

# 손 지시 명령을 통한 임의의 카메라 배치에서의 가전기기 위치 인식

양승은<sup>1</sup>, 도준형<sup>2</sup>, 장효영<sup>3</sup>, 정진우<sup>4</sup>, 박광현<sup>5</sup>, 변증남<sup>6</sup>  
한국과학기술원 전자전산학과<sup>1,2,3,5,6</sup>  
한국과학기술원 인간친화 복지 로봇시스템 연구센터<sup>4</sup>  
{seyang<sup>1</sup>, jhdo<sup>2</sup>, hyjang<sup>3</sup>, jinwoo<sup>4</sup>}@ctrsys.kaist.ac.kr  
akaii@robotian.net<sup>5</sup> zbien@ee.kaist.ac.kr<sup>6</sup>

## Home Appliance Position Recognition through Hand Pointing Command for Arbitrary Camera Location

Seung-Eun Yang<sup>1</sup>, Jun-Hyeong Do<sup>2</sup>, Hyoyoung Jang<sup>3</sup>, Jin-Woo Jung<sup>4</sup>,  
Kwang-Hyun Park<sup>5</sup> and Zeungnam Bien<sup>6</sup>

Department of Electrical Engineering and Computer Science, KAIST<sup>1,2,3,5,6</sup>  
Human-friendly Welfare Robot System Engineering Research Center, KAIST<sup>4</sup>

### 요약

지능형 주거공간에서 손 지시 명령을 통하여 가전기기를 선택하거나 로봇에게 이동하여야 하는 장소를 알려 주기 위해, 기존의 시스템은 선택되는 대상 기기의 3 차원 절대 위치를 미리 알고 있어야 한다. 또한 카메라 위치가 변동되었을 경우, 카메라의 위치를 절대좌표계 기준으로 새롭게 측정해야 하는 불편함이 있다. 이를 해결하기 위해 본 논문에서는 팬/틸트 모듈을 가진 두 대의 USB 카메라를 임의의 위치에 배치하더라도, 두 번의 손 지시만으로 선택 대상이 되는 기기의 3 차원 위치를 파악하고 이를 동작시키는 방법을 다룬다.

제안하는 방법에서는 두 대의 카메라 간의 상대 좌표계를 형성하기 위해 각 카메라에 표식을 부착한다. 각 카메라에서 다른 카메라의 표식을 관찰하면 카메라 간의 거리 및 각도를 구할 수 있기 때문에, 하나의 카메라를 기준으로 3 차원 절대 좌표계를 자동으로 설정할 수 있다. 또한, 두 대의 카메라로 사용자의 얼굴과 손을 검출하면 얻어진 기준 좌표계에 대해 얼굴과 손의 3 차원 위치를 계산하고, 두 지점을 연결하는 방향 벡터를 구함으로써 사용자가 손으로 지시하는 방향을 찾는다. 따라서, 카메라를 임의의 위치에 두더라도 사용자의 손 지시 동작만으로 대상체의 차원 위치를 파악할 수 있게 된다.

개발된 시스템의 유용성을 검증하기 위해 각 가전기기의 위치를 제안한 방법으로 구하고 실제 위치와의 오차를 분석하였다. 제안한 방법은 두 대의 USB 카메라와 일반 PC 또는 마이크로 프로세서만으로 구현할 수 있기 때문에 비용이 적게 들고 실시간 처리가 가능하며 사용자의 환경에서 편리성을 높이는 등 많은 장점을 가진다.

**Keywords:** 3 차원 위치 인식, 손 지시 방향, 임의의 카메라 배치

### 1. 서론

첨단 기술의 발전과 더불어 인간의 삶의 질이 높아지면서 인간과 기계 사이의 인터페이스에 관한 연구가 활발히 이루어지고 있다. 특히 보다 인간 친화적이고 비용이 적게 드는 인터페이스의 개

발은 인간 중심적 관점과 경제적 관점에서 매우 중요한 연구 주제이다. 비전 시스템의 경우 카메라를 이용하여 영상을 획득한 후 중요한 영상정보를 이용하여 인터페이스를 제공할 수 있으므로 인간 친화적 인터페이스에 적합한 방식이라 할 수 있다[1]. 이 과정에서 가장 기초가 되는 것은 기준

좌표로 사용자의 위치를 알아내는 것이다. 정의된 좌표계를 기준으로 사용자를 검출하여 좌표 값을 획득하면, 사용자의 실루엣이나 중요한 신체의 부분을 검출 및 추적 하여 사용자의 명령을 인식할 수 있게 된다.

현재 한국과학기술원의 인간친화 복지 로봇시스템 연구센터(HWRS)에서 개발중인 지능형 주거공간은 노약자와 장애인을 대상으로 그들이 독립된 생활을 할 수 있도록 돕는 공간적 시스템이다. 소프트 리모컨 시스템은 팬/틸트 모듈을 갖춘 3대의 카메라를 이용하여 사용자의 얼굴과 손을 검출하고 미리 정의된 좌표계를 기준으로 3 차원 좌표를 계산한다. 사용자가 원하는 가전기기를 선택하기 위해서는 사용자의 손을 해당 기기 쪽으로 뺀 손과 얼굴의 방향벡터를 계산하여 그 연장선상에 있는 물체를 검출한다[2-6].



그림 1. 지능형 주거공간에 설치된 3대의 카메라

그러나 그림 1 에서 보듯 주거공간에 장착된 카메라는 고정이 되어 이동이 불가능하다. 따라서 카메라를 통한 사용자의 인식 영역에 한계가 있으므로 사용자의 명령 가능한 범위도 제한된다. 또한 고정된 절대좌표를 이용하므로 가전 기기의 3 차원 좌표를 따로 측정해야 한다.

이에 새로 제안하는 방식은 보다 경제적인 USB 카메라를 이용하고 임의의 위치에 있더라도 표식을 이용해 카메라간 상대좌표를 형성할 수 있다.

시스템에서 얼굴 및 손을 검출하고 추적하는 부분은 현재 시스템에서 사용하는 방식을 그대로 적용한다[6]. 검출된 얼굴과 손의 3 차원 좌표를 계산하면 얼굴과 손의 방향벡터를 서로 다른 두 지점에서 계산하여 두 벡터의 접점 혹은 수선의 중

점을 가전기기의 위치로 저장한다. 저장된 위치를 이용하여 손 지시 동작으로 가전기기를 선택할 경우 동작이 정확하지 않아 선택되지 않을 수 있는데, 이때 가장 가까운 위치에 있는 사물을 기준으로 떨어져있는 거리 정보를 제공하여 사용자의 동작을 수정하도록 한다.

2 절에서 시스템의 동작 환경을 설명하고, 3 절에서 표식을 이용한 두 카메라간 거리 및 각도 구하는 과정을 다룬다. 4 절에서는 얻어진 카메라간 거리 및 각도 정보를 이용하여 검출된 사용자의 얼굴과 손의 3 차원 좌표를 구하고 가전기기의 위치를 저장하는 방법을 살펴볼 것이다. 또한 손 지시명령을 통한 가전기기 선택과 잘못된 선택 시 사용자 명령을 수정하는 방법을 서술한다. 그리고 5 절에서 실제 지능형 주거공간에서의 실험 결과를 정리하고 6 절에서 결론 및 추후 연구과제를 살펴보기로 한다.

## 2. 시스템 동작 환경

시스템에서 사용하는 USB 카메라는 사용자 검출이 가능한 영역에서 두 카메라의 간격은 0.5m 이상으로 한다. 카메라간 간격이 너무 가까울 경우 스테레오 영상에서 얻는 정보가 제한적이기 때문에 오차가 크게 발생한다. 시스템 동작 시 손 지시 명령이란 대상물체의 중심이 손에 의해 가려지도록 물체 방향으로 손을 뺀 상태를 말한다. 또한 본 시스템은 가전기기뿐만 아니라 임의의 사물에 대해 적용 가능함을 밝혀둔다.

## 3. 두 카메라간 거리 및 각도 구하기

두 카메라가 임의의 지점에 설치되더라도 대상물체의 3 차원 좌표값을 계산하기 위해 그림 4 와 그림 5 에 보는 것과 같이 카메라간 거리 및 각도를 알아야 한다. 이를 알아내기 위해 표식을 이용하여 이번 절에서는 표식의 검출 과정과 표식을 이용한 두 카메라의 거리 및 각도를 구하는 과정에 대해 다루도록 한다.

### 3-1. 표식의 검출

표식을 통해 얻어야 하는 정보는 카메라간 거리 및 각도이다. 또한 사용하는 표식은 임의의 환경에서 쉽게 검출이 가능하고 필요한 정보를 얻을 수 있으며 크기에 제한을 두어야 한다. 이러한 특성을 고려하여 검출하기 쉬운 색상인 붉은색과 단순한 원으로 이루어진 표식을 사용한다. 그림 2에 표식의 인식 과정을 나타내었다.

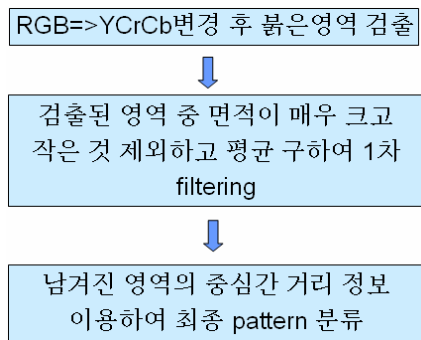
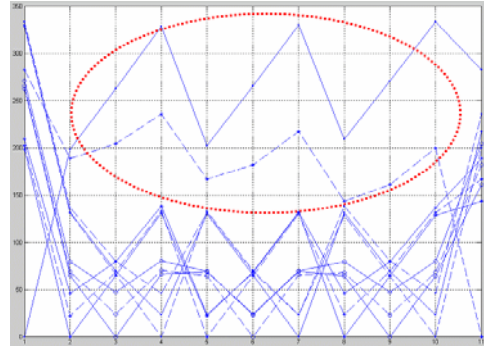


그림 2. 표식 인식 과정

USB 카메라를 이용하여 얻는 영상은 RGB이다. 그러나 RGB의 경우 조명에 대한 정보가 포함되어 있으므로 주위 환경에 영향을 크게 받는다. 따라서 획득하는 영상을 밝기 정보가 포함된 영상과 붉은색, 푸른색 중심의 영상으로 구성된 YCrCb로 변경한 후, 밝기정보에 해당하는 Y값은 버리고 Cr(붉은색 영역)과 Cb(푸른색 영역)로 분리한다. Cr과 Cb 영상에서 적절한 문턱값을 이용하여 표식의 붉은 영역을 검출한다. 이때 표식 외 붉은 영역이 함께 검출되기 때문에 이를 제거해 주어야 한다. 우선 검출된 영역의 넓이가 극단적으로 큰 것과 작은 것을 제외하고 평균을 구해 이 값으로 넓이가 평균에 가까운 것만 남기고 나머지는 제거한다. 1차 과정을 거친 후 각 붉은 영역의 중심간 거리 정보를 이용하여 특성이 다른 영역은 제거하고 원하는 영역만 검출할 수 있다. 그림 3(a)에서 위에 나타나는 두 그래프는 다른 그래프와 확연히 구분되는 것을 확인할 수 있다. 이런 특성을 이용하여 최종적으로 표식의 붉은 원만을 검출하게 된다.



(a) 검출된 붉은 영역간 거리 특성



(b) 최종 검출된 표식

그림 3. 붉은 영역의 중심간 거리특성 이용하여 표식 검출

### 3-2. 거리 및 각도 계산

검출한 표식을 이용하여 거리 및 각도 정보를 구하기 위해 우선 사용하는 카메라의 초점거리를 구해야 하는데 GML MatLab Camera Calibration Toolbox와 MATLAB™ (MATHWORKS, USA)을 이용하여 계산하였다[8].

표식에서 각 원의 중심간 거리는 알고 있으므로 초점거리를 알면 그림 4에서 보는 바와 같이 비례식을 이용하여 카메라로부터 표식까지의 거리를 알 수 있다.

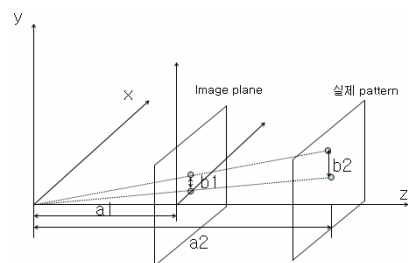


그림 4. 비례식을 이용한 거리 계산

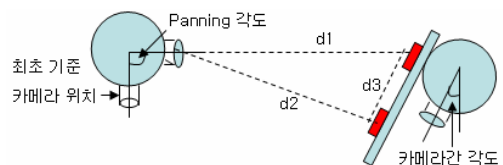


그림 5. 표식을 이용한 각도 계산

카메라로부터 표식까지의 거리를 구할 수 있게 되면 표식의 첫 번째 열과 마지막 열의 길이를 구해 그림 5 에서와 같이 하나의 삼각형을 형성할 수 있다. 이때 코사인 제 2 법칙을 이용하면 내각의 크기를 구할 수 있으므로 팬/틸트 변화를 함께 고려하여 두 카메라간의 각도를 구한다.

#### 4. 손 지시 명령을 이용한 가전기기 선택

##### 4-1. 3차원 좌표 계산

사용자가 가전기기를 저장하고 선택하기 위해 우선 사용자의 얼굴과 손의 3 차원 좌표를 구해야 한다. 손과 얼굴의 검출은 기존의 시스템에서 사용하는 방식을 그대로 적용하고 앞에서 구한 두 카메라간 거리 및 각도 정보를 이용하여 각 영상에서 나타나는 얼굴, 손의 중심 값을 절대좌표 기준으로 변경한다[6]. 각 카메라에서 획득한 영상의 값을 절대좌표 기준으로 변경한 후 그림 6 의 각 카메라의 중심  $C_{M,1}, C_{M,2}$  과 손 혹은 얼굴의 중심 점  $P_1, P_2$  을 지나는 선  $s_1, s_2$  를 구한다. 이때 두 선에 수직인 선분  $\overline{P_{M,1}P_{M,2}}$  의 중점  $P_M$  이 구하고자 하는 3 차원 좌표의 근사값이 된다[6,7].

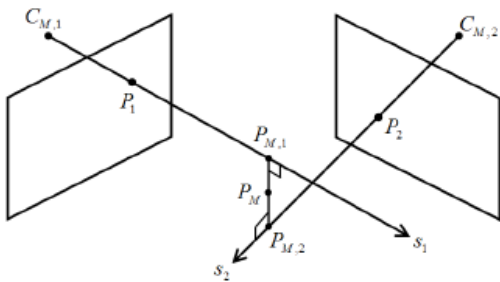


그림 6. 3차원 좌표 계산

##### 4-2. 방향벡터를 이용한 가전기기 위치 저장

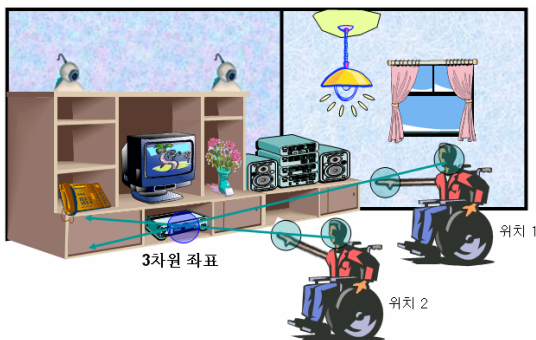


그림 7. 사용자에게 의한 가전기기 위치 저장

4-1 절에서 기술한 과정을 통해 사용자의 손과 얼굴의 3 차원 좌표 값을 구하면 이 값을 이용하여 얼굴을 시작점으로 하는 방향벡터를 구할 수 있다. 3 차원 좌표를 계산하는 방식을 적용하여 사용자가 그림 7 의 위치 1 에서 손 지시명령을 한 후 방향벡터를 저장하고 다른 위치로 이동하여 동일한 동작을 수행한다. 이때 두 대의 카메라는 팬/틸트 기능을 이용하여 이동하는 사용자를 추적하고 팬/틸트 변화에 따라 기준 좌표 관계를 수정해 주어야 한다.

##### 4-3. 손 지시 명령을 통한 가전기기 선택 및 잘못된 명령에 대한 정정

사용자가 손 지시 명령으로 가전기기를 선택하는 방법은 4-2 절의 가전기기 위치 저장 과정과 동일하다. 사용자의 얼굴과 손의 중심점을 찾아 방향벡터를 구한 후 그 연장선상에 있는 가전기기를 선택하게 된다. 우선 사용자의 얼굴 중심으로 부터 저장된 가전기기까지의 거리를 구한다. 3 개의 가전기기 위치가 저장되었다면 3 개의 거리를 구한 후 단위 방향벡터에서 해당 거리만큼 연장하고 해당하는 가전기기의 범위 내에 있는지 확인하여 선택을 한다. 그림 8 에서는 방향벡터를 비디오, 오디오, 전등까지의 길이만큼 각각 연장했을 때의 예를 나타낸다. 사용자가 선택하려는 가전기기가 아닌 경우 방향벡터를 연장하였을 때, 길이뿐만 아니라 방향에 차이가 나기 때문에 선택되지 않는다. 아래의 예에서는 비디오만이 해당 범위 안에 속하게 되므로 최종적으로 비디오가 선택된다.

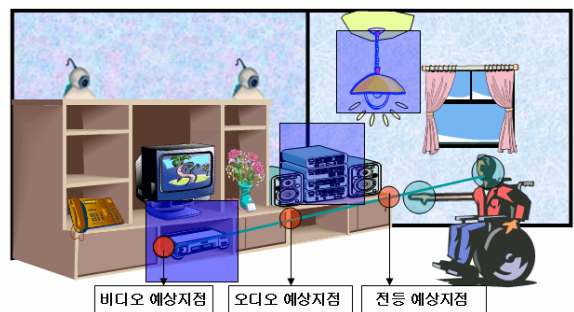


그림 8. 손 지시명령을 통한 가전기기 선택

그러나 구현하는 시스템의 경우 주위 환경에

민감한 영상 정보를 이용하고 사용자의 손 지시 명령 과정이 정확히 이루어지기 힘들기 때문에 사용자의 의도대로 가전기기가 선택되지 않을 수 있다. 이럴 경우 사용자의 명령 방향에서 가장 가까운 기기를 선택한 후 어느 방향으로 얼마큼 이동해야 하는지 지적해 주는 과정이 필요하다.

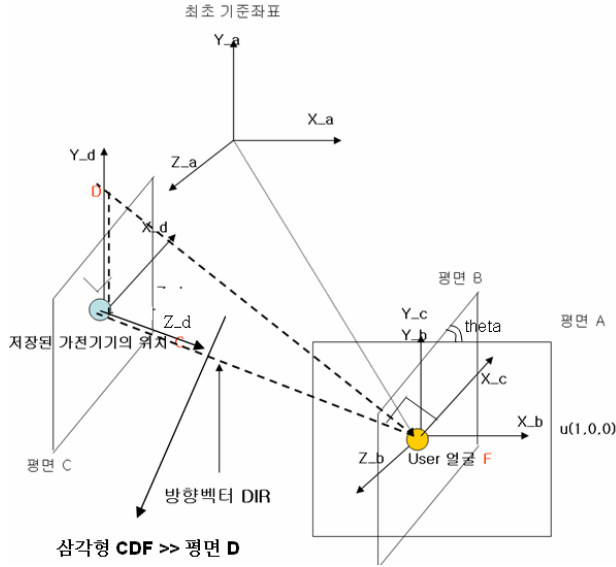


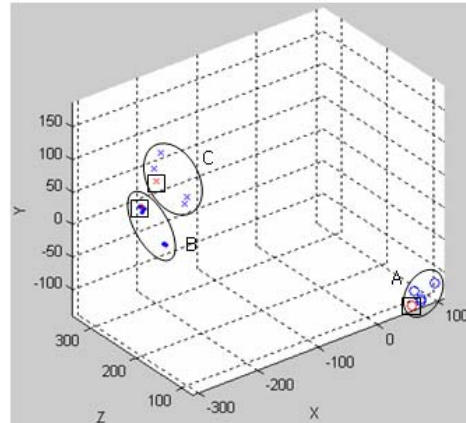
그림 9. 사용자의 잘못된 손 지시명령 수정을 위한 사용자 중심의 좌표 변환

사용자가 명령을 내리는 방향은 항상 변화하므로 수정해 주는 값 역시 사용자가 보는 방향을 기준으로 변환해 주어야 한다. 그림 9 에서 저장된 가전기기의 위치를 원점으로 하고 절대좌표의 y 축을 회전하여 새로운 좌표가 형성되는 것을 볼 수 있다. 이때 사용자의 손 방향 수정을 위해 시스템은 평면 C 상으로 손의 3 차원 위치를 사영하고 평면상에서의 나타나는 편차를 이용한다. 최초의 기준좌표로부터 저장된 기기 위치까지의 방향 벡터는 알고 있으므로 회전하는 각도를 구해야 한다. 사용자의 손 지시명령을 통해 형성되는 방향 벡터는 기준좌표의 y 축과 반드시 수직하지는 않지만, 이에 수직인 방향벡터가 포함된 가상의 평면 D (삼각형 DCF)를 생각할 수 있다. 그림에서 우리가 원하는 값은  $\theta$  이므로 평면 D 와  $u(1,0,0)$ 의 각도를 구하면  $90^\circ + \theta$ 의 값을 구할 수 있다. 따라서 최초의 기준 좌표에서 방향벡터와 변환 각을 이용하여 사용자가 바라보는 방향을 기준으로 새로운 좌표를 형성할 수 있다.

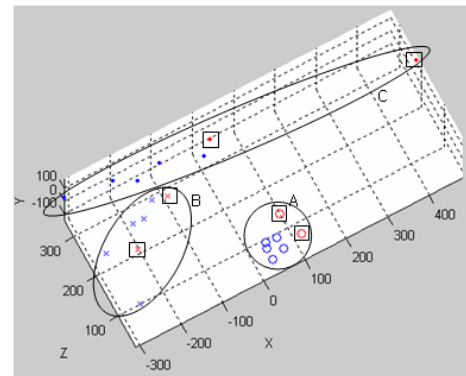
## 5. 실험 결과

### 5-1. 가전기기의 3차원 위치 저장

손 지시명령을 통해 가전기기를 선택함에 앞서 가전기기의 위치를 정확히 저장해야 한다. 따라서 사용자의 손 지시 명령을 통한 3 차원 위치 저장 결과를 그래프로 나타내었다.



(a) 실제 좌표와 계산된 결과 비교



(b) 2명의 사용자에 대한 좌표 계산

그림 10. 손 지시 명령을 통한 3차원 위치 저장

그림 10(a)의 경우 사각형으로 표시된 부분이 실제 기기의 위치이며 나머지 부분이 계산된 결과를 나타낸다. 이때 가전기기의 종류에 따라 A, B, C의 그룹으로 나누어 표시하였다. 그림 10(b)는 서로 다른 사용자에 대한 실험 결과를 나타내었다. 그래프에서 나타나는 특성을 보면 사용자에 따라 실험 결과에 편차가 나타나는 것을 알 수 있다. 또한 다수의 실험 결과, 실험 시마다 결과에 큰 차이가 나타났는데 이는 주위 환경(조명 등) 및 사용자의 명령 과정에서의 오차에 많은 영향을 받기 때문이다.

## 5-2. 손 지시 명령을 통한 기기 선택

저장한 위치에 손 지시 명령을 내렸을 경우 저장된 사물을 검출하는지 실험하기 위해 한번의 실험 과정에서 5 회씩 명령을 내렸으며 표 1의 결과는 총 10 번의 결과를 합하여 나타낸 것이다.

	기기 A	기기 B	기기 C
성공률	96%(48/50)	84%(42/50)	88%(44/50)

표 1. 사용자에게 의한 가전기기 선택 성공률

시스템에 의한 명령 수정의 정확도를 알아보기 위하여 기기가 선택되지 않았을 경우 가장 가까운 기기를 검출하는지 여부를 보고 수정하는 방향이 정확한지 표 2에 정리하여 나타내었다.

	최소거리 있는 사물 검출	수정 방향
성공률	89% (89/100)	83% (83/100)

표 2. 사용자 명령 수정의 정확도

## 6. 결론 및 추후 연구

본 논문에서는 두대의 USB 카메라를 임의의 위치에 배치하더라도 사용자의 손 지시 명령을 통해 특정 사물의 3 차원 위치를 저장하고 선택하는 문제를 다루었다. 카메라간 거리와 각도 정보를 얻기 위해 붉은색 원으로 이루어진 표식을 사용하였으며 카메라의 팬/틸트를 고려하여 3 차원 좌표를 계산하였다. 지능형 주거공간에서 각각 세 곳의 위치를 지정하여 손 지시 명령을 통해 위치를 저장하고 검출 실험을 한 결과 80% 이상의 성공률을 얻을 수 있었다. 그러나 사용자에게 따라 계산되는 3 차원 좌표에 차이가 나타나고 조명을 비롯한 주위 환경에 따라 큰 영향을 받으므로 이에 대한 보완이 필요하다. 또한 표식을 통한 거리와 각도 계산 과정에서도 오차가 나타나는데 이를 줄이기 위한 표식 선택과 정보획득 과정의 개선이 필요하다.

### 감사의 글

본 연구는 과학기술부/한국과학재단 우수연구센터육성사업의 지원으로 수행되었음(R11-1999-008).

본 연구는 BK21 산학과제 프로그램의 지원으로 수행되었음

## 참고문헌

- [1] Matthew Turk, "Computer Vision in the Interface," *Communications of the ACM*, vol. 47. no. 1, pp. , January 2004
- [2] Jun-Hyeong Do, Jung-Bae Kim, Kwang-Hyun Park, Won-Chul Bang and Z.Zenn Bien, "Soft Remote Control System using Hand Pointing Gesture," *International Journal of Human-friendly Welfare Robotic Systems*, vol. 3, no. 1, pp. 27-30, March 2002
- [3] Hyoyoung Jang, Jun-Hyeong Do, Jinwoo Jung, Kwang-Hyun Park, and Z. Zenn Bien, "View-invariant Hand-posture Recognition Method for Soft-Remocon System," *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2004)*, Alberat, Canada, pp.295-300 , September 28-October 2, 2004
- [4] Z. Zenn Bien, Kwang-Hyun Park, Jin-Woo Jung and Jun-Hyeong Do, "Intention Reading is Essential in Human-friendly Interfaces for the Elderly and the Handicapped," *IEEE Transactions on Industrial Electronics*, vol. 52, no. 6, pp. 1500-1505, 2005
- [5] Dimitar H. Stefanov, Z. Zenn Bien, and Won-Chul Bang, "The Smart House for Older Persons and Persons With Physical Disabilities: Structure, Technology Arrangements, and Perspectives," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 12, no. 2, pp. , June 2004
- [6] Jun-Hyeong Do, Hyoyoung Jang, Sung Hoon Jung, Jinwoo Jung and Zeungnam Bien, "Soft Remote Control System in the Intelligent Sweet Home," *International Conference on Intelligent Robots and Systems*, Edmonton, Canada, pp. 2193-2198, August 2-6, 2005
- [7] M. Kohler, "Vision Based Remote Control in Intelligent Home Environments," *Proceedings of 3D Image Analysis and Synthesis*, Erlangen, Germany, pp. 147-154, November 18-19, 1996
- [8] Vladimir Vezhnevets, *GML Matlab Camera Calibration Toolbox*, Graphics & Media Lab, Dept. of Computer Science, Moscow State University, November 1, 2005