

비디오 카메라를 이용한 감시 장치에서 그림자의 제거

Removing Shadows for the Surveillance System Using a Video Camera

김 정 대* 도 용 태**
(Jungdae Kim Yongtae Do)

Abstract – In the images of a video camera employed for surveillance, detecting targets by extracting foreground image is of great importance. The foreground regions detected, however, include not only moving targets but also their shadows. This paper presents a novel technique to detect shadow pixels in the foreground image of a video camera. The image characteristics of video cameras employed, a web-cam and a CCD, are first analysed in the HSV color space and a pixel-level shadow detection technique is proposed based on the analysis. Compared with existing techniques where unified criteria are used to all pixels, the proposed technique determines shadow pixels utilizing a fact that the effect of shadowing to each pixel is different depending on its brightness in background image. Such an approach can accommodate local features in an image and hold consistent performance even in changing environment. In experiments targeting pedestrians, the proposed technique showed better results compared with an existing technique.

Key Words : Video surveillance , Shadow pixel detection, Foreground image

1. 서론

비디오에 의한 감시(VS: Video Surveillance)는 특정 관심 지역을 관찰할 수 있도록 설치된 카메라의 비디오 영상을 보안이나 교통 관리 등 주어진 목적에 맞게 자동 처리하는 컴퓨터 시각 기술로, 근래에 와서 많은 연구자들의 관심을 끌고 있다. VS의 주요한 연구주제는 차량이나 사람 등 동적 물체를 비디오 영상내에서 검출하고(moving object detection)[1], 그 물체의 종류를 분류하고(object classification)[2], 동작을 분석하고 이해하는(activity analysis and understanding)[3] 것들이 된다.

일반적으로 VS의 첫 번째 절차는 전경(foreground)의 추출인데, 추출된 전경에는 관심의 대상이 되는 동적 물체 외에도 이의 그림자까지 포함되게 된다. 이 때 그림자는 물체를 본래의 형태와는 다르게 보이게 하므로, 분류나 인식과 같이 후행되는 절차들에서 오류를 유발할 수 있다. 이러한 이유로 영상내 그림자 제거를 위한 기법이 다양하게 연구되어 왔다. Prati 등[4]은 기존의 기법들을 분류하고 비교 실현한 결과, Cucchiara 등[5]이 제안한 컬러정보를 이용한 DNM 기법이 여러 가지 상황에서 가장 안정적으로 좋은 결과를 보였음을 보고한 바 있다.

저자 소개

* 學生會員 김정대 : 情報通信工學科 制御學科 碩士課程
** 正會員 도용태: 電子情報工學部 制御學科 教授 · 工博

본 논문에서는 실외의 비구조화된 환경내 보행자를 대상으로 한 VS 카메라로부터 얻어진 영상의 전경에서 그림자 화소를 판정함으로서 이를 제거하기 위한 새로운 기법이 제안된다. 사람은 차량과는 달리 영상내에서 차지하는 면적이 작아 해상도가 낮고, 고자유도로 인해 형태의 가변성이 크며, 이동 경로가 비예측적이라는 특성에 기인한 기술적 난점 때문에 연구의 필요성이 높다. 본 논문의 기법은 각 화소에 따라 다른 기준으로 그림자를 판정한다는 특징을 지닌다. 즉, 그림자에 기인한 화소 밝기의 강하가 화소가 배경에서 가지는 밝기에 의존적인 사실을 확인하고, 이를 그림자 화소의 판정에 이용한다. 제안된 기법은 전체 화소에 대해서 획일적인 그림자 판정기준을 적용하는 기존의 기법에 비해서 높은 공간적 그리고 시간적 적용성을 기대할 수 있다.

2. 그림자 화소의 특성

그림자 화소의 특성을 알아보기 위해 학교 캠퍼스내 사람의 영상을 다양한 시간 및 환경에 따라 획득하여, 그 변화를 알아보았다. 그림 1은 웹 카메라와 일반 CCD 카메라를 이용하여 학교 캠퍼스내에서 흔히 볼 수 있는 잔디밭과 시멘트 영역에 그림자가 드리워졌을 때와 그렇지 않을 때의 특성을 나타낸 것이다. 그림에서 알 수 있듯이, 그림자영역은 RGB(Red-Green-Blue)값들 모두가 낮아지고, 이는 HSV(Hue-Saturation-Value) 컬러공간에서 V값으로 대표될 수 있다. 그밖에 H값은 상당히 불안정하며, 특히 그림자영역에서 심했다. S값은 그림자에 의한 변화가 상대적으로 작았다.

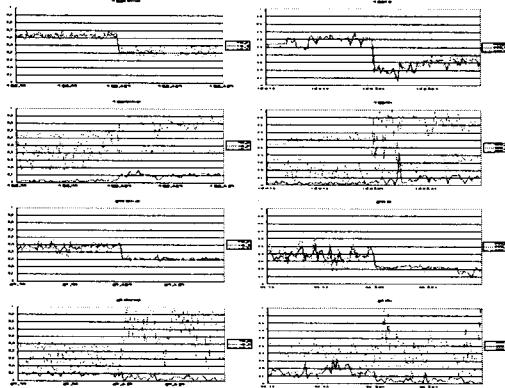


그림 1 웹 카메라 영상에서의 그림자 특성 - 시멘트영역 RGB(좌상), HSV(좌중상), 잔디영역RGB(좌중하), HSV(좌하) - 과 일반 CCD카메라 영상에서의 그림자 특성 - 시멘트영역RGB(우상), HSV(우중상), 잔디영역RGB(우중하), HSV(우하)

3. 그림자 제거를 위한 판정기법

Cucchiara 등[5]에 의한 그림자 판정방법은 화소기반의 DNM 기법으로 HSV 컬러공간에서 수행되었다. 구체적으로, 비디오 영상의 t번째 프레임에서 전경내 임의 화소 p 는 식(1)과 같은 3가지 조건을 모두 만족할 경우 그림자 영역 sh에 속한다고 판정하였다

$$p \in sh \text{ if } C_1 \wedge C_2 \wedge C_3 \quad (1.a)$$

단,

$$C_1: a \leq \frac{I'_V(p)}{B'_V(p)} \leq \beta \quad (1.b)$$

$$C_2: |I'_S(p) - B'_S(p)| \leq \tau_S \quad (1.c)$$

$$C_3: \min(|I'_H(p) - B'_H(p)|, 360 - |I'_H(p) - B'_H(p)|) \leq \tau_H \quad (1.d)$$

단, $I(p)$ 는 화소 p 의 추출된 전경에서의 값이고, $B(p)$ 는 화소 p 의 배경에서의 값이며, 아래 첨자는 화소의 H,S,V의 값을 지칭하기 위해 사용되었다. 또 a, β, τ_S, τ_H 는 판정을 위해 사용된 임계치들이다.

본 논문에서는 2장에서의 분석을 기반으로 화소의 V값만을 이용한 판정법을 제안한다. 우리는 기준의 기법과는 달리 영상내 국부적인 특성이 고려될 수 있고, 그 기준은 환경의 시간상 변화에도 유효할 수 있도록 하고자 하였다. 즉, 임의 화소는 특정한 밝기값을 가지고 있으며, 이 값은 시간에 따라 변화할 수 있다고 할 때, 갑자기 어떤 형태로 이 값이 강하하게 되면 그림자가 드리워졌다고 판단하는 것이다. 제안된 방식은 그림 2(좌)와 같이 각 카메라의 영상 화소의 밝기값, 즉 V값이 그림자에 의하여 강하하는 형태의 분포를 얻고, 이를 그림 2(우)와 같이 균사화 함으로서 실현하였다. 두 그림에서 화소의 배경에서의 V값 $B_V(p)$ 를 가로축으로 하고, 그림자에 의해서 그 값이 낮아진 정도, 즉 $B_V(p) - I_V(p)$ 를 세로축으로 하였다. 단, 각 값들은 최대값을 1로 정규화 하였다. 여기서 그림자에 의한 밝기의 강하특성은 그림자가 드리워지기 전의 밝기값에 의존적이다.

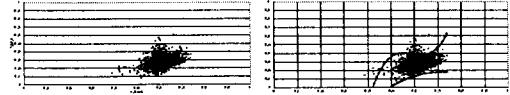


그림 2. (좌)는 그림자에 의한 밝기변화 특성, (우)는 곡선에 의한 모형화

그림 2에서 볼 수 있듯이 우리는 그림자에 의한 밝기 변화의 분포를 식 (2)와 같이 표현되는 3차 곡선으로 균사화한다. 즉, 배경에서 B_V 의 밝기값을 가지는 화소의 현재 영상 프레임에서의 밝기값이 I_V 일 때, B_V 에서 I_V 로의 강하가 미리 정의된 3차 곡선 (2.a)과 곡선 (2.b)사이에 있으면 그림자로 판정한다.

$$D_{Cmax} = k_1 B_V^3 + k_2 B_V^2 + k_3 B_V + k_4 \quad (2.a)$$

$$D_{Cmin} = l_1 B_V^3 + l_2 B_V^2 + l_3 B_V + l_4 \quad (2.b)$$

여기서 k 는 밝기 강하값이 같은 화소값들의 평균(u)에 30를 더한 값을 균사화한 식의 계수이고, l 은 30를 뺀 값을 균사화한 식의 계수이다.

4. 결과

기준의 Cucchiara기법[5]과 본 논문에서 제안한 그림자 화소 판정기법을 실제 영상에 적용하여 그 결과를 비교하여 보았다. 그림 3은 그 결과를 보여준다. 그림에서 회색은 전경 화소를 의미하며, 검은 색은 적용된 기법에 의하여 그림자로 판정된 화소를 의미한다.

정량적으로 기법의 성능을 평가하기 위한 척도로 식 (3)과 같이 정의되는 검출률(detection rate) η 와 구분률(discrimination rate) ξ 를 사용하였다[4]. 여기서 TP_S (true positive shadow)는 그림자 화소가 그림자 화소로 바르게 판정된 경우의 수, TP_F (true positive foreground)는 동적 물체에 속하는 화소가 동적 물체에 속하는 화소로 바르게 판정된 경우의 수, FN_S (false negative shadow)는 그림자 화소가 동적 물체에 속하는 화소로 잘못 판정된 경우의 수, FN_F (false negative foreground)는 동적 물체에 속하는 화소가 그림자 화소로 잘못 판정된 경우의 수, 그리고 \overline{TP}_F 는 동적 물체에 속하는 화소 수에서 동적 물체에 속하나 그림자로 판정된 화소의 수를 뺀 수를 나타낸다.

$$\eta = \frac{TP_S}{TP_S + FN_S} \quad (3.a)$$

$$\xi = \frac{\overline{TP}_F}{\overline{TP}_F + FN_F} \quad (3.b)$$

시간에 따른 환경의 변화가 기법의 성능에 어떠한 영향을 주는지 알아보기 위하여, 그림 3의 영상에서 결정된 그림자 판정기준인 식 (4)를 그림 4의 영상에 그대로 적용하여 보았다. 두 그림은 각각 전날 오후와 익일 오전에 촬득된 영상들이다. 두 그림에 식 (3)의 성능평가 척도를 적용한 결과는 표 1과 같다. 그림과 표에서 알 수 있듯이 기준의 Cucchiara의 기법과 비교하여 제안된 기법은 환경이 변화하는 경우에도 성능을 유지하는 경향이 상대적으로 큼을 알 수 있다.

$$D_{Cmax} = 91.575 B_V^3 - 162.61 B_V^2 + 95.85 B_V - 18.367 \quad (4)$$

$$D_{Cmin} = -0.8503 B_V^3 - 1.7767 B_V^2 + 3.8644 B_V - 1.3689$$

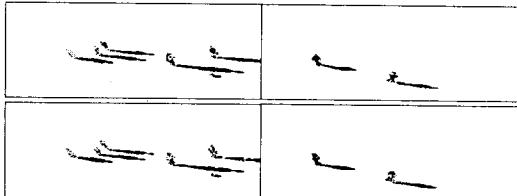


그림 3. 그림자 화소 검출의 예. 회색은 추출된 전경이며, 흑색은 그림자로 판정된 화소이다:
 (좌상) 웹카메라 영상에서 Cucchiara 기법의 결과,
 (좌하) 웹카메라 영상에서 제안된 기법의 결과,
 (우상) 일반 CCD영상에서 Cucchiara 기법의 결과,
 (우하) 일반CCD영상에서 제안된 기법의 결과

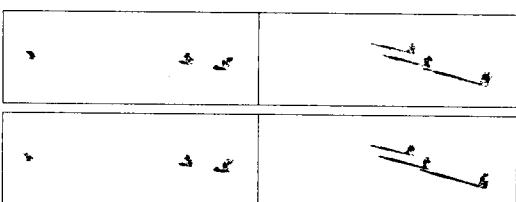


그림 4. 그림자 화소 검출의 다른 예. 그림 3의 영상에서 결정한 판정기준을 그대로 적용하였다:
 (좌상) 웹카메라 영상에서 Cucchiara 기법의 결과,
 (좌하) 웹카메라 영상에서 제안된 기법의 결과,
 (우상) 일반CCD 영상에서 Cucchiara 기법의 결과,
 (우하) 일반CCD 영상에서 제안된 기법의 결과

5. 결론

본 논문에서는 비디오 영상의 전경중 그림자를 제거하기 위한 판정기법을 제안하였다. 제안된 기법은 화소의 밝기값이 그림자에 의해 강하될 때, 그 강하 정도가 해당 화소의 배경 밝기값에 의존적임을 이용하였다. 배경 영상의 화소별로 밝기값이 다르며, 환경의 변화에 따라 배경이 적응적으로 갱신된다고 할 때, 제안된 기법에서는 영상의 환경에 따른 변화가 자연스럽게 판정에 적용되는 장점이 있다.

제안된 기법은 크게 시멘트와 잔디로 구성된 학교의 캠퍼스내 사람의 영상을 대상으로 그 성능이 평가되었다. 기존의 기법[5]과 비교하였을 때, 제안된 기법은 피관측 지역의 특성과 환경의 변화에 둔감한 결과를 보였다. 특히 그림자 화소의 검출률에서 제안된 성능의 장점이 두드러졌으며, 전경 물체와 그림자의 구분률에서는 비슷한 성능을 보였다. 일반적으로 VS 장치가 장시간동안 중단 없이 운용되며, 자동적으로 동적물체를 추출하고 해석해야 함을 고려하면, 제안된 기법은 현실적으로 효용가치가 높다.

표 1 그림 3과 그림 4의 결과의 정량적 비교; P는 제안된 기법, C는 Cucchiara의 기법

구분	장면	기법	TPs	TP _F	FN _S	FN _F	TP _F	η	ξ
웹캠	오후	P	1340	678	289	158	678	0.82	0.81
	그림3	C	1344	647	285	189	647	0.83	0.77
	오전	P	127	459	49	202	459	0.72	0.69
	그림4	C	111	436	65	225	436	0.63	0.66
일반	오후	P	442	231	56	170	231	0.89	0.58
	그림3	C	336	245	162	156	245	0.67	0.61
	오전	P	674	401	139	285	401	0.83	0.59
	그림4	C	449	479	364	207	479	0.55	0.70
CCD	오후	P	442	231	56	170	231	0.89	0.58
	그림3	C	336	245	162	156	245	0.67	0.61
	오전	P	674	401	139	285	401	0.83	0.59
	그림4	C	449	479	364	207	479	0.55	0.70

참 고 문 헌

- [1] H. Fujiyoshi and A. J. Lipton, "Real-time human motion analysis by image skeletonization," in *Proc. IEEE Workshop Application of Computer Vision*, pp. 15-21, 1998.
- [2] L. Wixson and A. Selinger, "Classifying moving objects as rigid or non-rigid," in *Proc. DARPA Image Understanding Workshop*, pp. 341-347, 1998.
- [3] I. Haritaoglu, et al, "W4: who? when? where? what? a real time system for detecting and tracking people," in *Proc. IEEE Int'l Conf. Automatic Face and Gesture Recognition*, pp. 222-227, 1998.
- [4] A. Prati, et al, "Analysis and detection of shadows in video streams: A comparative evaluation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 571-576, 2001.
- [5] R. Cucchiara, et al, "Detecting objects, shadows and ghosts in video streams by exploiting color and motion information," in *Proc. IEEE Int'l Conf. Image Analysis and Processing*, pp. 360-365, 2001.