

스마트 홈을 위한 사용자 위치추적 시스템에 관한 연구

최순용^o 최종화 신동규 신동일
세종대학교 컴퓨터공학과
{artjian^o, com97, shindk, dshin}@gce.sejong.ac.kr

Research on User Location Tracking System for Smart Home

Soonyong Choi^o Jonghwa Choi Dongkyoo Shin Dongil Shin
Dept of Computer Science, Sejong University

요 약

스마트 홈 환경에서 사용자가 원하는 바를 능동적으로 서비스 해주기 위해서는 사용자의 위치정보가 반드시 있어야 한다. 본 논문에서는 4대의 카메라를 사용하여 사용자의 위치인식을 한다. 기존의 카메라를 이용한 위치인식 알고리즘들은 단순한 물체와 사용자를 구분하는 방법을 제시하지 않았다. 본 논문에서는 교사학습 알고리즘을 사용하여 사용자를 판별해내는 방법을 제시한다. 단순한 물체와 사용자를 구분하기 위해서 교사학습 알고리즘을 사용한다. 또한 사용자의 위치좌표 외에 어떤 종류의 서비스를 받을 수 있는 위치에 있는지를 파악한다.

1. 서 론

스마트 홈(Smart Home)의 연구는 상황인식(Context Awareness) 미들웨어 관한 연구와 상황인식을 위한 구성요소인 컨텍스트(Context)에 관한 연구로 크게 나누어 볼 수 있다. 상황인식 미들웨어에 대한 연구 방향은 사용자의 행동패턴을 분석하여 별도의 명령 없이도 능동적으로 홈의 가전 및 설비들을 제어하는 쪽으로 나아가고 있다. 행동패턴 분석에는 인공지능과 데이터 마이닝 등의 방법들이 주로 사용되고 있다. 컨텍스트에 대한 연구는 위치, 표정, 제스처, 음성 등의 인식의 정확성을 높이는데 초점이 맞추어져 있다. 또한 각 컨텍스트 센서장비를 실생활에 적용하기 위한 노력이 계속되고 있다.

스마트 홈의 연구에서 사용자의 위치는 중요한 컨텍스트로서 다루어진다. 사람의 행동패턴을 분석하기 위해서는 '누가(Who)', '언제(When)', '어디서(Where)', '무엇을(What)', '어떻게(How)', '왜(Why)'의 정보들이 필요하다[1]. 사용자의 위치는 여기에서 '어디서'에 해당하는 컨텍스트로서 없어서는 안 될 부분이다.

위치인식 분야에서는 아직까지 이렇다 할 표준이 제시되지 않고 있다. 그 이유는 스마트 홈 프로젝트에 따라 홈의 구성이 다르기 때문에 모든 홈의 구성에 맞는 표준의 위치인식 방법을 정하기가 어렵기 때문이다. 그럼에도 불구하고 위치인식에 관한 연구에서 카메라를 이용한 방법이 많이 사용되고 있다. 카메라는 다른 센서장비에 비해 설치가 편리하고 사람의 몸에 별다른 장치를 소지하지 않아도 되는 장점이 있다. 또한 홈의 보안을 위한 감시 장치로서 이용[2], 사용자의 제스처 인식[3] 등 위치인식 외의 부가적인 서비스를 동시에 제공할 수

있다. 이러한 장점들이 다수의 프로젝트에서 위치인식을 위해 카메라를 선택하는 이유이다.

카메라를 이용한 위치인식은 해결해야 할 여러 가지 과제들을 가지고 있다. 첫째, 빛의 세기의 변화를 반영할 수 있어야 한다. 둘째, 사람의 이동과 물체의 이동을 구분할 수 있어야 한다. 마지막으로, 가전 및 가구의 재배치를 할 경우나 이사를 했을 경우에 간단한 조작만으로 원활한 위치인식을 할 수 있어야 한다.

현재까지 많은 스마트 홈 프로젝트들이 진행되었거나 진행 중에 있으며 사용자의 위치인식은 각 프로젝트에서 가장 중요한 컨텍스트로서 자세히 다루고 있다. 각 스마트 홈 프로젝트들을 살펴 본 바에 의하면 위치인식을 위해 어떠한 센서장비를 사용하였는가에 따라서 시스템을 분류할 수 있다. 크게 카메라를 사용한 것과 그렇지 않은 것으로 나누어 볼 수 있는데, 카메라를 사용하지 않은 것으로는 압력센서, 적외선 등을 이용한 방법들이 있다.

Microsoft의 Easy Living에서는 거실에서 다수의 사용자의 위치를 알아내기 위해 2대의 컬러 스테레오 카메라(color stereo camera)를 이용하였다[4]. 스테레오 이미지는 사용자의 위치를 알아내기 위해, 컬러 정보는 사용자 간의 식별을 위해 사용되었다. 하지만 EasyLiving에서는 물체의 이동을 고려하지 않았고 소파에 앉아있는지를 알기위해 압력센서를 사용하여 알고리즘의 한계를 보여준다. MIT의 House_n에서는 천장에 카메라를 설치하여 작은 방에서 다수의 사용자 위치를 추적하는 방법을

제시하였다[5]. 천장에 카메라를 설치하였기 때문에 넓은 지역에서는 더 많은 카메라를 사용해야 하고 사용자와 다른 물체를 구분하기가 어려운 단점이 있다. Percept Room 프로젝트에서는 벽의 상층부에 8대 벽의 중간 높이에 8대 총 16대의 카메라를 설치하였다. 각 카메라에 비춰진 사용자가 존재할 수 있는 영역들의 교집합을 구해서 사용자의 위치를 알아내는 방법을 사용하였다[6]. Percept Room에서는 실내의 가전과 사람의 위치와의 관계를 고려하지 않았기 때문에 상황에 맞는 서비스를 하기 힘든 점이 있다.

본 논문의 구성은 다음과 같다. 2장에서는 본 논문에서 제시하는 위치인식 시스템의 구성을 보인다. 3장에서는 위치인식을 위해 사용되는 내부 알고리즘에 대해서 알아보고, 4장에서는 논문의 결론 및 향후 연구 방향에 대해서 고찰한다.

2. 위치인식 시스템의 구조

우리의 스마트 홈 시스템[7]에서는 사용자의 위치 및 생체정보를 기반으로 하여 사용자가 원하는 바를 자동으로 대신해주는 서비스를 제공한다. 스마트 홈 시스템에서 사용자의 위치는 적절한 서비스를 공급할 수 있도록 하는 중요한 컨텍스트이다. 예를 들어, 사용자가 화장실에 있는데 거실의 TV를 켜는 일은 없어야 한다. 또한 사용자가 거실의 TV를 켜 놓은 채 방에서 잠들었다면 자동으로 TV의 전원을 오프해 주는 서비스를 한다면 편리할 것이다. 본 논문에서 제시하는 위치인식 시스템은 스마트 홈에서 위치인식 센서로서의 역할을 담당한다.

위치인식 시스템에서는 총 4대의 네트워크 카메라를 사용하고 있다. 위치인식 시스템의 전체 구성은 그림 1과 같이 되어있다. 주요 구성요소는 다음과 같다.

- Camera Handler. 주기적으로 네트워크 카메라로부터 이미지를 가져오는 역할을 주로 한다. 또한 카메라의 동작여부를 체크하며 네트워크 전송오류의 검출 및 복구, 이미지의 해상도 및 선명도의 조정 등 카메라에 관한 모든 관리를 한다.
- Moving Object Detector. 입력 이미지와 배경 이미지의 차이 정보를 이용하여 이동하는 객체의 영역을 검출한다. Position Recognizer로는 검출한 영역과 함께 배경 처리된 전체이미지를 전송하고, Object Classifier로는 검출된 영역만을 전송한다.
- Object Classifier. 포착된 이동하는 객체가 사람인지 아닌지 판별을 담당한다. 이를 가능케 하기 위해 교사학습 알고리즘을 사용한다. Position Recognizer로는 Moving Object로 검출된 영역이 사람인지 아닌지

판별한 정보를 전송한다.

- Position Recognizer. 4대의 카메라로부터 전송된 이미지를 바탕으로 사람으로 구분된 영역의 절대위치 좌표를 계산한다. 또한 주요 가구 및 가전과 사용자와의 가까운 정도를 판단하기 위한 User Proximity Process가 존재한다. 원활한 위치인식을 위해 배경갱신(Background Update)을 위한 알고리즘을 포함한다.
- Fixed Object Manager. 주요 가구 및 가전의 실내에서의 위치와 카메라 배경 이미지와의 매칭을 설정하기 위한 모듈이다.

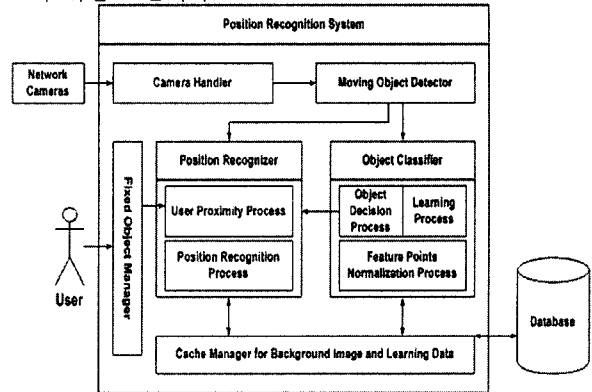


그림 1 위치인식 시스템의 구조

3. 위치인식 시스템의 내부 알고리즘

3.1 객체의 구분

본 논문에서는 사용자와 물체를 구별하기 위해서 교사학습 방법 중에 하나인 SVM(Support Vector Machine)[8]을 사용하였다. 그림 2는 사용자와 물체를 분류하기 위한 SVM을 이용한 객체 분류 알고리즘의 순서도이다.

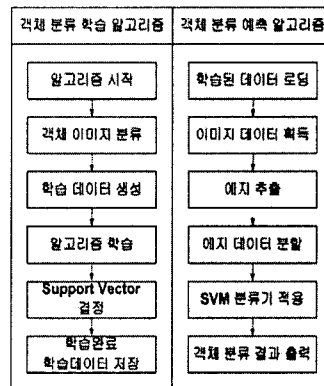


그림 2 객체 분류 알고리즘의 순서도

3.2 위치인식 알고리즘

본 논문에서의 위치인식은 두 가지의 형태의 결과를 산출한다. 하나는 일반적으로 생각할 수 있는 실내에서의 사용자의 절대위치 좌표이다. 다른 하나는 상대적인

위치로서의 권역정보이다. 권역정보는 스마트 홈 시스템에서 제어하는 가전과 의미 있는 가구(소파나 책상 같이 사용자가 위치했을 때 특정한 행동을 보이는 가구들)에서 얼마나 가까운 곳에 위치하는 가를 말한다. 스마트 홈 시스템에서는 절대위치 좌표도 중요하지만 'TV 주변', '창문 주변', '책상 앞' 등의 특정 가전 및 가구를 중심으로 한 권역 정보가 사용자의 위치로서 좀 더 중요한 의미를 갖는다.

사용자의 절대위치 좌표를 알아내기 위해 Moving Object Detector로 부티의 검출 데이터와 Object Classifier로 부티의 사람인지 아닌지의 판단 데이터가 사용된다. 절대위치 좌표의 계산은 실루엣(Silhouette)을 이용한 방법을 바탕으로 하고 있다[9][10].

각 카메라에서 비출 수 있는 사용자가 존재할 수 있는 영역을 사용자 투영영역 $C_k(k=1,2,3,4; \text{Camera No.})$ 라 정의하고 4대 카메라의 사용자 투영영역을 모두 합한 영역을 F 라 정의한다. F 의 식은 (1)과 같다.

$$F = C_1 \cup C_2 \cup C_3 \cup C_4 \quad (1)$$

사용자가 존재하지 않는 영역을 제거 영역 D_k 라고 하고 사용자가 투영된 영역을 T_k 라 정의하면 정의하면 T_k 의 식은 (2)와 같이 표현할 수 있다.

$$T_k = C_k - D_k \quad (2)$$

사용자의 위치 P 는 각 카메라에서 투영된 영역들의 교집합에서 식 (3)과 같이 정의할 수 있다.

$$P = T_1 \cap T_2 \cap T_3 \cap T_4 \quad (3)$$

권역정보의 추출을 하기 위해서는 먼저 각 카메라의 배경영상에서 사용자가 주로 이용하는 가구 및 가전들을 지정한다. 각 가구 및 가전의 설정이 마무리 되면, 실시간으로 전송되는 카메라 이미지에서 사용자와 설정된 가구 및 가전 이미지와의 상대적인 거리를 측정하여 어떤 가구 및 가전에 가까운지를 추출한다.

4. 결론

본 연구에서는 4대의 네트워크 카메라를 사용하여 실내에서 사용자와 물체를 구분할 수 있는 방법과 사용자의 절대위치 좌표를 알아내는 방법을 제안하였다. 이 외에도 사용자의 권역정보를 파악할 수 있는 알고리즘을 제시하여 카메라의 대수를 줄이고 위치인식의 성능향상을 도모하였다.

앞으로의 연구에서는 다중 사용자를 인식하고 위치인식에 쓰이는 카메라의 대수를 줄여서 최종적으로 한 대의 카메라로도 사용자 위치 컨텍스트를 추출할 수 있는 알고리즘의 연구가 필요하다.

참고 문헌

- [1] Sejie Jang, Woontack Woo, "Ubi-UCAM: A Unified Context-Aware Application Model", LNCS Volume 2680, pp. 178-189, August 2003.
- [2] Wren C.R., Azarbayejani A., Darrell T., Pentland A.P., "Pfinder: real-time tracking of the human body", Pattern Analysis and Machine Intelligence, IEEE Transactions on Volume 19, Issue 7, pp. 780 - 785, July 1997.
- [3] Fukuda T., Suzuki, T., Kobayashi F., Arai F., Hasegawa Y., Negi M., "Seamless tracking system with multiple cameras", Industrial Electronics Society, 2000. IECON 2000. 26th Annual Conference of the IEEE, pp. 1249-1254, 2, Oct. 2000.
- [4] Krumm J., Harris, S., Meyers B., Brumitt B., Hale M., Shafer S., "Multi-Camera Multi-Person Tracking for EasyLiving", IEEE Workshop on Visual Surveillance, July 2000.
- [5] Khalaf R., Intille S. S., "Improving multiple people tracking using temporal consistency," Massachusetts Institute of Technology, Cambridge, MA, MIT Dept. of Architecture House_n Project Technical Report, 2001.
- [6] Tominaga M., Hongo H., Koshimizu H., Niwa Y., Yamamoto K., "Estimation of human motion from multiple cameras for gesture recognition", Pattern Recognition, 2002. Proceedings. 16th International Conference on Volume 1, pp. 401 - 404, 11-15 Aug. 2002.
- [7] Jonghwa Choi, Dongkyoo Shin, Dongil Shin, "Research and implementation of the context-aware middleware for controlling home appliances", Consumer Electronics, IEEE Transactions on Volume 51, Issue 1, pp. 301 - 306, Feb. 2005.
- [8] K.M. Lin, C.J. Lin, "A study on reduced support vector machines", Neural Networks, IEEE Transaction on, Vol.14, Issue:6, pp.1449-1459, 2003.
- [9] Takashi Matsuyama, "Cooperative Distributed Vision: Research Achievements and Future Directions", The 7th Symposium on Sensing via Image Information(SSII2001), pp. 187-198, June. 2001.
- [10] Shogo Hamasaki, Hiroyuki Yoshida, Shinichi Shigenaga, "High Speed 3D Shape Reconstruction from Multiple Silhouettes", The 7th Symposium on Sensing via Image Information(SSII2001), pp. 59-64, June. 2001.