

비전을 이용한 움직임 물체 추출

김세진*, 탁명환*, 전철환*, 주영훈*, 박진배**
 군산대학교 전자정보공학과*, 연세대학교 전기전자공학과**

Movement Object Extraction Using Vision System

Se Jin Kim* · Myung Hwan Tak* · Chil Hwan Jeon* · Young Hoon Joon* · Jin Bae Park**
 School of Electronics & Information Engineering, Kunsan National University*
 Department of Electrical & Electronic Engineering, Yonsei University**

Abstract - 본 논문에서는 spatial gradient를 이용한 강인한 물체 추출 방법을 제안한다. 제안한 방법은 먼저 복잡한 환경과 다양한 빛의 변화에 의해 나타나는 여러 값 등을 해결하기 위해 기존에 제안된 입력 영상과 기준 영상에서 밝기와 색 성분을 이용하여 최초 배경을 제거한다. 배경을 제거한 다음, 그림자로 인식되어 전경 영역에 추가된 부분을 RGB 칼라 모델과 정규화 된 RGB 칼라 모델을 이용하여 제거하고, HSI 칼라 모델을 이용하여 불필요한 정보 값을 갖는 영역을 제거한다. 마지막으로, 배경으로 인식되어 전경으로부터 제거된 부분을 입력 영상의 공간상 정보인 spatial gradient와 HSI 칼라 모델을 이용하여 복구하는 방법을 제안한다. 마지막으로, 복잡하고 다양한 실내·외 환경에서의 실험을 통해 그 응용 가능성을 증명한다.

1. 서 론

오늘날 인간과 로봇사이의 자연스러운 상호작용을 위해 시각을 기반으로 한 사용자 의도 및 행위 인식에 대한 연구가 활발히 진행되고 있다. 이와 같은 기술로 인간-컴퓨터 간(Human-Computer Interaction)의 상호작용 기술이 있다. HCI기술은 인간과 컴퓨터의 상호 작용을 위해 인간의 시각 기능을 컴퓨터에 이식하여 컴퓨터 비전에 기반 한 사용자의 의도 및 행위를 인식하는 기술이다[1][2]. HCI기술 중 하나로 현재 가장 많이 연구 개발 중인 이동 객체(물체)분할이 있다. 이동 객체 분할 기술에서 일반적으로 물체를 배경으로부터 추출하는데 이용되는 배경 제거 기법으로 “연속적인 입력 영상의 누적된 픽셀 값의 평균, 중간, 차이 값을 이용하는 방법”[3]과, “기준 영상과 현재 입력 영상의 경계 픽셀 값의 차를 이용하는 방법”[4], 그리고 “기준 영상과 현재 입력 영상과의 차를 이용하는 방법”[5]등이 있다. 위 방법들은 가장 일반적인 배경 제거 기술임에도 불구하고, 특별한 문제점을 가지고 있다. 바로 전경을 추출할 경우 전경 영역 내에 그림자 또는 불필요한 값이 포함되어 전경부분이 삭제되는 것이다. 이와 같은 문제점을 해결하기 위해 기존의 제안된 방법으로는 입력 영상과 기준 영상에서 밝기와 색 성분을 이용하여 배경 제거하는 방법을[6] 제안(qualified RGB)하였고, 다른 방법으로는 Gaussian모델을 이용하여 전경을 추출하는 방법 등을 제안하였다. 하지만, 위의 방법론만으로는 확실한 전경을 얻을 수 없었다. 본 논문도 역시 이러한 문제점을 해결하기 위하여 제안한다. 본 논문에서는 그림자로 인식되어 추가된 부분과 배경으로 인식되어 삭제된 부분을 제거 및 복구 한다. 그림자 부분은 단순 RGB칼라 모델과 정규화 된 RGB칼라 모델[7]을 이용하여 제거한다. 그 후, 공간상 정보(Spatial gradient)[8]와 HSI칼라 모델[9][10]을 Morphology기법에 적용시켜 전경으로부터 제거된 부분을 복구한다. 본 논문은 다음과 같이 구성되어 있다. 2장에서는 물체 추출 시스템을 설명하고, 3장에서는 제안하는 알고리즘 기법을 설명한다. 4장에서는 실험 내용을, 마지막으로 5장에서 결론을 맺는다.

2. 본 론

2.1 물체 추출 시스템

본 논문의 시스템은 크게 세 단계로 구분하여 설명할 수 있다. 첫 번째 단계로 단순한 RGB칼라 모델과 정규화 된 RGB칼라 모델을 기존의 제안된 객체 분리 방법에 적용시켜 배경 모델을 생성한다. 그 후 생성된 배경 모델을 통해 현재 입력된 영상으로부터 그림자를 제거한 초기 전경을 추출한다. 두 번째 단계로 RGB칼라 모델을 변형 시켜 얻어진 HSI 칼라 모델로 빛의 변화에 대한 강인한 특성을 이용하여 입력 영상으로부터 불필요한 정보 값을 제거한 전경을 추출한다. 세 번째 단계로 Gradient기법을 이용하여 전체 영상의 기본 실루엣을 추출하고, 구해진 그 실루엣에서 복구 하고자 하는 이동하는 물체영역만을 바운딩 박스를 이용하여 제한한다. 그 후, 제한된 영역에 대해서만 초기 전경과 HSI칼라 모델을 이용하여 전경을 얻는다. 위 두 모델을 Morphology

(dilatation)기법에 적용시켜 바운딩 박스로 제한한 영역에 대해서 재차 복구한다. 결론적으로, 세 단계를 모두 OR 연산하여 움직이는 물체에 대한 정확한 영역 및 위치를 추출 한다.

2.2 공간상 정보를 이용한 강인한 물체 추출

기존의 전경 영역 분리 방법을 사용하여 전경을 추출하였음에도 불구하고 전경으로 추출된 영역 중에서 배경으로 인식되어 전경으로부터 제거된 부분이 있음을 그림 1을 통해 알 수 있다. 이와 같은 문제를 해결하기 위해서 본 논문에서는 공간상 정보를 이용한 강인한 물체 추출 알고리즘을 제안 한다. 제안하는 알고리즘은 세 가지 부분의 가정 하에 복구 및 제거 연산하여 정확한 전경 영역을 추출한다. 첫 번째, 전경 영역의 기준이 되는 값은 기존의 배경 제거 기법을 통해 얻어진 초기 전경 영역을 신뢰 할 수 있는 값으로 가정한다. 두 번째 공간상 정보인 Spatial gradient로 구해진 전체 입력 영상의 실루엣은 초기 전경 영역으로부터 제거 된 부분이 전부 포함되어 있다는 가정을 한다. 세 번째, 갑작스러운 빛의 변화에 강인함을 갖는 칼라 모델로 RGB칼라 모델을 변형시켜, 빛의 변화를 무시하고 단순히 RGB칼라 값만을 비교 할 수 있게 되면 그림자 영역과 불필요한 정보 값을 제거 하는데 효과적임을 가정한다.

2.2.1 RGB 칼라 모델 - 초기 배경 모델 생성

먼저, RGB칼라 모델만을 이용하여 배경 모델을 생성한 수식은 (1)과 같다. 식 (1)에서 Video는 연속된 N개의 영상의 배열을 의미한다. $Video^i(x)$ 는 연속된 N개의 배열인 Video에서 i번째 위치의 영상 x의 픽셀 값이고, $Video^z(x)$ 은 배경으로 그 중 선택된 픽셀 값을 의미한다. $\sigma(x)$ 는 N개의 전체 영상에서의 x 위치의 표준편차를 나타내고, $\lambda(x)$ 는 N개의 전체 영상의 위치 x의 픽셀 중간 값을 나타낸다. $\min(x)$, $Max(x)$ 는 최소값과 최대값을 의미하며, $gap(x)$ 는 이웃하는 영상간의 위치 x 픽셀 값 중 그 픽셀간의 차이가 최대임을 나타낸다.

$$\begin{aligned} \min(x) &= \min_z \{ Video^z(x) \} \\ Max(x) &= Max_z \{ Video^z(x) \} \\ gap(x) &= Max_z \{ |Video^z(x) - Video^{z-1}(x)| \} \\ \lambda(x) &= Avr_z \frac{1}{2} [Max(x) + \min(x)] \end{aligned} \tag{1}$$

여기서 $Video^z(x)$ 는 고정된 픽셀처럼 분류되어진다. 영상의 x 위치에서 중간 값 $\lambda^z(x)$ 와 현재 픽셀 값 $Video^z(x)$ 의 차이가 표준편차 $\sigma^z(x)$ 의 2배 보다 작은 경우 배경 픽셀로 결정된다. 이렇게 배경 픽셀로 결정된 픽셀들을 이용하여 식 (1)과 같이 초기 배경 모델을 생성한다.

$$B(x) = \begin{cases} 0, & background \left\{ \begin{aligned} (I^t(x) - \min(x)) < kd_u \\ \vee (I^t(x) - Max(x)) < kd_u \end{aligned} \right\} \\ 1, & foreground \text{ otherwise.} \end{cases} \tag{2}$$

초기 배경 모델 $[\min(x), Max(x), gap(x)]$ 을 이용하여 현재 입력된 영상으로부터 움직이는 물체를 식 (2)에 분리해 낼 수 있다. 식 (2)에서 $I^t(x)$ 는 현재 시간 t에 대한 입력 영상에서 x위치의 픽셀 값을 나타낸다. 상수 k는 전경부분과 배경부분의 경계치를 의미하며, 상수 k의 값이 커질수록 분리되는 전경 영역이 적어지게 되고, 반대로 상수 k의 값이 적은 경우는 상수 k의 값이 큰 경우 보다 많은 전경 영역을 분리해 낸다.

2.2.2 정규화 된 RGB 칼라 모델 - 그림자 제거

정규화 된 RGB칼라 모델은 픽셀 값을 밝기 성분으로 정규화 하기 때문에 현재 입력된 영상에서 픽셀 간의 밝기 변화에 관계없이 색상의 차이만을 비교할 수 있다. 만약 $R+G+B \neq 0$ 고 $r=g=b=0$ 이 아니면, 정규화 된 RGB칼라 모델은 식(3)과 같이 표현된다.

$$r = \frac{R}{R+G+B}, g = \frac{G}{R+G+B}, b = \frac{B}{R+G+B} \quad (3)$$

위와 같은 정규화 된 RGB칼라 모델에서는 색상 차이만을 비교 하는 것이 가능하게 된다. 기준영상의 원색과 현재 입력 영상의 원색을 1:1 비교가 가능하기 때문에 조명에 의한 변화에 명도차이는 고려하지 않아도 된다.

2.2.3 HSI 칼라 모델 - 불필요한 정보 값 제거

RGB칼라 모델을 조명에 덜 민감한 HSI칼라 모델로 변환하여 이용한다. 다음의 식 (4)는 RGB칼라 공간을 HSI칼라 공간으로 변환한 수식이다.

$$I = \frac{1}{3}(R+G+B) \quad (4)$$

$$S = 1 - \frac{3}{(R+G+B)} [\min(R, G, B)]$$

$$H = \cos^{-1} \left[\frac{\frac{1}{2} [(R-G) + (R-B)]}{\sqrt{(R-G)^2 + (R-B)(G-B)}} \right]$$

위의 수식을 이용하여 현재 입력 영상으로부터 불필요한 정보 값을 제거 한다.

2.2.4 물체 추출 영역 복구(spatial gradient)

연속적인 영상 정보에 있어서 기본적인 두 가지 정보는 영상의 공간상 정보(spatial gradient, SG)와 영상의 시간상 정보(temporal gradient, TG)이다. 두 가지 정보 중 공간상 정보는 입력 영상에 모든 윤곽선 정보를 가지고 있는 반면 전경 영역의 윤곽선만을 구분하기가 힘들다. 따라서, 이 단점을 최소화 하고 모든 윤곽선 정보를 가지는 장점을 최대화 하기 위해 Bounding Box을 이용하여 복구 하고 자 하는 영역의 범위를 한정한다. 연속적인 영상 정보를 시간 t 에 대해서 다음 식(4)과 같이 정의한다.

$$I(x, y, t) = [I_r(x, y, t), I_g(x, y, t), I_b(x, y, t)]^T \quad (4)$$

여기서, x, y 는 입력영상의 픽셀 정보, r, g, b 는 입력영상의 RGB 정보이고, $I_r(x, y, t), I_g(x, y, t), I_b(x, y, t)$ 은 시간 t 에 대해서 입력 영상의 RGB 픽셀 정보이다. 공간상 정보는 다음 식(5)과 같이 계산된다.

$$I_s = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right] = \|I_s\| \quad (5)$$

여기서, I_s 는 각각 입력영상의 공간상 정보, 시간상 정보이다. 그러나 일반적으로 에지 정보는 정수 형태를 가지나 상기 정보는 실수 형태의 값을 가진다. 따라서 이를 정수 형태로 변경시켜 사용한다.

3. 실험 및 결과 고찰

본 실험에서 사용된 영상은 320*240 크기의 24bit 칼라 영상이고 프레임 속도는 15frame/sec 이다. 최초 RGB 칼라 모델과 정규화 된 RGB 칼라 모델을 기존의 제안된 배경 제거 기법에 적용시켜 초기 전경 영역을 추출 하였다. 추출된 결과는 그림 1을 통해서 확인할 수 있다.



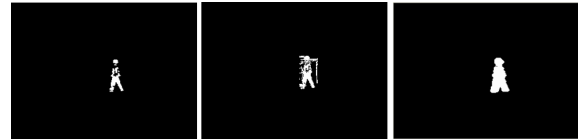
<그림 1> 전경 영역 추출 영상 : (a) 기준 영상 (b) 입력 영상 (c) RGB 칼라 모델 + 정규화 된 RGB 칼라 모델

위와 같이 기존에 제안된 방법을 통해 초기 전경 영역을 구한 후, SG 기법을 이용하여 입력 영상에 대한 전체 실루엣을 구한다. 구해진 실루엣에서 복구 하고자 하는 움직이는 물체 영역을 바운딩 박스로 영역을 제한한다. 그림 2(b)은 구해진 실루엣과 바운딩 박스로 제한된 영역을 나타낸다. 그림 2을 통해, 초기 전경 영역으로부터 제거된 부분이 모두

포함되어 있음을 확인할 수 있다. 따라서 두 번째 가정된 내용이 만족함을 알 수 있다. 이로써, 제한하는 알고리즘을 이용하여 이동 물체의 전경 영역을 완벽하게 복구 할 수 있다.



<그림 2> 바운딩 박스 안쪽의 실루엣에 대한 전경 영역이 차지하는 비율 : (a)전경 영상 (b) Spatial gradient의 실루엣 영상 (c)실루엣에 대한 모폴로지(팽창)방법+ HSI 칼라 모델을 이용하여 제거된 전경 영역을 복구하는 영상



<그림 3> 제안하는 알고리즘을 이용한 전경 추출 : (a) 기존의 전경 추출 방법을 이용한 전경 영상 (b) 바운딩 박스 안에 Spatial gradient 로 얻은 실루엣과 (a)를 OR 연산한 영상 (c) 제안한 알고리즘을 이용한 결과 영상

4. 결 론

본 논문에서는 SG 기법을 이용한 강인한 물체 추출 방법을 제안하였다. 이를 위해 기존에 제안된 전경 영역 분리 방법인 RGB 칼라 모델과 정규화 된 RGB칼라 모델을 이용하여 그림자를 제거한 1차 전경을 추출하였다. 그리고 HSI 칼라 모델을 이용하여 불필요한 정보 값을 제거한 2차 전경을 추출하였다. 위의 1, 2차 추출 방법을 통해 추출된 전경을 제안하는 시스템에 적용시켜 전경 영역의 대한 신뢰성을 높였으며, 정확한 위치와 형태를 추출 하였다. 마지막으로, 본 논문에서 제안한 알고리즘은 복잡하고 다양한 실내·외 환경에서의 실험을 통해 그 응용 가능성을 증명하였다.

[참 고 문 헌]

- [1] D. M. Gavrilu, L. S. Davis, "Towards 3D model based tracking and recognition of human movement: a multi view approach", Int Workshop on Face and Gesture Recognition, Vol. 162479, pp. 272-277, 1995. 6.
- [2] V. I. Pavlovic, R. Sharma, and T. S. Huang, "Visual Interpretation of Hand Gestures for Human Computer Interaction: A Review", IEEE, Transaction on PAMI, Vol. 19, No. 7, pp.677-695, July, 1997.
- [3] I. Haritaoglu, D. Harwood, and L. S. Davis, "W4 : Real-Time Surveillance of People and Their Activities," in Proc. IEEE Transactions On Pattern Analysis and Machine Intelligence, vol. 22, no. 8, pp. 809-822, 2000. 8.
- [4] P. L. Rosin, "Thresholding for Change Detection" Brunel University, technical report ISTR, pp. 274-279, 1998. 1.
- [5] K. S. Bhat, M. Satharishi, and P. K. Khosla, "Motion detection and segmentation using image mosaics," in Proc. IEEE Int. Conf. Multimedia and Expo, vol. 3, pp. 1577-1580, 2000. 6.
- [6] P. Blauensteiner, H. Wildenauer, A. Hanbury, and M. Kampel, "On Colour Spaces for Change Detection and Shadow Suppression" Computer Vision Winter Workshop, pp. 117-123, 2006. 2.
- [7] G. D. Finlayson, B. Schiele, and J. L. Crowley. "Comprehensive Color Image Normalization". In 5th European Conference on Computer Vision, Vol. 1, pp. 475-490, 1998.
- [8] M. H. Kim, Y. H. Joo, and J. B. Park "Gesture Extraction for Ubiquitous Robot-Human Interaction" The Institute of Control, Automation, and Systems Engineers, KOREA Vol. 11, No. 12, pp. 1062-1067 2005, 12.
- [9] C. Garcia and G. Tziritas, "Face Detection Using Quantized Skin Color Regions Merging and Wavelet Packet Analysis", IEEE Trans. on Multimedia, Vol. 1, No. 3, pp. 264-277, 1999, 9.
- [10] G. J. Jang, I. S. Kweon. "Robust Objects Tracking Using an Adaptive Color Model" IEEE Trans. International Conference on Robotics & Automation. Vol. 2, pp. 1677-1682, 2001. 5.