

學術發表 VIII

人工知能 神經網 및 퍼지 分野

Hue 영상을 기반한 손 영역 검출 및 추적

권화중 · 이준호

성균관대학교 전기·전자 및 컴퓨터 공학부

Hand Region Segmentation and Tracking Based on Hue Image

Kwon, Hwajung and Yi, Juneho

School of Electrical and Computer Engineering, Sungkyunkwan University

hjkweon@ece.skku.ac.kr, jhyi@yurim.skku.ac.kr

ABSTRACT

Hand segmentation and tracking is essential to the development of a hand gesture recognition system. This research features segmentation and tracking of hand regions based the hue component of color. We propose a method that employs HSI color model, and segments and tracks hand regions using the hue component of color alone. In order to track the segmented hand regions, we only apply Kalman filter to a region of interest represented by a rectangle region. Initial experimental results show that the system accurately segments and tracks hand regions although it only uses the hue component of color. The system yields near real time throughput of 8 frames per second on a Pentium II 233MHz PC.

1. 서론

사람들은 의사 전달 방법으로 말을 사용하며 때로는 자기의 의사를 정확하게 전달 또는 강조하기 위하여 손짓과 얼굴 표정을 변하게 하는 '제스처'를 취한다. 이러한 제스처를 인간과 컴퓨터의 상호작용(HCI: Human-Computer Interface)에 사용한다면 사람들은 기존의 물리적인 도구를 통하지 않고 자연스러운 제스처를 통하여 컴퓨터와의 자연스러운 상호작용이 가능할 것이다. 인간과 컴퓨터와의 상호작용을 위한 제스처로서는 얼굴의 움직임이나 표정에 의한 제스처 보다는 손에 의하여 행해지는 제스처가 적합할 것이다. 인간과 컴퓨터와의 상호 작용을 위해, 손에 의하여 행해지는 제스처를 인식하기 위해서 먼저 수행이 되어져할 것은 인식의 대상이 되는 손 영역을 배경으로부터 분리하고 추적하는 것이다.

손 영역을 검출하여 추적하는 방법은 크게 다음의 네가지 방법들로 대별된다. 첫 번째 방법은 칼라 영상을 이용^[1,2]하는 방법인데, HSI, YIQ, 정규화된 RGB칼라 모델 등을 이용하여 살색과 유사한 색을 영상에서 찾아 손 영역을 검출하게 된다. 그러나 이 방법은 보통의 그레이 영상보다 3배이상의 데이터가 필요하므로 처리량이 많아지게 된다. 두 번째 방법은 Optical flow이나^[3], 차 영상^[4]과 같은 프레임간의 차이를 이용하는 방법인데, 이 방법은 그레이 영상만을 이용해 데이터량이 줄어든다는 점은 있으나 손 영역이 움직이지 않는다면 손 영역을 검출하기는 어렵다는 점이 있다. 세 번째 방법은 모델을 이용하는 방법^[5]으로서, 그레이 영상만으로 충분하고 손 영역이 움직이지 않아도 손 영역을 검출할 수 있으나, 손모양에 따라 여러 가지 사용되는 모델이 필요하다는 점과 처리 시간이 길다는 것이 단점이 된다. 마지막 방법으로는 위의 세가지 방법을 복합적으로 사용하는 방법으로, 컬러 정보와 동영상을 복합해서 사용하는 경우 등이 있다^[6,7].

본 논문은 손 제스처 인식을 위한 실시간 손 영역검출 및 추적을 위한 알고리즘을 제안한다. 제안한 방법은 HSI 칼라 모델 기반의 칼라 정보를 이용하여 피부색에 해당하는 손 영역을 얻어낸다는 점이 기존의 방법과는 동일하나, 단지 Hue 영상만으로 손 영역을 검출한다는 점이 Hue와 Saturation을 모두 사용한 다른 방법과 구별된다. 미리 얻어진 손 영역의 Hue영상 histogram의 빈도수를 정규화한 후, 입력 영상에 재대입함으로써 살색에 해당하는 영역만을 그레이 영상으로 표현하게 된다. 얻어진 손 영역은 Kalman 필터를 사용하여 추적하게 된다.

본 논문의 구성은 다음과 같다. 2절에서는 손 영역 검출에 대하여 소개하며, 3절에서는 손 영역 추적에 대하여 설명한다. 4절에서는 시스템 구현 및 실험 결과를 기술하였다.

2. 손 영역 검출

본 절에서는 복잡한 배경에서 손 영역에 해당하는 영역을 검출하는 방법에 대하여 기술한다. 본 논문에서 제안한 방법은 기존의 칼라 정보를 이용하여 피부색에 해당하는 손영역을 얻어낸다는 점이 기존의 방법과는 동일하나, 3개의 칼라 영역 모두를 사용하는 것이 아니라 오직 1개의 칼라 영역만을 사용하게 된다.

손 영역에 해당하는 부분을 검출하기 위하여 먼저, 살색에 대한 정보를 이용하는 것이 필요하다. 살색에 대한 정보를 이용하기 위해서 YIQ, 정규화된 RGB, HSI 칼라 모델등 다양한 칼라 모델이 있으나, 본 논문에서는 기본적으로 HSI 칼라 모델을 사용하게 된다. HSI 칼라 모델은 Hue(색상), Saturation(채도), Intensity(명도)로 구성되어 있으며, Hue와 Saturation은 인간의 지각과 관계되며 Intensity는 빛의 세기와 관계되어 영상처리 알고리즘을 개발하는데 유용하다.

본 논문이 제안한 방법은 미리 얻어진 손 영역의 Hue 영상 histogram의 빈도수를 정규화한 후, 입력 영상에 재대입함으로써 살색에 해당하는 영역을 그레이 영상으로 표현하게 된다는 것이다. 본 논문이 제안하는 알고리즘의 pseudo code는 다음과 같다.

```

for i:=0 to 255 do begin
Normalized_Skin_Hue_Histogram[i]:
=(255/Max_Histogram_Frequency_Value)*Ski_
Hue_Histogram[i];
end;

for row:=0 to MAX_ROW do begin
for column:=0 to MAX_COLUMN do begin
Output_Image[row, column]:
=Normalized_Skin_Hue_Histogram[Input_Hue_I
mage[row, column]];
end;
end;
end;

```

미리 얻어진 손 영역의 Hue 영상의 histogram은 그림 1와 같으며, 그것의 정규화된 histogram의 그림은 그림 2와 같다.

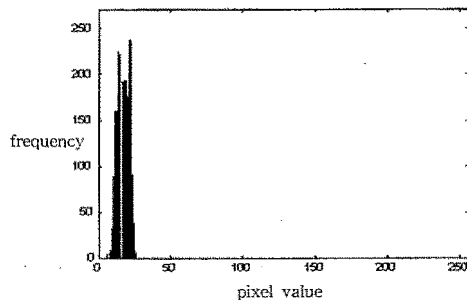


그림 1. 손 영역 Hue 영상의 histogram

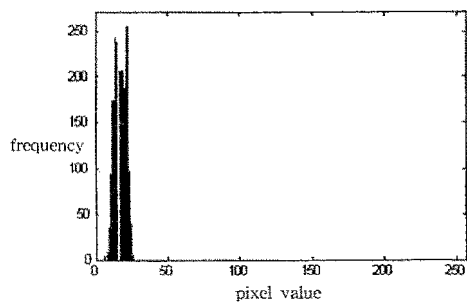


그림 2. 손 영역 Hue 영상의 정규화된 histogram

그림 3의 입력 영상으로부터 본 논문이 제안한 방법을 적용한 영상이 그림 4에 나타나 있으며, 그림 4를 erosion과 dilation을 취한 후 이진화된 영상은 그림 5에 나타나 있다. 그림 6은 살색과 유사한 색을 지닌 영역에서도 손 영역을 검출 할 수 있음을 보여 준다.



그림 3. 입력 영상



그림 4. 제안한 방법을 적용한 영상



그림 5. 이진 영상



그림 6. 살색과 유사한 색을 지닌 영역에서 손 영역을 검출한 경우

3. 손 영역 추적

입력 영상 전체를 처리하는 것이 아니라 관심의 대상인 특정 영역을 처리하도록 관심 영역 (Region Of Interest: ROI)을 제안한다면 계산 속도를 향상시킬 수가 있다. 본 논문에서는 관심의 대상이 되는 손영역을 추적하기 위하여 Kalman 필터^[8,9]를 사용하였다.

Kalman 필터는 다음 프레임에서 움직이는 특징점의 위치와 불확정성을 예측하는 재귀적인 알고리즘이다. Kalman 필터는 필터의 구조가 선형으로 구성되어 있으며, 선형 시스템에서 모든 변수들의 분포가 가우시안이고 부가되는 noise들이 white noise일 경우에는 그 특성이 최적을 지닌다고 알려져 있다^[10]. Kalman 필터는 system 모델과 measurement 모델로 구성되어 있다. 논 논문에서 사용되는 Kalman 필터의 system 모델은 다음과 같이 나타낼 수 있다.

$$\begin{pmatrix} x_{k+1} \\ y_{k+1} \\ v_{k+1,x} \\ v_{k+1,y} \end{pmatrix} = \begin{pmatrix} 1 & 0 & \Delta t & 0 \\ 0 & 1 & 0 & \Delta t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ v_{k,x} \\ v_{k,y} \end{pmatrix} + \omega_k$$

여기서 x_k 와 y_k 는 위치를 나타내며, $v_{k,x}$ 와 $v_{k,y}$ 는 속도를 ω_k 는 시스템 잡음을 모델링하는 평균이 0인 가우시안 random process를 나타낸다.

또한 measurement 모델은 다음과 같이 나타낼 수 있다.

$$\begin{pmatrix} x_k^m \\ y_k^m \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ v_{k,x} \\ v_{k,y} \end{pmatrix} + e_k$$

여기서 e_k 는 measurement noise를 모델링하는 평균이 0인 가우시안 random process를 나타낸다.

4. 시스템 구현 및 실험 결과

SAC-410ND 삼성 칼라 카메라와 Matrox Meteor프레임 그래픽 카드가 장착된 Pentium II 233MHz 컴퓨터를 사용하여 실험을 하였다. 입력 영상의 크기가 640×480 픽셀로 거의 실시간인 8 fps로 처리가 가능하였다.

Kalman 필터를 사용하여 추적의 대상이 되는 특징점으로는 손 영역의 중점을 사용하였다. 그림 7과 그림 8은 결과 영상으로 연속된 프레임의 일부를 나타내고 있다. 그림에서 큰 사각형 영역은 Kalman 필터를 사용하여 얻어진 손 영역이 위치할 예상 영역(즉, ROI)이 되며, 작은 사각형 영역은 실제로 얻어진 손 영역이다.



그림 7. 결과 영상 (1)



그림 8. 결과 영상 (2)

그림 9와 그림 10은 100 프레임 동안의 x 좌표와 y좌표 상에서의 Kalman 필터를 사용하여 예상되는 특징점(손의 중심)과 실제로 얻어진 특징점의 위치를 나타낸다. 두 그림에서 예상되는 특징점과 실제로 얻어진 특징점 사이의 에러가 작음을 알수 있다.

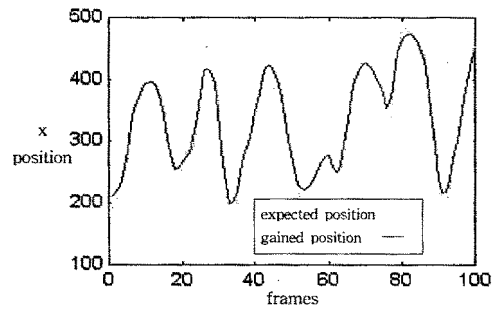


그림 9. x 좌표상에서 예상된 feature point 위치와 실제 얻어진 feaute point 위치

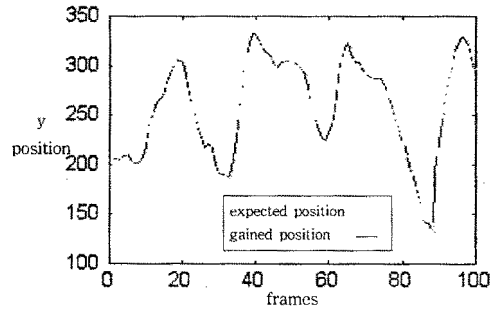


그림 10. y 좌표상에서 예상된 feature point 위치와 실제 얻어진 feaute point 위치

5. 결론

본 논문은 제스처 인식을 위한 손 영역 검출 및 추적을 위한 방법을 제안하였다. 손 영역을 검출하기 위하여 HSI 칼라 모델의 Hue 영상으로도 충분히 손 영역이 검출 가능하다는 것을 실험적으로 보여 주었다. 3개의 칼라 영역을 사용하여 처리하는 것보다 1개의 칼라 영역을 사용함으로써 계산량이 감소함을 확인 할수 있다. 손 영역 추적 방법으로는 Kalman 필터를 사용하였고 관심영역을 제안하여 거의 실시간으로 손 영역 추적이 가능하였다.

본 논문에서 제안한 방법은 비전에 기반한 인간과 컴퓨터와의 인터페이스에 적용이 가능할 것이다.

6. 참고 문헌

- [1] 윤호섭, 소정, 배창석, 전병태, 민병우, "손동작 제스처 인식을 이용한 윈도우 관리기 제어", 제 9회 영상처리 및 이해에 관한 워크샵 발표논문집, pp.264-269, 1997.
- [2] 이현규, 김호연, 김진영, "은닉 마르코프 모델을 이용한 연속 제스처 spotting", 제 9회 영상처리 및 이해에 관한 워크샵 발표논문집, pp. 114-119, 1997.
- [3] 송효섭, 양윤모, "Optical Flow를 이용한 손의 움직임 및 영역 추출", 한국정보과학회 '98 봄 학술 발표논문집(B), pp. 642-644, 1998.
- [4] 이형지, 이우규, 정재호, "수화용 손 제스처 인식을 위한 특징 벡터 추출", 제 11회 신호처리합동학술대회, pp. 895-898, 1988.
- [5] B. Moghaddam and A. Pentland "Probabilistic Visual Learning for Object Detection", The 5th International Conference on Computer Vision, 1995.
- [6] 조오영, 김형곤, 고성제, 안상철, 인터랙티브 가상환경을 위한 손 제스처 인식 시스템, 대한 전자공학회 36-s(4), pp. 70-82, 1999.
- [7] K. Imagawa, S. Lu and S. Igi. "Color-Based Hands Tracking System for Sign Language Recognition", in Proc. of IEEE International Conference on Automatic Face and Gesture Recognition, Killington, pp. 151-156, 1996.

- [8] E. Trucco, A. Verri, Introductory techniques for 3-D Computer vision, Prentice Hall, pp. 328-332, 1998.
- [9] G. Welch and G. Bishop "An Introduction to the Kalman Filter", Tech. Report TR 95-041 Department of Computer Science, Univ. of Carolina, 1995.
- [10] A. Gelb, "Applied Optimal Estimation", M.I.T. Press, 1974.