

# 단일 연속 영상을 이용한 모델 기반의 3 차원 손 인식

백창현, 김진형  
한국과학기술원 전산학과

## Model based 3D Hand Posture Recognition on Monotonic Image Sequence

Chang Hyun Baek, Jin Hyung Kim  
Dept. of Computer Science, KAIST

### 요약 :

3 차원 손모양 인식은 카메라에서 얻어진 손의 영상을 이용하여 이 손이 어떠한 3 차원적 모양을 하고 있는 것인가를 판단하는 것이다. 본 논문에서는 인체의 일부부인 손이 가진 구조적 특성과 영상에서 얻어진 손끝과 손톱의 위치 정보를 결합하여 손의 3 차원적 모양을 인식할 수 있었다. 지금까지의 연구에서 손의 구조적 특성은 손의 3 차원 모양인식에 성공적으로 이용 되어 왔다. 이러한 손의 구조적 특성과 손끝, 손톱의 정보를 결합하여 모든 손가락 관절의 꺾임각을 계산할 수 있었다.

현재의 인식 시스템은 손에 어떠한 표식을 하거나 번거로운 장갑을 이용하는 방식이 아닌, 맨손의 영상을 초당 약 4 프레임의 속도로 인식한다. 또한 여기서 제안된 시스템은 기존의 연구에서는 불가능했던 손가락이 겹칠 때 발생하는 겹침 현상을 손톱을 찾아서 해결하였다.

### 1. 서론

인간의 손을 이용한 입력 수단으로는 장갑장치(Glove Device)를 주로 사용하고 있다. 이러한 장갑 장치는 익숙하지 않은 사용자에게 거부감과 불편함을 느끼게 한다.

최근에는 맨손을 입력수단으로 하고자 하는 연구가 진행되고 있다. 이들 연구는 손의 영상을 카메라로 입력 받아 손의 모양을 인식한다. 손 모양 인식은 영상에 존재하는 손이 어떠한 모양인가를 판단하는 것으로, 얻어진 손영상에서 손의 영상이 미리 정의된 형태 중의 어떤 것인가를 판단하는 손 형태 인식과 손이 어떤 모양을 하고 있는가를 판단하는 손 구조 인식의 2가지로 나눌 수 있다. 손 형태 인식은 손의 위치와 모양을 파악하기 위하여 2차원적인 모델을 이용한다. 이에 반하여 손 구조 인식은 손의 영상에서 손의 각 관절이 가지는 꺾임각을 모두 획득하여 손의 완전한 3차원적인 구조를 재구성하는 것을 목표로 한다.

본 논문에서는 인체가 가지는 손의 구조적 특성을 이용하여 카메라에서 얻어진 손의 연속 영상으로부터 손의 3차원적 구조를 복원하고자 한다. 이를 위해 손이 가진 구조적 특성과 인체가 가진 손의 움직임의 특성을 골격모델에 적용하여 인식에 필요한 정보의 양을 최소화 하면서 손톱과 손끝의 정보만을 이용하여 손 구조 인식을 수행 하였다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 손 인식 방법들과 손의 겹침(Occlusion) 문제들을 해결하고자 했던 시도들을 설명하고 3장에서는 본 논문에서 제안하는 손의 구조적 정보를 이용한 손 구조 인식에 대하여 설명한다. 4장에서는 제안한 손의 인식 방법을 카메라에서 입력 받은 영상에 적용한 실험 결과를 보인다. 끝으로 결론 및 앞으로의 연구 방향을 5장에서 제시한다.

### 2. 연구배경

손 형태 인식은 영상에서 손의 분리와 분리된 손의 형태 판별로 이루어 진다. 손 분리에 중점을 둔 연구로는 Triesch[1]의 연구가 있다. 손의 형태 판별에 중점을 둔 연구로는 Andrew[2]와 Yuntao[3]의 연구가 있다.

손 구조 인식의 목적은 손이 나타내는 3 차원적 정보를 얻고자 하는데 있다. 주로 손이 단일 배경 위에 있다고 가정하고 손의 윤곽을 추출하였다. 손의 윤곽을 이용하여 손의 3차원 적인 정보를 얻는 방법은 손의 구조적인 정보의 이용여부에 따라 2 가지로 나눌 수 있다.

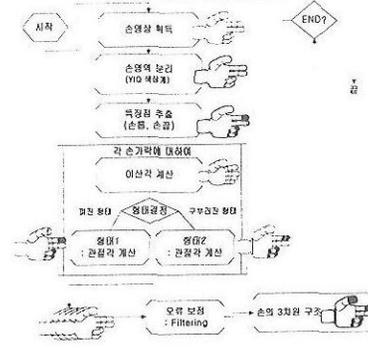
손의 구조적인 정보를 사용하지 않는 연구로는 통계적 모델을 이용한 Tony[5]의 연구를 들 수 있다. 손의 구조적인 정보를 이용한 연구로는 Regh[6]와 Lee[7]의 연구가 있다. Regh는 손의 윤곽선 정보와 마디의 길이 정보를 이용하여 손의 3 차원 정보를 획득하였

고, Lee는 손의 골격 모델을 이용하여 모델과 손영상을 비교하여 손의 모양을 인식하였다. 이때는 추적하고자 하는 손의 구조적 정보(손가락 관절의 길이, 손의 크기)가 미리 알려져 있다고 가정하였다. Lee의 골격모델을 이용한 3 차원 손 모양의 추적은 인체가 가지는 손의 구조적인 제한과 운동정보를 이용한 방법이다. Lee의 방법은 손 모델에서 얻어진 특징과 손 영상에서의 차이를 최소화하는 방법으로 좋은 결과를 얻었지만, 시간을 너무 많이 소모하는 단점이 있었고 맨손이 아닌 표식된 장갑을 착용하여 인식하였다.

아무런 표식이 없는 맨손을 3차원적으로 추적하기 위해서는 손의 영상에서 손가락을 식별해야 한다. 손가락에서 발생하는 겹침 현상은 손가락의 식별을 어렵게 한다. 겹침 현상을 해결하기 위해 Regh[8]는 연속영상에서 손가락의 모양변화를 관찰하여 다음 영상에서의 가려짐을 복원하는 기법을 제안하였다. 그러나 주먹을 쥐고 있을 경우와 같이 손가락의 완전한 겹침 현상이 지속되는 경우는 이러한 방법으로 해결이 불가능했고, 복상도가 높기 때문에 모든 손가락의 겹침 현상을 동시에 극복하기 어려웠다.

### 3. 손 구조 인식 방법의 제안

본 논문에서의 손 모양의 인식은 영상에서 손가락 관절각을 모두 계산하여 손의 3 차원적 구조를 재구성하는 것으로, 이는 손의 길이 정보와 손목의 위치가 미리 등록되고 손바닥의 방향이 카메라를 향하는 환경에서 수행된다. 영상에서 손쉽게 손을 분리하기 위하여 단일 배경에서 인식한다. 본 논문에서는 손의 구조적 정보와 손의 골격 모델을 이용하여 손인식을 수행하는 방법을 다음과 같이 제안한다.

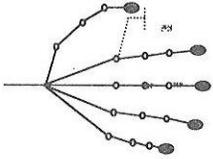


[그림 1 : 인식시스템 전체 구성도]

손의 RGB 영상을 YIQ 색상공간으로 변환하고 1,0 값의 임계치를 이용하여 손 영역을 찾는다. 얻어진 손영역에서 손끝과 손톱의 위치를 추출한다. 각 손가락의 특징점이 추출되면 이들 손가락의 관절각을 계산한다. 손 영상에서 모든 관절의 각이 계산되면 이전 영상에서의 인식 결과를 참조하여 filtering을 통해 오류를 보정한다.

3.1 손의 골격 모델

손의 구조적 특성을 이용하기 위한 모델로 손의 골격구조를 본 반 골격 모델을 이용한다 이 골격모델은 Lee[7]의 연구에 사용된 모델로 손의 골격구조와 동일하다.

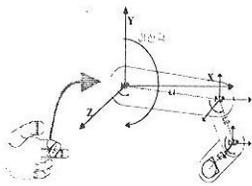


손가락 하나에는 3개의 관절이 존재하며 각 관절은 각각 의학 용어의 약자인 MP,PIP,DIP로 표기된다.

[그림 2 : 손의 골격 모델]

3.2 손의 구조적 특성.

손은 인체의 일부분이므로 움직임에 제한을 가지고 있다. 이러한 움직임 제한에는 손가락 관절의 꺾임과 이산의 가능 여부와 꺾임각과 이산각의 범위, 꺾임과 이산이 일어나는 방향이 있다[Lee7].



손가락에서는 접힘과 벌어짐의 2가지 움직임이 가능하다. 손가락이 접혀질 때 나타나는 관절의 꺾임은 MP,DIP,PIP의 모든 관절에서 발생한다. 벌어짐은 MP에서 발생하며, 이때의 손가락은 중지를 기준으로 산개하면서 벌어진다. 이산각(Abduction Angle)은 손가락이 벌어진 각을 의미한다.

[그림 3 : 손가락의 움직임]

손가락 관절은 인체의 한 부분으로 그 움직임에는 한계가 있다. MP, PIP,DIP의 각 관절이 변화할 수 있는 각의 범위는 표 1과 같다.

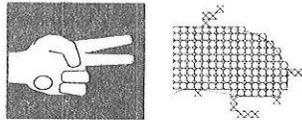
	엄지계의		엄지	
	Min	Max	Min	Max
MP	0	70	30	45
PIP	0	120	0	40
DIP	0	70	0	80
이산각	0	20	0	30

[ 표 1 : 손가락 관절각의 범위 ]

또한 손가락의 근육 구조에 의해 DIP와 PIP의 관절의 꺾임각은 2/3의 비율을 가진다.

3.3 손 영역 분리

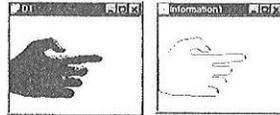
손 영역 분리는 손의 색상을 이용한다. 본 논문에서는 영상의 RGB 색상을 YIQ 색상으로 변환[4]하여 I,Q 값의 범위를 지정하여 손 영역으로 추출한다. I,Q 값의 범위를 지정하여 추출된 손 영역에는 잘못 추출된 픽셀들도 포함된다. 따라서 이들 추출된 픽셀들을 제거해야 정확한 손의 윤곽(Contour)을 얻을 수 있다. 손이 아닌 영역으로 잘못 판단 되는 내부 픽셀 에러는 손의 외곽지역을 하나의 연결 부분(Connected Component)으로 탐색하여 손의 외곽지역으로 찾아진 부분을 제외한 나머지 부분을 모두 손 영역으로 판단하여 제거한다. 손으로 잘못 판단 되는 외부 픽셀 에러는 각 픽셀별로 이웃한 손영역 픽셀과의 연결도(Connectivity)를 이용하여 제거 여부를 판단한다.



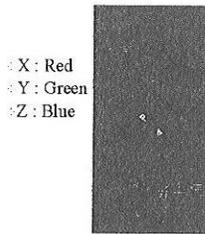
[그림 4 : 내부, 외부 픽셀 에러 제거]

3.4 특징점 추출

인식에 사용되는 특징점은 손끝과 손톱의 2가지가 있다. 손끝 부분은, 손가락이 퍼진 상태에서는 손가락의 끝부분이 화면 상에 나타나는 부분이며 손가락이 접혀진 상태에서는 손가락 마디의 끝부분이 돌출되어 나타나는 부분이다. 이 손끝 부분에서는 X 좌표가 지역적 극대값을 가진다.



[ 그림 5 : 손끝의 검출 예제 ]



손톱은 색상을 이용하여 식별한다. 그림 6은 손톱과 다른 부분의 픽셀(Pixel)을 선택하여 이들 픽셀의 RGB 값을 X,Y,Z 좌표의 값으로 전환하여 3차원 공간 상에 나타난 결과이다.

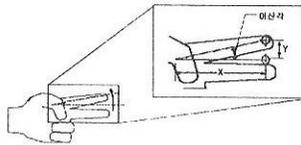
[ 그림 6 : 손톱과 非손톱 영역의 RGB 계에서 분포 ]

본 논문에서는 색상을 이용하여 손톱을 구분하기 위해 MLP(Multi Layer Perceptron)을 이용하였다. 신경망을 손톱의 RGB 값과 손톱이 아닌 부분의 픽셀의 RGB 값들로 훈련시켰다. 손영상의 각 픽셀들을 MLP에 적용하여 손톱인가 아닌가를 판단한다. 잘못 판단된 픽셀들을 제거하기 위하여 연결도를 이용하여 3이하인 픽셀은 에러 픽셀로 간주하여 제거한다. 에러 픽셀들이 제거된 후, 손톱 픽셀들을 레이블링(Labeling)기법을 이용하여 연결부분(Connected Component)으로 분리한다. 각 연결부분의 밀집도(Compactness), 크기(Size), 가로세로비를 계산하여, 손톱인가를 최종 판단한다.

3.4 손가락 관절의 이산각 계산

손가락 관절의 이산각은 손가락을 벌일 때 발생하는 손가락 사이의 각이다. 이산각은 손가락의 첫 관절(MP)에서 손끝의 위치까지의 가로/세로의 탄젠트(Tangent)값을 이용하여 계산한다.

$$Y = X \cdot \tan(\theta)$$



[ 그림 7 : 이산각의 계산 ]

3.5 손가락 형태 결정

손가락 관절의 꺾임각을 계산하기 위하여 각 손가락의 모양을 구부러진 정도에 따라 2가지 형태로 나누었다. 각 손가락의 형태 구분은 영상에서의 손톱의 존재여부와 손끝의 위치에 의해 결정되며, 이에 따라 이용되는 손가락의 움직임 특성과 영상에서 얻을 수 있는 정보가 달라진다.

형태 1은 손가락이 퍼져 있는 형태로 중간 관절(PIP)의 값이 일정 수준을 넘지 않았을 경우이다. 이 때는 손영상에서는 손끝의 위치를 얻고, 손의 형태와 구조적 특성에서  $\theta_{MP} = 0$ ,  $\theta_{PIP} = 2/3 \theta_{PIP}$ 의 꺾임각 제한을 이용한다. 형태 2는 손가락이 구부러진 형태이다. 형태 2에서는 손가락의 MP의 관절의 움직임이 자유롭다. 이 때는 손영상에서 손끝(손마디의 돌출 부분)의 위치와 손톱의 위치를 얻고, 손의 구조적 특성에서  $\theta_{MP} = 2/3 \theta_{PIP}$ 의 꺾임각 제한을 이용한다.

3.6 손가락 관절의 꺾임각 계산

앞에서 정의된 손가락의 형태에 따라 2가지 경우로 나누어 손가락 관절의 꺾임각을 계산한다.

형태 1은 손가락이 퍼진 형태로, 손가락의 첫번째 관절(MP)의 꺾임각이 0으로 고정되고, 손가락의 관절은  $\theta_{MP} = 2/3 \theta_{PIP}$ 의 비율을 가진다. 영상에서 얻어진 손끝의 위치를 이용해 그림 8과 같이 손가락 관절의 꺾임각을 계산한다.

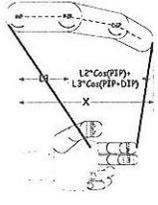


$$X = L_1 + L_2 \cdot \cos(\theta_{MP}) + L_3 \cdot \cos(\theta_{MP} + 2/3 \theta_{PIP})$$

- $L_1$  : 첫째 마디(마리등롱)
- $L_2$  : 둘째 마디(마리등롱)
- $L_3$  : 셋째 마디(마리등롱)
- $\theta_{MP} = 0$ ,  $\theta_{PIP} = 2/3 \theta_{PIP}$

[ 그림 8 : 형태 1에서의 손의 각 계산 ]

형태 2는 손가락이 구부러진 형태이다. 이 단계에서는 손톱의 위치와 손끝이 나타내는 위치 정보를 영상에서 얻는다. 손가락이 접혀진 상태에서  $\theta_{DIP} = 2/3 \theta_{PIP}$ 의 상관관계를 이용하여, 형태 2에서의 관절의 각임각을 그림 9와 같이 계산할 수 있다.



$$\begin{aligned} X_1 &= L_1 \cdot \cos(\theta_{IP}) \\ X_2 &= L_2 \cdot \cos(\theta_{IP} + \theta_{PIP}) + L_3 \cdot \cos(\theta_{IP} + \theta_{PIP} + 2/3 \theta_{PIP}) \\ \theta_{PIP} &: \text{첫째 마디 (미리 등록)} \\ &- L_2 : \text{둘째 마디 (미리 등록)} \\ &- L_3 : \text{세째 마디 (미리 등록)} \\ &- \theta_{DIP} = 2/3 \theta_{PIP} \end{aligned}$$

[ 그림 9 : 형태 2 에서의 손의 각 계산 ]

3.7 연속영상을 이용한 Filtering

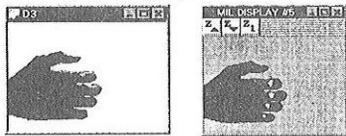
손을 인식하는 과정에는 손톱 추출의 오류와 손가락 형태 판단의 오류와 같은 여러 오류들이 나타나며, 그로 인하여 연속 영상에서의 추적된 손의 움직임이 부드럽지 않게 나타난다. 본 논문에서는 관절의 꺾임각을 Kalman Filtering 하여 오류를 보정하였다. Kalman Filter 는 물체가 선형으로 움직인다고 가정하고 선형모델을 통해 추적한다. 긴 시간 동안의 손가락의 움직임은 하나의 선형 모델로 표현할 수 없으므로 짧은 구간으로 나누어 선형(Linear)으로 근사하였다.

4. 실험 및 결과

실험은 단일 배경을 가진 손 영상에 대하여 수행하였다. CCD 카메라에서 입력 받은 영상은 Image Grabbing Board 로 전송이 되고, 손 영역 분리 과정을 거쳐 필요한 특징점을 추출한다. 영상입력을 위해 Matrox Meteor Image Grabbing Board™를 사용하였다. 실험은 Pentium Pro™ PC에서 수행하였다.

4.1 MLP를 이용한 손톱 추출 결과

손톱을 추출하기 위해서 Multi Layer Perceptron 을 이용하였다. 훈련은 손톱과 비손톱의 구분이 어려운 부분에서 데이터를 집중적으로 모아 훈련시켰다. 모두 1700 개의 데이터를 이용하여 훈련시켰으며, 훈련 데이터에 대한 분리 성공률은 85.2%였다. 손의 영상에 MLP를 적용하여 여러 픽셀들을 제거한 결과는 그림 10 와 같다.



[ 그림 10 : MLP를 적용한 손톱 추출 결과 ]

4.2 손 구조 인식 결과

그림 11 은 전체적인 인식 결과로 영상에서 손의 구조를 인식한 결과를 골격 모델에서 표현하고 있다. 기존의 방법에서는 불가능했던 손가락이 구부러진 상황에서도 손의 구조를 잘 인식하고 있다.

4.3 연속 영상의 인식 결과

연속영상에서 평균적인 인식 속도는 약 4 Frame / s 였다. 손의 움직임을 Kalman Filter 를 이용하였을 경우는 관절 각임각 계산의 오류를 보정되어 손의 동작이 훨씬 부드럽게 표현되었음을 인식 결과에서 확인할 수 있었다.

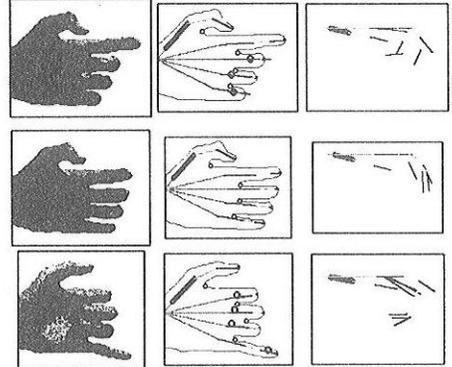
5. 결론

본 논문에서는 연속 영상에서 손의 3 차원적 구조 인식을 위한 방법을 제안하였다. 제안한 방법은 손의 구조적 특성에 추출된 손톱과 손끝의 위치 정보를 추가하여, 손가락 관절의 꺾임각과 이산각을 계산함으로써 이루어진다.

손톱의 정보를 이용하여, 기존의 맨손을 이용한 인식에서는 불가능했던 손가락의 접힘에 의해 발생하는 겹침 현상을 효과적으로 극복하였으며, 연속 영상에 filtering 을 적용하여 인식에서 발생

하는 오류를 최소화 하였다.

본 방법론은 몇 가지 제약을 지니고 있다. 우선, 손바닥의 방향이 카메라를 향하고 있어야 하는 제약을 두어 손목이 자유롭게 움직이지 못하고 있어 이를 극복하기 위한 연구가 수행되어야 하겠다. 다음으로 손 마디의 길이 정보가 미리 등록되어 있어야 손



13.a 손영상 13.b 인식결과-점선 13.c 인식결과-측면  
[ 그림 11 : 손 인식 결과 ]

의 구조를 인식할 수 있다. 이를 해결하기 위하여 손 마디의 길이정보를 자동적으로 획득할 수 있는 방법이 필요하다.

6. 참고 문헌

- Jochen Triesch , "Robust Classification of Hand Postures against Complex Background", In Proc. Second International Conference on Automatic Face and Gesture Recognition, Pages 170- 175 1996
- Andrew D. Wilson & Aaron F. Bobick, "Configuration States for the Representation and Recognition of Gesture", In Proc. of International Workshop on Automatic Face and Gesture Recognition, Pages 129-134 1995
- Yuntao Cui and John Weng, "Learning Based Hand Sign Recognition", In Proc. International Workshop on Automatic Face and Gesture Recognition, Pages 201-206 1995
- 고일주, 최형일, "손동작 인식을 이용한 컴퓨터 시각 기반 사용자 인터페이스의 구현", In Proc. 영상처리 및 이해에 관한 워크샵, Pages 340-345 1997
- Tony Heap and David Hogg, "Toward 3D Hand Tracking using a Deformable Model", In Proc. International Conference on Automatic Face and Gesture Recognition, Pages 140-145 1996
- J. Rehg and T. Kanade. "Visual tracking of high DOF articulated structures: An application to human hand tracking". In Proc. 3th ECCV, Pages 329-338 1993
- Lee JinTae, "Model based Analysis of Hand Posture", IEEE CG&A 77-86 1995
- James M. Rehg, "Visual Tracking of Self-Occluding Articulated Objects", University of Carnegie Mellon, CMU-CS-94-224, 1994
- 김영민, "유사특성을 가진 물체와의 겹침 현상을 고려한 손 추적", M.S. thesis, KAIST 1996
- Korean, J.U. "Techniques for generating the goal directed motion of articulated structures", IEEE CG&A 2.9 1982
- Paul R Kalata, "The Tracking Index: A generalized parameter for  $\alpha$ - $\beta$  and  $\alpha$ - $\beta$ - $\gamma$  Target Trackers", IEEE TR on Aerospace and electronic systems vol AES 20, NO2, 1984
- David F.Rogers & J. Alan Dams Second Ed. Of Mathematical Elements for Computer Graphics, McGRAW-HILL, Pages 101-140. 1990
- Mohinder S. Grewal & Angus P. Andrews . Kalman Filtering . Theory and Practice Prentice hall, Pages 108-116 1993
- Christopher Brown, "Tutorial on Filtering, Restoration, and State Estimation", TR534, University of Rochester, 1995
- Park Chan Jong, "A Kinematic Simulation System for Human Hand", M.S Thesis, KAIST 1985