

소그룹 기반 분류에 의한 손자세 인식에 대한 연구

장효영, 김대진, 김정배, 변중남
한국과학기술원 전자전산학과 전기및전자공학전공

Study on Hand Pose Recognition Using Decomposed Approach with Subgroup-based scheme

Hyoyoung Jang, Dae-Jin Kim, Jung-Bae Kim and Zeungnam Bien
Division of Electrical Engineering,
Dept. of Electrical Engineering and Computer Science, KAIST
E-mail: hyjang@ctrsys.kaist.ac.kr

Abstract

본 논문에서는 손자세 인식을 위해 손 영상을 소그룹으로 나누고 최종적으로 소그룹 내에서 개별 모델로 분류하는 다단계 접근 방식을 취한다. 이 방식은 처음부터 모든 특성치들을 다 구하여 기존에 가지고 있는 모델 모두와 비교하는 대신, 먼저 소그룹으로 분류 후에 해당 소그룹 내의 모델만을 대상으로 비교 연산을 수행한다. 따라서 계산량을 크게 줄일 수 있을 뿐 아니라, 확장이 용이하며, 각 소그룹 별로 특성화된 처리를 할 수 있으므로 효율적인 인식기의 구현이 가능하다.

I. 서론

최근의 컴퓨터 기술 발전과 함께 새로운 사용자 인터페이스 방식에 대한 연구가 활발하다. 그 중 하나로, 비전 기반의 제스처(Gesture) 인식 기술은 근본적으로 별도의 장치에 대한 사용자 접촉 배제를 추구하므로 사용자에게 기존 방식보다 편안한 기기 조작 환경을 제공한다. 또한 제스처에 의한 명령 방식은 일상생활에서도 많이 사용되고 있으므로 기기 조작에 대한 이질감을 줄여주며, 때에 따라 음성 등 다른 의사 소통 수단 이상의 표현력을 가진다. 특히 손 제스처(Hand Gesture)인식 기술은 3 차원 정보 표현에 대해 적합하므로 근래의 가상 현실 관련 연구와 함께 그 수요가 증가하고 있다.

이와 관련한 선행연구로, [1]에서 3 차원 공간 상에서 행해지는 수신호(Hand-signal) 인식을 위한 수신호 데이터베이스(Database) 구성과 수신호 인식 방법에 대해 소개한 바 있다. 기존 방식에서는 템플릿 정합(Template Matching) 결과를 이용한 수신호 데이터베이스와의 비교과정을 통해 대략적인 손의 3 차원 특징 추정치를 얻는다. 이러한 과정은 각 손가락 구분을 위한 특정한 마커(Marker)의 부착 없이 취득된 윤곽선 정보만으로 이루어진다. 그러나 기본적으로 템플릿 정합을 이용하므로 손이 가지는 높은 자유도와 보는 측면에 따른 윤곽선(Silhouette) 변화, 그리고 교합(Occlusion) 현상 등에 대해 취약하고 계산에 비교적 많은 시간을 필요로 한다는 단점을 갖는다.

본 논문에서는 이와 같은 문제점을 해결하기 위한 대안으로 소그룹(Subgroup) 기반 분류에 의한 손자세(Hand Posture) 인식 방법에 대해 다루고자 한다.

먼저 2 장에서 [1]의 시스템에 대해 간략히 소개하고 3 장에서는 소그룹 기반 분류 방식에 대해 다룬다. 이후 4 장에서 실험 결과를 보이고 5 장에서 연구 결과에 대해 결론짓도록 하겠다.

II. 기존 시스템의 구성[1]

대상으로 하는 시스템은 자동차 환경에서 사용되는 수신호인식을 목적으로 한다. 시스템은 크게 세 부분으

로 구성된다. 즉, 두 대의 카메라로 취득한 입력 영상으로부터 손 영역을 검출하고 일차적인 분석을 통해 추정치를 얻어내는 단계, 데이터베이스를 통해 대략적인 손 모델 변수를 추정하는 단계, 그리고 추정된 변수를 세부적으로 조정하고 인식 결과를 출력하는 부분이 그것이다. 그림 1에 시스템 처리 과정을 보였다.

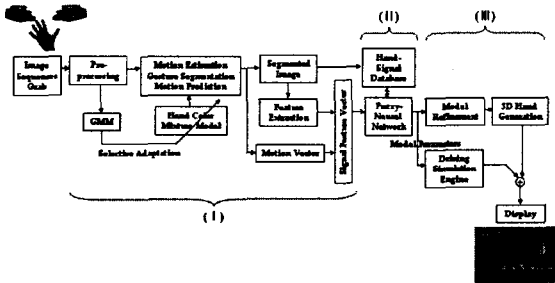


그림 1. 전체 시스템의 처리 과정

((I) 손 영역의 검출과 일차적인 분석, (II) 수신호 데이터베이스를 이용한 모델 변수 추정, (III) 모델 변수 세부 조정 및 인식 결과 출력)

인식 대상이 되는 수신호는 자동차 운전 중 빈번하게 사용되는 수신호들 중 선별된 총 50 종의 수신호이다. 이들은 32 종의 손자세(Hand Pose)와 5 종의 손 궤적(Hand Trajectory)으로 구성된다.

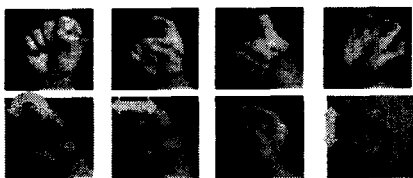


그림 2. 수신호의 예

손 궤적 인식을 위해서는 시간과 함께 공간적인 변화가 있는 패턴의 인식에 적절하다고 일반적으로 알려진 HMM(Hidden Markov Model)을 이용하였으며, 손자세 인식을 위해 데이터베이스를 이용한 대략적 추정치를 이용하는 방식을 제안하였다. 별도의 장갑장치나 마커(Marker) 등을 사용하지 않을 경우에는 손가락의 구분이 어려우므로 인식에 제한이 따르게 된다.

입력된 영상으로부터 추출한 손영역은 조명 변화에 따라 다른 색상분포를 가지게 되며, 대상이 되는 손 자체의 색 또한 사람마다 다르다. 따라서 일반적인 환경

에서 만약 손영역의 추출에 성공하였다고 한다면 손자세를 인식하는 과정에서 손영역의 실루엣이 매우 중요한 역할을 수행하게 된다(조명이 변화하는 환경에서의 손영역 추출에 대해서는 [1]에서 상세히 기술하였으며 여기서는 다루지 않는다.)

그러나 손은 손가락의 자세 변화를 통한 형태의 변형이 다양할 뿐 아니라, 하나의 형태라 하여도 보는 각도에 따라 무수히 많은 형태로 드러나기 때문에 단순히 윤곽선만으로 손가락을 구분 짓는 것은 매우 어려운 문제이다. 이와 같은 문제로 인해, 관련된 많은 연구들은 제한된 환경을 전제로 하거나, 관찰 각도를 고정시킨 매우 적은 수의 손자세만을 인식 대상으로 한다[4, 5, 6].

따라서 이에 대한 해결 방법으로 별도의 수신호 데이터베이스에 대표 손자세와 해당 손자세의 평균 손가락 각도를 값으로 갖는 모델 변수를 정의하여 참조하는 방식을 제안하였다(그림 3).

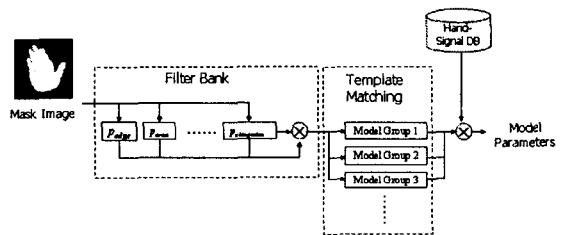


그림 3. 일차적인 모델 변수 추정 과정

우선 입력 영상에서 추출된 손영역을 대표 손자세 32 개와 템플릿 정합하고 그 중 가장 큰 값을 갖는 손자세에 대해 정의된 손자세 변수를 참조한다. 이와 같은 과정이 두 대의 카메라로부터 입력된 영상에 대해 각각 행해지게 되고 최종적으로 추정된 변수에 근거하여 손자세 변수를 세부 조정된 결과를 인식에 이용한다.

III. 소그룹 기반 분류에 의한 확장

II장에서의 방식은 기본적으로 일차 추정을 위해 템플릿 정합을 이용하므로 다음과 같은 두 가지 문제점을 갖는다.

첫째로 템플릿 정합 과정에서 계산적으로 많은 비용이 요구된다. 더욱이 각각의 정합 과정은 손의 회전 가능성을 위해 다중 상관을 요구하며, 하나의 추출 손영역당 기 정의된 손자세 종류만큼의 템플릿 정합을 행

해야 한다.

둘째로 II 장에서도 설명한 바와 같이 손의 윤곽선은 특정 형태로 정의하기가 어렵다. 기본적인 형태의 정의가 어렵고, 더군다나 입력 영상에서의 손의 크기 및 방향(Orientation)이 손이 움직이면서 계속해서 변화하며, 경우에 따라서는 손이 다른 각도에서 영상화되거나 자체 교합이 일어날 수 있다. 이와 같은 경우에 템플릿 정합은 신뢰할 수 없는 결과를 산출하게 된다.

따라서 이 두 가지 문제점의 해결을 위해 소그룹 기반 분류에 의한 계층적으로 손자세를 인식하도록 한다. 이 방식에서는 입력된 손 영상을 소그룹으로 나누고 최종적으로 소그룹 내에서 개별 모델로 분류하는 접근 방식을 취한다(그림 4).

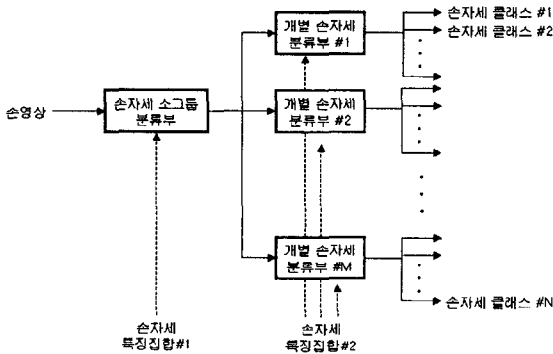


그림 4. 소그룹 분할기와 개별 손자세 분류기

이 방식의 이점은 계산량을 크게 줄이면서도 손자세에 따른 특성화된 인식 과정을 적용 가능하며, 인식 대상 손자세의 확장이 용이하다는 것이다. 즉, 처음부터 모든 특성치들을 다 구해서 기존에 가지고 있는 모델 모두와 정합하는 대신에 먼저 가능성 있는 소그룹으로 분류하고 해당 소그룹 내의 모델만을 대상으로 이후 연산을 수행하므로 연산에 요구되는 시간이 크게 줄어든다. 또한, 가능성 있는 그룹으로 분류하고 다시 그 그룹 내에서 특성화된 처리를 반복하는 계층적인 인식 구조를 통해 각 소그룹에 대해 특성화된 손자세 인식 과정의 적용이 가능하며, 새롭게 인식의 대상이 되는 손자세를 추가할 경우, 인식부에 대해 전체적으로 수정하는 대신에 유사한 소그룹에 대상 손자세를 추가하기만 하면 된다.

IV. 실험

인식 대상으로 하는 손자세는 총 32 개이며, 상호 유사도에 따라 총 12 개의 손자세 소그룹으로 분류된다(그림 5). 두 손자세 영상의 비교를 위해 각 손영역 추출 영상을 손의 장축 방향과 영역 크기에 대해 정규화하여 사용한다. 두 영상 간의 상호 유사도는 (식 1)과 같이 표현한다.

$$S = 1 + \frac{n(I_a \bar{\oplus} I_b) - n(I_a \oplus I_b)}{2n_T}, \quad i = 1, \dots, 12 \quad (\text{식 1})$$

I_a, I_b : 비교하고자 하는 정규화된 손 영상

$n(I)$: 영상 I 의 모든 픽셀 값의 합

n_T : 전체 영상의 픽셀 수

(가로 픽셀 수 × 세로 픽셀 수)

$I_a \bar{\oplus} I_b$: 영상 I_a 와 I_b 의 배타적 논리 부정합

(Exclusive NOR)

$I_a \oplus I_b$: 영상 I_a 와 I_b 의 배타적 논리합

(Exclusive OR)

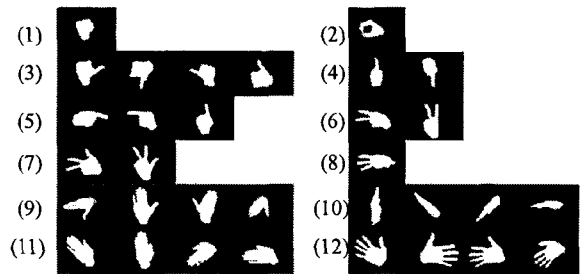


그림 5. 손자세 소그룹

실험에서 사용한 손자세 특성은 장축 방향각(θ), 손영상의 장/단축 길이비(E), 불룩 정도(C)이다[7].

표 1에 각 개별 손자세 추출을 위한 특징값을 나타내었다.

실험을 위해서 32 개 손자세 각각에 대해 5 인의 사람으로부터 각 6 개의 데이터 집합을 얻었다. 즉, 하나의 손자세에 대해 30 개의 데이터가 존재하며 전체 손자세에 대해서는 총 960 개의 데이터를 사용하였다. 이들 중 임의로 선택한 480 개의 데이터를 학습에 사용하고 나머지는 테스트를 위해 사용하였다. 표 2에 실험 결과를 나타내었다.

표 1. 개별 손자세 분류 특징

소그룹	사용된 손자세 특징	손자세 수
(1)	E, C	1
(2)	E, C	1
(3)	θ, E	4
(4)	θ	2
(5)	θ, E	3
(6)	θ	2
(7)	θ	2
(8)	E, C	1
(9)	θ, E	4
(10)	θ	4
(11)	θ	4
(12)	θ, E	4

참고문헌

[1] 장효영, 김대진, 김정배, 변증남, “3 차원 공간상의 수신호 인식 시스템에 대한 연구,” 대한전자공학회 하계종합학술대회 논문집 제 25 권 제 1 호, pp. 215-218, 2002.

[2] G. L. Turbin, An introduction to matched filtering, IRE Trans. Inf. Theory, IT-6, No. 3, pp. 311-329, June 1960.

[3] W. K. Pratt, Digital Image Processing, 2nd ed., Wiley-Interscience, New York, 1991.

[4] R. Hamdan, F. Heitz, and L. Thoraval, Gesture localization and recognition using probabilistic visual learning, CVPR, vol. 2, pp. 98-103, 1999.

[5] M. Kohler, Special topics of gesture recognition applied in intelligent home environments, Proc. Of the Gesture Workshop, pp. 285-296, 1997.

[6] A. Utsumi and J. Ohya, Multiple-hand-gesture tracking using multiple cameras, CVPR, vol. 1, pp.473-478, 1999.

[7] R. C. Gonzalez and R. E. Woods, Digital Image Processing, Addison-Wesley, 1992.

표 2. 개별 손자세 인식률

수신호	인식률(%)	수신호	인식률(%)	수신호	인식률(%)
1	100	12	100	23	100
2	86.67	13	100	24	100
3	93.33	14	100	25	100
4	100	15	100	26	93.33
5	100	16	93.33	27	93.33
6	93.33	17	100	28	100
7	86.67	18	100	29	100
8	93.33	19	100	30	100
9	93.33	20	93.33	31	100
10	86.67	21	100	32	93.33
11	86.67	22	100	-	-
				합 계	96.46

V. 결론 및 추후 과제

본 논문에서는 손자세 인식을 위해 손 영상을 소그룹으로 나누고 최종적으로 소그룹 내에서 개별 모델로 분류하는 다단계 접근 방식을 취한다. 이와 같은 방식은 비교적 간단한 구조로 개별 손자세에 대한 특성화된 처리 및 확장이 용이하다는 장점을 갖는다.

추후 연구의 진행은 크게 손자세 특성의 추가와 각 인식 단계의 최적화를 목적으로 할 예정이다.