

가상 터치스크린 시스템을 위한 TOF 카메라 기반 3차원 손 끝 추정

김민욱[†] · 안양근^{††} · 정광모^{†††} · 이철우^{††††}

요약

TOF(Time of Flight) 기술은 물체의 3차원 깊이 정보 추출을 가능케 하는 기술 중의 하나이다. 하지만, TOF의 카메라의 출력인 깊이 영상을 이용한 물체의 3차원 위치 추출은 몸이나 손 등 크기가 큰 물체의 경우에는 비교적 신뢰성 있는 결과를 얻을 수 있는데 비하여 크기가 작은 손 끝의 경우에는 신뢰할 수 있는 값을 얻기 힘들다. 본 논문에서는 TOF 카메라에서 육면체 손 모델을 이용하여 수정된 손의 영상에서의 위치 정보와 팔 모델을 이용하여 손 끝의 3차원 좌표를 추정한다. 제안된 방법으로 실험을 한 결과 TOF 카메라의 깊이 영상만을 사용하여 인식한 손 끝의 3차원 위치정보와 비교하여 훨씬 더 향상된 결과를 얻을 수 있었다.

키워드: 가상 터치스크린, TOF, 손 끝 추적, 손 모델링, 영상 처리

3D Fingertip Estimation based on the TOF Camera for Virtual Touch Screen System

Min-Wook Kim[†] · Yang-Keun Ahn^{††} · Kwang-Mo Jung^{†††} · Chil-Woo Lee^{††††}

ABSTRACT

TOF technique is one of the skills that can obtain the object's 3D depth information. But depth image has low resolution and fingertip occupy very small region, so, it is difficult to find the precise fingertip's 3D information by only using depth image from TOF camera. In this paper, we estimate fingertip's 3D location using Arm Model and reliable hand's 3D location information that is modified by hexahedron as hand model. Using proposed method we can obtain more precise fingertip's 3D information than using only depth image.

Keywords : Virtual Touch Screen, TOF, Fingertip Tracking, Hand Modeling, Image Processing

1. 서론

오랫동안 키보드와 마우스는 컴퓨터의 입력도구로 사용되었다. 하지만 최근에는 터치스크린이 개발되어서 실제 모니터 화면의 정보를 보고 화면 속의 버튼이나 스크롤 등을 손가락이나 펜을 이용하여 조절함으로써 정보를 컴퓨터에 입력할 수 있게 되었다. 터치스크린은 평소 키보드 자판이나 마우스에 익숙하지 않은 유아들이나 노인, 몸이 불편한 장애인들의 컴퓨터 이용을 편리하게 해준다.

하지만, 현재의 터치스크린은 물리적 성질을 이용한 것이기 때문에 그 크기가 한정되어 있으며, 설치 위치가 제한되어 있다. 그리고 기술적으로는 광투과율이 최대 86%에 그치고 있고, 진동에 대해서 약점을 가지고 있으며, 가격이 비싸다는 단점을 가지고 있다[1].

이와 같은 터치스크린의 단점을 보완하기 위하여 개발된 것이 가상 터치스크린이다. 가상 터치스크린은 기존의 터치스크린에서 요구되는 물리적인 막이 없이, 카메라나 센서만을 이용하여 손이나 다른 물리적인 물체로부터의 터치 제스처를 인식하는 기술이다. 따라서 가상 터치스크린 기술은 기존의 터치스크린의 단점을 보완할 수 있고 3DTV 등 3D 디스플레이를 조작하는 도구로 사용될 수 있기 때문에 차세대 실감형 입력 인터페이스로 활용할 수 있다.

현재까지 가상 터치스크린의 연구는 그리 많이 이루어지지 않았다. 김형준 등[1]은 두 대의 카메라를 이용하여 각각의 카메라의 사상된 좌표 값의 집합의 정합에 의하여 입력

* 본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업(NIPA-2010-C1090-1011-0008)과 지식경제부의 "차세대 지능형 정보통신기술개발사업"의 연구결과로 수행되었음.

† 준회원: 전남대학교 전자컴퓨터 공학부 박사과정

†† 정회원: 전자부품연구원(KETI) 선임연구원

††† 정회원: 전자부품연구원(KETI) 실감정보플랫폼 연구센터 센터장

†††† 정회원: 전남대학교 전자컴퓨터공학부 교수(교신저자)

논문접수: 2010년 2월 11일

수정일: 1차 2010년 5월 14일

심사완료: 2010년 6월 7일

물체의 3차원 위치를 검출하고, 3차원 터치점과 임의로 지정된 가상막의 영역을 이용하여 이동, 클릭, 더블 클릭, 드래그 등 이벤트를 정의하여 가상 스크린을 구성하였다. Eunjin Koh 등[2]은 스테레오 카메라를 이용하여 가상 터치스크린을 구성하였다. 그리고 실제 AR(Augmented Reality)기술을 이용하여 마커와 HMD(Head Mounted Display)를 이용하여 가상 터치스크린 막을 실제 시각적으로 볼 수 있게 하였다. 다음으로, 스테레오 카메라의 깊이 정보를 이용하여 추정된 손의 3차원 위치를 추정하고 사전에 미리 지정한 HMD와 손 사이의 거리와 비교하여 터치 여부를 결정한다.

하지만, [1]의 경우 제안한 시스템의 물리적 구조는 수식에 의해 고정되어 있기 때문에 시스템의 미세한 물리적인 이동에도 시스템의 구조를 모두 바꿔야 한다는 단점을 가지고 있다. 그리고 2대의 카메라를 사용하지만 스테레오 알고리즘을 사용하여 깊이 정보를 추출하는 것이 아니라 단순히 두 대의 카메라의 사상된 좌표의 집합을 3차원 공간상의 특정 지점에 대응시키는 방법을 사용하고 있다. 이는 추정된 3차원 위치를 주관적으로 대응시킴으로써 얻어진 점이 손 끝 지점이라고 신뢰하기 어렵다. [2]의 경우, 손의 3차원 위치를 추정할 때, 스테레오 카메라를 이용하였다. 스테레오 카메라의 시차(Disparity) 정보를 이용하여 깊이 정보를 검출하는 것은 물체의 3차원 거리는 객관적으로 추정할 수 있지만, 근거리에서는 3차원 정보 검출이 어렵다는 점과 조명의 변화 등 주변 환경에 민감하다는 점 등의 한계가 있다.

근거리 내의 손의 깊이정보를 추출하는 경우, TOF 카메라는 스테레오 카메라보다 강건하게 3차원 위치의 검출이 가능하다[14]. 하지만, 손 끝 검출의 경우에는 영상에서 차지하는 픽셀 수가 극히 작아서 정확한 위치를 검출하기 어렵다. 따라서 손 끝의 3차원 위치를 추정하기 위해 손의 정확한 위치를 추적하는 방법이 필요하다. 이 방법에는 크게 3가지 방법이 있다[3]. 뉴턴의 운동법칙과 주성분 분석(Principal Component Analysis)을 기반으로 변형이 가능한 손 모양 모델을 사용[4]하는 방법이다. 하지만, 이러한 방법은 손의 정확한 포즈를 추정할 수 없다는 단점이 있다. 다음으로, 입력 영상의 특징 공간과 손의 움직임에 대한 두 공간 사이의 매핑을 이용하는 외형 기반 방법[5, 6]이다. 이러한 접근법은 두 공간 사이의 매핑 함수를 학습하기 힘들다는 단점을 가지고 있다. 마지막으로, 입력 영상과 3차원 모델 투영 영상과의 정합을 추정하는 3차원 모델 기반 방법이다. 3차원 모델과 입력 영상과의 대응점을 찾기 위해 영상에서 손가락 끝[7-9], 윤곽[10], 실루엣[7, 11, 12] 등과 같은 다양한 특징들이 사용가능하다.

본 논문에서는 TOF 카메라를 사용하여 얻은 손 영역에 대한 3차원 무게 중심점의 위치 및 손의 회전각을 검출하고, 육면체 모델을 이용하여 손의 무게중심점의 3차원 위치 및 회전각을 보정함으로써 손 끝의 3차원 위치를 정확하게 추정한다. 그 결과 단순히 임계치를 이용하여 손 끝을 추정할 때보다 현저하게 개선된 결과를 얻을 수 있었다.

본 논문의 다음과 같이 구성되어 있다. 2장에서는 TOF

카메라를 이용하여 손의 3차원 위치 추정하는 과정을 설명한다. 그리고 3장에서는 육면체 손 모델 기반 손 위치 보정 및 손 끝 위치 추정 방법을 상세히 설명한다. 4장에서는 보정된 손 위치 추정 결과와 손 끝의 위치 추정 결과를 제공한다. 마지막 5장에서는 결론을 내린다.

2. TOF 카메라 기반 3차원 손 위치 추정

2.1 깊이 영상(Depth Image)

TOF 카메라는 카메라 내부에서 발산한 적외선 광분자가 대상 물체에 부딪쳐서 되돌아오는 시간을 카메라 내부에 센서가 인지한 후, 측정된 시간을 대상과의 거리로 변환하여 깊이 영상을 생성한다. TOF 카메라의 특성을 이용하면 2m 이내의 근거리에서의 물체의 3차원 깊이 정보를 비교적 정확하게 얻을 수 있다[13]. 그리고 TOF 카메라는 적외선을 이용하기 때문에 조명이나 주변 환경의 변화에도 강건한 결과를 얻을 수 있다. (그림 1)은 TOF의 입력 영상과 그에 따른 깊이 영상이 보여주고 있다. 이 영상은 그레이영상으로써, 카메라로부터 거리가 가까울수록 깊이 영상은 밝은색을 나타내고, 거리가 멀수록 어두운 색을 갖는다[14].

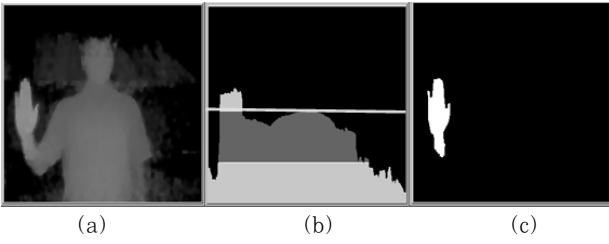


(그림 1) TOF 카메라의 입력 영상과 깊이 영상 (a) 카메라 입력 영상 (b) TOF 깊이 영상

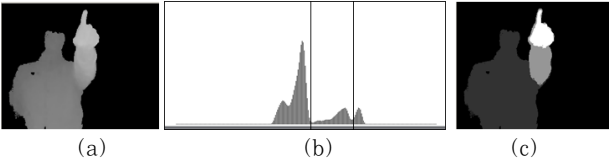
2.2 임계치 설정 및 객체 분리

카메라로부터 떨어진 거리를 나타내는 깊이 영상에 임의의 임계치를 설정해줌으로써 임의의 거리 안에 있는 물체의 카메라로부터 떨어진 거리를 추정할 수 있다. (그림 2)는 깊이 영상과 그에 따른 x-depth 맵을 이용하여 객체를 검출하는 것을 나타낸다. x-depth 맵은 횡축은 깊이 영상의 x축 해상도 값을 나타내고, 종축은 깊이 영상에서의 횡축의 값에 따른 깊이 영상의 배열 중 가장 큰 값을 선택하여서 표현한 영상이다.

(그림 2)-(c)와 같이 x-depth 맵에 임계치를 적용하면 특정 거리 안에 있는 객체를 추출할 수 있다. x-depth 맵을 이용하면 그의 임계치에 따라 단순히 특정 거리에 안에 있는 물체의 픽셀의 값을 추출하는 것에 그친다. 하지만, 깊이 영상의 히스토그램을 이용하면 (그림 3)-(c)와 같이 몸과 팔, 그리고 손을 분리할 수 있게 된다. 이는 깊이 히스토그램을 미분하여 그 미분된 이미지에서 지역 극대값과 지역 극소값을 각각 객체 분리의 값으로 설정하여 몸과 팔, 손의



(그림 2) x-depth 맵을 이용한 객체 검출 (a) 깊이 영상 (b) 깊이 영상의 x-depth Map (c) x-depth map을 이용한 객체 검출



(그림 3) 깊이 히스토그램을 이용한 객체 분리 (a) 깊이 영상 (b) 깊이 영상의 히스토그램 (c) 깊이 히스토그램에 의한 객체 분리

각 부분을 검출하고 분리하는 방법이다. 이는 깊이 영상에 특정 임계치를 사용하여 객체를 추출하는 것보다 향상된 결과를 보였다.

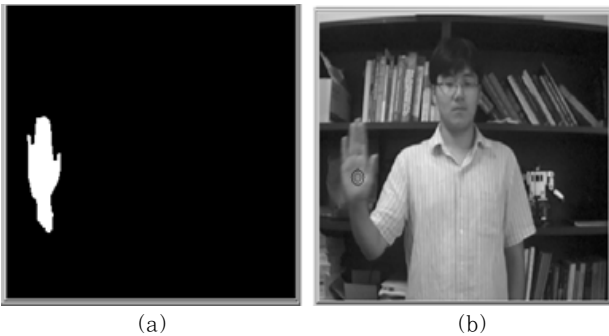
2.3 객체의 중심점 검출

검출된 객체의 중심 모멘트를 구함으로써 객체의 중심점을 얻을 수 있다. 모멘트는 식 1과 같다.

$$M_{x_{or},y_{or}} = \sum_{x,y} (I(x,y) \cdot (x - x_c)^{x_{or}} \cdot (y - y_c)^{y_{or}}) \quad (1)$$

그리고 객체의 중심점은 식 2와 같이 표현 가능하고, 그림 4)와 같이 나타난다.

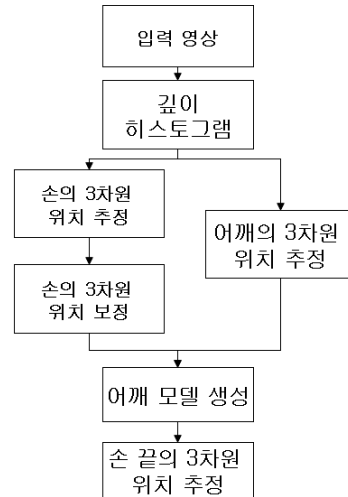
$$\mu_{x,y} = \left(\frac{M_{1,0}}{M_{0,0}}, \frac{M_{0,1}}{M_{0,0}} \right) \quad (2)$$



(그림 4) 객체의 무게 중심점 추출 (a) 분리된 객체 (b) 객체의 무게 중심점 추출

3. 모델 기반 손 위치 보정 및 손 끝 위치 추정

이 장에서는 2장에서 추정된 손의 3차원 중심점과 육각형 손 모델을 이용하여 영상에서 얻어진 손의 중심점을 보정하고 그에 따른 손 끝의 3차원 위치를 추정한다. (그림 5)는 모델 기반 손 위치 보정 및 손 끝 추정 시스템의 흐름을 보여준다.

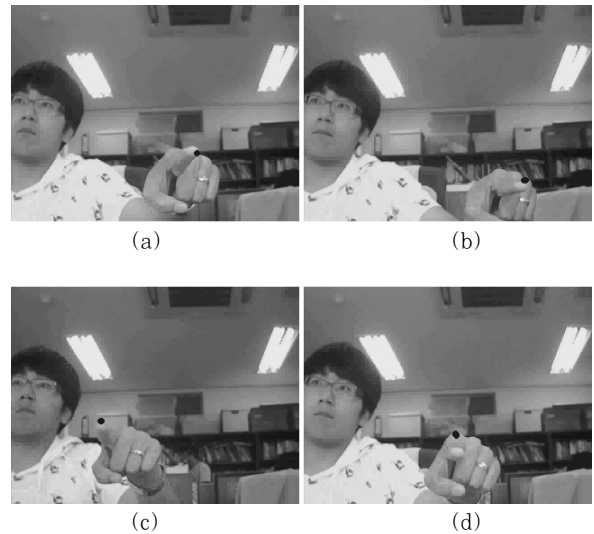


(그림 5) TOF 카메라 기반 3차원 손 끝 추정 시스템

3.1 깊이 영상을 이용한 손 끝 추정의 한계

2장에서 기술하였듯이 깊이 영상에 임계치를 설정하면 원하는 관심 객체의 영역만을 추출할 수 있다. 하지만, 손 끝이나 손가락의 경우에는 전체 이미지에서 차지하는 영역의 크기가 상대적으로 작기 때문에, 객체를 추출하는 데에는 한계가 있다.

(그림 6)는 임계치 설정만을 이용하여 손 끝 영역을 추출



(그림 6) 임계치 설정을 이용한 손 끝 영역 추출 (a), (b), (c) : 정상적인 검출 (d) : 잘못된 검출

한 결과이다. 실험에서는 실시간 환경에서 임계치은 깊이 영상 중에서 가장 높은 값을 설정하여 카메라에서 가장 가까운 인식되는 지점을 찾도록 하였다. 그 결과, (그림 6)처럼 손 끝이 인식되는 프레임도 있었지만, 오인식 되는 프레임도 있어서 임계치 설정만으로 손 끝의 위치를 추정하는 것은 한계가 있음을 알 수 있다.

3.2 팔 모델기반 3차원 손의 위치 추정

TOF 카메라의 깊이 정보만을 이용하여 손 끝의 위치 추적하는 것은 한계성을 갖는다. 따라서 본 논문에서는 실제 동작자와 같은 길이를 갖는 3D 팔모델을 이용한다. 어깨로부터 팔 길이만큼 공간에 위치해 있는 손의 3차원 위치를 추정하고, 그 기술기에 따라 손 끝의 위치를 추정하는 방법이다. 먼저 영상의 한 점과 3차원 공간 상의 한 점 사이의 상관관계를 추정해야한다.

실제 카메라에 투영된 이미지의 좌표 (x, y) 로부터 실제 3차원 상에 존재하는 실제 손의 좌표 (X, Y, Z)로 향하는 벡터를 추정할 수 있다. (그림 7)은 핀홀 카메라의 전역좌표에서의 3차원 객체의 중심점 (X, Y, Z)가 이미지 평면 I(영상에 표기)에 투영되는 과정을 보여주고 있다. 즉, 이미지 평면 I의 좌표 (x, y)로부터 객체의 중심점으로 향하는 벡터를 구할 수 있다[15-17].

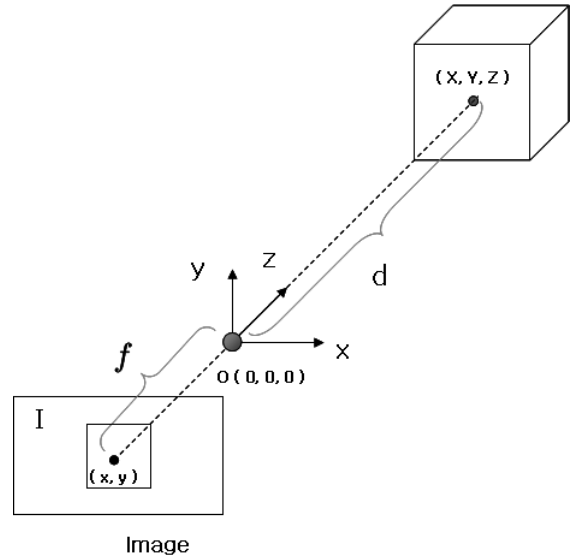
이 때, 핀홀 카메라 모델로 가정하면 (그림 8)과 같이 객체 중심점의 깊이 값 Z를 TOF 카메라에서 물체까지의 깊이 정보 d라고 간주할 수 있다. f를 초점거리를 의미하며, 삼각형의 비례식에 의해 손 끝의 3차원 좌표는 식 3과 같이 표현 가능하다.

$$(X, Y, Z) = \left(\frac{xd}{f}, \frac{yd}{f}, d \right) \tag{3}$$

팔의 길이는 항상 같은 길이이고, 팔은 어깨 위치를 중심으로 회전하기 때문에 팔의 이동은 (그림 7)과 같이 구 모양으로 회전하게 된다. 따라서 어깨의 위치를 원점으로부터 C만큼 떨어져있는 경우, 손의 3차원 위치점은 다음과 같다.

$$X^2 + Y^2 + (Z - C)^2 = r^2 \tag{4}$$

그리고 팔의 길이를 L이라고 한다면, 영상에 투영된 손의



(그림 7) 핀홀 영상으로부터 3차원 객체의 위치 추정

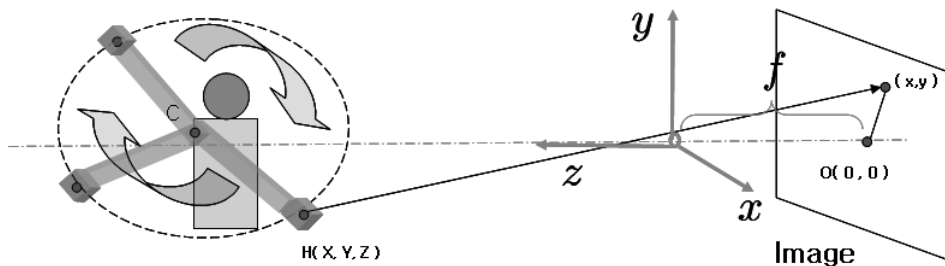
중심점 (x, y)의 3차원 위치는 식 3을 식 4에 대입하여 식 5를 얻을 수 있다.

$$\left(\frac{xd}{f} \right)^2 + \left(\frac{yd}{f} \right)^2 + (C-d)^2 = L^2 \tag{5}$$

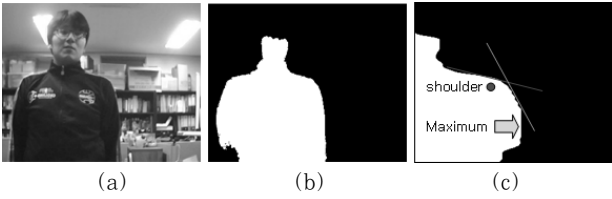
3.3 어깨의 위치 검출

팔 모델을 이용하여 3차원 손의 위치를 추정할 경우 중심점인 어깨의 위치를 결정하는 것이 중요하다[18]. (그림 8)과 같이 사용자가 Z축으로 C의 위치에 있다고 가정했을 경우, 사용자의 키와 체형에 따라 사용자의 어깨의 위치는 가변적으로 변하게 돼서 사용자가 변경되거나 사용자의 위치가 변경될 때는 어깨의 위치, 즉 팔 모델의 중심점은 변한다.

어깨의 위치는 깊이 영상을 이진화한 실루엣 영상으로부터 얻을 수 있다. 실루엣 영상을 높이를 축으로 하여 히스토그램을 그리면 어깨로부터 시작하여 골반까지의 부분에서 가장 히스토그램 값이 큰 한 점이 검출된다. 그 한 점으로부터 머리쪽으로의 히스토그램의 값 중 급격하게 히스토그램 값이 감소되었다가 점차 완화되는 부분, 즉 2차 미분 값이 0이 되는 부분은 목과 어깨 사이의 한 점이다. 그리고 급격한 변화가 시작되는 부분은 어깨라고 할 수 있다. (그림



(그림 8) 팔의 구이동에 따른 3차원 위치 추정



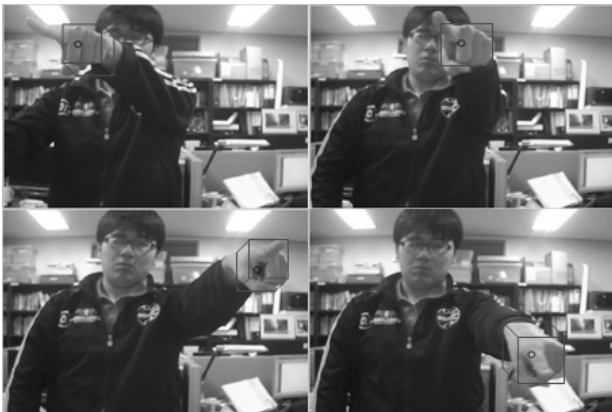
(그림 9) 어깨 위치 검출 과정 (a) 입력 RGB 영상. (b) 몸통의 실루엣 영상 (c) 실루엣 영상의 x-y축 히스토그램에 의한 어깨 위치 검출

9)는 어깨 위치 검출 과정을 나타낸 그림으로 입력 영상과 깊이가 영상으로부터 사용자의 실루엣 영상을 구하고, 그에 따라 어깨를 구하는 과정을 나타낸다.

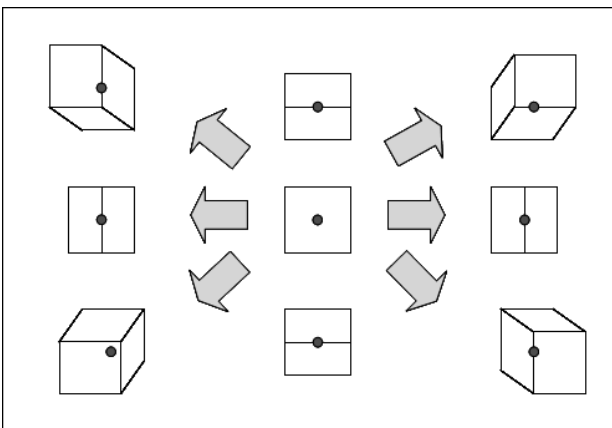
3.4 손 모델을 이용한 손의 중심점 보정

2장에서 추정된 손의 이미지에 투영된 중심점은 (그림 9)에서와 같이 카메라를 통해 보이는 손의 모양에 따라 달라진다. (그림 10)과 (그림 11)은 이와 같이 손의 움직임에 의한 위치 변화에 따른 육면체 손의 모델의 가시면의 변화와 그에 따른 중심점을 나타낸다.

손이 주먹 쥌 상태에서 영상의 정중앙에 위치하면 손의 주먹의 앞부분만이 가시면이 된다. 그러나 손의 3차원 위치



(그림 10) 손의 위치에 따른 손의 중심점의 변화



(그림 11) 손 모델의 가시면 변화와 그에 따른 중심점의 변화

의 변화에 따라 가시면도 변하게 되고, 그에 따라 영상에 투영된 손의 무게 중심점도 이동한다. 정확한 2차원 영상에서의 손의 중심점을 알기위해 손의 위치에 따른 변화에 관계없이 손 위의 동일한 지점을 찾아야한다.

손의 움직임은 앞으로 지면과 평행하게 뻗은 자세를 기본 자세라고 정의하고, 어깨를 중심으로 X, Z축을 축으로 θ_x, θ_z 만큼 회전변환을 할 경우, 회전하기 전 (x, y, z) 는 식 6에 의해 (x', y', z') 으로 나타낸다.

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_x & -\sin\theta_x \\ 0 & \sin\theta_x & \cos\theta_x \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (6)$$

(그림 12)와 같이 육면체 모델의 움직임에 따라 가시면은 1~3개가 될 수 있고, 각 면들의 회전된 3차원 지점에 각 가시면의 면적을 가중치로 갖는 점을 영상의 투영된 손의 중심점이라고 할 수 있다. (그림 12)는 가시면이 3개인 경우, 회전 후 3개의 가시면의 중심을 나타낸 것이다. 이 때, 가시면의 면적을 각각 A, B, C라고 한다면 모델의 영상에 투영될 손의 중심점 P_M 은 식 7과 같이 표현 가능하다.

$$P_M = \frac{A \cdot P_A + B \cdot P_B + C \cdot P_C}{A + B + C} \quad (7)$$

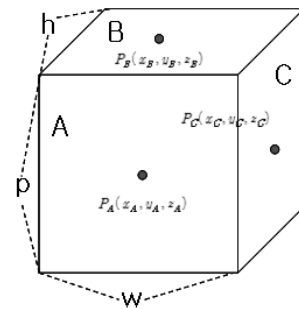
각각의 가시면의 넓이 A, B, C는 회전각 θ_x, θ_z 의 식으로 나타낼 수 있고, 육면체 모델의 가로, 세로, 높이 길이를 각각 w, h, p라고 한다면, 식 8, 9, 10로 표현된다.

$$A = w \cdot p \cdot \cos\theta_x \cdot \cos\theta_z \quad (8)$$

$$B = w \cdot h \cdot \sin\theta_x \cdot \cos\theta_z \quad (9)$$

$$C = h \cdot p \cdot \cos\theta_x \cdot \sin\theta_z \quad (10)$$

영상에서의 객체의 중심점은 모델의 영상에 투영될 손의 3차원 중심점 P_M 을 2차원 영상 평면에 투영한 점이기에 때문에 P_M 을 2차원으로 투영한 점 P_x 은 P_M 을 (x_m, y_m, z_m) 이라고 가정하고 f를 카메라 렌즈의 초점거리라고 하면 식 11로 표현되고, 이 수식은 2.3절에서 검출했던 객체의 중심 모멘트와 같다.



(그림 12) 회전 후, 육면체 모델



(그림 13) 어깨를 원점으로 한 손의 회전 이동

$$P_{\mu} = \left(\frac{f \cdot x_m}{z_m}, \frac{f \cdot y_m}{z_m} \right) = \mu_{x,y} \quad (11)$$

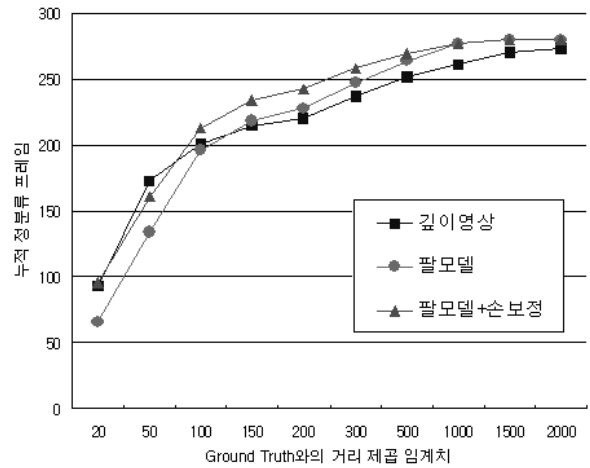
4. 실험

실험은 TOF 카메라의 깊이 영상만을 이용하여 손 끝을 찾는 방법, 팔 모델만을 이용하여 손 끝을 찾는 방법, 그리고 팔 모델과 육면체 손 모델을 이용하여 손 중심점을 보정하여 손 끝을 찾는 방법으로 각각 진행된다. 실험의 입력으로는 3DVSystems의 TOF 카메라인 ZCAM의 320×240 해상도를 갖는 깊이 영상과 컬러 영상이 이용되었다.

실험은 각각 280장의 컬러 영상과 깊이 영상으로부터 팔 모델과 육면체 손 모델을 사용하여 3차원 손 끝을 찾는 것으로 이루어진다. 3차원 손 끝 검출 실험의 신뢰도를 위하여 검출된 3차원 손 끝의 위치를 2차원 영상 평면에 투영해 보고 참 값(Ground Truth)와 비교하였으며, (그림 15)는 OpenGL 그래픽 툴을 이용하여 시각적으로 확인할 수 있게 하였다.



(그림 14) 팔모델과 육각형 손 모델을 이용한 손 끝 찾기



(그림 15) 임계치에 따른 누적 손 끝 정분류 프레임 수

(그림 14)에서 별의 위치점은 팔 모델의 원점인 어깨의 위치이고, 사각형은 깊이 영상에 팔 모델을 사용하여 얻어진 손의 위치점이고, 삼각형은 육각형 손 모델을 이용하여 개선된 손의 위치점이다. 그리고 원은 개선된 손의 위치로부터 계산된 3차원 손 끝 점이 2차원 영상 평면으로 투영된 점이다.

성능 검사는 수작업으로 얻은 손 끝의 픽셀의 위치 참 값(Ground Truth)의 데이터와 각 알고리즘의 손 끝의 위치 추정 결과값 사이의 오차를 계산하여 비교 평가하였다. 비교 대상 알고리즘은 깊이 영상만을 이용하여 얻은 손 끝 위치 검출 알고리즘, 팔 모델을 이용하여 얻은 손 끝의 추정 알고리즘, 팔 모델과 육각형 손 모델을 이용하여 얻은 손 끝의 추정 알고리즘이다. 각각 알고리즘의 검출 결과와 참 값과의 거리 차에 임계치를 설정하여 정분류된 손 끝의 프레임수를 합산한 결과와 각 알고리즘의 결과의 MSE (Mean Square Error)를 계산한 결과를 통해 성능 평가를 수행하였다.

(그림 15)는 임계치에 따른 누적 손 끝 정분류 프레임 수를 나타낸다. 전반적으로 팔모델과 손 보정을 거친 실험이 가장 좋은 결과를 낸다. 하지만, 참 값과의 거리 제곱의 임계치가 50인 경우에는 깊이 영상만을 이용한 손 끝 검출이 가장 좋은 결과를 보여주고 있다. 이것은 깊이 영상만을 이용한 손 끝 검출의 경우, 모델을 사용하지 않고, 입력 영상에서 바로 손 끝을 검출하기 때문에 몇몇의 경우에는 모델을 이용한 추정보다 비교적 정확하게 손 끝의 위치점을 얻을 수 있다.

<표 1>은 영상에서의 2차원 위치 참 값과 3가지 손 끝 추적 방법으로부터 얻은 3차원 손 끝 점을 2차원 영상 평면으로 투영한 점과의 거리의 MSE(Mean Square Error)를 구한 것이다. 결과적으로 <표 1>과 같이 팔 모델과 손 위치 보정을 해서 손 끝을 검출한 결과가 깊이 영상만을 이용한 결과보다 훨씬 더 좋은 결과를 얻었다.

〈표 1〉 손 끝 추적 방법 별 오류

실험 방법	MSE
깊이 영상만을 이용	219.9
팔 모델 이용	137.3
팔 모델 이용 + 손 위치 보정	107.0

5. 결론 및 향후 연구

본 논문에서는 TOF 카메라의 깊이 영상과 팔 모델, 육면체 손 모델을 이용하여 손 끝의 3차원 위치 검출하는 방법을 제안하였다. 기존의 TOF 카메라와 물체와의 거리, 즉 깊이 영상만을 이용하는 경우에는 카메라의 해상도와 손의 위치에 따라서 추정된 손 끝의 3차원 위치는 비교적 많은 오류를 포함하고 있으며, 특정 영상에서는 완전히 다른 부분을 검출하는 경우도 있었다.

깊이 영상으로부터 어깨의 지점을 구하여 원점으로 설정하고, 구 모형 회전 변환을 갖는 팔 모델을 생성한다. 이미 사진에 알고 있는 영상에서 손의 중심점과 카메라의 내부 파라미터인 초점거리를 이용하여 3차원 손의 위치를 추정한다. 하지만, 3차원 손의 위치를 영상에서 추정하고, 손의 중심점은 손의 위치에 따라 중심점이 변하므로 손을 육면체 모델로 가정한다. 회전변환에 따른 모형의 변화를 추정하여 손의 위치에 독립적으로 손의 일정한 지점의 3차원 위치를 추정한다. 그리고 손의 3차원 위치에 따른 x, z 축을 기준으로 하는 회전 자유도 θ_1, θ_2 를 구함으로써 손 끝의 3차원 위치를 추정하고, 그 결과 깊이 영상만을 이용하여 손 끝을 검출하였을 때보다 개선된 결과를 얻을 수 있었다.

제안된 방법은 사용자가 Z축으로 일정한 거리에 고정되어 있어야 한다는 제약을 가지고 있어서 사용자와 움직임에 따라서 어깨의 위치, 즉 팔 모델의 원점 위치를 재설정해 주어야 한다. 따라서 향후 가변적인 사용자의 위치에도 정확하게 손 끝의 3차원 지점을 검출할 수 있는 연구가 필요하다.

참 고 문 헌

[1] 김형준, "게임 응용을 위한 가상 터치스크린 시스템", 한국게임학회논문지, 제6권, 제3호, pp.77-86, 2006.
 [2] Eunjin Koh, Jongho Won, and Changseok Bae, "Vision-based Virtual Touch Screen Interface", ICCE 2008, LasVegas, USA, 2008.
 [3] 석홍일, 이지홍 and 이성환, "3차원 손 모델링 기반의 실시간 손 포즈 추적 및 손가락 동작 인식", 정보과학회논문지, 제35권, 제12호, 2008.
 [4] A. Heap and D. Hogg, "Improving Specificity in PDMs using a Hierarchical Approach," Proc. British Machine Vision Conference, Essex, UK, Vol.1, pp.80-89, Sept. 1997.
 [5] R. Rosales, S. Sclaroff, and V. Athitsos, "3D Hand Pose Reconstruction using Specialized Mappings," Proc. 8th IEEE

International Conference on Computer Vision, Vancouver, Canada, Vol.1, pp.378-385, July, 2001.
 [6] Y. Wu and T. Huang, "View-Independent Recognition of Hand Postures," Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, South California, USA, Vol.2, pp.88-94, June, 2000.
 [7] J. Deutscher, A. Blake, and I. Reid, "Articulated Body Motion Capture by Annealed Particle Filtering," Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, South California, USA, Vol.2, pp.126-133, June, 2000.
 [8] N. Shimada, Y. Shirai, Y. Kuno, and J. Miura, "Hand Gesture Estimation and Model Refinement using Monocular Camera Ambiguity Limitation by Inequality Constraints," Proc. 3rd IEEE International Conference on Automatic Face and Gesture Recognition, Nara, Japan, pp.268-273, 1998.
 [9] Y. Wu and T. Huang, "Capturing Articulated Human Hand Motion: A Divide-and-Conquer Approach," Proc. 7th IEEE International Conference on Computer Vision, Kerkyra, Greece, Vol.1, pp.606-611, 1999.
 [10] B. Stenger, A. Thayananthan, P. Torr, and R. Cipolla, "Hand Pose Estimation Using Hierarchical Detection," Proc. European Conference on Computer Vision, Lecture Notes in Computer Science, Prague, Czech Republic, Vol.3058, pp.105-116, May, 2004.
 [11] J. Kuch and T. Huang, "Vision based Hand Modeling and Tracking for Virtual Teleconferencing and Telecollaboration," Proc. 5th International Conference on Computer Vision, Cambridge, USA, pp.666-671, June, 1995.
 [12] J. Lee and T. Knui, "Model-based Analysis of Hand Posture," Proc. IEEE Computer Graphics and Application, New York, USA, Vol.15, No.5, pp.77-86, 1995.
 [13] G.J. Iddan and G. Yahav, "3D IMAGING IN THE STUDIO (AND ELSEWHERE...)," Proceedings of SPIE, Vol.4298, pp.48-55, 2001.
 [14] 김민욱, 안양근 and 이철우, "다중신경망과 3차원 시각정보에 기반한 가상스크린 시스템," 멀티미디어학회 춘계학술대회, 제12권, 제1호, p.87, 2009.
 [15] 양정석, "모델 기반 헤드포즈 인식에 의한 응시 위치 인식", 석사학위논문, 전남대학교 대학원, 2002.
 [16] Yi Ma, Stefano Soatto, Jana Kosecka and S. Shankar Sastry, An Invitation to 3-D Vision From Image to Geometric Models, Springer, 2006.
 [17] 주우석, "OpenGL로 배우는 컴퓨터 그래픽스," 한빛미디어, 2006.
 [18] Meskers C.G.M, Vermeulen H.M., Groot de J.H., Helm van der F.C.T. and Rozing P.M.. "3D shoulder position measurements using a six-degree-of-freedom electromagnetic tracking device," Clinical Biomechanics, Vol.13, pp.280-292, 1998.



김민욱

e-mail : purelover7@naver.com
2008년 전남대학교 전자컴퓨터공학부(학사)
2010년 전남대학교 전자컴퓨터공학부(석사)
2010년~현재 전남대학교 전자컴퓨터공학부 박사과정
관심분야: 영상처리, 컴퓨터 비전, 제스처 인식, 패턴 인식



이철우

e-mail : leecw@chonnam.ac.kr
1986년 중앙대학교 전자공학과(학사)
1988년 중앙대학교 전자공학과(공학석사)
1992년 동경대학교 전자공학과(공학박사)
1992년~1995년 이미지 정보과학연구소 수석연구원 겸 오사카 대학 기초공학부 협력연구원
1995년 리츠메이칸대학 정보공학부 특별초빙강사
1996년~현재 전남대학교 전자컴퓨터공학부 교수
관심분야: 컴퓨터 비전, 지능형 휴먼 인터페이스, 디지털 콘텐츠, 컴퓨터그래픽스



안양근

e-mail : ykahn@keti.re.kr
1995년 한국해양대 제어계측공학과(학사)
1997년 한국해양대 제어계측공학과(공학석사)
1997년~2000년 한국 전력 연구원(KEPRI) 연구원
2000년~2004년 해동텔레콤 연구원
2004년~현재 전자부품연구원(KETI) 선임연구원
관심분야: 영상 처리, 임베디드 시스템, 펌웨어, 공간투영



정광모

e-mail : jungkm@keti.re.kr
1990년 광운대학교 전자공학과(학사)
2002년 광운대학교 전자공학과(공학석사)
2006년 광운대학교 전자통신공학과(공학박사)
2000년~현재 전자부품연구원(KETI) 실감정보플랫폼 연구센터 센터장
관심분야: 마이크로프로세서