

# 3차원 손 모델을 이용한 비전 기반 손 모양 인식기의 개발

## Development of a Hand-posture Recognition System Using 3D Hand Model

장효영, 변증남

(Hyoyoung Jang, Zeungnam Bien)

**Abstract** - Recent changes to ubiquitous computing requires more natural human-computer(HCI) interfaces that provide high information accessibility. Hand-gesture, i.e., gestures performed by one or two hands, is emerging as a viable technology to complement or replace conventional HCI technology. This paper deals with hand-posture recognition. Hand-posture database construction is important in hand-posture recognition. Human hand is composed of 27 bones and the movement of each joint is modeled by 23 degrees of freedom. Even for the same hand-posture, grabbed images may differ depending on user's characteristic and relative position between the hand and cameras. To solve the difficulty in defining hand-postures and construct database effective in size, we present a method using a 3D hand model. Hand joint angles for each hand-posture and corresponding silhouette images from many viewpoints by projecting the model into image planes are used to construct the database. The proposed method does not require additional equations to define movement constraints of each joint. Also using the method, it is easy to get images of one hand-posture from many viewpoints and distances. Hence it is possible to construct database more precisely and concretely. The validity of the method is evaluated by applying it to the hand-posture recognition system.

**Key Words** : hand-posture recognition, hand-posture database, 3D hand model, human-computer interface

### 1. 장 서론

손 제스처는 일상생활에서 사람들 간에 사용되는 주요 상호 작용 수단으로, 물건을 지시하는 등과 같은 단순한 것에서부터 감정의 표현을 위한 것이나 수화 등에서와 같은 대화를 목적으로 하는 복잡한 행동까지 다양하게 활용되고 있다. 또한 소음이 심한 공사 현장이나, 수중에서는 음성을 대체하는 수단으로써 사용되기도 하며, 음성만으로는 표현에 제한이 존재하는 방향이나 사물의 지시 등과 같이 공간적인 정보 표현을 위한 수단으로 장점을 갖는다. 이와 같은 특징들로 인하여 1990년대 초반 이후 가상현실 분야를 포함한 여러 분야에 대하여 손 제스처를 이용한 사용자 인터페이스를 적용하고자 하는 시도가 계속되어 왔다[1][2][3][4].

이와 같은 손 제스처에 기반을 둔 사용자 인터페이스에서 대개 손 운동은 단순한 형태로 정의된다. 그러나 반면, 손 모양의 경우는 손 운동과 결합하여 동일한 손 운동에 대해서도 명령 간에 구분을 두고 명령을 다양화하기 위해 사용된다.

손 모양을 인식하는 시스템을 개발하기 위해서는 크게 두 가지 문제를 고려해야만 한다. 첫째는 관찰 방향에 독립적인 손 모양 인식 문제이다. 즉, 동일한 손 모양에 대해서는 어떤 방향에서 관찰하더라도 인식 결과가 항상 동일하게 나와야 한다는 것이다. 또 다른 문제는 손 자체의 높은 자유도로 인하여 손의 변형 가능한 형태가 매우 다양하다는 것이다. 이로 인해 사용자 독립적인 손 모양 인식을 위해서는 시스템

구축을 위해 손의 다양한 변형 형태를 모두 고려해야만 할 필요성이 생긴다.

손 모양 인식의 방법은 크게 3차원 모델에 기반을 둔 방법과 형상 특징을 이용한 방법으로 나뉜다.

전자는 3차원 손 모델을 이용하여 손의 구조적인 정보, 즉 각 관절의 꺾임을 예측하고자 하는 방식이다[5][6]. 손의 구조 자체는 관찰 방향에 대해 독립적이므로, 이 방식을 통해 관찰 방향에 독립적인 손 모양 인식을 기대할 수 있다. 그러나 3차원 모델을 각 사용자에 대해 조정되어야 하므로 사용자 독립적인 시스템을 만들기가 쉽지 않다. 또한 손의 구조적 정보를 예측하는 문제는 수학적으로 잘 정립되지 않은(ill-posed) 문제이며, 많은 기존 방법들은 신뢰성 높고 안정적인 특징 추출을 전제로 하나, 실제로는 자체 가림(self-occlusion) 현상으로 인하여 특징 추출이 항상 안정적인 수는 없게 된다.

형상 기반의 손 모양 인식기에서는 손 모양에 대한 다양한 영상을 통해 인식을 학습시킨다[7][8][9]. 이 방식의 경우 모델 기반 방식보다는 사용자 독립적인 시스템 개발이 용이하나, 특징 추출과 데이터베이스 구축에 어려움이 존재한다. 이 범주에 속하는 기존의 연구들은 대부분 특징 추출[8][9] 및 선택[7]에 대해 치중하고 있으며, 데이터베이스 구축에 대해서는 미약하다.

본 논문은 3차원 손 모델을 통한 손의 구조적 정보와 2차원 형상 특징을 동시에 활용하는 손 모양 인식기 개발에 대하여 다룬다.

## 2. 장 3차원 손 모델

데이터베이스 구축 과정에서는 특정 손 모양의 영상 특징 정보와 각도 정보 및 손 모양 식별자가 입력으로 주어진다. 손 모델은 이 때 각 손가락 마디의 각도 정보와 2차원 형상 특징을 동시에 추출하기 위하여 사용된다. 카메라를 이용한 손 모양 데이터 수집의 경우, 손가락의 각도 변경을 정확하게 통제할 수가 없으며 동시에 다양한 측면에서의 데이터를 취득하는데 무리가 따른다. 3차원 손 모델을 이용할 경우에는 이러한 문제들을 손쉽게 해결 가능하며, 보다 정확하고 견고한 데이터베이스 구축이 가능하다. 개발한 3차원 손 모델은 각 손가락 당 4 자유도를 가진다. 그림 1에 사용한 3차원 손 모델을 나타내었으며, 그림 2에 손 모델을 이용하여 표현한 다양한 손 모양의 예를 나타내었다.

인식 대상이 되는 손 모양의 수는 총 32종이며, 이는 각 손가락의 상태를 편 상태와 굽힌 상태의 두 가지로 구분지어 만들 수 있는 총 경우의 수에 해당한다. 이를 위하여 데이터베이스 구축 단계에서는 손가락의 각도를 좀 더 세분화하여 추출한 107개의 손 모양에 대하여 주변의 2507개 시점에서 추출한 손 형상 정보와 동시에 이에 해당하는 손가락 관절각도 정보를 추출한다.

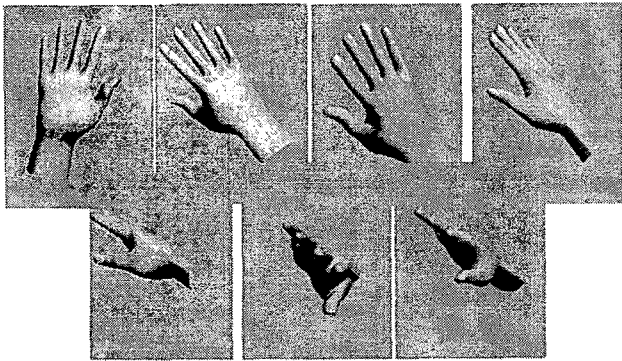


그림 1. 3차원 손 모델



그림 2. 손 모델을 이용한 손 모양 예시

## 3. 장 손 모양 인식기

손 모양 인식기 구축을 위해서는 손 모양의 영상 특징 정보와 각도 정보의 층 구조로 구축되는 데이터베이스와 구축된 데이터베이스 내의 탐색 과정으로 정의되는 OSS-Net(Object Shape and Structure Network)을 이용하였다[10]. OSS-Net의 구조를 그림 3에 나타내었다. OSS-Net에서는 주어진 손 모양 영상 특징 정보와 각도 정보 간 유사도를 정의하는 과정을 장기적인 관점에서 최적인 해를 찾아가는 경로 탐색 과정으로 보고 주어진 데이터 집합에 Q-학습을 적용한다.

특정 손 모양의 영상 특징 정보와 각도 정보 및 손 모양 식별자에 대하여 기존에 데이터베이스 내에 존재하는 영상 특징정보 중 유클리드 거리 측면에서 새롭게 입력된 영상 특징 정보와 가장 유사한 것을 찾는다. 만약 이 둘 사이의 유클리드 거리가 임계치보다 작을 경우, 입력된 영상 특징 벡터  $x_i$ ,  $x_j$ 와 가장 유사한 데이터베이스 내의 기존 영상 특징  $x_k$  그리고 그 다음으로 유사한 기존 영상 특징  $x_{k-1}$ 의 평균  $x_k'$ 로  $x_k$ 를 대체한다. 기존 데이터베이스 내의 영상 특징 중 새롭게 입력된 영상 특징  $x_i$ 와 유사한 것이 없을 경우에는  $x_i$ 가 데이터베이스에 등록된다.

$x_i$ 가 데이터베이스에 등록되거나  $x_k$ 가  $x_k'$ 으로 대체될 경우, 이와 함께 입력된 각도 정보  $y_i$  역시 기존에 데이터베이스에 존재하는 각도 정보와의 비교를 통해서 동일한 과정을 통해 새롭게 등록되거나 기존 데이터베이스 내의 가장 유사한 각도 정보를 갱신하는 데에 쓰인다.

이와 같은 과정에서 동시에 입력된 영상 특징과 각도 정보는 연관성 1로 초기화되고, 기존 데이터가 수정된 경우, 수정된 기존 데이터는 새롭게 입력된 <영상 특징 정보, 각도 정보, 식별자>의 영상 특징 정보 및 각도 정보 간 연관성 및 식별자를 상속하게 된다. 그 결과로서 데이터베이스 내 영상 특징 정보와 각도 정보 간에는 다대다 관계가 형성되게 되며, 각 영상 특징 및 각도 정보에는 그에 연관된 복수개의 손 모양 식별자가 동시에 기록되게 된다. 이와 같이 함으로써, 동일한 손 영상 정보라 하더라도 관측 방향을 달리하여 관측된 서로 다른 손 모양일 수 있음을 반영하게 된다.

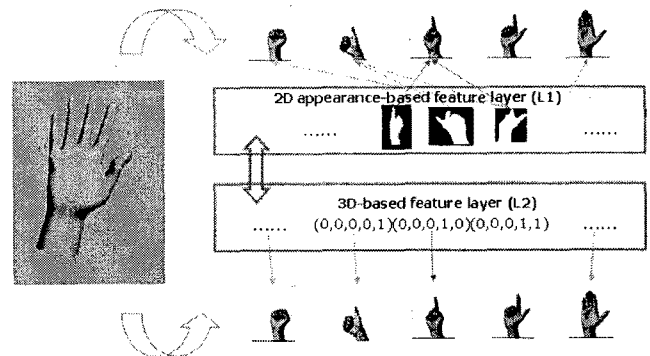


그림 3. OSS-Net의 구조

#### 4. 장 실험 및 결론

학습을 위해서는 107개의 32종의 손 모양 그룹으로 묶고, 이들 각각에 대하여 임의의 100개 시점에서 추출한 영상 정보와 손가락 각도 정보를 이용하였다. 또한 임의로 300개 시점에서 추출한 영상을 통해 테스트를 수행하였다. 인식 결과를 그림 4에 나타내었다.

손은 높은 자유도를 가진 대상으로서 변형 가능성이 커서, 비전에 기반하여 손을 인식하고자 할 경우에 자체 겹침(self-occlusion) 현상과 관찰 시점에 따른 영향을 고려하여야 한다. 본 논문에서는 3차원 손 모델을 이용하여 데이터 획득의 과정을 체계화하였으며, 이를 OSS-Net에 적용하여 손의 형상정보 뿐만이 아니라 관절각도 정보까지도 인식 과정에 활용하는 손 모양 인식기 개발의 방법을 제시하였다. 제안한 방법을 통해 실시간 처리가 가능하고 관찰 시점 변화에 강한 인식기를 설계할 수 있다.

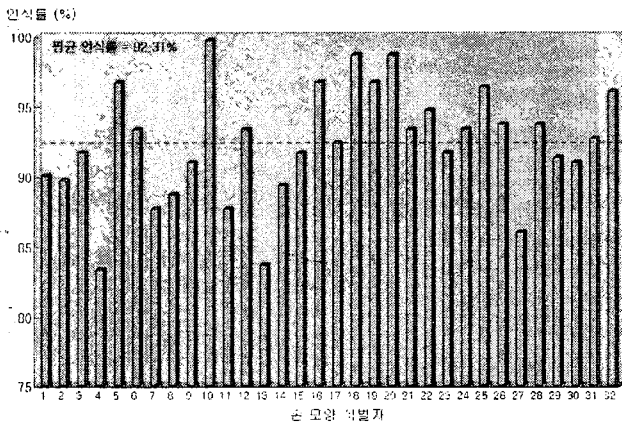


그림 4. 인식 실험 결과

#### 참 고 문 헌

[1] M.B. Waldron, S.W.Kim, "Isolated ASL Sign Recognition System for Deaf Persons", IEEE Transactions on Rehabilitation Engineering, vol. 3, no.3, 1995.

[2] Jung-Bae Kim, Kwang-Hyun Park, Won-Chul Bang and Z. Zenn Bien, "Continuous Gesture Recognition System for Korean Sign Languagebased on Fuzzy Logic and Hidden Markov Model", Proc. of FUZZ-IEEE, 2002

[3] V.I. Pavlovic, R. Sharma, and T.S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review", IEEE PAMI, 19(7):677-695, 1997.

[4] D. Geer, "Will gesture recognition technology point the way?", IEEE Computer, 37(10):20-23, 2004.

[5] C. Nolker, H. Ritter, "Illumination independent recognition of deictic arm postures", Proc. of 24th Annual Conf. of the IEEE Industrial Electronics

Society, Germanym pp.2006-2011, 1998.

[6] N. Sgunada, et al., "Hand gesture estimation and model refinement using monocular camera - ambiguity limitation by inequality constraints", Proc. of the 3rd Conf. on Face and Gesture Recognition, 1998.

[7] Y. Cui, J. Weng, "Hand sign recognition from intensity image sequences with complex background", Proc. of IEEE Conference on Computer Vision and Pattern Recognition, pp.88-93, 1996.

[8] F. Quek and M.Zhao, "Inductive learning in hand pose recognition", IEEE Automatic Face and Gesture Recognition, 1996.

[9] J. Triesch, C.Malsburg, "Robust classification of hand postures against complex background", Proc. of the 2nd Conf. on Face and Gesture Recognition, 1996.

[10] Hyoyoung Jang, Jin-Woo Jung and Zeungnam Bien, "A Study on Vision-based Robust Hand-Posture Recognition by Learning Similarity between Hand-posture and Structure", LNCS 4222 Advances in Natural Computation Part 2, pp. 550-559, 2006.