

사용자 개입을 통한 효과적인 배경 모델 생성 기법

*김형민, **이재훈, **박종일¹, ***김유경, ***김광용
 *한양대학교 전자컴퓨터통신공학과, **한양대학교 컴퓨터공학부, ***한국전자통신연구원
 *hmkim@mr.hanyang.ac.kr, **ljh0128@hanyang.ac.kr, ***jipark@hanyang.ac.kr,
 ****yk.kim@etri.re.kr, ****kwangyk@etri.re.kr

The Effective Background Modeling Method by User Intervention

*Hyungmin Kim, **Jae Hoon Lee, **Jong-Il Park,
 ***Yookyung Kim, ***Kwang-yong Kim
 *Department of Electronics and Computer Engineering, Hanyang University
 **Division of Computer Science and Engineering, Hanyang University
 ***Intelligent Convergence Media Research Dept, ETRI

요 약

객체를 추적하는 기술은 컴퓨터 비전 분야에서 활발히 연구되고 있는 분야 중 하나이다. 그 중 고정된 단일 카메라를 이용한 객체 추적 기술은 비디오 감시(Surveillance) 등에서 활용되고 있다. 고정된 카메라 환경에서 객체를 추적하는 방법 중 배경 모델링(Background Modeling)을 이용한 방법은 간단하면서도 널리 사용되는 방법 중 하나이다. 객체의 움직임이나 특징을 분석하여 배경 모델을 생성한 후 배경 정보를 이용하여 전경을 분리하면 쉽게 객체를 추출할 수 있다. 그러나 객체의 움직임이 적은 경우 해당 영역에서의 배경 모델은 정확하게 생성될 수 없다. 배경 모델을 학습하는 동안 객체가 충분히 움직이면 이런 문제를 해결할 수 있으나 객체가 움직이기 전까지는 오류가 지속된다. 이런 문제를 해결하기 위해 본 논문에서는 인페인팅(Inpainting)을 이용하여 움직임이 적은 영역을 보정하여 정확한 배경 모델을 생성하는 방법을 제안한다. 배경 모델을 생성한 후 객체로 식별할 수 있는 후보 영역을 식별한다. 선정된 영역들 중 사용자가 객체로 판단되는 영역을 선택하여 해당 영역에 대해 인페인팅으로 화소값 및 가중치들을 보정한다. 보정된 영상으로 배경 모델링을 수행하면 움직임이 적은 영역에 대해서도 효과적으로 배경 모델을 생성 할 수 있다.

1. 서론

카메라를 이용하여 객체를 검출하고 추적하는 기술은 컴퓨터 비전 분야에서의 중요한 기술로 오래전부터 연구되어 왔고 현재도 활발히 연구되고 있는 기술이다. 그 중 고정된 카메라 환경에서의 객체 추적 방법은 비디오 감시 분야 등에서 활용되고 있다. 고정된 카메라로 획득한 영상에서 전경과 배경을 분리하면 객체를 검출 할 수 있다. 전경과 배경을 분리하기 위해서는 영상에서 움직임이나 특징 변화를 분석하여 배경 정보를 생성하는 배경 모델링 방법이 사용된다. 배경 모델을 생성하기 위해서는 흔들림에 따른 동적인 변화나 조명이나 그림자와 같은 밝기 변화에 대한 고려가 필요하다. 이러한 문제를 해결하기 위해 다양한 연구가 진행되고 있다. 대표적인 배경 모델링 방법으로는 가우시안 모델을 이용하는 GMM(Gaussian Mixture Model) [1], 픽셀 정보의 변화를 분석한 코드북(Codebook) 모델 [2] 등이 있다. GMM 방법은 k 개의 가우시안 확률 밀도 함수로 배경 픽셀 분포를

모델링하여 배경모델을 생성한다. 그러나 이 방법은 가우시안 확률 밀도 함수에 의존적이고 급격한 조명 변화나 동적인 배경에 민감한 단점이 있다. 코드북 모델은 픽셀단위로 코드북을 생성하여 픽셀의 변화정보를 분석하는 방식으로 배경 정보를 생성하는 방법이다. 이렇게 생성된 배경 정보는 급격한 조명 변화나 동적인 배경에도 강인한 특징을 갖는다. 이와같은 방식들로 생성된 배경 정보들은 주로 객체의 조명 변화나 움직임을 분석하여 변화가 적은 부분을 배경 모델로 생성하게 된다. 그러나 배경 모델을 생성하거나 학습을 할 때 조명변화나 움직임이 없는 객체의 경우에는 객체가 배경 모델로 분류될 확률이 크다. 이렇게 생성된 배경 모델을 이용하여 전경 객체를 추출하게 되면 정지되어 있거나 움직임이 적었던 객체 주변의 배경 모델에서는 정확한 전경 객체를 추출 할 수 없게 된다. 전경 객체를 추출하면서 배경 모델을 갱신하는 동안 객체가 움직이면 이런 문제를 해결 할 수 있지만 배경 모델이 갱신되는 동안에는 오류가 계속 발생하게 되고 객체의 움직임이 발생해야지만 해결될 수 있다. 본 논문에서는 이러한 문제를 해결하기 위해 객체로 인식될 수 있는 후보군을

¹ 교신저자

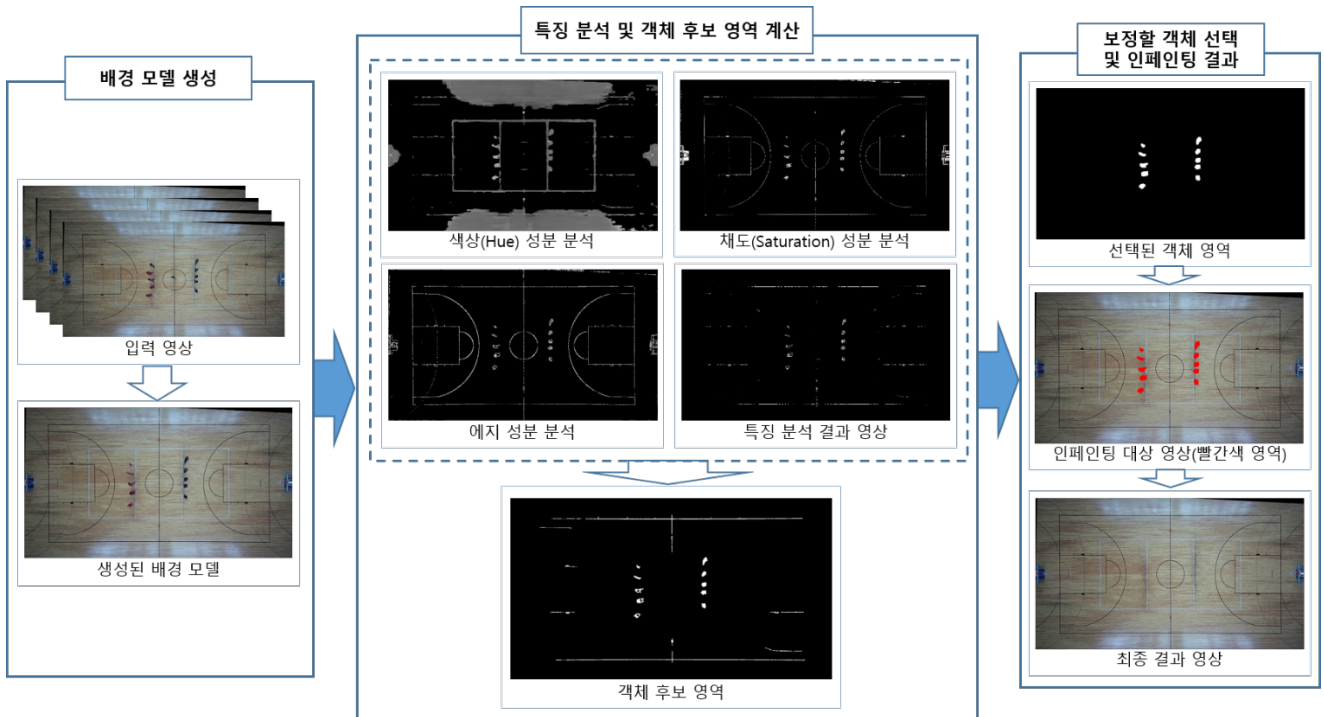


그림 1. 제안한 방법의 전체 흐름도

선정하여 사용자가 객체로 판단되는 영역을 선택한 후 선택된 영역에 대해서 인페인팅으로 보정함으로써 움직임이 없는 객체 영역에 대해서도 정확한 배경 모델을 생성하는 방법을 제안한다.

본 논문의 구성은 다음과 같다. 2 장에서는 배경 모델링과 인페인팅 및 이를 이용하여 제안한 알고리즘에 대해 설명하고, 3 장에서는 제안한 방법에 대한 실험 및 수행결과와 분석을 나타내었다. 마지막으로 4 장에서는 결론을 맺는다.

2. 움직임이 적은 객체 제거를 위한 배경모델 생성 기법

그림 1 은 제안한 기법의 전체 흐름도를 나타낸다. 실내 경기장이나 큰 대로변과 같이 배경 환경이 복잡하지 않은 환경을 대상으로 배경 모델을 생성하게 되면 생성된 배경 모델의 색상 정보나 형태(에지, 텍스처 등)의 특징 변화가 심하지 않고 일정하게 변하게 된다. 그러나 정지된 객체가 있었던 부분에서는 이러한 특징들의 변화가 심하게 발생한다. 이러한 특징을 이용하여 HSV 색정보와 에지 정보를 이용하여 객체를 할 수 있는 후보 영역을 계산한다. 계산된 후보 영역들 중 사용자가 실제 객체라 판단되는 영역을 선택한 후 인페인팅으로 해당 영역에 대해 주변의 배경 모델 정보(화소값, 가중치 등)들로 보정하여 배경 모델을 생성하게 되면 정지된 객체가 있던 영역에 대해서도 효과적으로 배경 모델을 생성할 수 있다.

2.1 배경 모델링

본 논문에서는 컬러 정보와 형태 정보를 함께 활용하여 배경 모델을 생성하는 방법 [3]으로 배경 모델을 생성한다.

LBP(Local Binary Patterns)는 텍스처가 풍부한 환경에서는 강인한 특징을 보이지만 변화가 단조로운 텍스처 환경이나 그림자 영역 등에서는 효율적이지 못하다. 이러한 단점을 보완하기 위해 그림 2 와 같은 광도에 불변하는 컬러 모델 (Photometric Invariant Color)을 함께 이용한다.

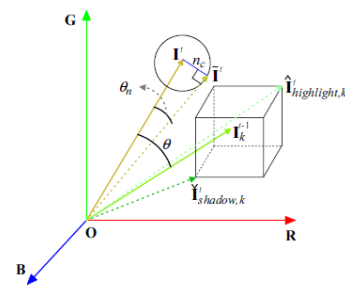


그림 2. 광도에 불변하는 컬러 모델

여기서 I^t 는 t 프레임에서의 RGB 벡터를 나타내고 $\hat{I}_{highlight,k}$ 는 최대 밝기값, $\hat{I}_{shadow,k}$ 는 최소 밝기값(그림자), n_c 는 노이즈 성분, θ 는 이전 영상의 RGB 벡터의 차이이다.

각 화소에 대해 식 1 과 같은 모드를 구성하여 배경 모델을 생성한다.

$$m_k = \{I_k, \hat{I}_k, \hat{I}_{shadow,k}, LBP_k, w_k, \hat{w}_k, L_k\}, k = 1, \dots, K. \quad (1)$$

LBP_k 는 학습된 LBP들 중 해당 모드의 평균 LBP값을 나타내고 $w_k \in [0,1]$ 는 배경으로 선택될 확률에 대한 가중치, \hat{w}_k 는 이전 가중치들 중 가장 큰 가중치를 나타낸다. 각 화소별로 모

드를 계산하여 가중치가 가장 큰 모드, 즉 배경일 확률이 가장 클 화소값을 이용하여 배경 영상을 생성하게 된다.

2.2 보정 영역 설정

2.1 절에서 생성된 배경 영상을 이용하여 입력 영상과의 차영상을 구하면 객체만을 분리할 수 있다. 그러나 그림 4(b)에서와 같이 움직임이 적은 객체가 있을 경우 정상적인 배경 영상을 생성 할 수 없다. 객체의 움직임이 적어 오류가 발생한 영역을 보정하기 위해 보정할 영역을 계산한다. 배경 환경이 복잡하지 않다는 가정하에 그림 1의 특징 분석 및 객체 후보 영역 계산 부분과 같이 배경 모델 영상의 특징을 분석한다. 색상과 형태 변화를 분석하기 위해 배경 모델의 RGB 영상을 HSV로 변환한 후 식 2와 같이 임계치 이상의 값들을 이용하여 영상을 생성한다.

$$I_{hsv}(x,y) = \begin{cases} 0 & I_{hsv} < th \\ otherwise & \end{cases} \quad (2)$$

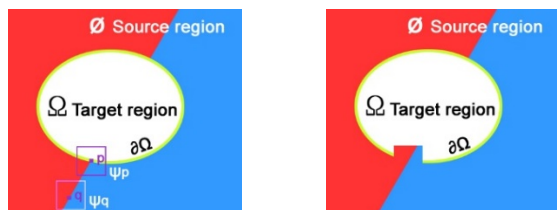
Sobel 에지 검출기를 이용하여 에지 영상을 생성한 후 식 2를 통해 생성된 영상과 공통된 영역을 계산한다. 계산된 영역을 객체 후보 영역으로 설정하고 사용자가 실제 객체라고 판단되는 영역을 선택한다. 사용자에게 의해 선택된 영역들에 대해 인페인팅으로 보정한다.

2.3 영상 인페인팅

영상 인페인팅은 영상에서 훼손된 부분을 복원하거나 불필요한 부분을 제거하고 제거된 부분에 대해서 자연스럽게 복구하는 기술이다. 본 논문에서는 2.2절에서 사용자에게 의해 객체라 선택된 영역에 대해서 Criminisi[4][5] 등이 제안한 방법을 이용하여 복원한다. 그림 3은 인페인팅 방법을 나타낸다.



(a) 복원할 영역 선택 (b) 우선순위가 높은 패치 선택



(c) 선택된 패치와 유사한 패치탐색 (d) 유사한 패치 복사

그림 3. 인페인팅 진행 과정

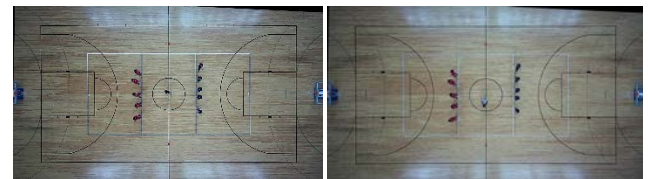
사용자에게 의해 복원할 영역이 설정되면 외곽선위의 점을 중심으로 하는 영역 중 가장 높은 우선순위를 갖는 패치를 찾는다. 선택된 패치와 유사한 패치를 탐색하여 가장 유사도가 높은 패치를 복사함으로써 설정된 영역을 복원한다.

3. 실험 결과 및 분석

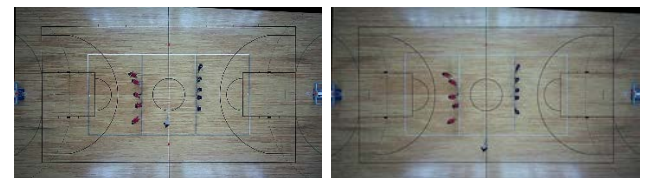
실험 영상은 농구 경기장 천장에 카메라를 고정하여 촬영한 영상을 사용하였다. 그림 4는 실험 영상으로 영상 중앙의 객체(심판)를 제외한 객체들의 움직임이 거의 없는 것을 알 수 있다. 기존의 방법을 이용하여 배경 모델을 생성하면 그림 5의 (b)와 같이 객체의 움직임이 있는 중앙의 객체가 있던 영역들에서는 정상적으로 배경 모델이 생성이 되나 선수들이 서 있는 영역에 대해서는 객체가 있던 영역들에 대해서도 배경 모델로 생성이 된다. 이렇게 생성된 배경 모델로 전경 객체를 검출하면 그림 5의 (e)와 같이 실제 객체가 검출이 되지 않는 것을 확인할 수 있다. 제안한 방법을 이용하여 그림 5의 (c)와 같이 배경 모델을 생성한 경우 그림 5의 (f)와 같이 모든 객체가 검출되는 것을 확인할 수 있다.



(a) 1 번째 영상 (b) 100 번째 영상



(c) 150 번째 영상 (d) 175 번째 영상



(e) 200 번째 영상 (f) 225 번째 영상

그림 4. 배경 모델을 생성할 영상들

4. 결론

본 논문에서는 배경 모델 생성 시 움직임이 적은 객체가 있는 영역을 사용자의 개입을 통해 보정하여 정확한 배경 모델을 생성하는 방법을 제안하였다. 배경 모델을 생성한 후 영상의 특징 변화를 분석하여 움직임이 적은 객체 영역 후보군을 생성한다. 후보군들 중 사용자가 객체로 판단한 영역을 선택하여 인페인팅으로 해당 영역을 보정한 후 배경 모델을 생성하면 움직임이 적은 객체가 있는 영역에 대해서도 정확한 배경 모델은

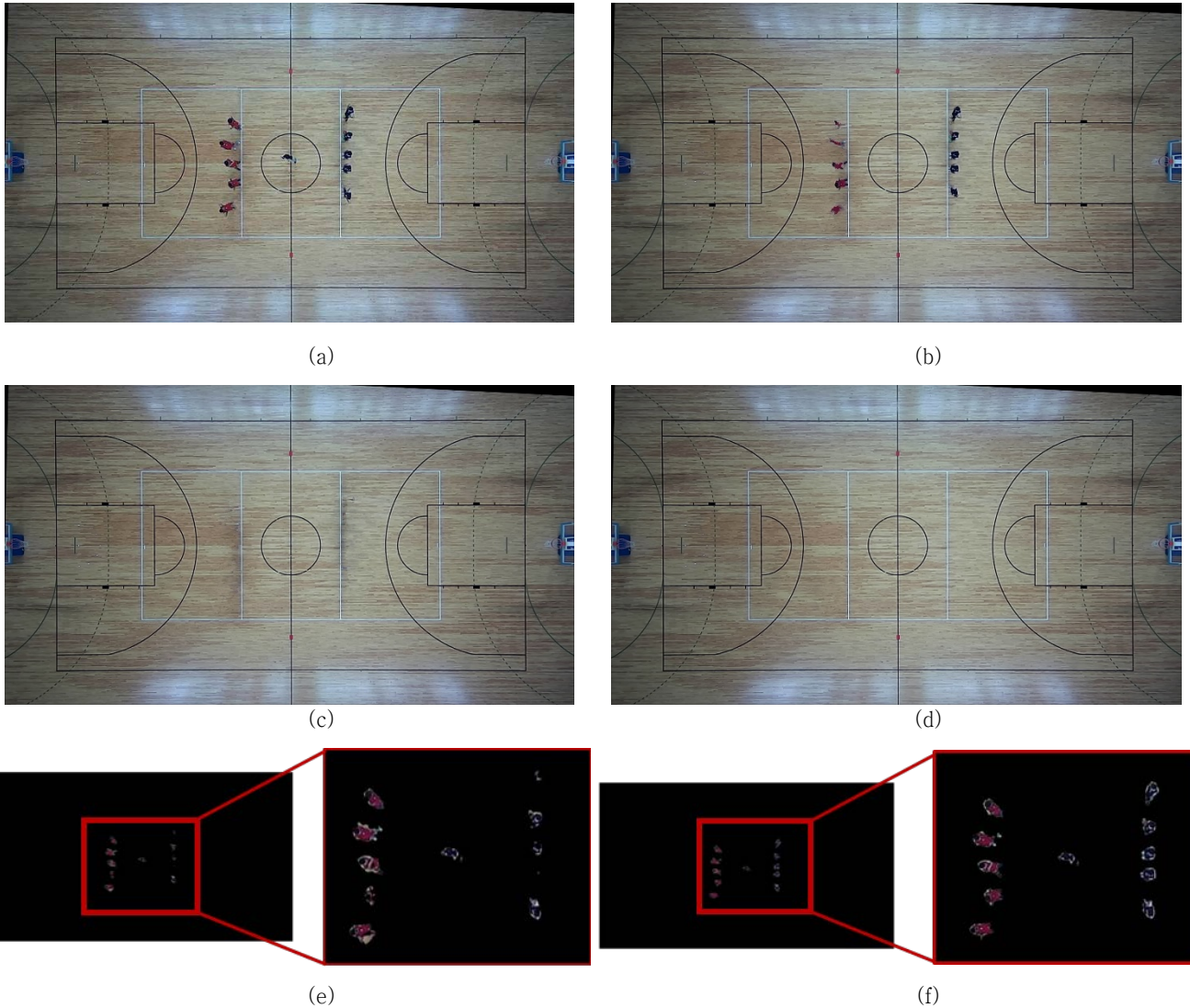


그림 5. (a) 입력 영상, (b) 움직임이 적은 객체를 고려하지 않은 배경 모델 생성 결과 영상, (c) 제한한 방법으로 생성된 배경 영상, (d) 배경 영상(Ground Truth), (e) (b)의 배경 모델을 이용하여 전경 객체를 분리한 결과 영상, (f) (c)의 배경 모델을 이용하여 전경 객체를 분리한 결과 영상.

생성할 수 있는 것을 확인할 수 있었다. 사용자의 개입 없이 학습 등을 통해 자동으로 후보 영역을 설정하는 연구를 진행할 계획이다.

감사의 글

본 연구는 미래창조과학부 및 정보통신기술연구진흥센터의 정보통신·방송 연구개발사업의 일환으로 수행하였음. [R0101-16-0293, 방송용 영상 인식 기반 객체 중심 지식융합 미디어 서비스 플랫폼 개발]

참고문헌

[1] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking" in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol. 2, pp.246 -252, 1999.

[2] K. Kim, T. H. Chalidabhongse, D. Harwood and L. Davis "Real-time foreground-background segmentation using codebook model", Real-Time Imaging, vol. 11, pp.172 -185 2005.
 [3] Yao, Jian, and Jean-Marc Odobez. "Multi-layer background subtraction based on color and texture." IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2007.
 [4] A. Criminisi, P. Perez and K. Toyama, "Region Filling and Object Removal by Exemplar-based Image Inpainting," IEEE Trans. Image Processing, vol.13, no.9, pp. 1200-1212, 2004.
 [5] 이승훈, et al. "프로젝션 기반의 감소현실 시스템." 컴퓨터그래픽스학회논문지 13.2 (2007): 52-57.