

Kinect Depth 카메라를 이용한 마우스 커서의 위치 선정에 관한 연구

A Study On Positioning Of Mouse Cursor Using Kinect Depth Camera

구 봉 회*, 이 승 호**

Bong-Hoe Goo*, Seung-Ho Lee**

Abstract

In this paper, we propose new algorithm for positioning of mouse cursor using fingertip direction on kinect depth camera. The proposed algorithm uses center of palm points from distance transform when fingertip point toward screen. Otherwise, algorithm use fingertip points. After image preprocessing, the center of palm points is calculated from distance transform results. If the direction of the finger towards the camera becomes close to the distance between the fingertip point and center of palm point, it is possible to improve the accuracy of positioning by using the center of palm point. After remove arm on image, the fingertip points is obtained by using a pixel on the long distance from the center of the image. To calculate accuracy of mouse positioning, we selected any 5 points. Also, we calculated error rate between reference points and mouse points by performed 500 times. The error rate results could be confirmed the accuracy of our algorithm indicated an average error rate of less than 11%.

요약

본 논문에서는 Kinect Depth 카메라에서 손가락의 방향을 이용한 마우스 커서의 위치 선정에 관한 새로운 방법을 제안한다. 제안하는 방법은 손가락의 방향이 카메라를 향해 있을 경우는 거리변환을 이용한 손의 중점좌표를 이용한다. 손가락의 방향이 카메라를 향해 있지 않을 경우에는 손가락의 끝점 좌표를 이용한다. 손의 중점좌표는 진처리를 거친 영상의 거리변환 결과에서 가장 밝은 픽셀의 좌표를 이용하여 구한다. 손가락의 방향이 카메라를 향해 있을 경우에는 손가락의 끝점 좌표와 손의 중점 좌표 사이의 거리가 가까워지므로 중점좌표를 이용하여, 위치 선정의 정확도를 향상시킬 수 있게 된다. 손가락의 끝점 좌표는 손 영역만을 추출한 영상에서 팔 부분을 제거하여 영상의 중심에서 가장 멀리 떨어져있는 픽셀을 이용하여 구한다. 본 논문에서 제안하는 커서 위치 선정의 정확도를 측정하기 위해 임의의 5개의 기준 좌표를 선정하고 기준 좌표와 측정된 마우스 좌표사이의 오차를 500회 반복하여 측정하여 에러율을 측정한 결과, 평균 11 % 미만의 오차율을 나타내어 제안한 방법의 정확도를 확인할 수가 있었다.

Keywords: Kinect, Depth Image, Human Interface, Finger Detection, Distance Transform

* Department of Electronics&Control Engineering,
Hanbat
National University, wantrnqhdghl@naver.com,
010-3137-4221

★Department of Electronics&Control Engineering,
Hanbat National University, shlee@cad.hanbat.ac.kr,
010-8807-9876 Corresponding author
Manuscript received Oct. 5, 2014; revised Oct. 30,
2014 ; accepted Nov. 7. 2014

1. 서론

Kinect camera는 구조적으로 크게 Depth camera와 RGB camera 등의 2가지 카메라를 내장하고 있다. Kinect camera로부터 얻을 수 있는 영상정보는 일반적인 RGB 영상과 Depth 영상을 얻을 수 있다. Kinect camera를 이용함으로써 얻을 수 있는 가장 큰 장점은

번거로운 장치를 손으로 쥐거나 사용할 필요 없이 Kinect camera 상에 비치는 몸짓 혹은 손짓을 인식하여 동작할 수 있다는 점이다. 초기 Kinect camera의 목적은 Xbox 게임 콘솔용으로 개발된 입력장치였으나 현재 개발자를 위해 Microsoft Kinect API를 공개함으로써 활발히 연구가 진행 중이다. 주된 연구 분야 [1],[2],[3]는 물체 추적과 인식, 인체 행동 분석, 손 제스처 분석, 실내 3D Mapping 등으로 크게 4가지로 나누고 있다. 손 제스처 분석[4],[5],[6]에서는 RGB 컬러 영상과 Depth 영상의 이진화, 병합 등을 통해 손의 영역만을 추출하고 그 영역 내에 손가락의 끝점을 구해 손의 끝점이 그리하는 문자를 학습된 데이터를 이용하여 판별한다. 이런 손 제스처 분석과 관련하여 많은 연구가 진행되고 있고, Kinect camera를 활용하여 현존하는 개인용 PC의 입력 장치를 대체하는 연구도 활발하다. 이러한 대체 입력 장치를 구현하거나, 손 제스처를 이용한 증강현실을 구현[2] 할 때에 Kinect camera의 영상에서 손의 중점 좌표나 손가락의 끝점 좌표의 정확도[7]는 매우 중요하게 작용한다. 그러나 손가락의 방향을 이용한 마우스 커서의 위치 선정에서는 손가락의 끝점 좌표보다는 마우스 커서의 위치 선정 정확도가 더 중요하다. 본 논문에서 제안하는 방법은 손가락의 방향이 카메라를 향해 있을 경우는 거리변환[8]을 이용한 손의 중점좌표를 이용한다. 손가락의 방향이 카메라를 향해 있지 않을 경우에는 손가락의 끝점 좌표를 이용한다. 손의 중점좌표는 전처리를 거친 영상의 거리 변환 결과에서 가장 밝은 픽셀의 좌표를 이용하여 구한다. 손가락의 방향이 카메라를 향해 있을 경우에는 손가락의 끝점 좌표와 손의 중점 좌표 사이의 거리가 가까워지므로 중점좌표를 이용하여, 위치 선정의 정확도를 향상시킬 수 있게 된다. 손가락의 끝점 좌표는 손 영역만을 추출한 영상에서 팔 부분을 제거하여 영상의 중심에서 가장 멀리 떨어져있는 픽셀을 이용하여 구한다.

2. 제안하는 방법

본 논문에서 제안하는 방법은 손의 형태 추출을 위한 영상획득 후의 전처리 과정, 손가락 및 손 중점 좌표 획득 과정, 마우스 좌표 계산 과정 등의 3가지의 정으로 구성되어진다.

2-1. 전처리 과정

전처리 과정은 손가락 끝점 및 손 중점 좌표 획득을 수행하며, 이 과정을 통하여 최종적으로는 손의 영

역만이 있는 영상을 구하게 된다. 전처리 과정은 첫 번째로 Kinect camera 원본 Depth 영상으로부터 최대값 및 최소값을 구하게 된다. 최대값 최소값은 Depth 영상의 정규화를 수행하기 위해 필요하다. 최대값과 최소값은 영상 전체의 픽셀 값들을 비교하면서 그림 1과 같은 과정을 거쳐 구한다.

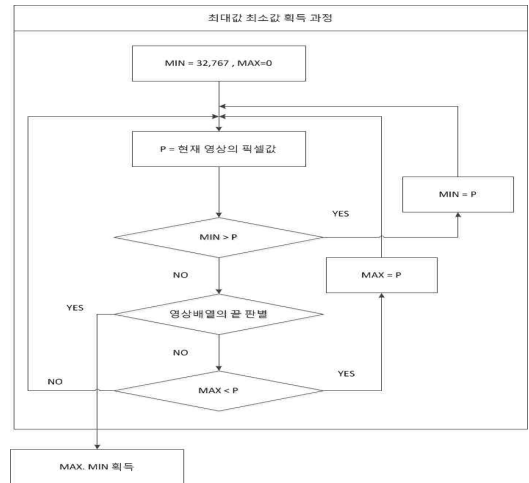


Fig 1. Procedure of obtaining maximum value and minimum value

그림 1. 최대값과 최소값 획득 과정

두 번째로 정규화, 이진화 과정을 거친다. 정규화는 그림 1에서 구한 최대값, 최소값을 이용하여 1차, 2차 이진화 영상에서 손 영역 영상을 추출하기 위해 수행한다. 정규화 과정은 다음과 같다. 그림 2의 Depth Image 원 영상은 각 픽셀 당 부호 없는 16비트의 크기를 가지는 영상이다. 따라서 원본 영상은 전반적으로 노이즈를 포함한 낮은 픽셀 값들로 이루어진 형상을 알아 볼 수 없는 영상이기 때문에 최소값과 최대값을 구한 후, 정규화 과정을 통해 식별 가능한 영상으로 처리한다. 원 영상의 픽셀 값을 A(x,y), 원 영상의 최대값을 max, 최소값을 min, 정규화된 영상의 픽셀 값을 B(x,y)라고 한다면 식 (1)과 같이 표현될 수 있다. 식(1)에서 32,767이란 값은 부호 없는 16비트 변수의 값의 최대값이며, 이 값은 구현 환경에 따라 바뀔 수 있다.

$$B(x,y) = A(x,y) / (\max - \min) \cdot 32,767 \quad (1)$$

한편, 그림 3은 정규화된 영상을 나타내고 있다.



Fig 2. Original depth image

그림 2. 원 Depth 영상



Fig 3. The normalized image

그림 3. 정규화 된 영상

정규화된 영상은 카메라와 화면상의 물체간의 거리가 가까울수록 낮은 픽셀 값들을 가지게 된다. 사용자는 마우스를 사용할 시 화면을 향해 손을 지시하므로 따라서 카메라로부터 가장 가까운 거리의 물체를 손으로 가정하고 이진화를 수행한다. 원본 영상의 픽셀 값을 $A(x,y)$ 로 하였을 경우에, 실험을 통해 카메라로부터 0.6m ~ 1m 사이 거리의 사람의 전신을 대략적으로 뽑아 낼 수 있는 이진화 문턱치 값은 식 (2)와 같은 범위이다. 이 범위에 값이 존재할 경우 따로 할당된 8비트, 1 채널 영상의 빈 공간에 255의 값을, 범위 이외의 값일 경우에는 0의 값을 저장하여 1차적으로 8비트 이진화 영상을 만든다. 그림 4는 1차 이진화된 영상을 나타내고 있다.

$$9000 < A(x,y) < 25000 \quad (2)$$



Fig 4. First processed binary image

그림 4. 1차 이진화된 영상

1차 이진화 영상에서 손의 영역만을 추출하기 위해

2차 이진화를 수행한다. 1차 이진화 과정 중에서 식 (2)의 범위 내에 있는 픽셀 값에서 가장 작은 값을 추출한 값을 \min , 원 영상의 픽셀 값을 $A(x,y)$ 이라고 할 때 2차 이진화 경계값은 식 (3)과 같다. 이 경계 조건은 손의 위치가 바뀌어도 대략적인 손의 형태 영상을 추출할 수 있다. 손 영역을 추출하기 위해 필요한 2차 이진화 영상 결과는 그림 5과 같다.

$$\min < A(x,y) < \min + 2000 \quad (3)$$



Fig 5. Second processed binary image

그림 5. 2차 이진화 영상

2-2. 손가락 끝점 및 손 중점 좌표획득 과정

손가락 끝점 및 손 중점 좌표획득 과정은 다음과 같다. 첫 번째로 손가락 중점을 구하기 위해 손 영역 영상에서 Distance Transform[8]을 수행한다. Distance Transform을 통한 중점 좌표 획득은 불규칙한 물체의 형상에도 비교적 정확한 손의 중점 좌표를 획득할 수 있다. 그림 6은 손 영역 영상을 Distance Transform을 수행하고 8비트 1채널 영상으로 정규화된 영상을 나타내고 있다.



Fig 6. The image performed distance transform

그림 6. Distance Transform이 수행된 영상

손의 중점은 손 영상의 외곽선 픽셀들 간의 거리가 서로 멀리 떨어져 위치하고 있다. Distance Transform

을 수행한 결과를 살펴보면 손의 중점은 영상에서 가장 밝은 값을 가진다. 따라서 결과영상에서 가장 밝은 픽셀 값을 찾으면 손의 중점이 된다. 손의 중점 좌표가 결정이 되면 손의 중점 좌표로부터 가장 멀리 떨어져있는 좌표를 손가락의 끝점으로 가정한다. 그러나 화면의 중심 좌표로부터 손가락이 멀어질수록 손 영역과 함께 팔도 영상에서 검출이 되어 팔 안쪽으로 손의 끝점으로 오인하여 인식될 수 있다. 그림 7은 팔이 검출되어 오인된 손의 끝 지점을 나타내고 있다.(파란색 점 : 오인하여 인식된 손의 끝점, 붉은 점 : 손의 중점)



Fig 7. In wrong recognition case of fingertip point

그림 7. 팔 방향으로 손끝이 잘못 인식된 경우

두 번째로 원활한 손가락 끝점 좌표 획득을 위해 팔 부분을 제거한다. 본 논문에서는 영상의 중심으로부터 손의 중점까지를 반지름으로 하는 원을 그려 팔 부분을 제거하였다. 그림 8은 검은색 원이 그려져 팔 부분이 제거된 영상을 나타내고 있다.



(a) 흰색 원이 검은색 원이 그려질 부분 (b) 팔 부분이 삭제된 영상
(a) The part of black circle (b) The removed arm image

Fig 8. Image that removed arm
그림 8. 팔 부분을 제거한 영상

2-3. 마우스 좌표 계산 과정

마우스 좌표 계산 과정에서는 손가락 끝점 및 손 중점 좌표획득 과정을 거쳐 획득한 손의 중점과 손의 끝점 좌표를 조건에 따라 선택한다. 그림 9와 같이 획득

된 2개의 좌표를 이용하여, 2개의 좌표 중에서 50 픽셀 조건에 따라 하나를 선택하여 마우스 커서 위치를 선정하는 좌표로 사용한다.

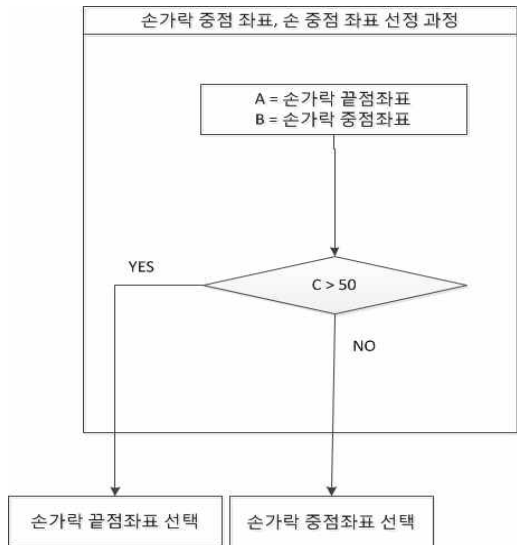


Fig 9. Selecting one point by condition
그림 9. 조건에 따른 하나의 좌표 선택

손가락 끝점 좌표와 손 중점 좌표 차의 절대 값이 50픽셀 이하일 경우에 손 중점 좌표를, 50픽셀 이상일 경우 손 끝점 좌표를 마우스 커서 위치 선정에 사용하는 이유는 다음과 같다. 손가락의 끝점이 카메라를 향해 있을 경우에 손의 형태는 그림 10과 같이 손의 중점으로부터 가장 멀리 떨어진 손의 끝점을 이용해 마우스 커서 위치 선정에 사용하기에는 매 프레임마다 불규칙한 형태를 보이고 잦은 떨림이 발생한다. 따라서 손의 중점과 손가락 끝점의 좌표가 일정 수치로 가까워질 때(50픽셀 이하) Distance Transform을 수행하여 획득한 손의 중점 좌표를 마우스 커서 위치 선정을 위한 좌표로 사용한다. 이는 불규칙한 물체의 형태에서도 비교적 손의 중점 좌표의 위치를 유지하기 때문이며 손의 끝점 좌표와 손의 중점 좌표의 위치가 비슷하기 때문이다. 또한 Depth Image의 전 해상도를 이용하여 선택된 좌표를 마우스 포인트의 위치 선정에 사용하면 화면의 모서리 위치에 가까이 마우스 포인트를 위치시킬 때 어려움이 있다.

마우스 커서의 위치 계산에 사용되는 좌표는 식 (4)와 같이 표현할 수 있다. 식 (4)에서 640x320 Depth 영상, 컴퓨터 해상도 1280x1024를 기준으로 dix는 Depth 영상 해상도의 가로 픽셀의 길이, diy는 세로의

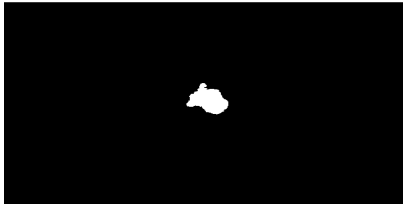


Fig 10. The image that fingertip toward screen
그림 10. 손가락 끝점이 화면을 향한 영상

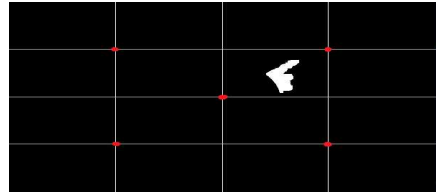


Fig 11. The crossed line by 5 points
그림 11. 5개의 좌표를 교차하여 그려진 선

길이, px, py는 Depth 영상의 손의 중점 좌표 혹은 손가락의 끝점 좌표이다.

$$\begin{aligned} x &= px / (dix/8 \cdot 7 - dix/8) \cdot 1280 \\ y &= py / (diy/8 \cdot 7 - diy/8) \cdot 1024 \end{aligned} \quad (4)$$

이렇게 구해진 x, y좌표는 컴퓨터의 해상도에 맞춘 식 (5)와 같이 컴퓨터 해상도 1280x1024를 기준으로 범위를 제한한다.

$$\begin{aligned} 0 &\leq x \leq 1280 \\ 0 &\leq y \leq 1024 \end{aligned} \quad (5)$$

3. 실험 결과

3-1. 실험 환경

본 실험을 진행한 환경은 Kinect camera가 설치된 Window 7 OS를 탑재한 i5 프로세서의 PC를 가지고 일반 실내 환경에서 진행하였다. 실험자와 Kinect camera간의 거리는 0.6m ~ 1m 사이로 제한하였다. 손가락의 끝점 좌표를 이용한 마우스 좌표와, 손의 중점 좌표를 이용한 마우스 좌표의 정확도 확인을 위해 그림 11과 같이 5개의 기준 좌표 (360, 256) (1080, 256) (360, 768) (1080, 768) (640, 512) 등을 선정하여 기준 좌표와 실제 손가락의 끝점 혹은 손의 중점 좌표를 이용하여 선정된 마우스 커서 좌표간의 오차를 각각 500개 씩 총 2500개의 수집한 좌표를 통해 계산했다. 각각의 좌표는 화면의 중앙의 1개의 좌표와 그 주변에 위치한 4개의 좌표이다.

3-2. 실험 결과

본 논문에서 제안하는 커서 위치 선정의 정확도를 측정하기 위해 임의의 5개의 기준 좌표를 선정하고 기준 좌표와 측정된 마우스 좌표사이의 오차를 측정하였다.

각각의 오차는 마우스의 좌표를 500회 반복하여 측정하였다. 그림 12, 13, 14, 15, 16는 기준 좌표 (360, 256), (1080, 256), (360, 768), (360, 256), (640, 512)에 대한 기준 좌표와 측정된 마우스 좌표사이의 오차를 나타내고 있다.

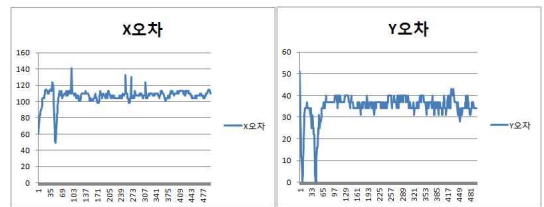


Fig 12. The error for between the reference points and the measured mouse points for the reference points (360, 256)

그림 12. 기준 좌표 (360, 256)에 대한 기준 좌표와 측정된 마우스 좌표사이의 오차

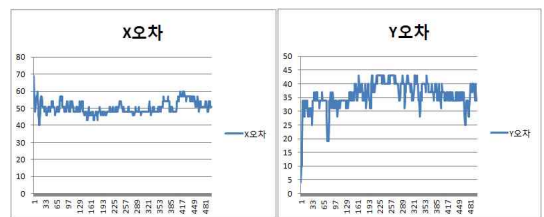


Fig 13. The error for between the reference points and the measured mouse points for the reference points (1080, 256)

그림 13. 기준 좌표 (1080, 256)에 대한 기준 좌표와 측정된 마우스 좌표사이의 오차

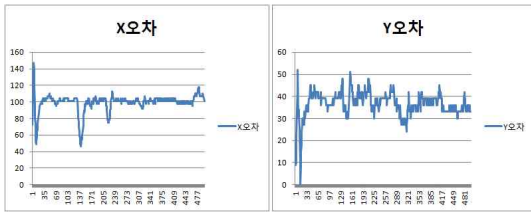


Fig 14. The error for between the reference points and the measured mouse points for the reference points (360, 768)

그림 14. 기준 좌표 (360, 768)에 대한 기준 좌표와 측정된 마우스 좌표사이의 오차

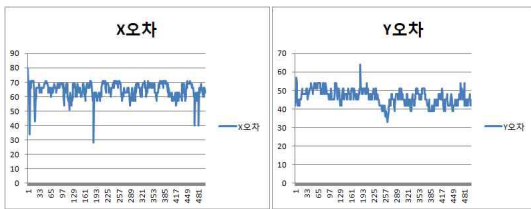


Fig 15. The error for between the reference points and the measured mouse points for the reference points (1080, 768)

그림 15. 기준 좌표 (1080, 768)에 대한 기준 좌표와 측정된 마우스 좌표사이의 오차

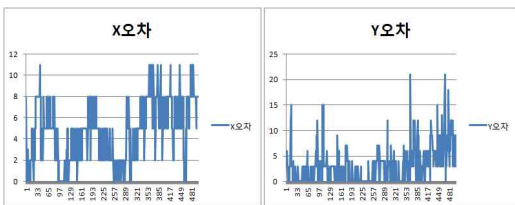


Fig 16. The error for between the reference points and the measured mouse points for the reference points (640, 512)

그림 16. 기준 좌표 (640, 512)에 대한 기준 좌표와 측정된 마우스 좌표사이의 오차

표 1, 2는 기준좌표 5개에 대해 측정된 마우스 좌표 사이의 오차율을 나타내고 있다. 표 1, 2에서 x좌표의 평균 오차율은 13.76%, 와 y좌표의 평균 오차율은

6.52%, 전체 좌표의 평균 오차율은 11% 미만으로 확인되어서 제안한 방법의 정확도를 확인할 수가 있었다.

Table 1. Each error rate for reference points

표 1. 기준좌표 5개에 대해 측정된 마우스 좌표사이의 오차율

기준좌표	오차율
(360, 256)	(29.6%, 10.6%)
(1080, 256)	(5.1%, 12.0%)
(360, 768)	(27.6%, 4.3%)
(1080, 768)	(5.9%, 5.6%)
(640, 512)	(0.6%, 0.1%)

Table 2. Error rate of between each measured mouse points for five reference points

표 2. 기준좌표 5개에 대해 측정된 마우스 좌표사이의 평균 오차율

x좌표의 평균 오차율	y좌표의 평균 오차율	전체 좌표 평균 오차율
13.76%	6.52%	10.14%

4. 결론

Kinect Depth 카메라에서 손가락의 방향을 이용한 마우스 커서의 위치 선정에 관한 새로운 방법을 제안하였다. 제안하는 방법은 손가락의 방향이 카메라를 향해 있을 경우는 거리변환을 이용한 손의 중점좌표를 이용하고, 손가락의 방향이 카메라를 향해 있지 않을 경우에는 손가락의 끝점 좌표를 이용하였다. 본 논문에서 제안하는 방법은 손의 형태 추출을 위한 영상획득 후의 전처리 과정, 손가락 및 손 중점 좌표 획득 과정, 마우스 좌표 계산 과정 등의 3가지의 과정으로 구성되어져 있다. 전처리 과정은 손가락 끝점 및 손 중점 좌표 획득을 수행하며, 이 과정을 통하여 최종적으로는 손의 영역만이 있는 영상을 구하였다. 손가락 끝점 및 손 중점 좌표획득 과정은 Distance Transform을 수행하고 1차와 2차 이진화된 영상을 추출하고 팔 부분을 제거한 영상을 추출하였다. 마지막으로 가락 끝점 및 손 중점 좌표획득 과정을 거쳐 획득한 손의 중점과 손의

끝점 좌표를 조건에 따라 마우스 커서 위치를 선정하게 되었다. 손가락의 방향이 카메라를 향해 있을 경우에는 손가락의 끝점 좌표와 손의 중점 좌표 사이의 거리가 가까워지므로 중점좌표를 이용하여, 위치 선정의 정확도를 향상시킬 수 있게 되었다. 본 논문에서 제안하는 커서 위치 선정의 정확도를 측정하기 위해 임의의 5개의 기준 좌표를 선정하고 기준 좌표와 측정된 마우스 좌표사이의 오차를 500회 반복하여 측정하여 에러율을 측정한 결과, 평균 11 % 미만의 오차율을 나타내어 제안한 방법의 정확도를 확인할 수가 있었다. 향후 연구 과제로는 단순히 손 끝점과 손 중점사이의 거리에 따른 좌표의 선택이 아닌 다른 접근방법의 연구가 필요하다. 단순히 손 끝점과 손 중점사이의 거리에 따른 좌표의 선택이 아닌 다른 방식의 접근방법이 필요하다.

References

- [1] Jungong Han, Ling Shao, and Dong Xu, Jamie Shotton, "Enhanced Computer Vision with Microsoft Kinect Sensor : A Review", Proc. IEEE TRANSACTIONS ON CYBERNETICS, Vol.43, No.5, pp.1318-1334, Oct, 2013.
- [2] Byeong-Ju Park, Jae-Heung Lee "High Efficient Viola_Jones Detection Framework for Real-Time Object Detection", Journal of IKEEE, Vol.18, No.1, pp.1-7, Jan, 2014.
- [3] Jae-Won Lee, Jee-Hoon Jung, Sung-Hoon Hong, "3D Multiple Object Detection and Tracking on Accurate Depth Information for Pose Recognition", Korea Multimedia Society, Vol.15, No.8, pp.963-976, Aug, 2012.
- [4] Xin Zhang, Zhichao Ye, Lianwen Jin, Ziyong Feng, and Shaojie Xu, "A New Writing Experience : Finger Writing in the Air Using a Kinect Sensor", Proc. MultiMedia IEEE, Vol.20, No.4 pp.85-93, Oct, 2013.
- [5] Junghwan Pyo, Namhyuk Kang, Jiwon Bang, Yongjin Jeong, "Hand Language Translation Using Kinect", Journal of IKEEE, Vol.18, No.2, pp.291-297, Jun, 2014.
- [6] Hanhoon Park, Junyeong Choi, Jong-Il Park, Kwang-Seok Moon, "A Study on Hand Region Detection for Kinect-Based Hand Shape", Journal of KOSBE, Vol.18, No.3, pp.393-400, May, 2013.
- [7] Marco Maisto, Massimo Panella, Luca Liparulo, Andrea Proietti, "An Accurate Algorithm for the Identification of Fingertips Using an RGB-D Camera", Proc. IEEE JOURNAL ON EMERGING AND SELECTED TOPICS IN CIRCUITS AND SYSTEMS, Vol.3, No.2, pp.272-283, Jun, 2013.
- [8] Junyeong Choi, Hanhoon Park, and Jong-il Park, "HAND SHAPE RECOGNITION USING DISTANCE TRANSFORM AND SHAPE DECOMPOSITION", 2011 18th IEEE International Conference on Image Processing, pp.3605-3608, Sep, 2011.

BIOGRAPHY

Bong-Hoe Goo (Student Member)



2010 ~ current : BS course of Electronics&Control Engineering, Hanbat National University

Seung-Ho Lee (Member)



1986 : BS degree in Electronic Engineering, Hanyang University
 1989 : MS degree in Electronic Engineering, Hanyang University
 1994 : Ph. D degree Electronic Engineering, Hanyang University
 1994 ~ current : Professor, Department of Electronics&Control Engineering, Hanbat National University