

# 스마트 테이블 구축을 위한 템플릿 매칭과 Optical flow를

## 이용한 손 트래킹

권오륜<sup>o</sup> 전준철

경기대학교 정보과학부

{kor5663<sup>o</sup>, jcchun}@kyonggi.ac.kr

### A Hand Tracking by Template Matching and Optical Flow for Smart Table

Ohryun Kwon<sup>o</sup> Junchul Chun

Dept of Information Science, Kyonggi University

#### 요 약

본 논문에서는 컴퓨터와 상호작용을 할 수 있는 스마트 테이블 구축을 위한 마우스를 대체하는 손의 움직임을 트래킹하는 방법을 제안하고자 한다. 테이블 위에 놓인 프로젝터와 카메라는 각각 프로젝터의 영상이 테이블에 투영이 되고 카메라로 테이블에 투영된 영상을 입력받게 된다. 이렇게 입력된 영상으로부터 손 검출과 트래킹을 통해 테이블을 이용하여 컴퓨터와의 상호작용을 할 수 있다. 먼저 영상은 손의 그림자를 포함하고 있기 때문에 그림자를 제거한 후 canny 에지 필터를 이용하여 손의 후보영역을 검출하게 된다. 검출된 후보 영역으로부터 템플릿 매칭을 이용하여 손 영역을 검출하고 검출된 영역으로부터 optical flow를 이용하여 손의 위치를 트래킹하게 된다.

#### 1. 서 론

최근 HCI(human computer interaction)에 대한 관심이 커짐에 따라 키보드나 마우스와 같이 기존의 사용자 인터페이스에 비해 보다 자연스럽고 편리한 사용자 인터페이스에 대한 연구가 활발히 진행되고 있다. 특히, 유비쿼터스 컴퓨팅 환경을 위해 사용자와 컴퓨터간의 상호작용을 하는데 장비에 의존하지 않고 편리하고 자연스러운 행동으로 인해 작업을 수행하기 위한 기술이 필요시 되고 있다.

HCI분야에서 비전 기반 손 검출 및 트래킹하는 것은 손의 움직임과 제스처를 이용하여 좀 더 자연스러운 방법으로 컴퓨터와 상호작용을 하기 위해 중요한 문제이다 [1]. 하지만 손은 매우 구조적이고 다양한 형상을 나타내기 때문에 검출하고 트래킹하는 것은 결코 쉬운 일이 아니다. 손을 검출하기 위한 연구로는 다음과 같은 것들이 있다. 첫째, 컬러 기반의 손 검출 방법이다 [2]. 이 방법은 피부의 고유한 색을 이용하여 손을 검출하는 방법으로서 쉽고 빠르게 배경으로부터 손을 검출할 수 있다. 하지만 복잡한 배경이나 피부색과 유사한 배경이 있을 경우, 조명의 조건에 따라 검출율이 현저하게 차이가 난다는 단점을 가진다. 둘째, 형상 기반 방법으로 PCA이나 신경망과 같이 손의 형상을 이용하여 손 검출하는 방법이다 [3]. 이 방법은 컬러 기반 방법보다 검출율이 정확하지만 학습 데이터에 따라 결과가 현저히 달라지거나 밝기값에 영향을 많이 받는 단점을 가지고 있다. 셋째, 템플릿 매칭을 이용한 방법으로서 컬러 기반 방법보다

좋은 결과를 나타내지만 템플릿과 비교하는데 연산 시간이 오래 걸린다는 단점을 가지고 있다.

본 논문에서는 스마트 테이블을 구축하기 위해 테이블 위에 프로젝터와 카메라가 위치하게 되며 프로젝터로부터 투영된 테이블로부터 손의 움직임을 이용하여 컴퓨터와 상호작용을 할 수 있게 한다. 이때 테이블에 나타나는 모습을 카메라로부터 영상을 입력하고 분석하게 된다. 이와 같은 환경에서 손의 위치를 트래킹하기 위해서는 먼저 손의 후보 영역을 검출한 후 검출된 손으로부터 트래킹을 시작하게 된다. 실시간으로 수행해야하기 때문에 계산량을 최소화하는 것이 중요하다. 따라서 최소한의 계산량으로 손을 검출하기 위해 에지 정보와 템플릿 매칭을 이용하게 된다. 먼저 에지를 이용하여 손의 후보 영역을 검출한다. 검출된 손의 후보 영역으로부터 템플릿 매칭을 이용하여 손 영역을 검출하게 되는데 이때 임계치를 이용하여 후보 영역에서 손의 유무를 파악하게 되며 임계치 이하이면 손을 검출하게 된다. 템플릿 매칭을 이용하여 손을 검출하게 되면 검출된 손 영역의 에지로부터 optical flow를 이용하여 손의 위치를 트래킹하게 된다.

이후 본론에서는 에지 정보와 템플릿 매칭을 이용하여 손을 검출하기 위한 방법과 검출된 손 영역으로부터 optical flow를 이용하여 트래킹하는 방법에 대해 설명한다. 또한 마지막으로 실험 결과와 향후 연구 방향에 대해 제시하도록 한다.

#### 2. 본 론

본 논문은 스마트 테이블 구축을 위해 카메라로부터 실시간으로 입력되는 영상으로부터 손 영역을 검출하고 트래킹하는 것이 목적이다. 테이블 위에 있는 프로젝터는 영상을 투영하게 되고 테이블에 투영된 영상을 카메라로 입력받고 분석하게 된다. 이러한 손의 움직임을 이

본 연구는 21세기 프론티어 연구개발사업의 일환으로 추진되고 있는 정보통신부의 유비쿼터스컴퓨팅네트워크원천기반기술개발사업의 지원에 의한 것이다

용한 스마트 테이블을 구축하기 위해 카메라로부터 입력된 영상에서 손 영역을 검출하고 손의 끝점을 트래킹하기 위한 방법은 그림 1과 같이 간략하게 설명할 수 있다.

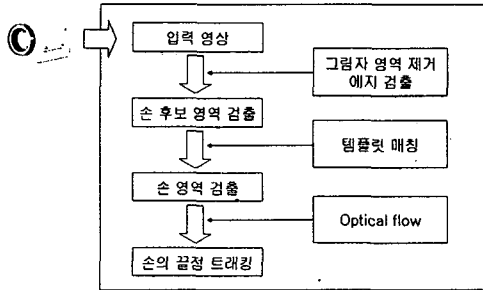


그림 1. 시스템 개략도

2.1 손 후보 영역 검출

먼저 입력된 영상으로부터 손의 후보 영역을 검출하게 된다. 이때 프로젝터의 빛에 의해 생기는 그림자 영역을 제거한 후 에지 필터를 이용하여 손의 후보 영역을 검출한다. 입력 영상의 배경이 복잡하지 않기 때문에 쉽고 빠르게 수행할 수 있는 에지 필터를 이용한다.

2.1.1 그림자 영역 제거

테이블 위에서 손이 움직이면 프로젝터로부터 투영된 빛 때문에 손의 그림자가 생기게 된다. 이때 에지 필터를 이용하게 되면 그림자에 의해 손 영역 보다 좀 더 큰 영역을 검출하게 된다. 이러한 오차를 없애기 위해 손의 그림자 영역을 제거하는 전처리 과정을 거친다. 사람의 피부색은 붉은색을 많이 가지고 있다. 따라서 영상의 RGB 컬러모델 중 R 컴포넌트를 이용하게 되는데 그림자의 R값이 손 영역보다 작다. 따라서 R값이 80보다 작은 경우에는 배경과 유사한 밝기값으로 변환시킨다. 그리고 미디언 필터를 이용하여 잡음을 제거하고 모폴로지 연산을 통해 그림자가 제거된 영상을 획득할 수 있다. 그림 2의 (a)와 (b)는 입력 영상과 그림자를 제거한 영상을 보여주고 있다.

2.1.2 에지 필터

입력 영상에서 그림자 영역을 제거한 후 손 후보 영역을 검출하기 위해 canny 에지 필터를 사용하게 된다. Canny 에지 필터는 여러 가지 에지 필터 중 가장 정확하게 정의된 연산자로서 에지 검출에 좋은 결과를 가진다. 에지 필터를 이용하여 영상의 에지를 검출하여 손의 후보 영역으로 결정하게 된다. 그림 2의 (c)와 (d)는 에지를 검출한 결과와 검출된 에지로부터 손의 후보 영역이 결정되는 것을 보여준다.

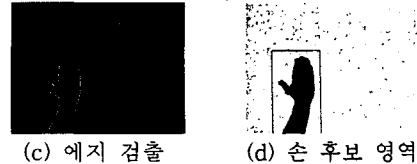
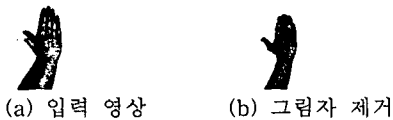


그림 2. 손 후보 영역 검출

2.2 손 영역 검출

손의 후보 영역으로부터 템플릿 매칭을 이용하여 손 영역을 검출한다. 템플릿 매칭은 손에 대한 표준 템플릿을 만든 후, 입력 영상과 비교하여 손을 검출하는 것이다. 이는 일반적으로 많이 사용하고 있는 방법이지만 템플릿을 만드는 것이 어렵고 밝기값 변화에 민감한 단점이 있다.

본 논문에서는 손 영역 검출을 위해 1) 전처리 단계, 2) 템플릿 생성, 3) 템플릿 매칭으로 크게 3가지 단계로 나눌 수 있다.

2.2.1 전처리

조명의 변화는 검출 성능을 좌우하는 중요한 요소가 될 수 있다. 따라서 정확한 손 검출을 위해서 전처리 단계로 조명을 보정하는 것이 반드시 필요하다.

본 논문에서는 전처리 과정으로 최소-최대 정규화 방법을 사용한다(식 1). 최소-최대 정규화 방법은 입력 영상을 새롭게 정의된 데이터 범위로 변형시켜주는 방법으로, 아주 어둡거나 밝은 입력 영상을 특정 밝기 영역 안에 존재할 수 있도록 변화시켜 줄 수 있다.

$$y = \left( \frac{y - \min_1}{\max_1 - \min_1} \right) (\max_2 - \min_2) + \min_2 \quad \text{식 (1)}$$

식 (1)에서  $\min_1$ 과  $\max_1$ 는 입력 영상의 최소, 최대 밝기 값이며,  $\min_2$ 과  $\max_2$ 는 변형하고자하는 영상의 최소, 최대 밝기값이다.

2.2.2 템플릿 생성

템플릿은 손 영역을 80\*80크기로 만든다. 손 영상 여러 개의 평균 영상을 템플릿으로 이용한다. 그림 3은 템플릿을 만들기 위한 영상과 템플릿 영상이다.



그림 3. 손 영상과 템플릿 영상

2.2.3 템플릿 매칭

생성된 템플릿과 손 후보 영역 밝기값 차이의 제곱의 합인  $L_2$  norm(식 2)을 이용하여 템플릿 매칭을 한다. 손이 항상 테이블에 위에 있는 것이 아니기 때문에 손 후보 영역에 언제나 손이 있다고 말할 수 없다. 따라서  $L_2$  norm의 총합이 임계치보다 클 경우 손 후보 영역에는 손이 존재하

지 않는다는 것을 나타낸다. 임계치보다 작은 영역에서  $L_2$  norm의 총 합이 가장 작은 영역을 손으로 검출하게 된다. 그림 4는 손을 검출하는 결과를 보여준다.

$$Error = \sum_{i=0}^{79} \sum_{j=0}^{79} \sqrt{(I_{ij} - T_{ij})^2} \quad \text{식 (2)}$$



그림 4. 손 검출 결과

### 2.3 손 트래킹

위의 과정에서 손이 검출되면 손 검출 과정은 중단하고 optical flow를 이용하여 트래킹 좌표인 손의 상단 끝점을 트래킹하기 시작한다. 손은 테이블을 위에서 내려다보았을 때 항상 아래쪽으로부터 들어온다고 가정한다. 따라서 트래킹 좌표의 y값은 손 영역에서 최대값이다. 또한 손 영역의 y값을 이용하여 x값을 구할 수 있다. 트래킹 좌표로부터 optical flow를 이용하여 손의 상단 끝점인 트래킹 좌표를 트래킹한다.

#### 2.3.1 Optical flow

Optical flow는 연속적인 이미지에서 두 프레임 사이의 밝기 패턴의 움직임을 추정하는 방법이다. 몇몇의 optical flow 알고리즘이 있는데 그 중에서 Lucas-Kanade가 가장 정확하고 노이즈에 강인하다[4][5]. 이 방법은 원래 스테레오 매칭을 하기 위한 방법이었지만 현재에는 주로 optical flow에 사용된다. 식 (3)은 optical flow의 기본적인 식을 보여준다.

$$I(x, t) = I(x + u, t + dt) \quad \text{식 (3)}$$

식 (3)에서  $I$ 는 밝기값,  $x$ 는 픽셀의 위치,  $u$ 는 픽셀의 변화량,  $t$ 는 시간을 나타낸다. 테일러급수 확장에 의해  $I$ 는 다음과 같이 나타낼 수 있다(식 4).

$$I(x + u, t + dt) = I(x, t) + \nabla I \cdot u + \frac{\partial I}{\partial t} \cdot dt \quad \text{식 (4)}$$

식 (4)에서  $\nabla I$ 는 공간상의 이미지 변화량을 나타낸다. 식 (3)과 식 (4)를 통해 optical flow 제약 공식을 이끌어 낼 수 있다(식 5).

$$\nabla I \cdot u + \frac{\partial I}{\partial t} \cdot dt = 0 \quad \text{식 (5)}$$

Lucas-Kanade 알고리즘은 해당 윈도우  $W$ 에서 식 (5)의 좌변 값을 최소화하는 것이다(식 6).

$$\sum_{x \in W} \omega^2(x) (\nabla I \cdot u + \frac{\partial I}{\partial t} \cdot dt) \quad \text{식 (6)}$$

여기서  $\omega : W \rightarrow \mathbb{R}$  은 가우시언 가중치 함수이다. 이와 같이 Lucas-Kanade 알고리즘을 통한 optical flow를 이용하여 손 끝점의 위치인 추적 좌표를 트래킹하게 된다. 그림 5는 optical flow를 이용한 트래킹 결과이다. 왼쪽으로부터 입력영상, 손 영역 검출, 트래킹 결과를 보여준다.

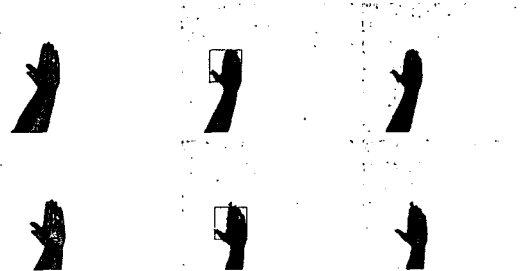


그림 5. optical flow를 이용한 트래킹

### 3. 결론 및 향후 연구

본 논문에서는 카메라로부터 입력받은 영상에서 손을 검출하고 트래킹에 대한 방법을 제시하였다. 에지 필터와 템플릿 매칭을 이용하여 손을 검출하고 optical flow를 통해 손을 추적하는 방법에 관한 연구이다.

향후 스마트 테이블을 만들기 위해 본 논문을 통해 제안된 방법을 이용할 것이다. 또한 스마트 테이블의 터치에 대한 이벤트를 처리하는 방법으로 FSR 압력 센서에 대한 연구를 할 것이다. 따라서 손 트래킹과 터치 유무를 파악함으로써 테이블을 통한 컴퓨터와 상호작용이 이루어질 수 있는 스마트 테이블에 대한 연구가 필요하다.

### 4. 참고 문헌

- [1] V. Pavlovic, R. Sharma, and T. Huang, "Visual interpretation of hand gestures for human computer interaction: a review", IEEE Trans. PAMI, Vol. 19, No. 7, pp.677-695, 1997.
- [2] K. Imagawa, S. Lu and S. Igi, "Color-based Hand Tracking System for Sign Language Recognition", Proceedings of International Conference on Automatic Face and Gesture Recognition, 1996.
- [3] J. Martin, J. Crowley, "An Appearance-Based Approach to Gesture-Recognition", Proceedings of the 9th International Conference on Image Analysis and Processing-Volume II, pp.340-347, 1997.
- [4] J.L. Barron, D.J. Fleet and S.S. Beauchemin, "Performance of Optical Flow Techniques", Int. Journal of Computer Vision, pp.43-77, 1994.
- [5] B.D. Lucas and T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision", DARPA Image Understanding Workshop, pp.121-130, 1981.