

## 3차원 공간상의 수신호 인식 시스템에 대한 연구

장 효 영, 김 대 진, 김 정 배, 변 증 남  
한국과학기술원 전자전산학과 전기및전자공학전공  
전화 : 042-869-5419 / 핸드폰 : 019-298-2745

### A Study on Hand-signal Recognition System in 3-dimensional Space

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@mail.kaist.ac.kr

#### Abstract

Gesture recognitions needed for various applications and is now gaining in importance as one method of enabling natural and intuitive human machine communication. In this paper, we propose a real time hand-signal recognition system in 3-dimensional space performs robust, real-time tracking under varying illumination. As compared with the existing method using classical pattern matching, this system is efficient with respect to speed and also presents more systematic way of defining hand-signals and developing a hand-signal recognition system. In order to verify the proposed method, we developed a virtual driving system operated by hand-signals.

#### I. 서론

일상 생활에서 컴퓨터의 활용도가 높아지고 가상현실에 대한 관심이 높아지면서 이에 따른 새로운 사용자 인터페이스 장치의 연구가 요구되고 있다. 과거 텍스트 기반 컴퓨터 환경에서 그래픽 사용자 인터페이스(Graphic User Interface) 환경으로의 발전 과정에서 기존의 키보드에 더하여 마우스가 주요한 인터페이스

장치로서 부각되었듯이, 3차원 공간상에서 현실 세계를 모사하는 가상 현실 기술의 개발은 이에 따른 새로운 사용자 인터페이스 장치의 개발을 요하는 것이다. 이러한 새로운 인터페이스 장치가 갖추어야 할 특징으로는 3차원 정보의 표현에 대한 적합성과 사용의 편의성, 환경에 대한 강인성, 그리고 충분한 확장 가능성을 들 수 있다.

본 논문은 이와 같은 관점에서 비전 기술을 이용한, 보다 범용적인 수신호 인식 시스템을 제시한다. 다양한 수신호의 정의가 가능하고 일상 생활에서 흔히 접할 수 있는 환경으로서 자동차 내부를 가정하여 전체 시스템의 구현을 보이고자 한다.

#### II. 전체 시스템의 구성

본 논문에서 제시하는 시스템은 두 대의 카메라로 취득한 입력 영상으로부터의 손 영역 검출 및 일차적인 분석을 하는 단계, 데이터베이스를 통한 모델 파라미터(Model Parameter)의 추정 단계, 그리고 추정된 파라미터의 세부 조정 및 인식 결과의 디스플레이 단계로 크게 세 부분으로 구성된다.

그림 1에 전체 시스템의 처리 과정을 보였다.

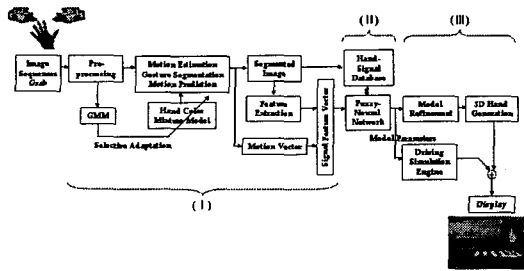


그림 3 전체 시스템의 처리 과정

- (I) 손 영역의 검출과 일차적인 분석
- (II) 수신호 데이터베이스와 모델 파라미터의 추정
- (III) 파라미터 세부 조정 및 인식결과와 디스플레이

### 2.1 손 영역 검출과 일차적인 분석

손 영역 검출 및 기초 분석 단계에서는 크게 3가지 처리 과정을 거친다. 전처리 과정, GMM(Gaussian Mixture Model)의 의해 표현된 손 색상 모델을 근거한 손 영역의 검출, 그리고 움직임 추정 및 제스처 추출(Gesture Segmentation) 과정이 그것이다.

전처리 과정은 RGB 컬러 스페이스를 HSI 컬러 스페이스로 변환하는 과정과 이후 영상처리를 위한 정상화(Normalize) 과정을 포함한다. 손 색상의 분포를 살펴보면 RGB 컬러 스페이스를 이용하는 것보다는 HSI 컬러 스페이스를 이용하는 것이 효과적으로 손 영역을 검출할 수 있다. 이는 손 색상에 대해 각 컬러 성분의 분포를 조사하였을 때 RGB 컬러 스페이스보다 HSI 컬러 스페이스에서 각 성분의 분산이 작기 때문이다.

따라서 손 영역의 검출을 위하여 HSI 컬러 스페이스를 이용하였으며, 조도 변화에 따른 영향을 최소화하기 위하여 이 중 명도에 해당하는 I(Intensity)축을 제외한 H(Hue:색상)와 S(Saturation:채도)만을 사용한다. 그러나 일차적인 후보 영역의 이미지의 경우 급격한 조명 또는 배경 변화가 있을 경우 신뢰도가 적어진다. 따라서 선택적으로 손 색상 정보를 갱신해주는 과정이 필요하다.

이를 위하여 본 논문에서는 GMM을 이용하였다. GMM을 이용한 영역 추출 방법은 기존의 고정된 임계치를 이용하는 방법과는 달리 가우시안 분포를 가지는 컴포넌트들의 합으로써 대상의 색상 분포를 정의한다.

$$p(X|O) = \sum_{j=1}^M p(X|j)\pi(j) \quad (식 1)$$

$\pi(j)$  : 픽셀  $x$  가  $j$ 번째 컴포넌트에 속할 확률

$$p(x|j) = \frac{1}{2\pi |\Sigma_j|^{1/2}} e^{-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1} (x-\mu_j)}$$

이 때 주어진 데이터 집합으로부터  $m$ 개의 가우시안 분포를 따르는 최적의 부분을 찾아주는 알고리즘으로서 EM(Expectation-Maximization) 알고리즘을 사용하였다. 또한 매 프레임마다 색상 정보를 갱신하게 될 경우, 그로 인하여 연산 시간의 지연이 야기되므로, 취득 영상의 색상분포로부터

$$T_t = v - \sigma \quad (식 2)$$

$v$ : 중앙값,  $\sigma$ : 표준편차

와 같이 특정 시간  $t$ 에서의 문턱 값을 계산하여 관측 데이터로부터 얻은 정규화 로그 우도(Normalized log-likelihood) 함수의 값보다 작을 경우에만 색상 정보를 갱신한다. 색상 정보의 갱신을 위하여 이렇게 추출된 손 영역의 이미지에는 많은 노이즈가 포함되어 있어 노이즈를 제거하고, 분리된 영역에 대하여 형태적인(Morphological) 확장(Dilation) 연산이 필요하다. 그림 2에 취득된 영상의 예를 보였다.



그림 2 손 영역 추출의 예

손 영역 추출 과정에서 외부에 특별한 제한을 두지 않는 대신에 손목 부분에 청색의 밴드를 착용하도록 하여 팔 영역으로부터 손 영역을 구분되도록 하였다.

### 2.2 수신호 데이터베이스 및 모델 파라미터 추정

손 영역의 추출 이후 모델 파라미터의 추정을 위한 과정으로서, 손의 실루엣과 그로부터 얻은 기초적인 특성 값들을 이용하여 해당 손동작에 가장 근접한 수신호 후보를 추정하게 된다. 단순한 템플릿 매칭에 앞서서 입력 영상으로부터 추출한 장축과 단축 길이의 비, 장축의 각도 등과 같은, 손 영역에 대한 수치적인 특성을 적용하여 비교의 대상이 되는 모델의 수를 줄였다.

이와 같이 기존의 정의된 수신호 데이터베이스를 참조하여 일차적인 모델 파라미터를 추정하는 과정은 다음의 두 가지 이유에서 필요하다.

첫째로 본 시스템은 손가락의 세부 위치를 추정함에 있어서 위치 판단의 근거가 되는 별도의 마커를 부착하지 않는다[1]. 따라서 윤곽선의 템플릿 매칭을 통하여 대략의 손 모양을 추정하고 그로부터 세부적인 형

태에 대한 대략의 수치 정보를 얻어, 이후 파라미터 조정 과정의 초기 값으로 삼고자 하는 것이다.

둘째로 위와 같이 윤곽선 이미지로부터 얻은 수치정보를 이용하여 현 손동작이 소속 가능한 모델 군을 정하여 그로부터 템플릿 매칭을 수행하고 그로부터 이후 세부적인 수치 조정의 초기 값을 얻음으로 인하여 연산 속도의 증가를 얻을 수 있다.

그림 3에 그 과정을 도식화하였다.

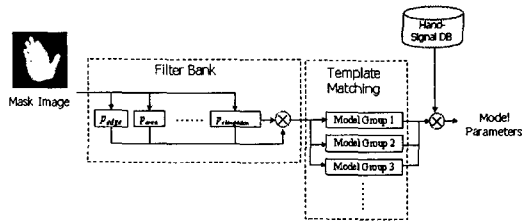


그림 3 일차적인 모델 파라미터 추정 과정

### 2.3 모델 파라미터 세부 조정과 인식 결과의 디스플레이

모델 파라미터의 세부 조정은 이전 단계에서 추정된 손의 개략적인 형태 정보를 초기 데이터로 하고 손동작의 강제 요인을 기반으로 이루어지게 된다. 손동작의 강제 요인이란 각 손동작의 형태를 취함에 있어서 자연스럽게 이루어지는 요소를 나타내는 것으로 다음에 그 예를 보였다[2].

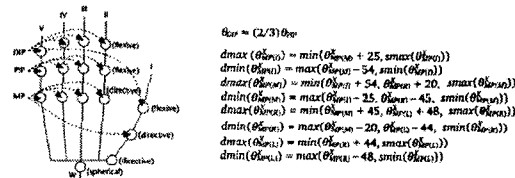


그림 4 손동작과 관련한 제약 조건의 예

이와 같은 제약 조건들을 이용하여 최종적으로 추정하고자 하는 모델 파라미터에 대해 그림 5에 나타내었다.

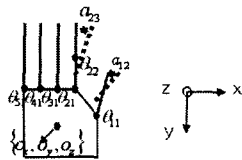


그림 5 모델 파라미터

템플릿 매칭 과정에 의해 추정된 수신호와 세부 조정된 모델 파라미터들을 기초로 하여 입력된 수신호를 인식하는 방법을 통해 조명 및 배경 변화에 대해 보다 강인하게 해당 수신호를 추정 가능할 뿐만 아니라, 이와 같은 방식은 결과적으로 데이터베이스의 형태로 기존의 정의된 수신호 외의 입력에 대해서도 모델 파라미터들을 추정 가능하므로, 이후 다른 수신호 인식 시스템으로의 확장을 가능하게 한다.

## III. 실험

### 3.1 실험 환경

본 논문에서는 제시한 수신호 인식 시스템의 구조를 검증하기 위하여 우선적으로 자동차 환경을 전제로 한 우측 한 손만을 사용하는 것으로 50종의 수신호를 정의하였다[3].

그림 6에 정의된 수신호의 예를 보였다.

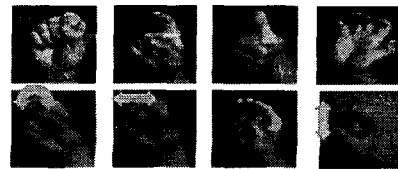


그림 6 정의된 수신호의 예

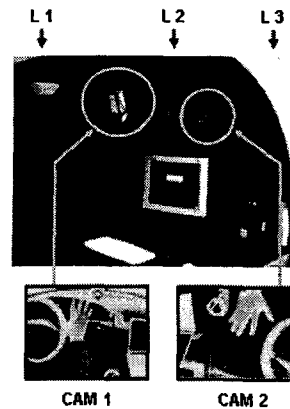


그림 7 실험 환경

그림 7에 실험을 위한 구조물과 카메라 및 조명의 배치를 보였다. 구조물의 크기 및 카메라의 배치는 실제 자동차의 크기 및 구조를 참조하여 결정하였다. 또한 세 개의 조명이 구조물의 후면(L1)과 정면 좌우(L2, L3)에 각각 배치되어 있다. 이들은 실제 자동차 환경

에서 취득한 영상으로부터 손 영역 추출을 어렵게 하는 요소인 조명 변화의 영향을 실험실 환경에서 확인하기 위한 것으로서, 다음의 표 1과 같은 밝기 범위를 보인다.

|              | CAM 1 | CAM 2 |
|--------------|-------|-------|
| 조명 無         | 50    | 50    |
| L1           | 1100  | 100   |
| L2           | 1200  | 70    |
| L3           | 120   | 170   |
| L1 & L2      | 2200  | 300   |
| L1 & L2 & L3 | 2500  | 500   |

표 1 실험환경의 조도

위 표에서 각 수치는 첫째 열에 기입된 조명의 스위치를 최대로 놓고 측정된 밝기의 평균값이다. 또한 둘째 열과 셋째 열은 각기 수신호 영역에서 CAM1과 CAM2를 향해 측정된 조도를 의미한다. 실제 맑은 날 1시~3시 경 일반 도로상 운행중인 자동차 환경에서의 조도를 측정하였을 때 도로 상에서 2000~3000lx 정도의 평균을 보이며, 터널 내에서는 70~270lx, 지하철도에서는 70~700lx의 평균 밝기를 보임에 근거하여, 위 실험 환경 내의 조작 가능한 조도 범위는 타당한 것임을 알 수 있다.

또한 두 대의 CCD 카메라를 동시 사용하여, 보다 복잡한 수신호에 대해서도 인식이 가능하도록 하였다.

### 3.2 실험 결과

그림 8에 입력 영상과 추출된 손 영상의 예를 보았다.

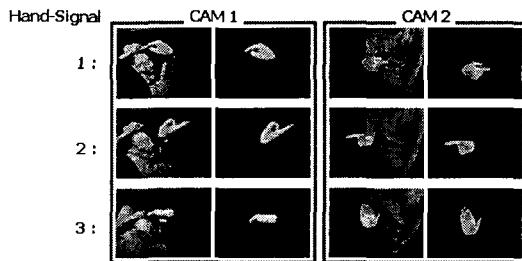


그림 8 두 대의 카메라 입력영상과 손 영역 추출 결과

위 그림으로부터 확인할 수 있는 바와 같이 두 카메라를 통해 입력받은 영상으로부터 추출한 손 영역은 각기 두 카메라 간 상당한 구분성을 보인다. 따라서 단순한 템플릿 매칭 과정과 이후 파라미터의 조정을 통해서 실루엣 이미지로부터 손의 형태 추정이 가능하다.

이렇게 얻어진 손의 파라미터 추정은 그림 9에 보인 것과 같은 형태의 GUI를 통하여 확인할 수 있다.

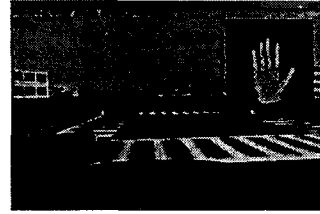


그림 9 결과 확인을 위한 GUI

## IV. 결론

본 논문에서는 높이고 조명 및 배경 환경에 대한 강인성과 응용 분야 확장 가능성을 위한 3차원 공간상의 수신호 인식 방식을 제시한다. 템플릿 정합과 수신호 DB와의 비교 과정을 통해 대략적인 손의 범주가 결정되면, 이후에는 수신호 DB를 참조하여 해당 범주에 대한 3차원 특징값들의 추정값을 얻게 된다. 이러한 과정은 각 손가락을 구분하기 위한 특정한 마커(Marker) 부착 없이 취득된 윤곽선 정보만을 이용하여 이루어진다.

GMM을 통한 손 윤곽선의 추출과 이후 수신호 DB 및 손 모델 개선 과정을 통한 형태 정합과정은 전체적인 시스템의 조명 및 배경 환경에 대한 강인성을 높이고, 결과적으로 수신호 DB에 포함된 50종의 수신호 외의 손동작에 대해서도 형태 추정을 가능하게 하며, 이러한 특성으로 인하여 본 시스템은 추후 다른 응용 분야로의 확장이 기대된다.

### 참고문헌(또는 Reference)

- [1] Iwai, Y., Watanabe, K., Yagi, Y., Yachida, M., "Gesture recognition by using colored gloves", IEEE International Conference on Systems, Man and Cybernetics, 1996, Page(s): 76-81 vol.1
- [2] Jintae Lee, Tosiya L. Kunii, "Model-based analysis of hand posture", IEEE International Conference on Computer Graphics and Applications, 1995, Page(s): 77-86
- [3] 장효영 외, "의사 결정 트리를 이용한 수신호 인식 시스템에 대한 연구", 정보과학회 HCI2002 학술대회