얻은 최적의 state와 mixture 에서의 인식률을 구하였다.

훈련된 모델의 실험상 수치를 살펴보면 위의 실험 1을 통해 N=9, M=4 인 것을 알 수 있었다. A의 예시는 아래 식(10)과 같고(소수점4째자리에서 반올림), B는 실험에 쓰인 특징인 각도 θ 와 길이 l 의 평균과 분산으로 표현된 Gaussian mixture model(GMM)을 각 state 별로 얻을 수 있었다.

$$A_j = \begin{pmatrix} 0.863 & 0.137 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.803 & 0.197 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.814 & 0.186 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.645 & 0.355 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.803 & 0.197 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.687 & 0.313 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.573 & 0.427 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.891 & 0.109 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.432 \end{pmatrix} \quad (10)$$

표 3. 기본 손동작 집합의 인식 결과와 Confusion matrix

Table 3. Recognition rate result of fundamental hand motion set and confusion matrix.

Confusion matrix for fundamental hand motions												
Accuracy %	U	R	L	D	UL	UR	DL	DR	LU	RU	LD	RD
U	98.8	0	0	0	0	0	0	0	1.2	0	0	0
R	0	97.6	0	0	0	0	0	1.2	0	1.2	0	0
L	0	0	98.8	0	0	0	1.2	0	0	0	0	0
D	0	0	0	97.6	0	0	1.2	0	0	0	0	1.2
UL	0	0	0	0	97.6	2.4	0	0	0	0	0	0
UR	0	0	0	0	0	100	0	0	0	0	0	0
DL	0	0	0	0	0	0	100	0	0	0	0	0
DR	0	0	0	0	0	0	0	100	0	0	0	0
LU	0	0	0	0	0	0	0	0	100	0	0	0
RU	0	0	0	0	0	0	0	0	0	100	0	0
LD	0	0	0	0	0	0	0	0	0	0	100	0
RD	0	0	0	0	0	0	0	0	0	2.4	0	97.6
Overall accuracy : 99.01%												

모션 프리미티브의 조합을 통해 정의된 기본 손동작 집합의 12가지 모델에 대한 총 인식률을 99.01%로 각 모델에 대한 인식률은 다음 표 3과 같다.

3. 토의

수신호 인식 task를 위한 최적의 mixture, state 평가 실험을 통하여 '기본 손동작 집합'의 최적의 state는 9개 이고 mixture는 4개 인 것을 확인하였다. 이 결과를 이용해 손동작 인식 task의 각 인식기 성능 평가 실험 환경을 조성하였다.

12개의 모델로 구성되어 있는 '기본 손동작 집합'의 인식률은 99.01%이다. 특정 손동작에 대해 큰 오인식을 보이지 않았고, 모든 손동작에 고르게 평균 0.09%의 낮은 오인식률을 보였으며, 특히 절반에 해당하는 6개의 손동작에 대해서는 100%의 인식률을 보였다. 이를 통해 제안한 모션 프리미티브를 기반으로 조합한 기본 손동작 집합이 구별성을 갖는 것을 확인할 수 있었다.

VI. 결 론

본 논문에서는 컴퓨터와 사람간의 상호작용을 위한 명령어 체계에 적용될 수 있는 복잡한 손동작인 수화와 단순한 손모양 사이의 중간 정도의 복잡도를 가지는 명령어의 중요성을 인지하여, 그것을 음성인식 구조에 적용하여 모션 프리미티브를 제안하였다. 그리고 그에 따른 조합을 통한 손동작 집합의 생성 방법을 모색하여 인식률 실험을 통해 성능을 검증하였다.

그 결과 모션 프리미티브 단위에서 접근하여 ‘기본 손동작 집합’을 만드는 방법이 99.01% 인식률을 보여 큰 구별성을 갖는 것을 검증할 수 있었다. 4가지의 단순한 모션 프리미티브의 조합에 따라 손동작이 생성되더라도 충분히 구별성이 보장되는 손동작 집합을 만들 수 있다는 것을 알 수 있었다.

이러한 구별성이 큰 기본 손동작으로 구성되어 있는 수신호 집합을 생성한다면 그 집합 또한 구별성을 가져 인식률을 유지하면서도 더 많은 수의 다양한 수신호 집합을 생성할 수 있을 것이다. 향후 연구에서는 제한한 모션 프리미티브를 바탕으로 더욱 복잡하고 많은 수의 수신호 집합을 만들고, 그것의 인식률, 즉 구별성을 보장하는 방법을 모색할 것이다.

VII. Acknowledge

Authors gratefully acknowledge that this research was support in part by SMG(WR080951) and in part by ONR (N000141010068)

참 고 문 헌

- [1] Schlömer, T. et al, “Gesture Recognition with a Wii Controller.”, *Proceedings of the Second International Conference on Tangible and Embedded Interaction (TEI’08)*, pp. 11-14 Bonn, Germany. Feb 2008.
- [2] Jacob, R.J.K., “The Use of Eye-Movements in Human-Computer Interaction Techniques - What You Look at Is What You Get.”, *Acm Transactions on Information Systems*, vol. 9, no. 2, pp. 152-169, Apr 1991
- [3] Jia, P., et al., “Head gesture recognition for hands-free control of an intelligent wheelchair.”, *Industrial Robot-an International Journal*, vol 34, no. 1, pp. 60-68, 2007
- [4] Park, H.S., et al., “HMM-based gesture recognition for robot control.”, *Pattern Recognition and Image Analysis, Pt 1, Proceedings*, vol 3522, pp. 607-614, 2005
- [5] Lee, J.T. and T.L. Kunii, “Model-Based Analysis of Hand Posture.”, *IEEE Computer Graphics and Applications*, vol. 15, no. 5, pp. 77-86, Sep 1991.
- [6] James, D. and S. Mubarak, “Recognizing Hand Gestures.”, *European Conference on Computer Vision*, pp. 331-340, Stockholm, Sweden, May 1994.
- [7] Starner, T. and A. Pentland, “Real-time American Sign Language recognition from video using hidden Markov models.”, *International Symposium on Computer Vision*, pp. 265-270, Coral Gables, FL, USA, Nov 1995.
- [8] Christian, V. and M. Dimitris, “Parallel hidden markov models for american sign language recognition.”, *International Conference on Computer Vision*, pp. 116-122, Kerkyra, Greece, Sep 1999.
- [9] Just, A. and S. Marcel, “A comparative study of two state-of-the-art sequence processing techniques for hand gesture recognition.”, *Computer Vision and Image Understanding*, vol. 113, no. 4, pp. 532-543, Apr 2009
- [10] Yanghee, N. and W. KwangYun, “Recognition of Space-Time Hand-Gestures using Hidden Markov Model.”, *ACM symposium on Virtual reality software and Technology*, pp. 51-58, Hong Kong, China, 1996.
- [11] Lee, H.K. and J.H. Kim, “An HMM-based threshold model approach for gesture recognition.”, *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no.10, pp. 961-973, Oct 1999.
- [12] Eickeler, S., A. Kosmala, and G. Rigoll, “Hidden Markov Model based continuous online gesture recognition.”, *Fourteenth International Conference on Pattern Recognition*, vol. 1, no. 2, pp. 1206-1208, Aug 1999
- [13] Kang, H., W.L. Chang, and K.C. Jung, “Recognition-based gesture spotting in video games.”, *Pattern Recognition Letters*, vol. 25, no. 15, pp 1701-1714, 2004.
- [14] Kortenkamp, D., E. Huber, and R.P. Bonasso, “Recognizing and interpreting gestures on a mobile robot.”, *Proceedings of the Thirteenth National Conference on Artificial Intelligence and the Eighth Innovative Applications of Artificial Intelligence Conference*, vol. 1, no. 2, pp. 915-921, 1996.
- [15] 박혜선, 김은이, 김향준, “제스처 인식을 위한 은닉 마르코프 모델”, 대한전자공학회 논문지 제 43권 CI편, 제 1호, 17-26쪽, 2006년.
- [16] Rabiner, L.R., “A Tutorial on Hidden Markov-Models and Selected Applications in Speech Recognition.”, *Proceedings of the Ieee*, vol. 77, no. 2, pp. 257-286, 1989.
- [17] Waldherr, S., R. Romero, and S. Thrun, “A gesture based interface for human-robot interaction.”, *Autonomous Robots*, vol. 9, no. 2,

pp. 151-173, 2000.

[18] Eddy, S.R., "Profile hidden Markov models.", *Bioinformatics*, vol. 14, no. 9, pp755-763, 1998.

[19] Soding, J., "Protein homology detection by HMM-HMM comparison.", *Bioinformatics*, vol. 21, no. 7, pp. 951-960, 2005

[20] Young, S., et al, "The HTK Book (for HTK version 3.4).", <http://htk.eng.cam.ac.uk>, 2009

[21] McNeill, D., "*Hand and mind : what gestures reveal about thought.*", Chicago: University of Chicago Press, pp. 117, 1992.

[22] Baum, L.E., "An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes.", *Inequalities*, vol. 3, no. 9, pp. 1-8, 1972.

[23] Rigoll, G., A. Kosmala, and S. Eickeler, "High Performance Real-Time Gesture Recognition Using Hidden Markov Models.", *Proceedings of the International Gesture Workshop on Gesture and Sign Language in Human-Computer Interaction*, pp. 69-80, Bielefeld, Germany, 1998,

[24] Dreuw, P., et al., "Speech Recognition Techniques for a Sign Language Recognition System.", *Interspeech*, pp. 705-708, Antwerp, Belgium, Aug 2007.

[25] L.Neitzel, R., N.S. Seixas, and K.K. Ren, "A review of crane safety in the construction industry.", *Applied Occupational and Environmental Hygiene*, vol. 16, no. 12, pp 1106-1117, 2001.

[26] "Law 5. The Referee", *Fédération Internationale de Football Association*, www.fifa.com, 2010.

[27] "Official Referee Hand Signals", *National Basketball Association*, www.nba.com, 2010

[28] "NFL Official Rule Book", *National Football League*, www.nfl.com, 2010.

[29] "Official Rules", *National Hockey League*, www.nhl.com, 2010.

저 자 소 개



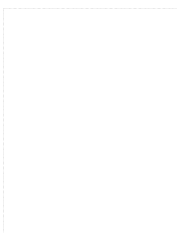
손 수 원(학생회원)
2010년 고려대학교 전기전자전파 공학부 학사 졸업
2010년~현재 고려대학교 전기전자전파공학과 석박사통합과정 재학중

<주관심분야 : 음성처리, 패턴인식>



배 정 훈(정회원)
2001년 고려대학교 전기전자전파 공학부 학사 졸업.
2003년 고려대학교 전자컴퓨터 공학과 석사 졸업.
2008년 고려대학교 전자컴퓨터 공학과 박사 졸업.

<주관심분야 : 음성인식, 빔포밍, 손동작 인식, human-computer interaction(HCI)>



양 철 종(학생회원)
2010년 고려대학교 전기전자전파 공학부 학사 졸업
2010년~현재 고려대학교 영상정보처리협동과정 석박사통합과정 재학

<주관심분야 : 영상처리, 모션 추적, 제스처 인식>



왕 한(학생회원)
2003년 ShenYang Ligong University 전자공학과 학사 졸업.
2007년 Northeastern University 전자공학과 석사 졸업.
2009년~현재 고려대학교 전자전파공학과 박사과정 재학

<주관심분야 : 영상처리, 신호처리, HCI>



고 한 석(정회원)-교신저자
1982년 Carnegie Mellon Univ., 전기공학과 학사 졸업
1986년 Maryland Colleague Park 시스템 공학과 석사졸업
1988년 Johns Hopkins Univ. 전기공학과 석사 졸업
1992년 Catholic Univ. of America 전기공학과 박사 졸업.
1995년~현재 고려대학교 전기전자전파공학부 교수 재직.
<주관심분야 : 음향신호처리, 영상신호처리, 멀티모달리티를 이용한 human-robot-interface>