

국부적인 조명변화와 복잡한 배경에 강인한 손 끝 좌표 추적

김 유 호, 김 중 선, 이 준 호
성균관대학교 전기·전자 및 컴퓨터공학부

Fingertip Tracking Robust to Local Illumination Changes and Cluttered Background

Yu-Ho Kim, Jong-Sun Kim, June-Ho Yi

School of Electrical and Computer Engineering, Sungkyunkwan University

요약

본 연구는 손의 동작변화로 인한 손 영역의 국부적인 조명변화와 복잡한 배경환경에서 손 영역의 검지좌표를 안정적으로 검출, 추적하여 마우스 포인터를 제어하는 평거 마우스 시스템을 제안하였다. 손의 동작변화로 인한 국부적인 조명변화에 강인한 손 영역 검출을 위한 적응적인 on-line 학습법을 제안하였으며 복잡한 배경에서도 안정적인 손 영역 추적이 가능하도록 칼만 트래킹과 차영상을 이용한 모션 세그멘테이션을 복합적으로 적용하였다. 실험결과 복잡한 배경과 손의 움직임에 상관 없이 검지 좌표를 안정적으로 추적 할 수 있었다.

1. 서론

부가적인 하드웨어 없이 사람의 손 동작을 컴퓨터와 인간사이의 상호작용을 위한 의사 전달 수단으로 사용하면 기존의 하드웨어 의존적인 인터페이스를 사용자 중심의 인간 친화적인 인터페이스로 재구성 할 수 있다. 그러나 복잡한 주위 환경으로부터 사람의 손 동작만을 정확히 추출하고 추적하여 의미 있는 입력정보로 사용하는 것은 어려운 일이다. 기본적으로 정확한 손 영역 분할을 위한 알고리즘이 필수이며 손 동작의 분석과 추적을 위한 빠른 처리속도 또한 무시할 수 없는 고려요소이다.

기존의 연구들은 크게 3차원 모델을 사용하거나 스테레오 비전을 이용하여 추출된 손의 3차원 기하학적

인 정보를 분석하는 방법[5-7]과 손의 2차원 패턴 정보를 분석하는 방법[1-4]으로 분류할 수 있다. 3차원 기하정보를 이용하는 방법은 스테레오 영상을 이용한 3차원 정보추출과 복잡한 손의 3차원 운동 모델을 구축하기 위한 많은 계산량이 필요하므로 고성능의 하드웨어가 필요하다는 단점이 있다. 2차원 정보를 이용하는 방법들은 사람의 피부색의 컬러분포를 사전 학습과정에서 계산하여 입력 영상 내 해당 피부 컬러 분포에 해당되는 영역만을 추출하는 방법을 적용하였으나 대부분 단일 배경이거나 조명 변화는 거의 없다고 가정하였다.

본 연구는 손의 동작변화로 인한 손 영역의 국부적인 조명변화와 복잡한 배경환경으로부터 손 영역의 검지좌표를 안정적으로 추적하여 윈도우 시스템제어를 위한 입력 값으로 사용하는 것을 목적으로 한다. 본 연구에서는 손의 동작변화로 인한 국부적인 조명변화에 강인한 적응적인 on-line 학습법을 제안하였으며 복잡한 배경에서도 안정적인 손 영역 추적이 가능하도록 칼만 트래킹과 차영상을 이용한 모션 세그멘테이션을 복합적으로 적용하였다. 실험결과 제안된 알고리즘은 Pentium III-450 시스템에서 640 X 480크기의 입력영상을 약 15fps로 처리가 가능하였다. 본 논문의 구성은 다음과 같다. 2절에서는 제안된 시스템의 전체 알고리즘을 설명하며 3절에서는 적응적인 on-line 학습법과 손 끝 좌표 계산을 위한 제안된 방법에 대하여 설명한다. 4절에서는 손 끝 좌표 추적을 이용한 실제적인 윈도우 시스템

템제가 결과를 보여준다.

2. 전체 시스템 개요

본 연구에서는 피부색의 컬러 분포를 이용한 손 영역 추출 알고리즘과 칼만 필터와 배경과 입력 영상과의 차 영상을 복합적으로 이용한 손 끝 추적 알고리즘을 실제 응용 기술에 적용한 평거 마우스 시스템을 제안하고 구현하였다. 그림 1.은 전체 시스템의 구성도를 나타낸다.

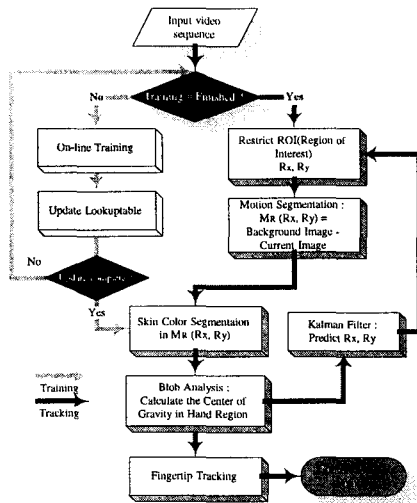


그림 1. 제안된 시스템

제안된 on-line 학습과정에서는 사용자의 피부색과 조명 환경에 최적화된 2차원 Hue와 Saturation 룩업테이블을 구성한다. on-line 학습과정 후 손 영역 분할은 칼만 필터에 의하여 예측된 손 영역의 관심 영역 내에서 배경과 연속된 입력영상과의 차 영상을 계산하고 차이가 나타난 영역에서만 HS 룩업테이블과의 비교를 통하여 이루어진다. 분할된 손 영역의 무게 중심의 좌표는 손 영역 추적을 위한 칼만 필터 알고리즘의 측정값으로 입력되어 다음 움직임 영역을 예측한다. 윈도우 시스템 제어를 위한 손 끝 좌표는 예측된 영역에서 분할된 손 영역의 검지 부분만을 제한하여 계산한다.

3. 손 동작 검출 및 추적

손 동작을 분석하고 추적하기 위해서는 기본적으로 정확한 손 영역 분할을 위한 알고리즘이 필수적이다. 그

러나 인종 혹은 사람에 따라서 피부 컬러 성분의 분포가 다르게 존재하며 같은 피부색일지라도 다른 조명 환경에 따라 컬러 분포가 다르게 나타날 수 있다. 또한 주위의 배경에 피부색과 유사한 컬러 분포를 갖는 물체가 존재한다면 정확한 영역 분할은 어렵다.

3.1 적응적인 on-line 학습법

본 논문에서는 물체 고유의 색을 나타내는 Hue와 고유의 색에 대한 순도를 나타내는 Saturation만을 이용하여 피부 영역의 컬러 분포를 확률적 모델링을 통하여 해석하였다. 만약 피부 영역에 대한 클래스를 W_s 라 하면 현재 입력 픽셀(X)이 피부 영역인지 아닌지를 판단할 수 있는 분별 함수[8]를 정의할 수 있다.((1) 식)

$$y_s(X) = P(W_s | X) \quad (1)$$

따라서 분별 함수 $y_s(X)$ 의 값이 클수록 현재 입력 픽셀이 사람의 피부일 가능성이 크다는 것을 의미한다. (1)식은 W_s 에 대한 *a posteriori* 확률로서 이를 모델링하기 위해서는 미리 학습과정에서의 모델링이 필요하다 (1)식에 Bayes의 법칙을 적용하면 다음 (2)식과 같이 표현된다.

$$P(W_s | X) = \frac{p(X | W_s) P(W_s)}{p(X)} \quad (2)$$

여기서 $p(X|W_s)$ 는 클래스 W_s 내에 입력 픽셀 X 가 존재할 확률을 나타내는 likelihood 함수로서 변수 Hue와 Saturation에 대한 2차원 gaussian 확률 밀도 함수로 모델링이 가능하다((3) 식). $p(X|W_s)$ 는 W_s 의 *a priori* 발생 확률로서 영상 내 피부 영역의 확률을 나타낸다. $p(X)$ 는 피부 영역의 클래스 W_s 와는 무관한 전체 입력 픽셀의 발생 확률이므로 클래스의 분류 결정에는 영향을 미치지 않는다. 최종적으로 분별함수 $y_s(X)$ 는 다음 (4)식과 같이 표현할 수 있다.

$$p(X|W_s) = (2\pi)^{-1} |\Sigma|^{-1/2} \exp\{-\frac{1}{2}(X - \mu_{hs})^T \Sigma^{-1} (X - \mu_{hs})\} \quad (3)$$

$$y_s(X) = \frac{1}{2} (X - \mu_{hs})^T \Sigma^{-1} (X - \mu_{hs}) - \frac{1}{2} \ln |\Sigma| - \ln(2\pi) + \ln P(W_s) \quad (4)$$

여기서 μ_{hs} 는 손 영역에 해당하는 Hue와 Saturation의 평균 벡터($m_{hs} = [m_H \ m_S]^T$)이고 Σ 는 손 영역의

Hue와 Saturation의 데이터 분포를 나타내주는 공분산 행렬을 나타낸다. 피부 영역에 대한 HS 록업데이터를 구성하는 방법은 다음과 같다.

사용자는 영상 내 위치한 일정한 영역에 손을 위치시켜 피부 영역에 대한 분별 함수를 생성한다. 마찬가지로 손 영역이 포함되지 않은 배경에 대한 분별함수 $y_B(X)$ 를 정의하고 두 클래스(피부 영역과 배경 영역)의 분별함수의 차를 이용하면 현재 입력픽셀이 피부 영역인지 아닌지를 결정할 수 있다.((5)식)

$$y(X) = y_S(X) - y_B(X)$$

$$X = \begin{cases} skin, & \text{if } y(x) \geq 0 \\ non_skin, & \text{if } y(x) < 0 \end{cases} \quad (5)$$

피부 영역으로 판단되어진 픽셀들의 HS값은 손 영역 검출을 위한 HS록업데이터의 인덱스로 구성된다. 또한 손의 동작 변화에 따른 Hue와 Saturation값의 변화를 포함하기 위하여 사용자는 일정 시간동안 손의 모양을 변화시키면서 위의 과정을 되풀이 하게 된다(그림 2).

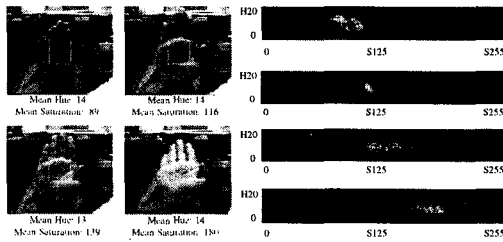


그림 2. 적응적인 학습법과 HS 컬러분포 변화

3.2 손 끝 좌표 계산 및 추적

관심이 되는 손 영역만을 다른 움직임과 배경으로부터 추출하는 방법은 다음과 같다. 1차적으로 관심이 되는 손 영역의 움직임을 칼만 필터를 이용하여 예측하여 1차 탐색 영역으로 제한한다. 1차 탐색 영역 내에서 입력 영상과 배경과의 차 영상을 구하면 손 영역의 정보만을 얻을 수 있다. 얻어진 손의 움직임 영역에 대하여 Hue와 Saturation의 2차원 록업데이터를 적용하게 되면 손 이외에 팔과 같은 부분은 제거되고 정확한 손의 영역만이 추출 된다. 추출된 손 영역은 blob 분석 과정을 거쳐 현재 손의 중심점과 가로 세로 폭이 계산되고 계산된 측정값은 손의 다음 상태 예측을 위한 측정값으

로 칼만 필터 알고리즘에 입력된다.

손 끝 좌표를 얻기 위하여 계산 영역을 전체 손 영역 중 검지 영역에만 한정시켜 손 끝 좌표를 얻는 방법을 제안하였으며 계산 방법은 다음과 같다. 분할된 이진 영상은 blob 분석을 거쳐 전체 손 영역의 x축 방향 길이와 y축 방향 길이를 계산하게 되고 손의 회전각도에 따라서 각 히스토그램(vertical or horizontal histogram)에 대한 백분율(히스토그램의 빈도수/축 길이 = 25%이내의 영역만을 검지로 제한)을 적용하여 계산한다. 손의 회전각은 2차 모멘트를 이용하여 계산하게 되고 만약 손의 회전각이 수직 축(90°)을 중심으로 각도의 편차가 ±45° 이내이면 손의 x축 길이에 대한 x축 방향 히스토그램(horizontal histogram)의 빈도수의 백분율(horizontal histogram frequency / length in the x direction)을 적용하여 검지 영역을 제한한다(그림 3-(b)).

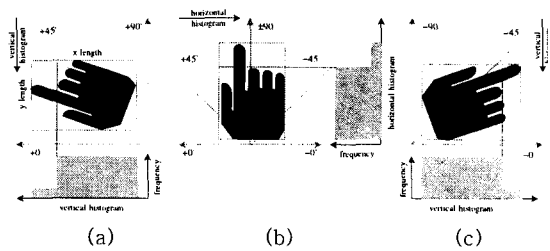


그림 3. 검지 영역 제한법과 히스토그램 분포

(a) scan order: top->down, left->right, (b) scan order: left->right, top->down, (c) scan order: right->left, top->down

수직 축에 대한 손의 회전각의 편차가 ±45° 이상의 범위일 때는 손의 y축 길이에 대한 y축 방향 히스토그램(vertical histogram)의 빈도수의 백분율(vertical histogram frequency / length in the y direction)을 적용하게 되는데 히스토그램의 scan order는 회전각의 부호에 따라 다르게 적용된다(그림 3-(b,c)). 윈도우 시스템 제어를 위한 손 끝 좌표는 제한된 검지영역의 1/6 지점에 대한 무게 중심 좌표로 결정한다. 제안된 본 방법은 손의 크기, 이동, 회전(±90° 이내)에 관계없이 안정적으로 손 끝 좌표를 검출할 수 있다.

4. 응용 실험

그림 4.는 손 끝의 움직임의 궤적에 따른 마우스 포인터의 움직임과 손 영역 검출 결과를 보여준다.

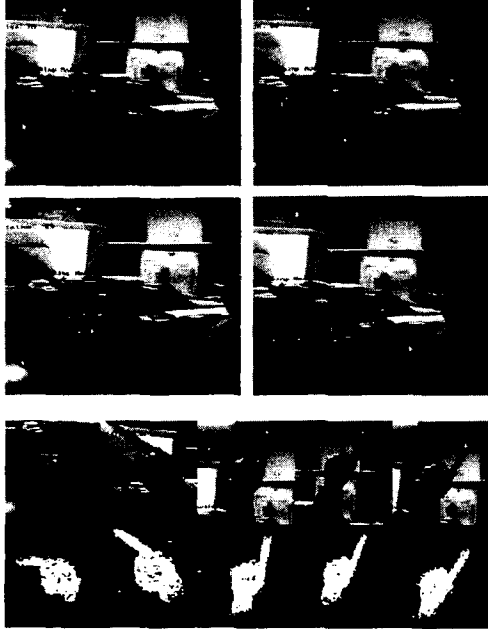


그림 4. 제안된 평거 마우스 시스템과 검출 결과

5. 결론 및 추후과제

실험 결과 제안된 알고리즘은 조명 변화와 복잡한 배경에 무관하게 안정적으로 손 영역을 검출, 추적할 수 있었으며 원활한 마우스 포인터 제어가 가능하였다. 따라서 일반적인 사용자 환경에서의 응용 기술 개발이 가능하리라 사료된다. 추후 손 동작 인식 알고리즘을 본 시스템에 적용하여 윈도우 시스템제어를 위한 완전한 핸드 마우스 시스템을 개발할 예정이다.

감사의 글.

본 연구는 정보통신연구진흥원 대학기초연구 지원사업의 연구비에 의하여 수행되었음.

참고문헌

[1] R. Kjeldsen and J. Kender, "Towards the use of Gesture in Traditional User Interfaces,"

Proceedings of International Conference on Automatic Face and Gesture Recognition, pp. 151-156, Killington, October, 1996.

[2] K. Imagawa, S. Lu and S. Igi, "Color-Based Hands Tracking System for Sign Language Recognition," Proceedings of International Conference on Automatic Face and Gesture Recognition, Killington, 1996.

[3] F. Quek, T. Mysliwiec, and M. Zhao, "FingerMouse: A Freehand Pointing Interface." Proceedings of International Conference on Automatic Face and Gesture Recognition, Zurich, pp. 372-376, 1995.

[4] S. Ahmad, "A Usable Real-Time 3D Hand Tracker," Conference Record of the Asilomar Conference on Signals, System and Computers, pp. 1257-1261, 1994.

[5] J. Segen and S. Kumar, "Shadow Gestures: 3D Hand Pose Estimation Using a Single Camera," CVPR99, vol. 1, pp. 479-485, Fort Collins, Colorado, June, 23-25, 1999.

[6] J. Rehg and T. Kanade, "Visual Analysis of High DOF Articulated Object with Application to Hand Tracking," CMU Tech. Report CMU-CS-95-138, Carnegie Mellon University, April, 1993.

[7] S. H. Kim, N. K. Kim, S. C. Ahn, and H. G. Kim, "Object Oriented Face Detection Using Range and Color Information," Proceedings of International Conference on Automatic Face and Gesture Recognition, pp. 76-81, 1998.

[8] Christopher M. Bishop, "Neural Networks for pattern Recognition," Clarendon Press, Oxford, 1995.