

Gesture Spotting by Web-Camera in Arbitrary Two Positions and Fuzzy Garbage Model

Seung-Eun Yang[†]

ABSTRACT

Many research of hand gesture recognition based on vision system have been conducted which enable user operate various electronic devices more easily. 3D position calculation and meaningful gesture classification from similar gestures should be executed to recognize hand gesture accurately. A simple and cost effective method of 3D position calculation and gesture spotting (a task to recognize meaningful gesture from other similar meaningless gestures) is described in this paper. 3D position is achieved by calculation of two cameras relative position through pan/tilt module and a marker regardless with the placed position. Fuzzy garbage model is proposed to provide a variable reference value to decide whether the user gesture is the command gesture or not. The reference is achieved from fuzzy command gesture model and fuzzy garbage model which returns the score that shows the degree of belonging to command gesture and garbage gesture respectively. Two-stage user adaptation is proposed that off-line (batch) adaptation for inter-personal difference and on-line (incremental) adaptation for intra-difference to enhance the performance. Experiment is conducted for 5 different users. The recognition rate of command (discriminate command gesture) is more than 95% when only one command like meaningless gesture exists and more than 85% when the command is mixed with many other similar gestures.

Keywords : 3D Position Recognition, Gesture Spot, Fuzzy Garbage Model, User Adaptation

임의 두 지점의 웹 카메라와 퍼지 가비지 모델을 이용한 사용자의 의미 있는 동작 검출

양 승 은[†]

요 약

각종 지능형 전자장비의 개발과 사용자 편의성 증대를 위해 영상기반의 손 동작 인식시스템이 다양하게 개발, 적용되고 있다. 손 동작 인식을 위해 손의 3차원 위치를 계산하고 오 동작 방지를 위해 명령 동작을 다른 유사동작과 구분하여 정확히 검출해야 한다. 본 논문에서는 설치가 쉽고 저렴한 비용으로 3차원 위치를 계산하는 시스템과 다양한 유사 동작 중 정의된 동작만을 검출해 내는 방법에 대해 다룬다. 팬/틸트 가능한 두 대의 USB 카메라와 표식을 이용하여 카메라를 임의의 위치에 두더라도 부착된 표식을 통해 자동으로 두 카메라간 상대위치를 구해 3차원 위치를 계산할 수 있다. 사용자의 명령 동작을 다른 유사 동작과 구분하기 위해 퍼지 가비지 모델을 개발 하였는데 퍼지 명령모델과 가비지 모델 두 가지를 이용하여 행동 인식에 대한 가변적 문턱 값을 구할 수 있다. 또한 두 단계의 적응 과정을 통해 각 사용자마다 다르게 나타나는 행동 특성 및 동일 사용자가 환경에 따라 다르게 나타내는 행동 특성을 반영 하여 성능을 개선한다. 개발된 시스템을 5명의 사용자를 대상으로 실험을 실시하였는데 명령 동작과 하나의 유사동작만 있을 경우 95% 이상, 다양한 유사동작이 혼재되어 있을 경우 85%이상의 인식률(명령 동작 검출)을 보였다.

키워드 : 3차원 위치 인식, 동작 검출, 퍼지 가비지 모델, 사용자 적응

1. 서 론

첨단 기술의 발전과 더불어 사용자 편의를 위한 인간 편의적 인터페이스에 관한 연구가 활발히 이루어지고 있다. 비전 시스템의 경우 사용자가 별도의 장비를 장착할 필요 없이 카메라를 이용하여 영상을 획득한 후 중요한 영상정보를 통해 인터페이스를 제공할 수 있으므로 대표적인 인간

[†] 정 회 원 : 한국항공우주연구원 선임연구원

논문접수: 2012년 7월 25일

수정일: 1차 2012년 9월 3일

심사완료: 2012년 9월 5일

* Corresponding Author : Seung-Eun Yang (seyang@kari.re.kr)



Fig. 1. Three Cameras in the intelligent sweet home

친화적 시스템으로 볼 수 있다[1]. 대부분의 기존 시스템의 경우 인식 대상의 3차원 좌표를 얻기 위해 고정된 위치에 카메라를 설치해 두고 있다. Fig. 1에서는 한국과학기술원의 인간친화 복지 로봇시스템 연구센터(HWRS)에서 개발한 지능형 주거공간의 소프트리모컨 시스템의 예로써 노약자와 장애인을 대상으로 팬/틸트 모듈을 갖춘 3대의 카메라를 이용하여 사용자의 얼굴과 손을 검출하고 손 지시 명령을 통해 가전기기를 제어할 수 있다[2-6].

그러나 Fig. 1 에서 보듯 주거공간에 장착된 카메라는 고정되어 이동이 힘들고 다시 배치 하더라도 초기 설정된 3차원 절대좌표를 다시 측정하여 설정해야 하는 어려움이 있다. 이에 제안하는 방식은 경제적인 USB 카메라를 이용하여 임의의 위치에 있더라도 표식을 이용해 자동으로 카메라간 위치 정보를 구해 상대좌표를 형성할 수 있다. 따라서 저렴한 시스템 구성이 가능하고 위치 계산 및 설정을 위해 전문가가 별도의 장비를 가지고 측정을 할 필요가 없으므로 설치 시간 및 비용을 절약할 수 있다.

정의된 손 명령동작을 인식하여 동작하는 시스템의 경우 그와 유사한 다른 동작을 명령동작으로 오인식 할 수 있고 이를 방지하기 위해 동작을 복잡하게 할 경우 사용자의 편의성을 떨어뜨리게 된다. 명령 동작을 최대한 단순하게 유지하면서 인식률을 높이기 위해 퍼지 가비지 모델을 제안하였다. 제안하는 방법은 인식하고자 하는 손동작에 대한 퍼지 모델과 그와 유사한 동작에 대한 퍼지 가비지 모델을 각각 구현하여 입력되는 사용자의 행동에 대한 각 모델의 출력 값을 비교하여 최종 결정을 내리는 시스템이다. 특히 사람의 행동은 시간 및 사용자에 대해 다양한 특성으로 나타나므로 확률 혹은 통계적 모델로 표현하는 것 보다 행동 특성의 관찰 결과를 쉽게 반영할 수 있는 퍼지 로직이 우수한 성능을 보일 수 있다[7].

본 연구의 목적은 정의된 손 동작을 이와 유사한 동작으로부터 구분하는 것이며 본문의 인식률은 목표 동작의 검출 능력을 의미한다. 정의된 동작으로는 사용자가 손을 아래에서 위로 움직이는 '위' 동작과 우측에서 좌측으로 움직이는 '왼쪽' 동작의 두 가지로 한다. '위' 동작의 경우 식사할 때 음식을 집어 입으로 가져오는 경우, '왼쪽' 동작의 경우 독서

중 책장을 넘기는 과정에서 유사 동작이 나타난다. 따라서 식사 동작과 독서 동작을 각 명령에 대한 가비지 동작으로 정의 한다. 또한 사용자의 얼굴과 손의 경우 붉은색 마스크와 장갑을 이용하여 각각의 영역을 인식하고 정확히 인식된 것으로 가정한다.

제안 방법의 인식률을 높이기 위해 두 단계의 적응 과정을 적용한다. 각 사용자에게 따라 다르게 나타나는 행동 특성을 반영하기 위해 다양한 경우를 고려하는 탐색 능력이 우수한 유전자 알고리즘(GA)을 이용한다. 또한 동일한 사용자라 하더라도 시간에 따라 행동 특성이 달라질 수 있는데 이를 실시간으로 반영하기 위해 최대 경사 법을 이용하여 시스템 적응을 실시한다.

2장에서 표식을 이용하여 두 카메라간 거리 및 각도를 구하고 이를 통해 인식 대상의 3차원 좌표 획득 과정에 대하여 살펴볼 것이다. 3장에서는 퍼지 가비지 모델을 이용한 사용자 행동 인식 방법과 두 단계 적응 과정을 통한 인식률 향상 방법에 대해 다루기로 한다. 4장에서 실제 사용자 데이터를 이용하여 제안한 방법의 시험 결과를 정리하고 5장에서 결론을 통해 본 논문을 마무리 하도록 하겠다.

2. USB 카메라를 이용한 3차원 위치 계산

2.1 표식의 검출

두 카메라가 임의의 위치에 설치 될 경우 인식대상의 3차원 위치를 계산하기 위해 두 카메라간 거리 및 각도를 알아야 한다. 이를 구하기 위해 검출이 용이한 6개의 붉은 원으로 이루어진 표식을 사용한다. Fig. 4의 하단에 두 카메라와 표식 부착 위치를 나타내었는데 초기화를 시작 하면 왼쪽의 카메라가 반 시계 방향으로 회전하면서 표식을 인식한다. 이때 표식의 정보를 이용하여 두 카메라간 거리 및 각도를 구하게 된다.

표식 인식 과정에서 카메라를 통해 얻는 RGB영상은 빛에 대한 정보가 포함되어 있으므로 주위 환경에 영향을 크게 받는다. 따라서 획득된 영상을 밝기 정보와 붉은색과 푸른색의 색상 정보로 구분된 YCrCb로 변경 한 후, 밝기 정보에 해당하는 Y값은 버리고 Cr(붉은색 영역)과 Cb(푸른색 영역)의 값을 비교하여 붉은 영역을 검출한다. 이때 표식 외

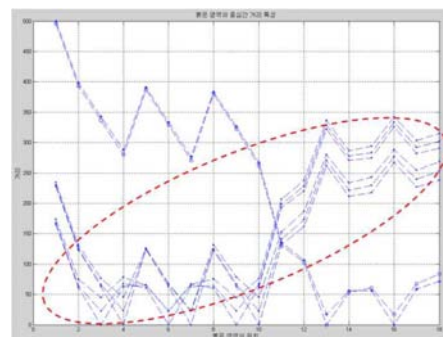


Fig. 2. Distance characteristics of red circles

붉은 영역이 함께 검출 될 수 있다. 이를 제거하기 위해 검출된 영역의 넓이가 극단적으로 큰 것과 작은 것을 제외하고 평균을 구해 이 값으로 넓이가 평균에 가까운 것만 남긴다. 이 과정을 거친 후 각 붉은 영역의 중심간 거리 정보를 확인해 보면 Fig. 2와 같이 표식의 영역과 그 외 영역이 구분 되는 것을 확인할 수 있는데 이러한 특성을 이용하여 최종적으로 표식의 영역만 검출하게 된다.

2.2 두 카메라간 거리와 각도 계산 및 3차원 위치 인식

사용자의 얼굴과 손의 3차원 위치 획득을 위해 두 카메라간 거리와 각도 정보를 알아야 한다. GML MatLab Camera Calibration Toolbox와 MATLAB™ (MATHWORKS, USA)을 이용하여 카메라의 초점거리를 계산하고 Fig. 3에서 두 카메라간 거리 a2를 비례식을 통해 구한다[8].

카메라로부터 표식까지의 거리를 구할 수 있게 되면 표식의 첫 번째 열과 마지막 열의 길이를 구해 Fig. 4 에서와 같이 하나의 삼각형을 형성할 수 있다. 이때 코사인 제2법

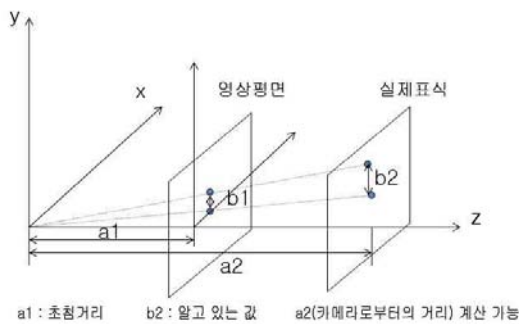


Fig. 3. Distance calculation

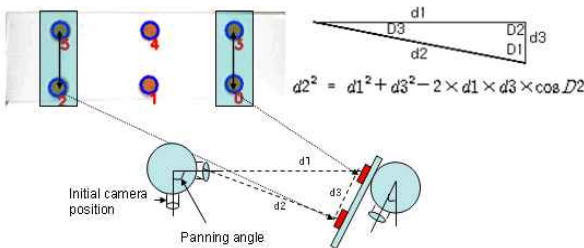


Fig. 4. Angle calculation from the marker

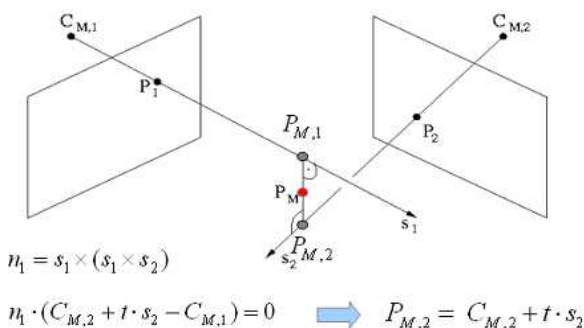


Fig. 5. 3-D position calculation

칙을 이용하면 내각의 크기를 구할 수 있으므로 팬/틸트 변화를 함께 고려하여 두 카메라간의 각도를 구한다.

두 카메라간 거리 및 각도 정보를 이용하여 각 영상에서 나타나는 얼굴 영역과 손 영역의 중점을 절대좌표 기준으로 변경한다[6]. 카메라에서 획득한 영상의 값을 절대좌표 기준으로 변경한 후 Fig. 5의 각 카메라의 중심 $C_{M,1}, C_{M,2}$ 와 손의 중심점 P_1, P_2 을 지나는 선 s_1, s_2 를 구한다. 이때 두 선에 수직인 선분 $\overline{P_{M,1}P_{M,2}}$ 의 중점 P_M 이 구하고자 하는 3차원 좌표의 근사값이 된다[6,8].

3. 퍼지 가비지 모델을 이용한 손 동작 검출

퍼지 가비지 모델을 이용한 사용자의 손 동작 검출 과정은 Fig. 6과 같다. 사용자의 행동 데이터를 이용하여 검출을 위한 특징을 선택한 후 인식 동작과 가비지 동작(명령을 제외한 유사동작)에 대한 2개의 퍼지 모델을 통해 인식을 수행한다. 또한 두 단계 시스템 적응을 통해 인식률을 높이는 과정이 포함되어 있다.

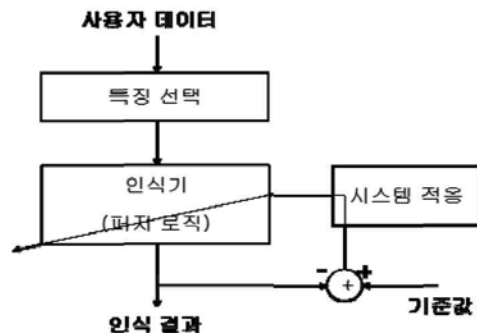


Fig. 6. Recognition system through fuzzy garbage model

3.1 특징 선택

사용자의 의미 있는 동작을 검출하기 위해서는 의미 있는 동작과 이와 유사한 동작을 구분할 수 있는 특징을 찾아야 한다. 우선 관찰을 통해 ‘위’ 명령어와 식사 동작을 구분하기 위한 특징으로 Table 1의 특징 후보를 구한다.

그러나 주어진 모든 특징을 이용하여 행동 인식기를 구현할 경우 정의해야 하는 퍼지 규칙의 수가 많아지고 전체적인 복잡도가 증가하게 된다. 따라서 중요도에 따른 특징을 선정할 필요가 있는데 본 연구에서는 분리도 행렬 (Separability Index Matrix, SIM)과 클래스의 분류능력을 평가하는 클래스 분리도(Classifiability)에 기반한 특징집합 선택 방법인 Separability Index Based Feature Selection Method (SIMF)를 사용하여 이를 수행한다[9].

각 특징의 중요도를 검출 해도 얼마나 많은 수의 특징을 사용해야 하는지에 대한 문제가 남는다. 4가지 퍼지집합을 갖는 소속 함수의 경우 사용하는 특징의 개수 n에 따라 정

Table 1. Feature candidate for gesture recognition

- 손 이동 시 손과 얼굴의 거리 변화
- 명령 동작 후 손과 얼굴의 거리
- 명령 동작 후 손의 정지 시간
- 손의 이동 거리
- 손의 이동 시간
- 손의 이동 속도
- 손 이동 궤적의 이심률
- 손 이동 속도의 평균
- 손 이동의 최대 주파수
- 손 이동 속도의 중앙값
- 손 이동 속도의 표준편차

의되는 퍼지 규칙의 수는 4^n 이 된다. 따라서 특징의 수가 증가하면 복잡도가 급격히 커진다.

적절한 특징의 수를 찾기 위해 사용된 특징 수에 따른 각 클래스간 분리 도를 측정하는데 이용되는 바타차리야 거리 (Bhattacharyya distance)를 구하여 Table 2에 정리하였다. Fig. 7을 보면 사용되는 특징의 수가 많아질 수록 두 클래스간의 분리 도는 높아지지만 정의되는 규칙의 수도 역시 급격히 증가하는 것을 알 수 있다. 이 경우 적절한 규칙의 정의가 어려워지고 전체적인 복잡도가 크게 증가하기 때문에 성능 개선과 복잡도 증가 비율을 고려하여 2개의 특징을 사용하기로 한다. 이때 선택된 특징은 Table 3, 그 분포는 Fig. 8과 같다.

Table 2. Bhattacharyya distance for different feature number

구분	1	2	3	4
명령/식사동작	0.00594	0.00784	0.00897	0.01222

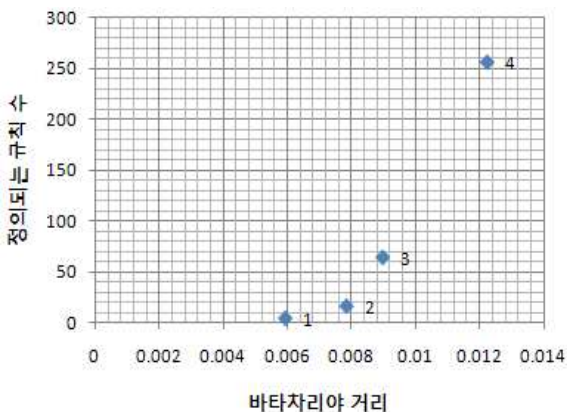


Fig. 7. Performance/cost analysis for feature number

Table 3. Selected feature for command recognition

- 손 이동 시 손과 얼굴의 거리 변화
- 명령 동작 후 손과 얼굴의 거리

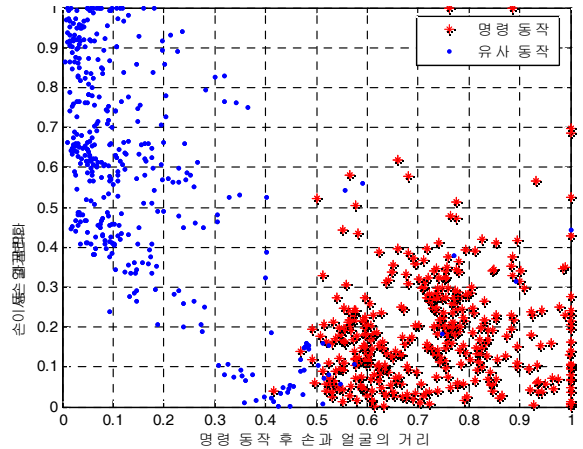


Fig. 8. Feature distribution for command recognition

사용한 특징 중 “손 이동 시 손과 얼굴의 거리 변화”는 사용자가 특정 동작을 취하기 전/후의 손과 얼굴의 거리를 각각 구하여 차를 구한 후 절대치한 것을 의미한다. 또한 “명령 동작 후 손과 얼굴의 거리”는 사용자가 동작을 완료한 후 계산한 손과 얼굴의 거리를 의미한다. 사용자 동작에 대한 데이터는 시작과 종료 시점에 맞추어 수동으로 저장하였다.

선정된 두 개의 특징을 이용하여 명령 동작을 인식하기 위해 Fig. 9와 같이 4가지의 퍼지 집합(ZO: Zero, PS: Positive Small, PM: Positive Medium, PB: Positive Big)을 갖는 소속함수를 사용하고 관찰을 통해 Table 4에 제시된 규칙을 정의한다. 이때 5명의 사용자에게 대해 명령 동작과 유사 동작의 비 퍼지화 값은 Fig. 10과 같다. 비 퍼지화 방식은 합중심법(Center of Sums Method)을 사용한다[10].

Fig. 10 에 도시된 바와 같이 명령 동작과 유사동작에 대한 인식기의 출력 값을 이용하여 사용자의 행동이 무엇인지 결정할 수 있다. 그러나 이 과정에서 하나의 문턱 값을 쓰게 되면 유사동작의 출력 값이 명령 동작의 출력 값보다 큰 경우가 발생 하므로 명령인 동작이 명령이 아닌 것으로, 또는 명령이 아닌 동작이 명령 동작으로 오인식 될 수 있다. 따라서 하나의 문턱 값을 사용함으로써 나타나는 문제를 해결하기 위해 가변적 문턱 값을 제시할 수 있는 퍼지 가비지 모델을 제안한다.

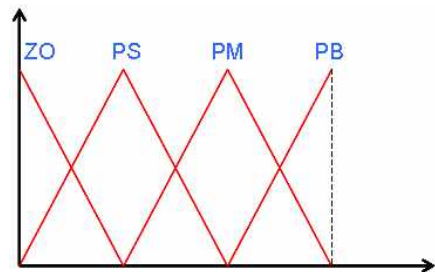


Fig. 9. Membership function for command recognition

Table 4. Fuzzy rules for up command recognition

- If A is ZO and B is ZO then Out is ZO
- If A is ZO and B is PS then Out is ZO
- If A is ZO and B is PM then Out is ZO
- If A is ZO and B is PB then Out is PS
- If A is PS and B is ZO then Out is PS
- If A is PS and B is PS then Out is ZO
- If A is PS and B is PM then Out is ZO
- If A is PS and B is PB then Out is PB
- If A is PM and B is ZO then Out is PB
- If A is PM and B is PS then Out is PM
- If A is PM and B is PM then Out is PB
- If A is PM and B is PB then Out is PS
- If A is PB and B is ZO then Out is PB
- If A is PB and B is PS then Out is PB
- If A is PB and B is PM then Out is PM
- If A is PB and B is PB then Out is PS

A: 손 이동 시 손과 얼굴의 거리 변화
 B: 명령 동작 후 손과 얼굴의 거리

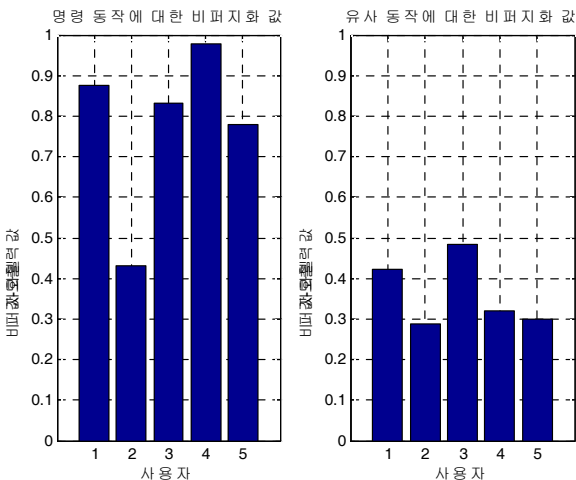


Fig. 10. Defuzzification value for command and garbage

3.2 퍼지 가비지 모델

퍼지 가비지 모델이란 인식하고자 하는 행동을 제외한 유사 동작의 특성을 반영하여 구현한 퍼지 모델을 말한다. 따라서 명령 동작에 대한 출력 값은 작고 유사 동작에 대한 출력 값은 크게 나타난다. 반대로 명령 동작에 대한 퍼지 모델의 경우 명령 동작에 대한 출력 값이 유사 동작에 대한 출력 값보다 크게 나타나게 된다. 따라서 사용자의 행동 데이터를 명령 모델과 가비지 모델에 입력한 후 각 모델의 출력 값을 비교하면 사용자의 행동이 어떤 동작인지 구분할 수 있다. 퍼지 가비지 모델을 이용한 행동 인식 방법을 Fig. 11에 정리하였다. Fig. 12에는 하나의 문턱 값을 사용했을 경우와 퍼지 가비지 모델을 사용한 경우를 도시 하였는데 가변적 문턱 값의 사용으로 오인식이 줄어드는 것을 알 수 있다.

퍼지 가비지 모델을 구현하기 위해 SIMF로 획득한 특징을 이용하여 유사 동작에 대한 규칙을 정의해야 한다. 가비지 동작에 대해 정의한 퍼지 규칙은 Table 5와 같다.



Fig. 11. Gesture recognition using fuzzy garbage model

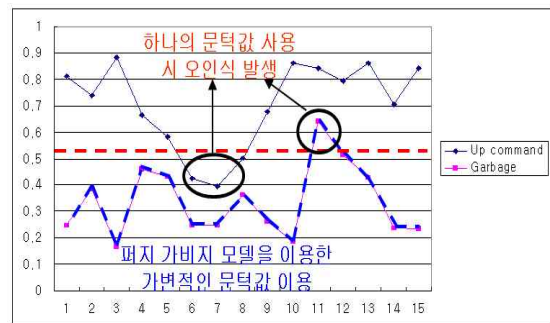


Fig. 12. Gesture recognition using variable threshold

Table 5. Fuzzy rules for garbage gesture recognition

- If A is ZO and B is ZO then Out is ZO
- If A is ZO and B is PS then Out is PM
- If A is ZO and B is PM then Out is PB
- If A is ZO and B is PB then Out is PB
- If A is PS and B is ZO then Out is PM
- If A is PS and B is PS then Out is PM
- If A is PS and B is PM then Out is PM
- If A is PS and B is PB then Out is PM
- If A is PM and B is ZO then Out is PS
- If A is PM and B is PS then Out is PS
- If A is PM and B is PM then Out is PS
- If A is PM and B is PB then Out is ZO
- If A is PB and B is ZO then Out is ZO
- If A is PB and B is PS then Out is PS
- If A is PB and B is PM then Out is ZO
- If A is PB and B is PB then Out is ZO

A: 손 이동 시 손과 얼굴의 거리 변화
 B: 명령 동작 후 손과 얼굴의 거리

두 퍼지 모델을 이용하여 인식 성능을 높이려면 사용하는 소속 함수를 최적화 할 필요가 있다. 사용자의 행동 특성은 다양하게 나타나기 때문에 최적화 값에 대한 탐색 능력이 우수한 GA를 이용한다[11]. 두 개의 입력과 하나의 출력으로 이루어진 퍼지 규칙에 대해 최적화 해야 하는 파라미터는 Table 6에 제시된 바와 같다. w 과 w' 은 소속 함수의 좌, 우측 길이 C 는 소속 함수의 중점을 나타낸다.

Table 6. Fuzzy parameters which will be applied for optimization

	X_1 (특징1)	X_2 (특징2)	Y (특징3)
ZO	w_{r1}	w_{r2}	w_r
PS	c_1, w_{l1}, w_{r1}	c_2, w_{l2}, w_{r2}	c, w_l, w_r
PM	c_1, w_{l1}, w_{r1}	c_2, w_{l2}, w_{r2}	c, w_l, w_r
PB	w_{l1}	w_{l2}	w_l

이때 사용한 최적화 함수는 수식 (1)과 같다. 학습하는 모든 데이터에 대해 올바른 결과로부터 잘못된 결과의 차를 평균한 값을 사용하여 최적화 한다. 이때 GA의 조건은 Table 7과 같으며 최적화된 소속함수는 Fig. 13에 도시하였다.

$$\frac{\sum_{i=1}^{N_A} \{X_A(i) - Y_A(i)\} + \sum_{i=1}^{N_B} \{X_B(i) - Y_B(i)\}}{N_A + N_B} \quad (1)$$

- N_A : 명령 동작의 총 학습 데이터 수
- N_B : 가비지 동작의 총 학습 데이터 수
- X_A : 명령 데이터를 명령 모델에 인가했을 때 출력
- Y_A : 명령 데이터를 가비지 모델에 인가했을 때 출력
- X_B : 가비지 데이터를 명령 모델에 인가했을 때 출력
- Y_B : 가비지 데이터를 가비지 모델에 인가했을 때 출력

Table 7. Condition for GA optimization

내용	사용 값
각 파라미터에 대한 인구	50
각 파라미터에 할당된 bits	20
돌연변이 (Mutation) 비율	0.02
교차 (Crossover) 비율	0.25
총 세대 수	10000

최적화를 위해 5명의 사용자 데이터를 모두 적용 하였으며 Fig. 13의 명령과 가비지 모델에 대해 최적화된 소속함수를 이용하여 인식 실험을 실시한 후 그 결과를 Fig. 14에 도시하였다. 두 그림은 사용자 1부터 5까지 행동 인식 과정에서 나타난 오 인식률을 나타내고 있으며 좌측 그림의 경우 하나의 문턱 값을 사용한 결과이고 우측 그림은 제한한 퍼지 가비지 모델을 이용한 실험 결과이다.

Fig. 14를 보면 퍼지 가비지 모델을 사용한 경우 오차가 하나의 문턱 값을 이용한 경우보다 작게 나타나는 것을 알 수 있다. 그러나 가비지 모델 적용 시 사용자 5의 경우 다른 사용자에 비해 오 인식률이 크게 나타난다. 이렇듯 특정 사용자에게 더 큰 문제가 나타나는 이유는 단일한 시스템을 서로 다른 사용자에게 적용했기 때문이다. Fig. 15의 경우 사용자 5에 대한 특징 분포를 나타내는데 5명의 사용자 모두의 특징 분포를 도시한 Fig. 8과 큰 차이가 나타나는 것을 알 수 있다. 이렇듯 행동 특성은 각 사용자 마다 다르게 나타나

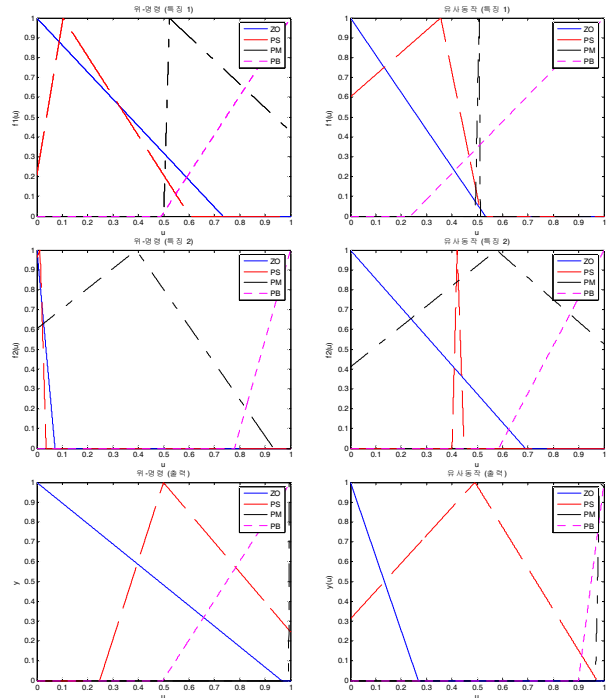


Fig. 13. Optimization using GA (L: Command, R: Garbage)

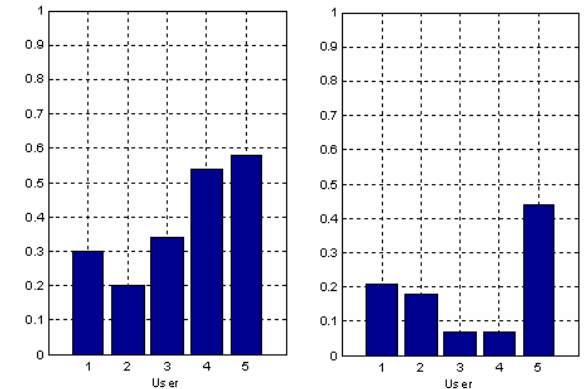


Fig. 14. Wrong recognition rate (L: Threshold value used, R: Fuzzy garbage model used)

기 때문에 이러한 차이점을 고려하여 인식률을 높이기 위해서는 사용자에 대한 시스템의 적응이 필요하다. 다음 장에서 사용자에 대한 시스템 적응 방법을 다루기로 한다.

3.3 두 단계의 시스템 적응

앞에서 살펴 보았듯이 각 사용자에 대해 다르게 나타나는 행동 특성(Inter-Personal Difference)이 인식률을 떨어뜨릴 수 있다. 이 문제를 해결하기 위해 특정 사용자에 대한 적응을 실시해야 한다. 이는 사용자 독립 시스템 (User Independent(UI) System)에서 사용자 의존 시스템 (User Dependant (UD) System) 전환으로 볼 수 있는데 일반적으로 학습 데이터가 충분히 많을 경우 사용자 의존 시스템이 독립 시스템에 비해 우수한 인식 능력을 보인다[12]. 첫 번

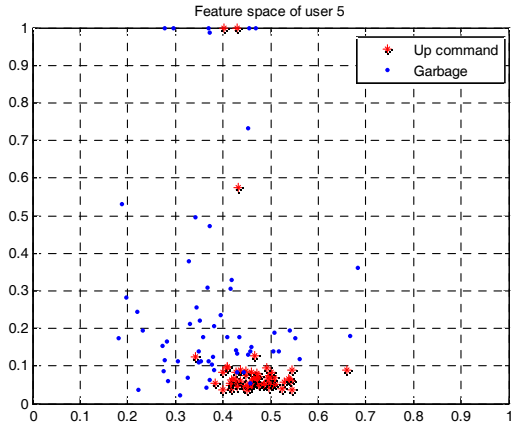


Fig. 15. Feature distribution of 5 users

제 단계의 시스템 적용은 인식기 동작 전에 실시 되므로 오프라인 배치 적용 방식을 사용하며 주어진 조건에서 GA가 우수한 성능을 발휘할 수 있으므로 이 과정에서 GA를 이용한다. GA에서 사용하는 조건은 3-2장의 최적화 과정과 동일하며 각 파라미터에 대한 초기값 역시 이때 얻어진 최적화 값을 사용한다.

다음 단계로 생각해야 할 것은 시간에 따라 변화하는 사용자의 행동 특성(Intra-Person difference)이다. 사람은 기계와 같이 항상 정확하고 동일하게 행동하지 않고 시간 및 환경에 따라 그 특성이 변할 수 있다. 시스템 사용 중 변화하는 행동 특성에 대한 시스템 적용은 실시간으로 수행되어야 하므로 온라인으로 가능한 방법을 이용해야 한다. 최대 경사법의 경우 주어진 함수에 대해 가장 빨리 감소하는 방향을 구할 수 있으므로 짧은 시간에 수렴 값을 구할 수 있다[13]-[15]. 따라서 두 번째 시스템 적용 과정에서는 최대 경사법을 이용하도록 한다.

시스템 적용 방식은 미리 정의된 사용자의 행동을 인식하여 각 파라미터의 갱신 규칙을 통해 파라미터 값을 수정하는 것이다. 그러나 사용자의 행동은 일시적으로 우연히 평소와 다른 특성이 나타날 수 있으므로 변화율을 크게 하면 잘못된 방향으로 적용이 이루어 지게 된다. 따라서 본 과정에서는 주어진 퍼지 규칙 중 해당 입력에 의해 사용된 규칙에 대한 퍼지 파라미터만 갱신하고 변화율을 작게 한다.

사용자의 행동에 대한 모델의 출력 값을 D 라고 할 경우 식(2)와 같이 오차 함수를 구할 수 있고 파라미터 갱신 방법은 식 (3)과 같다.

$$J = \frac{1}{2}(1-D)^2 \tag{2}$$

$$\alpha_{new} = \alpha_{old} - \eta \frac{\partial J}{\partial \alpha_{old}} \tag{3}$$

Fig. 16과 같이 출력에 대한 소속 함수를 나타낼 경우 이는 식 (4)-(6)과 같이 정의된다. 퍼지 출력 값 D 는 파라미터 w_l, w_r, c 로 계산되며 각 파라미터 갱신을 위해서는 J

를 해당 파라미터로 편미분 하면 된다. 따라서 파라미터 갱신 방법은 식(7)과 같이 정의되며 식(8)-(14)에 편 미분 과정을 정리하였다.

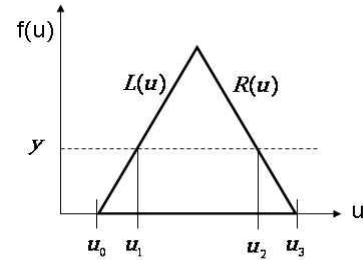


Fig. 16. Membership function for output

$$L(u) = \frac{u}{w_l} + \frac{1}{w_l}(w_l - c) = \frac{u}{c - u_0} - \frac{u_0}{c - u_0} \tag{4}$$

$$R(u) = -\frac{1}{w_r}(u - c) + 1 = \frac{u}{c - u_3} - \frac{u_3}{c - u_3} \tag{5}$$

$$\begin{aligned} u_0 &= c - w_l, & u_1 &= w_l \cdot y + c - w_l, \\ u_2 &= c + w_r - w_r \cdot y, & u_3 &= c + w_r \end{aligned} \tag{6}$$

$$\alpha_{new} = \alpha_{old} - \eta(1-D) \frac{\partial D}{\partial \alpha_{old}} \tag{7}$$

$$\frac{\partial D}{\partial \alpha_{old}} = \frac{\left(\sum_{i=1}^N \frac{\partial}{\partial \alpha_{old}} \int_0^1 u \cdot f_i(u) du \right) \times \left(\sum_{i=1}^N \int_0^1 f_i(u) du \right)}{\left(\sum_{i=1}^N \int_0^1 f_i(u) du \right)^2} - \frac{\left(\sum_{i=1}^N \int_0^1 u \cdot f_i(u) du \right) \times \left(\sum_{i=1}^N \frac{\partial}{\partial \alpha_{old}} \int_0^1 f_i(u) du \right)}{\left(\sum_{i=1}^N \int_0^1 f_i(u) du \right)^2}$$

$$S_u = \int_0^1 u \cdot f_i(u) du, \quad S = \int_0^1 f_i(u) du \tag{8}$$

$$\frac{\partial S_u}{\partial c} = \frac{(u_0 - u_1)^2}{2w_l} + y(u_2 - u_1) + \frac{(u_2 - u_3)^2}{2w_r} \tag{9}$$

$$\begin{aligned} \frac{\partial S_u}{\partial w_l} &= \frac{2u_1^2(y-1) - u_0^2 + u_1^2 - 2u_0u_1(y-1)}{2w_l} \\ &+ \frac{2u_1^3 + u_0^3 - 3u_0u_1^2}{6w_l^2} - y(y-1)u_1 \end{aligned} \tag{10}$$

$$\begin{aligned} \frac{\partial S_u}{\partial w_r} &= \frac{2u_2^2(1-y) - u_2^2 + u_3^2 - 2u_2u_3(1-y)}{2w_r} \\ &- \frac{2u_2^3 + u_3^3 - 3u_2u_3^2}{6w_r^2} + y(1-y)u_2 \end{aligned} \tag{11}$$

$$\frac{\partial S}{\partial c} = 0 \tag{12}$$

$$\frac{\partial S}{\partial w_l} = \frac{(u_1 - u_0)(2yw_l - u_1 + u_0)}{2w_l^2} + y(1 - y) \tag{13}$$

$$\frac{\partial S}{\partial w_r} = -\frac{(u_2 - u_3)(2yw_r + u_2 - u_3)}{2w_r^2} + y(1 - y) \tag{14}$$

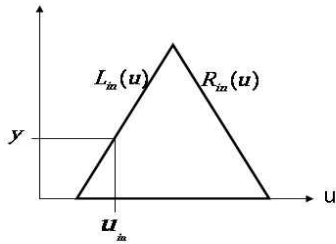


Fig. 17. Membership function for input

또한 입력에 대한 소속함수를 Fig. 17과 같이 나타낼 경우 소속 함수는 식(15), (16)과 같이 각 파라미터로 표현할 수 있다. 입력 소속함수의 경우 함수의 왼쪽이 사용될 경우와 오른쪽이 사용될 경우로 나누어 생각해야 한다. 이에 대해 두 가지 경우로 나누어 각 파라미터 갱신을 위한 편 미분 값을 Table 8에 정리하였다.

$$L_m(u) = \frac{u}{w_{in,l}} + \frac{1}{w_{in,l}}(w_{in,l} - c_{in}) \tag{15}$$

$$R_m(u) = -\frac{1}{w_{in,r}}(u - c_{in}) + 1 \tag{16}$$

온라인과 오프라인으로 이루어지는 시스템 적응 과정을 Fig. 18에 나누어 도시하였다. 위의 경우 특정 사용자에게 대한 적응 과정을 나타내고 아래 그림은 실제 행동 인식 과정에서 시간 및 환경에 따라 변화하는 사용자 특성을 반영한 적응 과정을 나타낸다.

Table 8. Partial derivative for input membership function

경우 1	경우 2
$\frac{\partial S}{\partial c_{in}} = -\frac{1}{w_{in,l}}(w_r + w_l)(1 - y)$	$\frac{\partial S}{\partial c_{in}} = \frac{1}{w_{in,r}}(w_r + w_l)(1 - y)$
$\frac{\partial S}{\partial w_{in,l}} = \frac{c_{in} - u_{in}}{w_{in,l}^2}(w_r + w_l)(1 - y)$	$\frac{\partial S}{\partial w_{in,l}} = 0$
$\frac{\partial S}{\partial w_{in,r}} = 0$	$\frac{\partial S}{\partial w_{in,r}} = -\frac{c_{in} - u_{in}}{w_{in,r}^2}(w_r + w_l)(1 - y)$
$\frac{\partial S_u}{\partial c_{in}} = -\frac{1}{2w_{in,l}}(u_2^2 - u_1^2)$	$\frac{\partial S_u}{\partial c_{in}} = \frac{1}{2w_{in,r}}(u_2^2 - u_1^2)$
$\frac{\partial S_u}{\partial c_{in}} = \frac{c_{in} - u_{in}}{2w_{in,l}^2}(u_2^2 - u_1^2)$	$\frac{\partial S_u}{\partial c_{in}} = 0$
$\frac{\partial S_u}{\partial w_{in,r}} = 0$	$\frac{\partial S_u}{\partial w_{in,r}} = -\frac{c_{in} - u_{in}}{2w_{in,r}^2}(u_2^2 - u_1^2)$

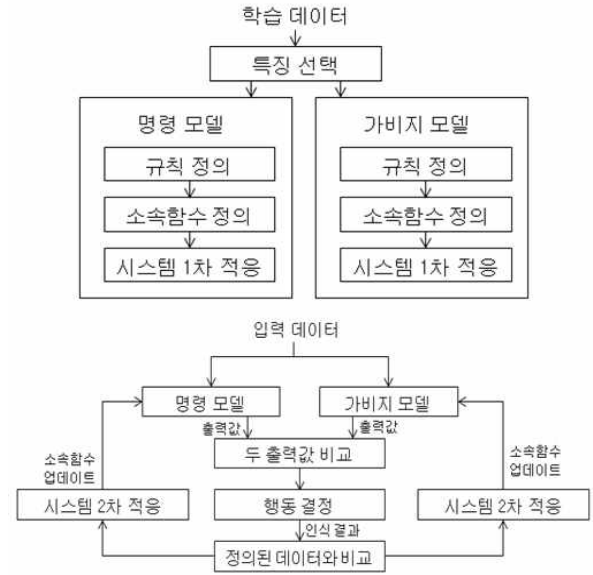


Fig. 18. Two stage user adaptation (Top: Offline-Batch, Bottom: Online-Incremental)

4. 실험 결과

실험을 수행하기 위해 사용자는 마스크와 장갑을 착용해야 하고 시스템에 대한 이해가 필요하기 때문에 20대 남학생 5명으로부터 아래와 같이 6가지 행동 데이터를 수집하였다.

- 명령 동작: ‘위’, ‘왼쪽’,
- 단일 가비지 동작: 식사, 독서 동작
- 다양한 유사 동작: (제약 없이 다음 동작이 포함된 다양한 행동 유도)
 - 손을 위, 아래로 흔드는 동작
 - 손을 좌, 우로 흔드는 동작

각 동작에 대해 75개의 데이터를 획득한 후 25개는 학습에 사용하고 50개는 테스트에 사용하였다. 하나의 문턱 값을 이용한 퍼지 모델, 퍼지 가비지 모델(UI: User Independent), 퍼지 가비지 모델을 각 사용자에게 대해 적응(UD: User Dependent)한 경우로 나누어 실험을 실시하였다. Table 9와 Table 11은 명령 동작과 이와 유사한 동작이 한 가지만 있는 경우, Table 10과 Table 12는 여러 유사동작이 섞여 있는 경우에 대한 결과이다. 왼쪽의 1~5까지는 각각의 피험자를 의미하고 RR은 Recognition Rate, FN은 False Negative error, FP는 False Positive error를 나타낸다.

각각의 실험 결과를 비교해 보면 가비지 모델과 시스템 적응을 수행한 경우 하나의 문턱 값만 사용한 퍼지모델에 비해 인식률이 크게 개선되는 것을 확인할 수 있다. 가변적 문턱 값을 사용하는 가비지 모델에서 명령 동작과 이와 유사동작이 하나만 있는 경우 명령이 아닌 동작을 명령 동작으로 잘못 인식하는 FP 오차가 크게 감소하였다. 이는 유사

Table 9. Up command recognition from eating garbage gesture

	퍼지모델 + 하나의 문턱값			가비지 모델 (가변적 문턱값)			가비지 모델 + 시스템 적용		
	RR	FN	FP	RR	FN	FP	RR	FN	FP
1	91	10	8	98	4	0	100	0	0
2	74	32	20	74	50	2	100	0	0
3	72	36	20	93	8	6	96	0	8
4	92	6	10	100	0	0	100	0	0
5	41	44	74	80	38	2	99	0	2

Table 10. Up command recognition from various garbage gestures

	퍼지모델 + 하나의 문턱값			가비지 모델 (가변적 문턱값)			가비지 모델 + 시스템 적용		
	RR	FN	FP	RR	FN	FP	RR	FN	FP
1	61	16	62	62	64	12	85	16	14
2	64	28	44	62	58	18	88	6	18
3	57	24	62	53	74	20	99	2	0
4	55	34	56	87	4	22	98	2	2
5	62	6	70	82	6	30	94	4	8

Table 11. Left command recognition from reading garbage gesture

	퍼지모델 + 하나의 문턱값			가비지 모델 (가변적 문턱값)			가비지 모델 + 시스템 적용		
	RR	FN	FP	RR	FN	FP	RR	FN	FP
1	49	26	76	95	8	2	97	6	0
2	87	22	4	85	26	4	98	2	2
3	75	36	14	94	10	2	99	2	0
4	78	20	24	97	4	2	100	0	0
5	83	4	30	95	0	10	100	0	0

Table 12. Left command recognition form various garbage gestures

	퍼지모델 + 하나의 문턱값			가비지 모델 (가변적 문턱값)			가비지 모델 + 시스템 적용		
	RR	FN	FP	RR	FN	FP	RR	FN	FP
1	61	48	30	78	8	36	88	2	22
2	56	42	46	58	2	82	96	2	6
3	82	16	20	78	0	44	95	4	6
4	45	36	74	45	82	28	98	4	0
5	45	78	32	90	0	20	97	0	6

한 동작 중에서 의미 있는 동작을 검출하는 능력이 크게 증가한 것을 의미한다.

그러나 명령 동작과 다양한 유사 동작이 섞여 있을 경우 단일 유사동작이 있는 경우에 비해 인식률이 떨어지는데

Fig. 19의 특징 분포를 통해서도 이를 확인할 수 있다. 이는 다양한 유사 동작들 중에서 명령 동작과 구분하기 위해 사용할 수 있는 공통적인 특성을 구하기가 어렵기 때문이다. 따라서 다양한 동작이 섞여 있는 경우에도 의미 있는 동작을 검출하기 위한 보완 방법이 필요하다.

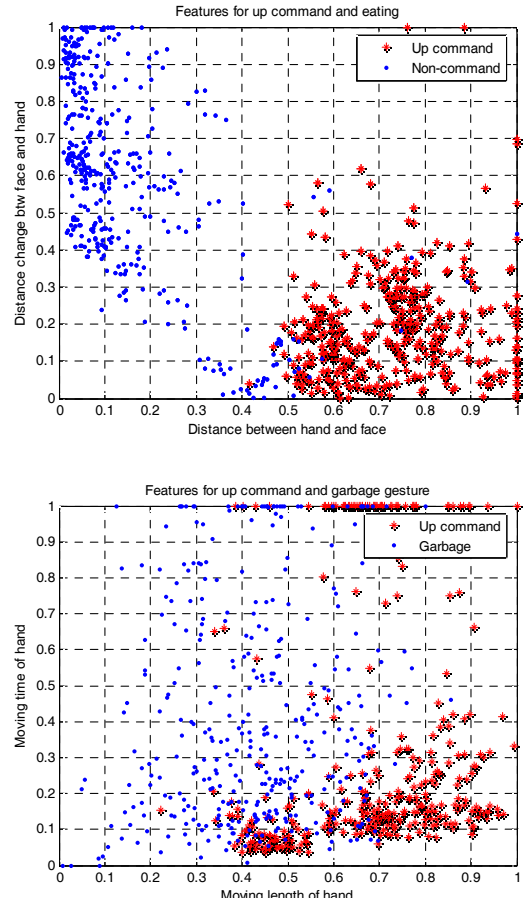


Fig. 19. Feature Distribution of command with one garbage gesture and various garbage gestures

5. 결론 및 추후 연구

본 논문에서는 두 대의 USB 카메라를 임의의 위치에 배치하더라도 사용자 손의 3차원 위치를 인식하는 방법과 정의된 손 동작과 유사동작을 구분하는 방법에 대해 다루었다. 카메라간 거리와 각도 정보를 얻기 위해 붉은색 원으로 이루어진 표식을 사용하였으며 각 영상에서 얻어진 위치정보를 이용하여 3차원 좌표를 계산하였다.

명령 동작과 이와 유사한 동작을 구분하기 위해 퍼지 가비지 모델을 제안하여 가변적 문턱 값을 사용하는 행동 인식 시스템을 개발 하였다. 각 사용자에게 따라 행동 특성이 다르게 나타날 수 있는데 GA를 이용한 사용자 시스템을 실시하여 문제를 해결 하였다. 그리고 동일한 사용자라 하더라도 시간에 따라 그 특성이 달라질 수 있는데 이에 대해서는 최대 경사 법을 이용하여 사용하는 퍼지 파라미터를

갱신하도록 하였다. 제안한 방법에 대한 실험 결과를 4장에서 살펴 보았는데 하나의 문턱 값을 사용하는 경우에 비해 전체적인 인식률이 크게 개선되는 것을 확인할 수 있었다.

그러나 사용자의 명령 동작이 한가지의 유사 동작이 아닌 다양한 동작과 섞여 있을 경우 인식 결과가 나빠지게 되는데 이 경우에도 명령 동작을 잘 검출할 수 있는 보완책이 필요하다. 또한 마스크 및 장갑 착용 등 시험 조건으로 인해 사용자의 행동 특성에 제한이 가해졌을 수 있는데 실험 환경을 개선하고 보다 많은 사용자에 대해 다양한 데이터를 사용하여 시스템 최적화를 할 필요가 있다.

참 고 문 헌

[1] Matthew Turk, "Computer Vision in the Interface," Communications of the ACM-Multimodal interfaces that flex, adapt, and persist, Vol.47, No.1, pp.60-67, 2004.

[2] Jun-Hyeong Do, Jung-Bae Kim, Kwang-Hyun Park, Won-Chul Bang and Z.Zenn Bien, "Soft Remote Control System using Hand Pointing Gesture," International Journal of Human-friendly Welfare Robotic Systems, Vol.3, No.1, pp. 27-30, 2002.

[3] Hyoyoung Jang, Jun-Hyeong Do, Jinwoo Jung, Kwang-Hyun Park, and Z. Zenn Bien, "View-invariant Hand-posture Recognition Method for Soft-Remocon System," in Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2004), Alberat, Canada, 2004, pp.295-300.

[4] Z. Zenn Bien, Kwang-Hyun Park, Jin-Woo Jung and Jun-Hyeong Do, "Intention Reading is Essential in Human-friendly Interfaces for the Elderly and the Handicapped," IEEE Transactions on Industrial Electronics, Vol.52, No.6, pp.1500-1505, 2005.

[5] Dimitar H. Stefanov, Z. Zenn Bien, and Won-Chul Bang, "The Smart House for Older Persons and Persons With Physical Disabilities: Structure, Technology Arrangements, and Perspectives," IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol.12, No.2, pp.228-250, 2004.

[6] Jun-Hyeong Do, Hyoyoung Jang, Sung Hoon Jung, Jinwoo Jung and Zeungnam Bien, "Soft Remote Control System in the Intelligent Sweet Home," in proceeding of the International Conference on Intelligent Robots and Systems, Edmonton, Canada, 2005, pp.2193-2198.

[7] Kevin M. Passino, Stephen Yurkovich, Fuzzy Control, ADDISON-WESLEY, 1998.

[8] M. Kohler, "Vision Based Remote Control in Intelligent Home Environments," in Proceedings of 3D Image Analysis and Synthesis, Erlangen, Germany, 1996, pp.147-154.

[9] 한정수, "새로운 특징 집합 선택 방법과 근전도 신호 인식에의 응용", 박사학위논문, 한국과학기술원, 2006.

[10] 변증남, "퍼지논리 제어", 홍릉과학출판사, 1997.

[11] T.K. Teng, J.S. Shieh, and C.S. Chen, "Genetic algorithms applied in online autotuning PID parameters of a liquid-level control system," Transaction of the Institute of Measurement and Control, Vol.25, No.5, pp.433 - 450, 2003.

[12] Hsin-Chia Fu, Hung-Yuan Chang, Yeong Yuh Xu, and H.-T. Pao, "User Adaptive Handwriting Recognition by Self-Growing Probabilistic Decision-Based Neural Networks," IEEE Transactions on Neural Networks, Vol.11, No.6, pp.1373-1384, 2000.

[13] Ahcene Habbi, Mimoun Zelmat, "An Improved Self-Tuning Mechanism of Fuzzy Control by Gradient Descent Method," in proceeding of the 17th European Simulation Multiconference ESM 2003, Nottingham, England, 2003, pp.43-47.

[14] R.J. Marks II, S. Oh, P. Arabshahi, T.P. Caudell, J.J. Choi and B.G. Song, "Steepest Descent Adaptation of Min-Max Fuzzy If-Then Rules," in proceeding of the IEEE/INNS International Joint Conference on Neural Networks, Beijing, China, 1992, pp.471-477.

[15] Payman Arabshahi, Robert J. Marks, Seho Oh, Thomas P. Caudell, J. J. Choi, and Bong-Gee Song, "Pointer Adaptation and Pruning of Min-Max Fuzzy Inference and Estimation," IEEE Trans. Circuits and systems-II: Analog and Digital Signal Processing, Vol.44, No.9, pp.696-709, 1997.



양 승 은

e-mail : seyang@kari.re.kr
 2005년 서강대학교 전자공학과(학사)
 2007년 한국과학기술원 전자공학과(석사)
 2007년~현 재 한국항공우주연구원
 선임연구원
 관심분야: 패턴 인식, 인공 지능