

# ubiHome 환경 제어를 위한 비전기반의 3차원 공간센서\*

홍동표, 우운택  
광주과학기술원 정보통신공학과 U-VR 연구실  
{dhong, wwoo}@kjist.ac.kr

## A Vision-based 3D Space Sensor for Controlling ubiHome Environment

Dongpyo Hong, Woontack Woo  
KJIST U-VR Lab.

### 요약

본 논문에서는 인간 중심의 새로운 주거환경(ubiHome)에서 거주자의 제스처를 파악하기 위한 인터페이스로 비전기반의 3차원 공간센서를 제안한다. ubiHome 환경은 거주자에게 필요한 서비스를 제공하기 위해서 다양한 센서(Vision, IR, RF 등)를 활용하여 환경맥락정보(SW1H)를 생성한 뒤, 거주자의 의도를 직,간접적으로 파악한다. 하지만 일반적인 비전기반의 사용자 행동 인식 시스템들은 복잡한 알고리즘을 사용하고 그에 따른 많은 계산량 때문에 실시간성이 결여되어 있다. 따라서 본 논문에서는 사용자의 제스처를 파악하기 위해 복잡한 인식 알고리즘을 사용하는 대신 간단히 감지된 공간센서만을 이용하여 사용자의 제스처를 판단한다. 제안된 공간센서는 향후 거주자의 주변 공간 변화에 따라 다양한 기능을 하는 센서로 활용될 수 있다.

Keyword : Background Subtraction, Space Sensor, ubiHome, Context

### 1. 소개

경제성장과 더불어 현대인들이 가정에서 보내는 시간이 증가되면서 지능형 가정(SmartHome)환경과 관련한 다양한 연구가 진행 중이다[1][2]. 기존의 지능형 가정의 연구들은 대부분 Home Automation의 형태이며, 이러한 지능형 가정 환경들의 단점은 거주자가 환경으로부터 서비스를 제공받기 위해서 거주자의 직접적인 환경 제어가 필요하다. 하지만 ubiHome 환경은 거주자에게 필요한 서비스를 제공하기 위해서 다양한 센서(Vision, IR, RF 등)를 활용해서 환경맥락정보(SW1H)를 생성한 뒤, 거주자의 의도를 직,간접적으로 파악한다[2]. 이러한 ubiHome 환경과 거주자간의 자유로운 상호작용을

위해서는 새로운 형태의 센서를 이용하는 인터페이스가 필요하다. 기존의 AwareHome (GATECH), EasyLiving (Microsoft)등에서 사용된 비전 시스템의 경우에는 거주 환경내에서 사용자의 위치정보를 획득하는데 사용되었다[2][3][4]. 그리고 일반적인 비전기반의 사용자의 행동 인식 시스템들은 복잡한 알고리즘과 그에 따른 많은 계산량 때문에 실시간성이 결여되어 있다[5][6][7][8]. 따라서 새로운 ubiHome 환경에 적합한 비전기반의 3차원 공간센서를 활용하는 방법을 제안한다.

제안된 비전기반의 3차원 공간센서는 Ishii의 "Tangible Bit"에서 제안한 "Ambient Media"처럼 주변 환경(공기, 소리, 물등)을 매개체로 하는 인터

\* 본 연구는 교육부의 BK21 사업과 ICU 디지털 미디어센터 지원에 의해 수행됨

페이스들중 하나이다[9]. 이는 사용자와 환경간의 상호작용에 있어서 불편하고 복잡한 제어 기기를 사용자로부터 제거하고, 사용자와 환경간의 관계를 보다 친숙하게 만들어 주는 역할을 한다. 따라서 제안된 비전기반의 3 차원 공간센서를 통해서 사용자는 자신의 자유로운 동작을 통해서 자신의 주변환경과 상호작용할 수 있다.

제안된 방법은 3 차원 카메라를 이용하여 다수의 영상으로부터 거주자의 3 차원 위치정보를 획득한 뒤, 주거환경으로부터 사용자만을 분리한다. 그리고 분리된 사용자를 중심으로 가상의 공간을 분할한다[10][11][12]. 이렇게 분할된 3 차원 가상 공간을 목적에 맞도록 공간센서로 지정한다. 거주자는 센서로 지정된 공간센서를 직접 접촉함으로써 자신 의도를 ubiHome 환경에 전달한다. 공간센서의 감지성능은 배경으로부터 거주자의 분리 정도와 영상으로부터 획득한 깊이 맵의 정확성에 비례한다.

제안된 비전기반의 3 차원 공간센서는 거주자의 주변 공간을 인터페이스로 활용함으로써 복잡한 센서 기기들을 제거하고 거주자의 행동에 자유로움을 제공한다. 또한 사용자의 제스처를 파악하기 위해 복잡한 행동 인식 알고리즘을 사용하는 대신 간단히 감지된 공간센서만을 판단해서 거주자의 제스처를 파악한다.

본 논문의 구성은 다음과 같다. 2 장에서는 제안된 시각기반 사용자 인터페이스에 대해 설명한다. 그리고 3 장에서는 제안한 시스템을 응용한 예를 보여주고, 4 장과 5 장에서는 실험결과와 결론 및 추후 과제에 대해 언급한다.

## 2. 시각기반 사용자 인터페이스

본 논문에서 제안한 비전기반의 3 차원 공간센서를 사용하기 위해서는 자연스러운 주변환경(사용자의 주거환경)으로부터 사용자만을 분리하는 일이 선행되어야 한다. 일반적인 배경분리 기법에는 RGB 컬러 공간만을 활용한 컬러밀도의 차이, 움직임은 물체의 속도차이, 혹은 스테레오 영상의 차이를 이용하는 방법등이 있다. 하지만 기존의 방법들에서는 각 방법들이 갖는 단점들이 존재하

며 또한 대표적인 단일 문턱값을 사용하는 단점이 있다. 따라서 본 논문에서는 일반적인 주거환경에서 사용자만을 분리하기 위해서 통계적 기반의 배경학습과 동적인 문턱값을 이용하는 객체분할(배경분할) 기법을 이용한다[10]. 그림 1 은 배경으로부터 사용자를 분리하는 알고리즘이다.

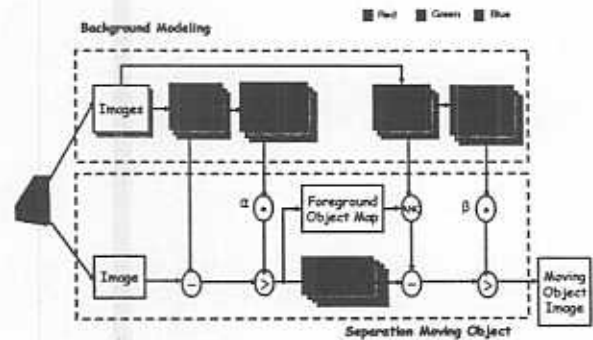


그림 1. 통계적 기반의 배경 분리 기법

그림 1 에서 M 과 S 는 각각 RGB 컬러모델에서 평균영상과 표준편차영상을 나타내며, m 과 s 는 각각 정규화된 rgb 컬러모델에서 평균영상과 표준편차영상을 나타낸다. 그리고  $\alpha$  와  $\beta$  는 각각의 표준편차에 대한 임의의 스케일 상수값이다. '-' 는 레퍼런스 영상과 현재 영상과의 차이를, '>' 는 두 영상간의 크기비교를 나타내는 연산이다.

제안된 객체분할 기법은 일반적인 RGB 컬러 공간뿐만 아니라 정규화된 rgb 컬러 공간도 활용하여 배경으로부터 1 차 분리된 물체의 그림자 부분을 제거 할 수 있다. 따라서 객체분할 과정을 통해서 획득된 정보와 3 차원 카메라로부터 획득한 깊이 맵 정보를 이용하여 사용자의 3 차원 위치정보를 얻는다. 제안된 비전기반의 3 차원 공간 센서를 생성하기 위해서 자연스러운 주거환경으로부터 사용자만을 분리하는 일은 기본적으로면서도 중요하다.

제안된 공간 센서 디자인은 사용자에게 자연스러운 움직임 제공과 정확한 사용자 제스처 감지를 위해서 중요하다. 사용자에게 자연스러운 움직임을 제공하기 위해서 공간 센서를 특정 위치에 고

정시키는 것 대신 사용자를 중심으로 동적으로 위치하도록 한다. 우선 사용자 세그멘테이션 과정을 통해서 획득한 사용자의 3 차원 위치 정보를 이용하여 사용자의 중심점( $H_c$ )을 다음과 같이 구한다.

$$H_c = \frac{1}{I_c} \sum_{i=1}^{I_c} D_c(i) \quad (1)$$

위 식에서  $D_c$  는 사용자 세그멘테이션 과정에서 획득된 사용자의 3 차원 위치 좌표를 나타내고,  $I_c$  는 분리된 사용자에 대응하는 좌표 값들의 갯수이다.

식(1)을 통해서 획득된 사용자의 중심점으로부터 공간 센서 생성에 필요한 좌표 정보를 다음과 같이 구한다.

$$S_x = S_y = S_z \equiv D_{Y-Top} - D_{Y-Bottom} \quad (2)$$

위 식에서  $S_x, S_y, S_z$  는 각각 사용자 중심으로 형성된 공간센서의 *Width, Height, Depth* 를 나타낸다. 그리고  $D_{Y-Top}, D_{Y-Bottom}$  는 각각 사용자 세그멘테이션과정을 통해서 획득된 영상에서 사용자가 차지하는 위쪽과 아래쪽 부분에 해당하는 좌표점들이다. 식(2)는 레오나르다다빈치의 "The Vitruvian man" 의 인체구성에 기반한다[11]. 식(1)과 식(2)를 통해서 제안된 공간 센서는 세그먼트된 사용자 주변에 그림 2 처럼 생성된다[12].

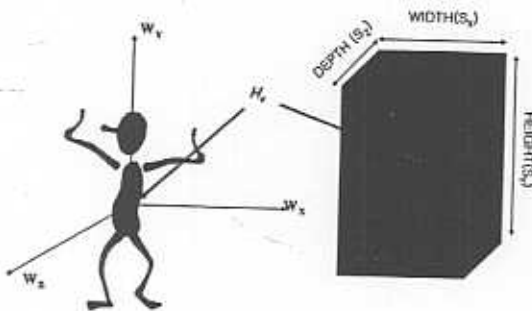


그림 2. 세그먼트된 사용자 중심의 동적 공간센서

제안된 비전기반의 3 차원 공간 센서의 정확한 사용자 제스처 감지와 사용자 제스처의 3 차원 위치 정보를 활용하기 위해서 사용자 주변에 생성된

가상의 공간을 그림 3 처럼 격자 형태의 분할하고 적절한 공간을 공간 센서로 지정한다

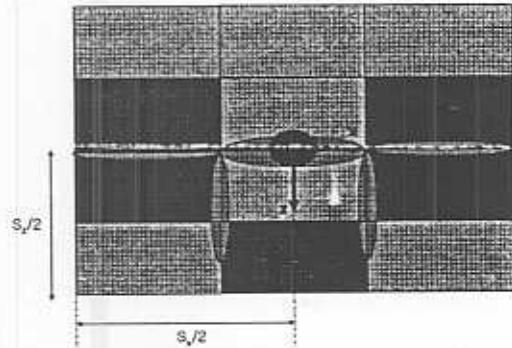


그림 3. 공간센서의 배치

그림 3 에서 SS1, SS2 그리고 SS3 는 공간 센서들 중에서 사용자의 제스처를 감지할 수 있는 영역이다. 그리고 이 영역을 사용자가 직접 접촉함으로써 공간 센서를 이용한다. 특히, SS1 과 SS2 영역에서는 사용자의 Z 축 방향의 움직임을 감지할 수 있다. 따라서 기존의 2 차 영상에 기반한 비전 시스템들이 갖는 한계를 극복하고, 보다 다양한 사용자의 제스처를 감지할 수 있다. 제안한 공간 센서는 사용자가 환경과 상호작용을 하기위해서 전형적인 GUI (Graphical User Interface) 보다 사용자의 의사를 효과적으로 전달할 수 있는 3 차원 인터페이스이다.

### 3. Case Study : ubiHome 환경 제어

본 논문에서 제안된 비전기반의 3 차원 공간센서를 ubiHome 환경에서 사용하기 위해서는 공간센서와 ubiHome 환경을 연결하는 인터페이스가 필요하다. 유비센서(ubiSensor)와 유비서비스(ubiService)로 구성된 ubiHome 환경에서는 유비 센서들로 사용자의 주변환경 정보를 수집한 뒤, 유비서비스에 수집된 환경 정보를 반영하여 사용자에게 적절한 서비스를 제공한다. 따라서 사용자의 주변환경 정보는 ubiHome 환경에서는 중요한 요소이다. 그림 4 는 ubiHome 환경에서 유비센서와 유비서비스로 구성된 모델이다[13].

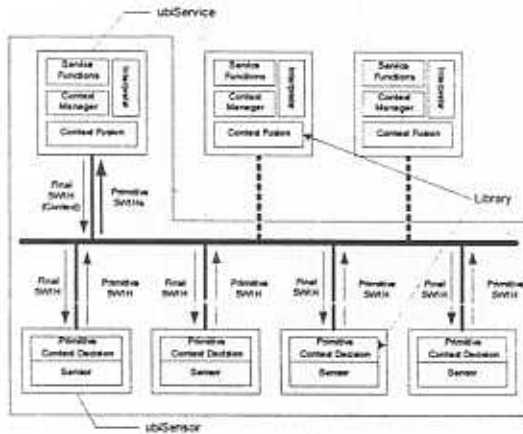


그림 4. 유비센서와 유비서비스 모델

그림 4 와 같이 유비센서는 모든 유비서비스가 사용할 수 초기 환경정보(primitive SWiH)를 생성함으로서 임의의 유비서비스와 연결이 가능하다. 또한 유비서비스는 여러 개의 초기 환경정보로부터 서비스를 제공하기 위한 최종 환경정보(final SWiH)를 결정한다. 이러한 의미에서 본 논문에서 제안하는 공간센서는 유비센서와 일맥상통한다.

본 논문에서 제안한 비전기반의 3 차원 공간센서의 응용으로 UbiHome 환경에서 Movie Player 를 제어하는 시스템에 적용하였다. 그림 4 에서 제안된 비전기반의 3 차원 공간센서는 유비서비스 혹은 다른 유비센서들로부터 필요한 정보(사용자, 영화제목 등)를 전달 받는다. 그리고 사용자가 공간센서로 지정된 영역을 접촉하게 되면, 사용자의 움직임 정보(Hand-up, Hand-down 등)가 유비서비스로 전달된다. 유비서비스는 공간센서뿐만 아니라 다양한 센서들로부터 전달된 정보들을 통합, 분석하여 Movie Player 응용프로그램에게 적절한 제어 정보를 전달한다. 그리고 제안된 공간센서로 Movie Player 를 제어하기위해서 간단한 상태를 공간센서에 두었다.

본 응용 시스템에서는 사용자에게 자신이 Movie Player 를 제어하고 있다는 것을 직관적으로 제공하기 위해서 시각적 피드백과 청각적 피드백을 동시에 사용자에게 제공한다. 따라서 사용자는 자신이 제어하는 것을 직관적으로 보면서 혹은 사운드를 통해서 들으면서 현재 자신이 제어하는 Movie Player 의 상태를 알 수 있다. 이는 사용자

계 보다 몰입감과 실제감을 제공한다.

#### 4. 실험결과

제안된 공간센서의 정확성은 기본적으로 객체분할 알고리즘의 강건함에 기인한다. 제안한 공간센서를 구성하기 위해서는 사용자의 3 차원 위치정보가 필요하다. 따라서 3 차원 카메라로부터 획득된 다수의 영상정보와 객체분할을 통해서 획득된 사용자 정보를 이용하여 사용자의 3 차원 위치정보를 추출한다. 다음은 제안된 객체분할 기법을 이용하여 사용자를 배경으로부터 분리한 실험결과이다.

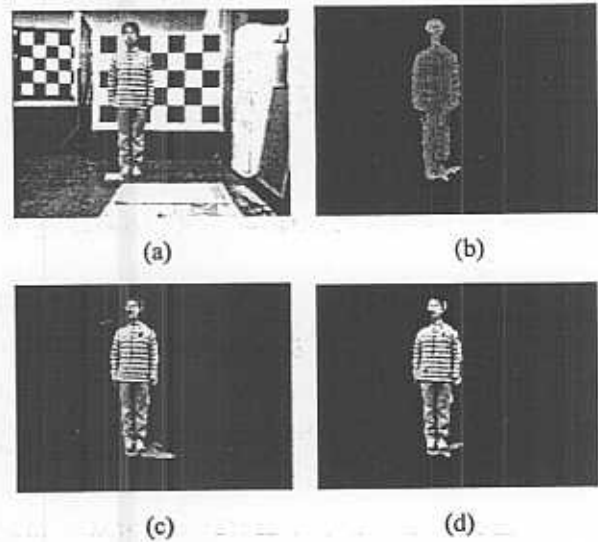


그림 5. 사용자 세그멘테이션

그림 5 의 (a)는 입력 영상, (b)는 세그멘테된 사용자의 깊이 맵 영상, (c)는 그림자를 포함한 사용자 세그멘테이션 영상, (d)는 그림자를 제거한 사용자 세그멘테이션 영상이다. 그림 5 의 (b)에서 보듯이 3 차원 정보를 포함하는 깊이 맵 영상에서 분할된 객체 정보를 이용하면 보다 깨끗한 3 차원 영역 정보를 획득할 수 있다.

제안하는 3 차원 비전기반의 공간센서는 환경과 사용자간의 상호작용을 보다 자연스럽게 효과적으로 수행할 수 있는 인터페이스로 사용된다. 편리한 인터페이스를 사용자에게 제공하기 위해서는 공간센서를 사용자 주변의 적절한 위치에 증강되도록 한다. 제안된 공간센서는 사용자의 포즈와 위치에 변화에 따라 동적으로 생성된다. 따라서 사

용자의 3 차원 위치 정보를 공간센서에 실시간으로 반영될 수 있도록 한다. 다음은 다양한 포즈에 따른 공간센서의 증강을 실험한 결과이다.

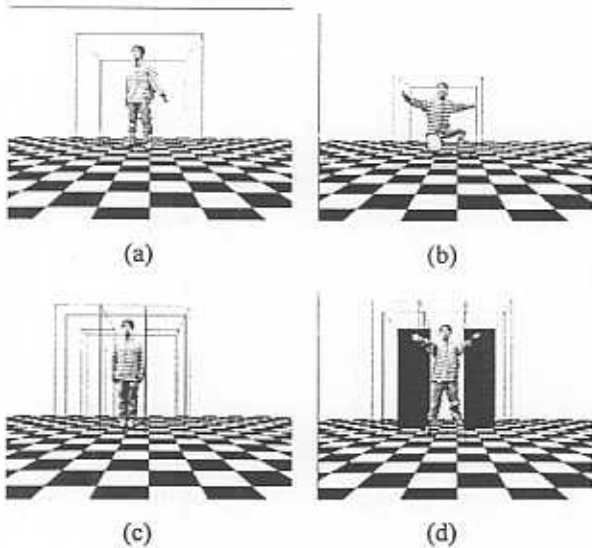


그림 6. 사용자의 포즈에 따른 공간센서의 증강

그림 6 의 (a)와 (b)는 사용자 포즈에 따른 공간센서의 증강을 나타낸 것이다. 그리고 (c)는 사용자 주변에 생성된 실제 공간센서와 감지되는 영역을 증강시킨것이다. (d)는 사용자가 공간센서를 접촉했을 때를 감지된 센서를 시각적으로 나타낸 것이다. 그림 6 에서 보듯이 제안된 공간센서는 사용자의 키 높이와 움직임 위치에 관계없이 사용자 주변에 증강된다.

제한한 공간센서의 정확성을 측정하기 위해서 사용자의 중심을 기준으로하여 다음 그림과 같이 10° 씩 움직여가며 측정하였다. 그리고 다음 표는 Movie Player 제어에 필요한 사용자의 동작과 제어 기능 그리고 공간센서의 상태 관계이다.

Movie Player Control	SpaceSensor	Movie Player State
Selection	SS1 (*)	INIT
Play	SS1 (*)	SELECTION
Pause	SS3 (*)	PLAY
Stop	SS2 (*)	PLAY
FF	SS1 (-)	PLAY
RW	SS2 (-)	PLAY
Volume Up	SS1 (↑)	PLAY
Volume Down	SS1 (↓)	PLAY

표 1. 공간센서와 Movie Player 제어 기능

표 1 에서 SS1, SS2, SS3 은 공간센서에서 감지되

는 영역을 나타내고, '\*'는 사용자가 공간센서를 접촉한 경우를 나타낸다. 그리고 '-' 는 사용자를 중심으로 뒤에서 앞으로 움직인 경우이고, '↑' 는 사용자를 중심으로 앞에서 뒤로 움직인 경우이다. 또한, '↓' 는 아래에서 위로 '↑' 위에서 아래로 움직인 경우를 나타낸다.

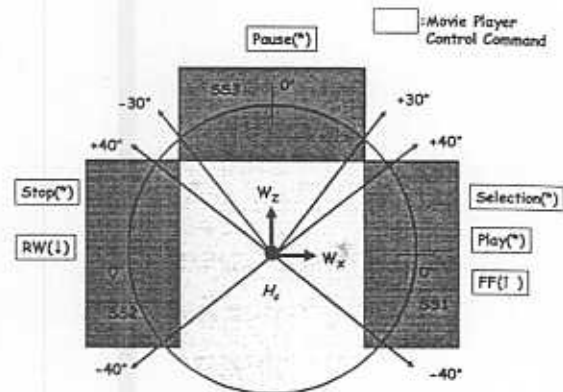


그림 7. 사용자 움직임에 따른 공간센서(SS1, SS2, SS3)의 감지영역

그림 7 에서 보듯이 사용자의 중심을 기준으로 SS1, SS2 센서의 경우 최대 40°, 최소 40° 정도의 범위 내에서는 모두 감지가 되었다. 그리고 SS3 센서의 경우 최대 30°, 최소 30° 정도의 범위 내에서 감지가 되었다. 측정된 결과에 따르면 사용자가 충분히 자연스럽게 움직일 수 있는 범위내에서는 모두 감지가 됨을 알 수 있다. 따라서 제안된 공간센서는 사용자에게 자연스러운 움직임을 제공할 수 있는 인터페이스이다.

## 5. 결론 및 추후과제

본 논문에서는 ubiHome 환경에서 사용자와 환경간의 상호작용을 보다 자연스럽게 해주는 인터페이스로서 비전기반의 3 차원 공간센서를 제안하였다. 제안된 공간센서는 사용자의 주변 공간을 인터페이스로 활용고 사용자의 의도를 파악하기위해 복잡한 인식 알고리즘을 사용하는 대신 감지된 공간센서만을 이용하여 사용자의 의도와 주변 환경 정보를 활용하여 사용자에게 적절한 서비스를 제공한다. 제안된 비전기반의 3 차원 공간센서는 ubiHome 환경에서 사용자의 제스처를 인식하는 센

서 기능뿐만 아니라 스토리텔링, 오락, 에듀테인먼트 분야등에서 사용자의 입력 기기로서의 활용가치도 있다.

#### ACKNOWLEDGEMENT

본 논문을 위해 ubiHome 테스트 베드에서 실험을 함께 진행한 장 세이, 오 유수, 류 승완에게 감사드립니다.

#### 참고문헌

- [1] M. Weiser, "The Computer for the 21st Century," *Scientific American*, pp. 94-104, Sep. 1991
- [2] S. Jang, S. Lee and W. Woo, "Research activities on Smart Environment," *Magazine of the KITE*, vol. 28, pp.1359-1371, Dec. 2001.
- [3] "A Context-based Infrastructure for Smart Environments" Anind K. Dey, Daniel Salber and Gregory D. Abowd. In *Proceedings of the 1st International Workshop on Managing Interactions in Smart Environments (MANSE '99)*, Dublin, Ireland, Dec. 1999
- [4] "The EasyLiving Intelligent Environment System", Shafer, S., Brumitt, B., and Meyer B. CHI Workshop on Research Directions in Situated Computing, April 2000
- [5] Soren Lenman, Lars Bretzner, Bjorn Thuresson, "Computer Vision Based Hand Gesture Interfaces for Human-Computer Interaction," Technical report TRITANA-D0209, CID-report, June 2002
- [6] Atsushi Nishikawa, Akio Ohnishi, Fumio Miyazaki, "Description and Recognition of Human Gestures Based on the Transition of Curvature from Motion Images," In *Face and Gesture Recognition*, pages 552-557, 1998
- [7] William T. Freeman, and Craig D. Weissman, "Television control by hand gestures", *IEEE Intl. Wrkshp. on Automatic Face and Gesture Recognition*, Zurich, June, 1995.
- [8] Markus Kohler. "System Architecture and Techniques for Gesture Recognition in Unconstraint Environments", In *Nadia Magnenat Thalmann, editor, International Conference on Virtual Systems and Multimedia VSMM'97*, pages 137- 146, University of Geneva, Switzerland, September 10-12th 1997.
- [9] H. Ishii, B. Ullmer, "Tangible Bits: Towards Seamless Interfaces between People, Bits and Atoms," in *Proceedings of Conference on Human Factors in Computing Systems (CHI '97)*, ACM Press, pp. 234-241, Mar. 1997
- [10] D. Hong and W. Woo, "Foreground Separation using Color Channel Characteristics," *KSPC 2002*, vol. 15, No. 1, pp. 299, Sep. 28, 2002.
- [11] <http://www.aiwaz.net/Leonardo/>
- [12] Woontack Woo, Namgyu Kim, Karen Wong and Makoto Tadenuma, "Sketch on Dynamic Gesture Tracking and Analysis Exploiting Vision-based 3D Interface," in *Proc. SPIE PW-EI-VCIP'01*, vol. 4310, pp. 656-666, Jan. 2001.
- [13] S. Jang, W. Woo, "Context-Aware Model for ubiHome," *HCI2003 (출판중)*, Feb. 2003.