

혼잡한 환경에 적합한 적응적인 배경모델링 방법

이광국[†], 송수한^{**}, 가기환^{***}, 윤자영^{****}, 김재준^{*****}, 김희을^{*****}

요 약

기존의 배경 모델링 방법은 배경 모델의 반복적 갱신(recursive update)으로 인해 배경보다 객체가 더 자주 등장하는 혼잡한 환경에서는 정확한 배경 모델링을 생성하기 어려운 문제를 지니고 있다. 본 논문은 이러한 기존 방법의 문제를 해결하기 위해 기존의 혼합 Gaussian 모델을 기반으로 하는 적응적 배경 모델링 방법을 제안한다. 제안한 방법은 영상 내 전경 영역의 비율에 따라 배경 모델의 학습 비율을 적응적으로 조절한다. 따라서, 혼잡 상황에서는 배경 모델의 갱신을 억제하여 배경 모델을 잘 유지시키는 것이 가능하다. 실험을 통해 제안한 방법이 일반적인 상황의 영상에서는 기존 방법과 유사한 정확도를 보이지만 혼잡한 상황에서는 기존 방법과 비교하여 배경 제거를 효과적으로 수행하는 것을 확인하였으며, 또 정확도 측정 결과 혼잡한 상황의 영상에서 기존 방법과 비교하여 F 값이 5-10% 가량 향상함을 확인하였다.

Adaptive Background Modeling for Crowded Scenes

Gwang-Gook Lee[†], Su Han Song^{**}, Kee-Hwan Ka^{***}, Ja-Young Yoon^{****},
Jae-Jun Kim^{*****}, Whoi-Yul Kim^{*****}

ABSTRACT

Due to the recursive updating nature of background model, previous background modeling methods are often perturbed by crowd scenes where foreground pixels occurs more frequently than background pixels. To resolve this problem, an adaptive background modeling method, which is based on the well-known Gaussian mixture background model, is proposed. In the proposed method, the learning rate of background model is adaptively adjusted with respect to the crowdedness of the scene. Consequently, the learning process is suppressed in crowded scene to maintain proper background model. Experiments on real dataset revealed that the proposed method could perform background subtraction effectively even in crowd situation while the performance is almost the same to the previous method in normal scenes. Also, the F -measure was increased by 5-10% compared to the previous background modeling methods in the video of crowded situations.

Key words: Video surveillance(영상 감시), background subtraction(배경 제거), background modeling(배경 모델링), Gaussian mixture model(Gaussian 혼합 모델)

※ 교신저자(Corresponding Author): 이광국, 주소: 서울시 성동구 행당동(133-791), 전화: 02)2220-0561, FAX: 02)2292-6316, E-mail: gglee@vision.hanyang.ac.kr
접수일: 2008년 1월 25일, 완료일: 2008년 3월 25일

[†] 준회원, 한양대학교 전자컴퓨터통신공학과 박사과정

^{**} 준회원, 삼성전자 디지털미디어총괄 디지털프린팅사업부
(E-mail: suhan.song@samsung.com)

^{***} 준회원, 한양대학교 전자컴퓨터통신공학과 석사과정
(E-mail: khka@vision.hanyang.ac.kr)

^{****} 준회원, 한양대학교 건축환경공학과 석·박사통합과정
(E-mail: nyx80@hanyang.ac.kr)

^{*****} 정회원, 한양대학교 건축공학부 교수
(E-mail: jjkim@hanyang.ac.kr)

^{*****} 중신회원, 한양대학교 전자통신컴퓨터공학부 정교수
(E-mail: wykim@hanyang.ac.kr)

※ 본 연구는 2007년도 한국건설교통기술평가원 연구비 지원 (06건설핵심D06) 에 의해 수행되었습니다.

1. 서 론

감시 카메라, DVR 등과 같은 영상 장비의 발달과 공공장소에서의 안전에 대한 관심의 증대로 인해 최근 영상 감시 시스템에 대한 수요가 증가하고 있으며, 이에 따라 지능형 영상 감시 시스템의 개발에 대한 요구도 꾸준히 증가하고 있는 추세이다. 지능형 영상 감시 시스템은 전적으로 인력에 의해서만 영상을 모니터링하던 기존의 감시 시스템과는 달리, 컴퓨터가 입력 영상을 분석하여 상황에 대한 인지를 자동으로 행함으로써 감시 시스템의 효과적인 운용을 가능케 한다. 이러한 지능형 영상 감시 응용의 예로는 객체 추적[1], 영상 내 군중의 밀도 측정[2,3], 통행량 측정[4], 이상 행동의 감지[5] 등이 있다. 그런데, 이러한 지능적 영상 감시의 다양한 응용들 가운데 많은 방법들은 영상으로부터 차량이나 사람과 같은 객체를 분리해내는 배경 제거(background subtraction)을 전처리 과정으로 이용한다. 배경 제거 방법에서는 초기에 주어진 영상에 의해 또는 연속된 일련의 입력 영상을 통하여 배경 모델을 생성하며, 생성된 배경 모델과 현재 영상의 차이에 의하여 배경을 제외한 객체만을 추출하게 된다. 이러한 배경 제거의 결과는 이후 전체 프로세스의 정확도에 큰 영향을 끼치기 때문에 정확한 배경 제거는 효과적인 지능적 영상 감시 시스템의 개발에 있어서 매우 중요한 문제이다.

지금까지 효과적인 배경 제거를 위해 많은 연구가 있어왔으며, 그 가운데 W4 시스템에서는 시간적 윈도우내의 최대값, 최소값, 연속한 픽셀 값의 차이의 최대값과 같은 몇 가지 순서 통계량을 이용하여 효과적으로 배경 제거를 수행하는 방법을 제안하였다[6]. 이 방법은 간단한 연산만으로 이루어져 실시간 처리가 가능하였다. 하지만, 이 방법은 회색조 영상을 대상으로 하고 있어서 복잡한 환경에서는 잘 동작하지 못할 수 있다. Koller 등은 Kalman 필터를 이용하여 배경 조명의 변화에 대응하는 방법을 제안하였다[7]. 이 방법은 시간이나 날씨에 따라 발생하는 영상의 점진적 변화에 효과적으로 대응할 수 있었으나, 배경 모델 생성에 있어서 픽셀을 선택적으로 이용하였기 때문에 새로운 물체가 배경으로 등장하는 것과 같은 배경의 급격한 변화에는 대응할 수 없었다. Pfinder 시스템에서는 각 픽셀 당 하나의 Gaussian

모델을 이용하여 배경을 모델링 하는 방법이 제안되었다[8]. 이 방법은 Gaussian 모델의 파라미터를 매프레임의 입력마다 반복적으로 갱신하여 조명의 변화뿐 아니라 배경의 변화에 적용할 수 있었다. 하지만, 이 방법은 정적인 배경만을 고려하였기 때문에 나뭇가지의 흔들림이나 회전문과 같이 계속적으로 발생하는 배경의 변화에 대해서는 효과적으로 이용되기 어려웠다. Grimson 등은 각 픽셀당 하나의 혼합 Gaussian 모델(GMM: Gaussian Mixture Model)을 이용하여 이러한 문제를 해결하였다[9]. 이 방법에서는 각 픽셀이 3-5개의 Gaussian 모델의 합산으로 모델링되기 때문에 나뭇잎의 흔들림, 물결의 출렁거림과 같이 지속적으로 발생하는 배경의 변화에도 대응하는 것이 가능하였다. 이 방법은 현재까지 제안된 배경 제거 방법 중 가장 널리 이용되고 있는 방법 가운데 하나이다.

또한 Grimson의 방법을 기반으로 하여 혼합 Gaussian 모델을 이용한 많은 배경 모델링 방법이 제안된 바 있다. KaewTraKulPong 등은 Grimson의 방법이 갖는 문제점 가운데 하나인 느린 초기화의 문제를 해결하기 위하여 초기 L 프레임의 입력까지는 충분 통계량(sufficient statistics)를 이용하여 배경 모델을 생성하며, 이후 L 프레임 이상의 입력이 발생하였을 때 기존의 online EM 방법을 이용한 배경 모델 갱신으로 전환하는 방법을 제안하였다[10]. 하지만, 이 방법은 충분한 프레임이 입력된 이후에 발생한 배경의 변화에 대해서는 여전히 기존의 Grimson 등의 방법과 마찬가지로 느린 적응 속도를 갖는 문제점이 있었다. Lee 등은 각 Gaussian 모델의 발생 회수에 따라 학습 속도를 제어하는 파라미터를 도입하여 배경 모델의 수렴 속도를 향상시켰다[11]. 이 방법은 배경에 변화가 발생했을 때는 충분 통계량을 이용한 방법과 유사하게 동작하여 빠른 수렴속도를 가지며, 배경 모델이 수렴될 수록 Grimson 등의 원래 배경 모델 갱신 방법인 반복적인 갱신(recursive update) 방법에 접근하는 특성을 가지고 있다. Tang 등은 Grimson 등의 방법을 간략화 하여 기존 방법과 유사한 성능을 유지하면서 실행 속도를 향상시킬 수 있는 방법을 제안하였다[12]. 이 방법은 Grimson 등의 방법에서 몇 가지 파라미터를 제거하고 간략화 함으로써 유사한 성능을 유지하면서도 속도를 1.6 배 향상시키는 것이 가능했다.

효과적인 배경 제거를 위해 이처럼 많은 방법들이 제안되었지만, 기존의 배경 제거 방법은 다수의 움직이는 물체가 계속적으로 발생하는 혼잡한 환경에서는 제대로 동작하기 어려운 문제가 있다. 이는 기존의 배경 제거 방법들에서는 환경의 변화에 적응하기 위하여 새로운 프레임의 픽셀 값을 배경 모델의 갱신에 계속적으로 이용하기 때문이다. 즉 재귀 필터(recursive filter)에 의해 배경 모델을 업데이트 하는 기존의 배경 제거 방법은 시간적 윈도우(temporal window)내의 화소값의 통계치를 이용하는 것으로 해석될 수 있으며, 따라서 배경이 영상에 등장하는 빈도보다 움직이는 객체가 시간적으로 더 많이 등장하는 혼잡한 상황에서는 배경 모델이 잘못 생성될 수 있다. 본 논문에서는 이러한 문제를 해결하기 위하여 혼잡 상황에서도 효과적으로 배경 제거를 가능하게 하는 배경 모델링 방법을 제안한다. 제안한 방법은 다중 Gaussian 혼합 모델을 이용한 Lee의 배경 모델링 방법을 기반으로 하여 영상 내 전경 영역의 비율에 따라 모델 갱신을 위한 학습 비율을 조정한다. 즉, 혼잡한 상황에는 배경 모델의 갱신을 억제시킴으로써 혼잡한 상황에서도 배경 모델을 효과적으로 유지하는 것이 가능하다.

본 논문의 구성은 다음과 같다. 2장에서는 제안한 방법의 기반이 되는 기존의 배경 모델링 방법을 설명하며, 3장에서는 기존의 방법을 개선한 제안된 배경 모델링 방법을 소개한다. 4장에서는 실험 결과를 보이며, 5장에서는 결론을 맺는다.

2. 기존의 혼합 Gaussian 모델 방법

본 논문에서 제안된 배경 모델 생성 방법은 기존의 혼합 Gaussian 모델 방법에 기반하고 있다. 따라서, 이 장에서는 우선 Grimson 등에 의해 제안된 기존의 혼합 Gaussian 모델 방법과 이를 개선하여 빠른 수렴속도를 갖도록 한 Lee의 방법을 소개한다.

2.1 Grimson의 방법

Grimson에 의해 제안된 배경 모델링 방법에서는 각 픽셀 위치에서 픽셀 값들의 확률분포 수식 (1)과 같이 K 개의 Gaussian 모델의 혼합으로 표현된다. $w_{i,t}$ 는 시간 t 에서 i 번째 Gaussian 모델의 가중치를 나타낸다.

$$I = \sum_{i=1}^K \omega_{i,t} * g(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (1)$$

수식 (1)의 $g(X_t, \mu_{i,t}, \Sigma_{i,t})$ 는 Gaussian 확률 밀도 함수를 나타낸다. X_t 는 픽셀값으로써 RGB 컬러공간 의 3차원 벡터이다.

$$g(X_t, \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{n/2} |\Sigma_{i,t}|^{1/2}} e^{-\frac{1}{2} (X_t - \mu_{i,t})^T \Sigma_{i,t}^{-1} (X_t - \mu_{i,t})} \quad (2)$$

또, $\mu_{i,t}$ 와 $\Sigma_{i,t}$ 는 각각 시간 t 에서 i 번째 Gaussian의 평균과 공분산행렬을 나타낸다. 계산 속도를 위하여 R, G, B 각 컬러채널은 동일한 분산을 갖는다고 가정되므로 공분산행렬의 대각 성분은 모두 같은 값을 갖는다.

K 개의 Gaussian 모델들 가운데 다음 수식 (3)의 조건을 만족하는 Gaussian 모델들이 배경 모델로써 이용된다. 이때 Gaussian 모델들은 fitness 값, w_k/ω_k 에 의해 미리 정렬되어 있으므로 가중치가 높고 분산이 적은 Gaussian 모델이 우선적으로 배경으로써 선택된다. 가중치가 높은 모델은 동일 시간 내에 더 많은 빈도로 발생하였으므로 배경일 가능성이 높으며, 분산이 적을수록 정적이므로 마찬가지로 배경일 가능성이 높게 된다. 임계치 T 는 배경이 최소한으로 차지해야 하는 시간적 비율을 의미한다.

$$\text{arg min}_b \left(\sum_{i=1}^K \omega_k > T \right) \quad (3)$$

Gaussian 모델 K 와 배경 모델 선택을 위한 임계치 T 의 선택은 배경의 특성에 의해 적절하게 이루어져야 하며, 배경이 계속적으로 변하는 환경에서는 더 큰 값을 선택하고, 배경의 변화가 비교적 적은 환경에서는 적은 값을 선택하여야 한다.

Grimson의 방법에서 배경 모델의 갱신은 on-line K-means 알고리즘에 의해서 이루어진다. 즉, K 개의 Gaussian 모델 가운데 정합(matching)이 발생한 첫 번째 Gaussian 모델이 다음의 수식 (4) - (7)에 의해서 갱신된다. 수식 (4)에서 α 는 학습 속도를 결정하는 학습 비율으로써 α 가 작으면 더 많은 시간 윈도우(temporal window)내의 분포를 이용하여 배경 모델이 생성되므로 시스템의 안정성이 증가하는 장점이 있으나 반대로 배경 모델의 수렴 속도가 느려지게 되는 문제가 있다. 하지만, 필요 이상으로 큰 α 를 사용하게 되며 배경의 변화에 배경 모델이 빠르게 적응

하는 것은 가능하지만, 배경 모델이 잡음에 민감해지고 너무 적은 수의 과거 값들만을 이용하게 되므로 시스템의 불안정해질 수 있다. Gaussian 모델의 파라미터는 수식 (8)에 나타난 것과 같이 학습 비율과 정합이 발생한 Gaussian 모델의 사후확률(posterior)의 곱으로 결정된다.

$$= (1-\alpha)\omega_{k,t-1} + \alpha(M_{k,t}) \quad (4)$$

$$= (1-\rho)\mu_{t-1} + \rho X_t \quad (5)$$

$$l) = (1-\rho^*) \cdot \sigma_{t-1}^2 + \rho(X_t - \mu_t)^2 \quad (X_t - \mu_t) \quad (6)$$

$$= \alpha \eta(X_t | \mu_k, \sigma_k) \quad (7)$$

2.2 Lee의 방법

Lee는 Grimson의 개선한 Gaussian 모델링 방법을 제안한 바 있다. Lee의 방법에서는 수식 (7)에 의해서 나타난 배경 모델의 학습 비율을 다음과 같이 수정하는 방법을 제안하였다.

$$= q_k \cdot \left(\frac{1-\alpha}{c_k} + \alpha \right) \quad (8)$$

여기에서 c_k 는 각 Gaussian 모델이 생성된 이후 발생한 정합시의 사후확률(posterior)의 총 합을 갖는다. 즉, 어떤 Gaussian 모델의 c_k 가 크다면, 이는 해당 Gaussian 모델을 갖는 픽셀이 발생한 횟수가 많았음을 의미한다.

Lee의 방법에서는 on-line EM을 사용했으므로, Gaussian 모델의 갱신이 하나의 Gaussian 모델에 대해서만 이루어지는 것이 아니라 매칭된 모든 Gaussian 모델에 대하여 사후확률의 비율에 따라 이루어지게 된다.

$$= p_k / \sum_{i=1}^L p_i \quad (9)$$

$$(10)$$

Lee에 의해 제안된 개선된 배경 모델 방법은 새로운 Gaussian 모델이 처음 생성되었을 때는 c_k 가 매우 작으므로 학습 비율 η_k 는 사후확률 q_k 에 가까운 값을 가지며, 따라서 충분통계량을 이용한 배경 모델링 방법과 유사하게 동작한다. 반대로 어떤 Gaussian 모델이 충분한 시간 동안 남아있게 되면 c_k 가 큰 값을 가지

므로 η_k 는 α 에 의해서 조절되며, 이는 기존의 on-line 알고리즘에 의해 모델이 갱신되는 효과를 갖는다. 즉, Lee에 의해 제안된 배경 모델링 방법은 배경의 변화에 대해 빠른 수렴 속도를 가지면서도 기존의 시스템과 같은 안정성을 갖게 된다. 그러나 Lee의 방법은 빠른 적응 속도로 인해 앞서 언급된 것과 같이 객체가 배경보다 자주 등장하는 혼잡한 상황에서는 객체를 배경으로 더 쉽게 잘못 인식할 우려가 있다.

3. 제안된 방법

이 장에서는 제안하는 개선된 혼합 Gaussian 모델 기반 배경 모델 생성 방법을 소개한다. 제안하는 배경 모델 생성 방법은 앞 장에서 소개된 Lee의 효과적인 혼합 Gaussian 모델 방법을 기반으로 하여 배경 모델의 학습 속도를 조절하는 학습 비율을 영상의 혼잡도에 따라 변경함으로써 혼잡한 상황에서도 효과적으로 배경 모델이 유지될 수 있도록 한다.

3.1. 혼잡도에 따른 학습 비율 조절

2장에서 소개된 기존의 배경 모델 생성 방법은 on-line 학습 알고리즘에 배경 모델을 반복적으로 갱신하게 된다. 이러한 반복적 갱신 방법은 최근 L 프레임 동안의 화소값들을 이용하여 배경 모델의 확률 밀도함수를 생성하는 방법으로 해석될 수 있다. 그런데, 배경보다 객체가 더 많은 빈도로 발생하는 혼잡 상황에서는 기존 방법에 의해 올바른 배경 모델을 생성하는 것이 어려워질 수 있다. 제안한 배경 모델링 방법에서는 혼잡도에 따라 학습 속도를 적응적으로 조절하여 이러한 문제를 해결하며 혼잡 상황에서는 배경 모델의 학습 속도를 낮춤으로써 생성된 배경 모델이 잘 유지될 수 있도록 하였다.

본 논문에서는 영상 내의 혼잡도를 계산하기 위하여 배경 분리로 얻어진 전경 영역의 비율을 이용하였으며, 분리된 전경 픽셀 가운데 이용이 빈번히 발생하는 영역의 픽셀에 더 높은 중요도를 부여하기 위해 수식 (11)과 같은 가중치 지도(weight map), $WROI$ 를 도입하였다.

$$ROI(i) = \frac{1-\alpha}{\sum_{i=0}^M \sum_{t=0}^L FG(i,t)} \quad (11)$$

가중치 지도는 정해진 초기 훈련 시간 동안 수식 (11)에 의해 계산된다. 수식 (11)에서 L 은 가중치 지도 계산을 위한 초기 훈련시간의 길이이며, M 은 영상의 전체 픽셀 개수를 나타낸다. $FG(i, t)$ 는 시간 t 의 프레임에서 i 번째 픽셀이 전경일 경우 1 을, 해당 픽셀이 배경일 경우 0 을 갖는다. 본 논문에서는 가중치 지도의 생성을 위한 초기 훈련 기간에는 기존의 혼합 Gaussian 기반 배경 제거 방법을 이용하였으나, 인접한 프레임의 차이와 같은 다른 방법이 이용될 수도 있다.

그림 1은 생성된 가중치 지도의 예이다. 가중치 지도는 훈련 기간 동안 영상 내의 각 픽셀 위치에서 전경 영역이 발생한 상대적인 빈도를 나타내며, 그림 1에 나타난 것과 같이 잦은 이동이 발생하는 출입문 주변에서 큰 값을 갖는 것을 알 수 있다. 가중치 지도 계산을 위해 본 논문에서는 기존의 배경 제거 방법을 이용하였으므로 마찬가지로 혼잡한 상황에서 전경 영역을 정확하게 획득하는 것은 어려울 수 있으나, 가중치 지도는 객체의 정확한 분리가 아니라 관찰 공간 내 이동량의 통계 획득이 목적이므로 이는 전체 시스템의 성능에 큰 영향을 끼치지 않는다.



(a)



(b)

그림 1. 가중치 지도의 예: (a) 실험 영상의 예, (b) 생성된 가중치지도

프레임 t 의 혼잡도 c 는 수식 (12) 와 같이 정의된다. 혼잡한 상황 하에서는 영상 내에 전경 영역이 넓게 발생할 것이므로, 혼잡도는 영상 내 전경 픽셀의 전체 개수로 계산될 수 있다. 이때 객체의 이동이 빈번히 발생하는 영역의 전경 픽셀에 중요도를 더 주기 위하여, 가중치 지도의 값을 전경 검출 결과에 곱해 주어 전체 가중합을 영상의 혼잡도로 이용한다.

$$\sum_{i=0}^L W(x_i(t)) \cdot F(x_i(t)) \quad (12)$$

제한된 방법에서는 혼잡도가 높은 상황에서는 배경 모델의 학습 속도를 낮추어 기존의 배경 모델을 가능한 유지시키기 위해 배경 모델의 학습 비율을 다음의 수식 (13)과 같이 정의하였다.

$$\alpha = \alpha_0 \cdot c \quad (13)$$

수식 (13)에서 α 는 기존의 배경 학습 비율이며, c 는 수식 (12)에 의해 정의되는 영상의 혼잡도, α_0 는 혼잡 상황에 따른 영상의 학습 속도 조절의 정도를 결정하는 상수이다. 영상 내에 이동 객체가 매우 적은 상황에서는 혼잡도가 0에 가까운 값을 가질 것이므로 수정된 배경 모델 학습 비율 α_1 는 원래의 배경 모델 학습 비율 α 에 가까운 값을 갖는다. 반대로 이동 객체가 많이 발생하는 영상이 혼잡한 상황에서는 c 값이 1에 가까워지며, 이때 α_1 는 0에 가까운 값을 가지게 되어 배경 모델의 갱신이 억제되어 유효한 배경 모델이 유지되게 된다.

3.2. 영상의 급격한 변화에 대한 대응

앞 절에서 설명된 혼잡도에 따른 배경 모델 학습 비율의 적응적 조절에 의해 이미 생성되어 있는 배경 모델을 혼잡한 상황에서도 배경 모델을 효과적으로 유지하는 것이 가능하다. 하지만, 영상의 급격한 변화에 의해 혼잡 상황이 아님에도 영상에 급격한 변화가 발생한 경우에 제안된 방법의 배경 모델 생성 방법은 잘못된 결과에 빠질 수 있다. 예를 들면, 실외 영상에서 구름이 태양을 지나감으로써 영상에 전체적인 그림자가 드리워지는 경우, 또는 실내 영상의 경우에 조명 스위치를 켜거나 끄므로 해서 영상에 급격한 변화가 발생 하는 경우가 이러한 예가 된다. 이렇게 영상에 급격한 변화가 발생하게 되면 영상의 대부분의 픽셀이 전경으로 분리되게 된다. 따라서,

영상의 혼잡도가 매우 높다고 판단되어 배경 모델의 갱신이 이루어지지 않아 실제로는 영상이 혼잡한 상황이 아님에도 잘못된 배경 모델이 계속 유지되게 된다.

본 논문에서는 이와 같은 잘못된 상황에서 배경 모델이 더 이상 갱신되지 않는 것을 막기 위해 경계 정보를 이용해 현재의 혼잡 상황이 조명의 급격한 변화에 의해 발생한 것인지 아닌지를 확인한다. 즉, 혼잡도가 매우 높음에도 영상 내에 실제 변화는 거의 발생하지 않는다면 이는 급격한 영상 변화가 발생한 것이므로 혼잡도가 높다 하더라도 배경 모델의 갱신이 빨리 이루어져 시스템이 안정된 상황에 도달할 수 있도록 한다.

영상의 혼잡도가 매우 높다고 판단되었을 때, 이를 실제 혼잡 상황인지 판단하기 위해 본 논문에서는 영상의 경계 정보를 이용해 영상 내에 움직임이 존재하는지를 확인하였다. 경계 정보는 조명의 변화에 영향을 받지 않는 장점 때문에 움직임의 확인을 위해 선택되었다. 우선 수식 (14)에 의해 영상 내에서 움직임이 존재하는 부분의 경계 픽셀이 검출된다. 경계 정보 검출을 위해 Canny 경계 검출자를 이용하였으며, $CE(i)$ 는 i 번째 프레임의 Canny 경계 검출 결과에서 i 번째 픽셀의 값을 나타낸다 [13]. 지난 m 프레임 동안 픽셀 i 의 값이 255로 유지되었다면, 이 픽셀은 정적인 경계로, 그렇지 않다면 이는 동적인 경계로 판단된다.

$$\begin{cases} 0, & \text{if } CE_{t-m}(i) = CE_{t-m-1}(i) = \dots = CE_i(i) = 255 \\ 1, & \text{otherwise} \end{cases} \quad (14)$$

그리고, 영상 내 움직임의 양은 수식 (14)에 의해 얻어지는 동적 경계 픽셀의 프레임 내 전체 합에 의해 계산된다. 영상의 혼잡도가 매우 높음에도 영상 내 동적 경계 픽셀의 개수가 임계치 이하 일 경우에는 시스템이 잘못된 결과에 빠졌다고 판단하여 회복 과정이 선언되게 된다. 이 회복 과정에서는 배경 모델의 학습 비율을 정상 상태보다 크게 해주어 빨리 올바른 배경 모델을 생성할 수 있도록 하며, 혼잡도가 회복 상태 종료 판단을 위한 임계치 이하로 다시 떨어진 후 학습 비율을 원래의 값으로 재설정한다.

그림 2는 조명 스위치 조작에 의해 영상에 급격한 조명 변화가 발생하는 상황에서 배경 모델링의 예이다. 그림에서 (a), (b), (c)는 각각 조명이 켜진 상태에

서의 각각 입력영상, 생성된 배경 모델, 전경의 예이며, (d), (e), (f)는 각각 조명이 꺼진 직후의 각각 입력 영상, 생성된 배경 모델, 전경의 예이다. 그림 (f)에 나타난 것과 같이 급격한 영상 변화로 인해 영상 전체에서 전경 영역이 발생한 것을 볼 수 있다. 그림 (g), (h), (i)는 각각 조명이 변한 이후로 200프레임이 지난 이후의 입력 영상, 생성된 배경 모델, 전경의 예이다. 급격한 조명의 변화가 발생했음에도 그림 (h)의 배경 모델과 같이 올바른 배경 모델로 복원된 것을 확인할 수 있다.

3.3. 그림자 제거

전경 분리 결과에 그림자 영역이 잘못 포함되어 있는 경우, 이는 객체의 정확한 경계 및 위치를 추정할 수 없게 하여 전체 영상 감시 시스템의 성능에 영향을 미치게 된다. 본 논문에서는 이러한 문제를 피하기 위하여 KaewTraKulPong의 시스템에 사용된 그림자 제거 방법을 이용하였다 [10]. 이 방법은 RGB 색상 공간을 이용하면서도 배경 모델과 입력 영상의 차이를 밝기 성분(brightness)과 색도 성분(chromaticity)으로 분리하여 계산할 수 있기 때문에 효과적인 그림자 제거가 가능하다.

KaewTraKulPong의 시스템에 이용된 그림자 제거 방법은 배경 모델과 입력 영상의 차이를 다음과 그림 3에 나타난 것과 같이 밝기 왜곡(brightness distortion)과 색도 왜곡(chromaticity distortion)의 두 가지 성분으로 나누어 계산한다. 배경 모델과 입력 영상의 픽셀 값의 차이 가운데 밝기 성분의 영향에 의해 발생한 밝기 왜곡 a 는 수식 (15)에 의해 계산되며, 색도 차에 의해 발생한 색도 왜곡 c 는 수식 (16)에 의해 계산된다. 여기에서 I 와 E 는 각각 입력 영상의 픽셀 값과 배경 Gaussian 모델의 평균값을 의미한다. RGB 컬러 공간의 원점 O 과 배경 모델의 색상 E 를 연결하는 직선은 색도선(chromaticity line)이라고 표현되며, 이 직선 상에 위치한 점들은 R, G, B값의 비율이 동일하므로 밝기값에 차이가 있을 뿐 동일한 색도를 갖는다. 영상에서 그림자에 의한 픽셀 값 차이가 발생할 때는 색상의 변화는 없이 밝기 값의 차이만 발생한다고 가정하게 되면, 그림자 영역에서는 원래의 배경 모델과 비교하여 픽셀 값이 이 직선을 따라서 변화한다. 따라서, 밝기 왜곡은 입력 영상의 픽셀 값과 색도선과의 거리를 최소화하는 색상

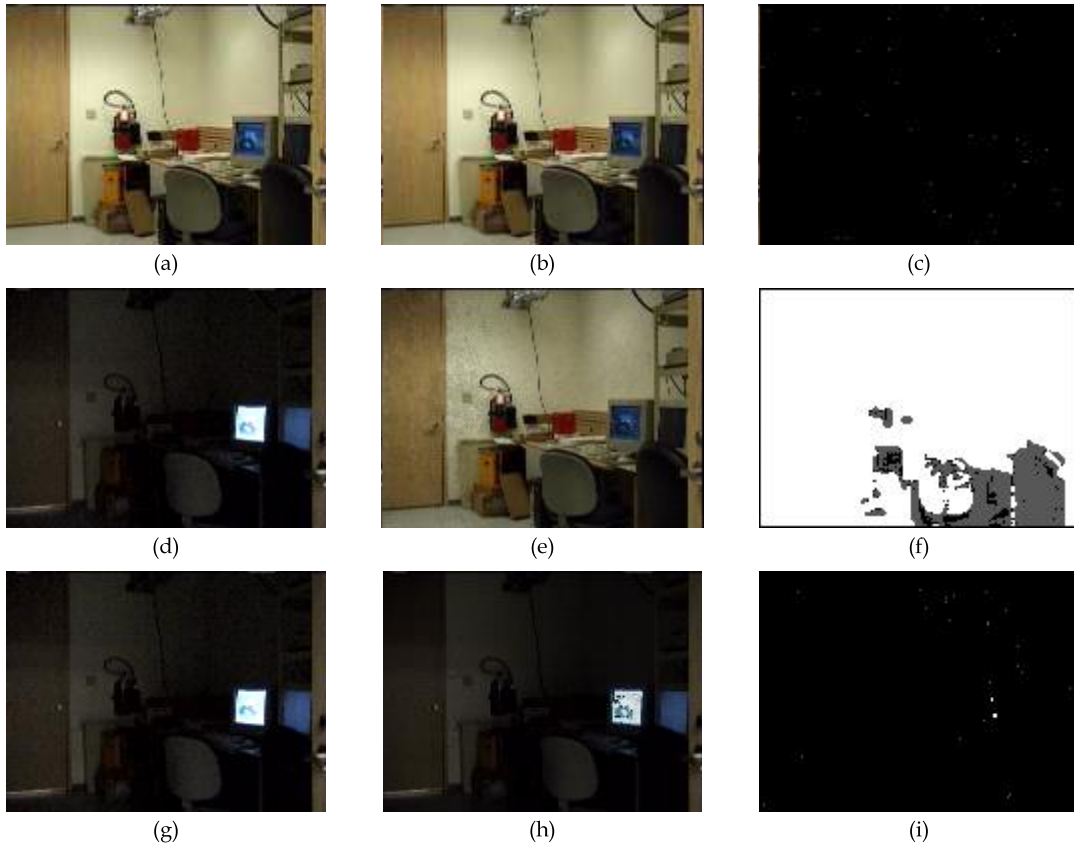


그림 2. 급격한 조명 변화의 예: (a), (b), (c) - 조명 변화 발생 전의 입력영상, 배경모델, 전경분리 결과, (d), (e), (f) - 조명 변화 발생 후의 입력 영상, 배경 모델, 전경 분리 결과, (g), (h), (i) - 200프레임 이후의 입력영상, 복원된배경 모델, 전경 분리 결과

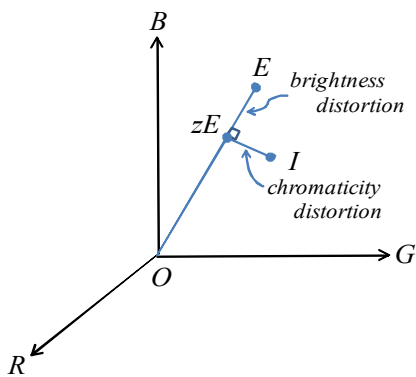


그림 3. 그림자 제거를 위한 RGB 컬러공간 내의 색상 모델

zE 와 배경 모델의 색상 E 와의 거리에 의해 수식 (15)와 같이 얻어질 수 있다. 그리고, 이때 색도 성분에 의해 발생한 색상 차는 zE 와 입력 영상의 픽셀 값간의 거리에 따라 수식 (16)과 같이 계산된다.

$$= \arg \min_z (I - zE)^2 \tag{15}$$

$$\|I - aE\| \tag{16}$$

그림자 검출은 전경 영역으로 검출된 모든 픽셀에 대하여 수행되며, 밝기 왜곡이 배경 모델의 표준편차 σ 와 비교하여 $a < k \cdot \sigma$ 의 조건을 만족하고, 색도 왜곡이 $0.7 < c < 1.0$ 의 조건을 만족하는 픽셀들을 그림자 영역에 속한 픽셀로 분류한다. 그림 4는 그림자 제거 결과를 보여준다. 그림 3의 (a)는 입력 영상이며, (b)는 그림자 제거 결과로써 영상에서 흰색은 전경 영역에 속한 픽셀들은, 회색은 그림자로 분류되어 전경에서 제거된 픽셀들을 나타낸다.

4. 실험 및 결과

제한한 배경 제거 알고리즘의 정확도 측정을 위해



(a)



(b)

그림 4. 그림자 제거의 예: (a) 입력영상, (b) 그림자 제거 결과

잘 알려진 PETS 2006 Benchmark 데이터[14]의 일부와 직접 촬영한 비디오 영상을 이용하였다. PETS Benchmark 데이터는 비교적 혼잡하지 않은 상황만이 포함되어 있으므로, 제안한 알고리즘의 효과를 입증하기 위해 혼잡한 상황을 포함하고 있는 비디오 데이터를 이용하였다. 직접 촬영한 영상은 Sony HDR-SR7을 이용하여 획득되었으며, 320x240의 해상도에 30fps로 인코딩 되었다. 그림 5는 실험 영상의 예를 보여준다. 그림 5 (a)는 PETS 2006 Benchmark 데이터의 예이며, (b)는 직접 촬영한 혼잡한 상황의 영상(Library)의 예이다. Library 데이터는 통행량이 많은 도서관의 출입구에서 촬영되었으며, PETS 데이터에 비해 시점이 가깝고 지속적으로 이동이 발생하기 때문에 정확한 배경의 생성에 어려움이 있다.

실험을 위하여 두 실험 영상 모두에 대하여 1,200 프레임 구간 동안 50프레임 간격으로 수작업을 통해 전경 영상을 생성하였다. 정확도 비교 대상으로는 Grimson의 방법[9], Lee의 방법[11], 그리고 Tang의 방법[12]이 이용되었다. Tang의 방법은 기존의 GMM을 이용한 배경 모델링 방법을 개선한 것으로 모든 색상 공간에 동일한 하나의 분산값만을 이용하



(a)



(b)

그림 5. 사용된 실험 영상의 예: (a) PETS 2006 benchmark data, (b) Library

었던 기존 알고리즘과 달리 각 색상 공간마다 서로 다른 분산값을 이용하여 더 정확한 배경 모델을 생성하면서, 동시에 기존 방법에서 불필요하다고 판단된 파라미터를 제거하여 더 빠른 수행 시간을 갖는 것으로 알려져 있다.

성능 비교를 위한 정확도 측정 방법으로는 F -measure를 이용하였다. F -measure는 수식 (17)-(19)에 의해서 정의된다. 여기에서 TP (true positive)는 전경으로 제대로 검출된 픽셀의 개수, FN (false negative)은 배경으로 잘못 분류된 전경 픽셀의 개수, FP (false positive)는 전경으로 잘못 분류된 배경 픽셀의 개수를 나타낸다.

$$recall = \frac{TP}{TP + FN} \quad (17)$$

$$precision = \frac{TP}{TP + FP} \quad (18)$$

$$recall + precision \quad (19)$$

그림 6과 그림 7은 각각 PETS 2006 영상과 Library 영상에 대한 정확도 측정 결과를 보여준다.

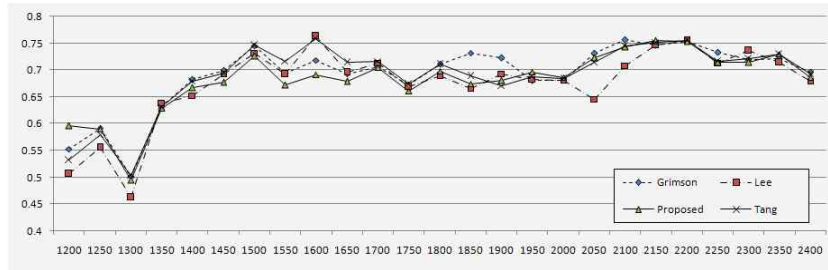


그림 6. PETS 2006 영상에 대한 정확도 측정 결과

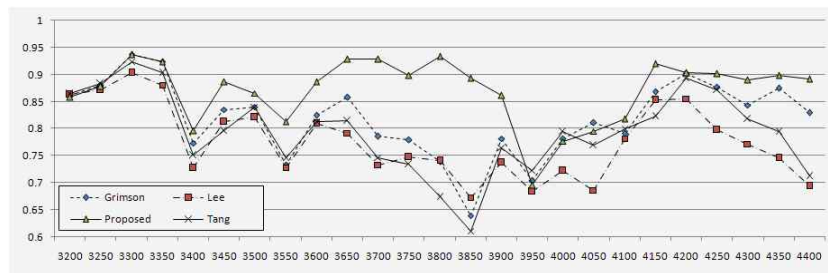


그림 7. 'Library' 영상에 대한 정확도 측정 결과

그림 6에 나타난 것과 같이 비교적 혼잡도가 적은 PETS 2006 영상에 대해서는 Grimson의 방법, Lee의 방법과 제안한 방법이 유사한 성능을 나타내었다. PETS 2006 영상에 대한 F -measure의 평균값은 Grimson의 방법, Lee의 방법, Tang의 방법, 그리고 제안한 방법에 대해 각각 0.69, 0.67, 0.69, 0.68로 나타났다. 즉, Grimson의 방법과 Tang의 방법이 가장 좋은 성능을 보여주었으나, 모든 방법에서 그 성능의 차이는 미미하게 발생하였다.

하지만 혼잡한 상황의 영상에 대해서는 그림 7에 나타난 것과 같이 제안한 방법이 기존의 방법과 비교하여 더욱 좋은 성능을 보이는 것으로 알 수 있었으며, 특히 매우 혼잡한 상황이 발생한 3,600에서 3,900 프레임 구간 내에서 성능 차이가 더욱 크게 발생하였다. Library 영상에 대한 F -measure의 평균값은 Grimson의 방법, Lee의 방법, Tang의 방법, 그리고 제안한 방법에 대하여 각각 0.81, 0.78, 0.79, 0.87로 제안한 방법이 가장 좋은 성능을 보여주었다. Lee의 방법은 Grimson의 방법을 수렴속도가 빠르게 개선한 것이므로 혼잡한 상황에서는 오히려 잘못된 배경을 생성할 가능성이 높아져서 가장 낮은 성능을 보였다. Tang의 방법은 색상 공간마다 서로 다른 분산값을 이용하여 더 정확한 배경 모델을 생성하지만, Lee

의 방법과 마찬가지로 빠른 수렴속도를 갖기 때문에 혼잡한 영상에서는 쉽게 잘못된 배경 모델이 생성되었다. 그림 7의 3,950 프레임에서 모든 방법이 매우 낮은 정확도를 보이는 것은 잘못된 배경 모델링의 결과가 아니라 해당 프레임에서 전경 영상이 극히 적게 발생한 것에 반해 제거되지 못한 그림자 영역이 상대적으로 컸기 때문이었다.

그림 8과 그림 9는 각각 PETS 2006 영상과 Library 영상에 대한 실험 결과의 예를 보여준다. 각각의 그림에서 (a), (b), (c)는 기존의 Grimson의 방법, (d), (e), (f)는 Lee의 방법, (g), (h), (i)는 Tang의 방법, (j), (k), (l)은 제안한 방법에 대해 각각 입력 영상, 배경 모델, 전경 분리 결과의 예를 나타낸 것이다. 그림 8에서 PETS 2006 영상에 대해서는 기존의 알고리즘과 제안된 알고리즘의 결과에 큰 차이가 나타나지 않음을 알 수 있다. 하지만, 그림 9의 (b), (e), (h)에 나타난 것과 같이 기존의 방법으로 생성된 배경 모델은 혼잡한 상황에서 전경 영역의 일부가 배경 모델로써 잘못 생성된 것을 확인할 수 있다. 배경 모델의 영상은 배경으로 분류된 Gaussian 분포의 평균값의 가중평균이므로 배경 영상 내에 어두운 픽셀은 배경보다 전경이 더 주요한 배경 모델의 성분으로 자리 잡은 것을 의미한다. 그림 (c), (f), (i)에서는

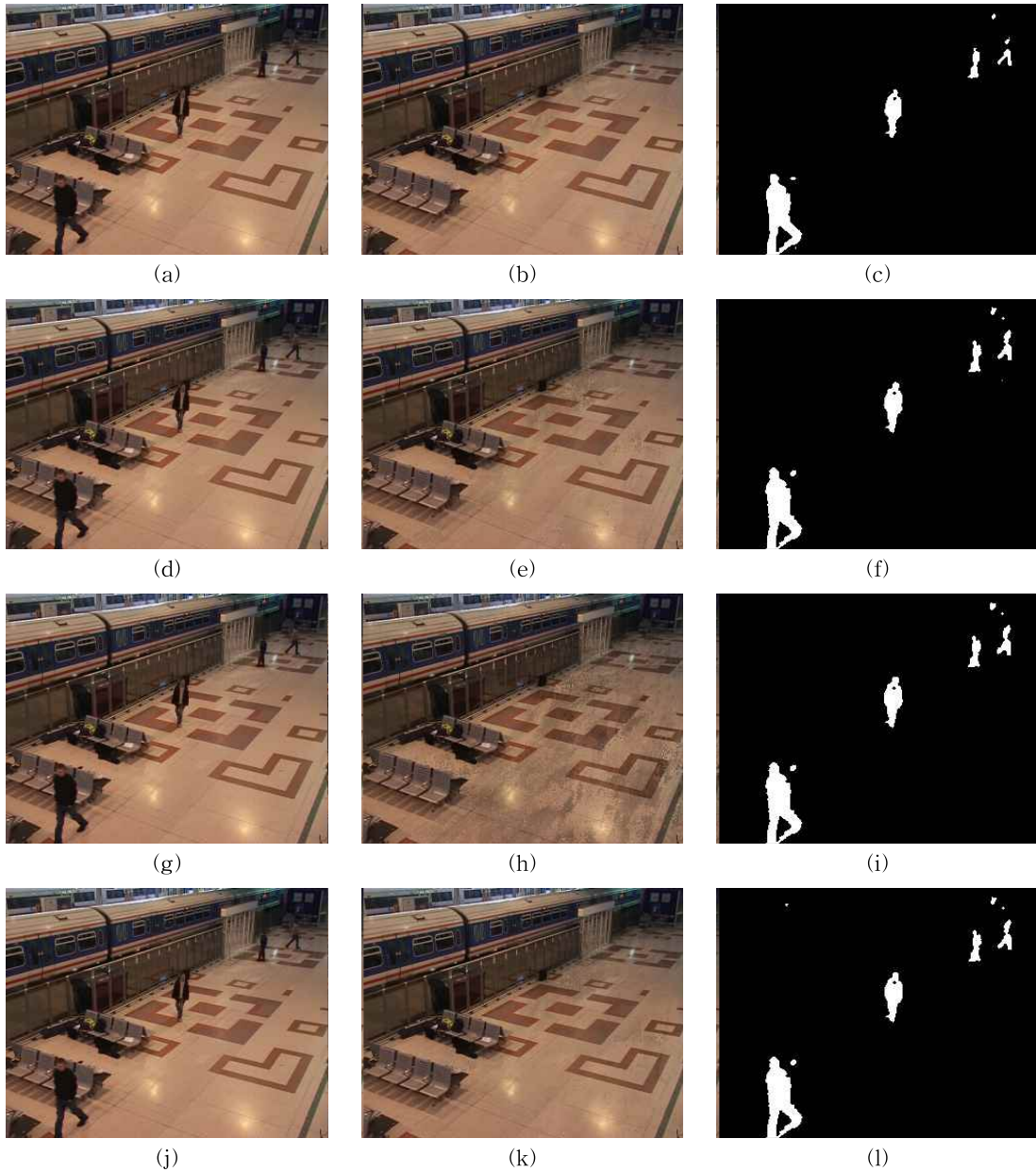


그림 8. PETS 2006 영상에 대한 실험 결과의 예: (a), (b), (c) - Grimson의 방법, (d), (e), (f) - Lee의 방법, (g), (h), (i) - Tang의 방법, (j), (k), (l) - 제안한 방법 (각각 (a) - 원영상, (b) - 배경모델, (c) - 전경분리 결과).

잘못된 배경 모델의 영향으로 전경 영역의 많은 부분이 배경으로 판단되어 검출되지 않은 것을 확인할 수 있으며, 제안한 방법의 결과에서는 그러한 예가 상대적으로 적은 것을 확인할 수 있다. Morphology 등에 의한 후처리를 통해 전경 분리 결과를 개선시키는 것이 가능하지만, 본 논문의 실험에서는 배경 모

델 생성 방법의 정확도를 비교하기 위해 그러한 별도의 후처리 과정은 적용하지 않았다.

5. 결 론

본 논문에서는 혼잡한 상황에서 효과적으로 배경

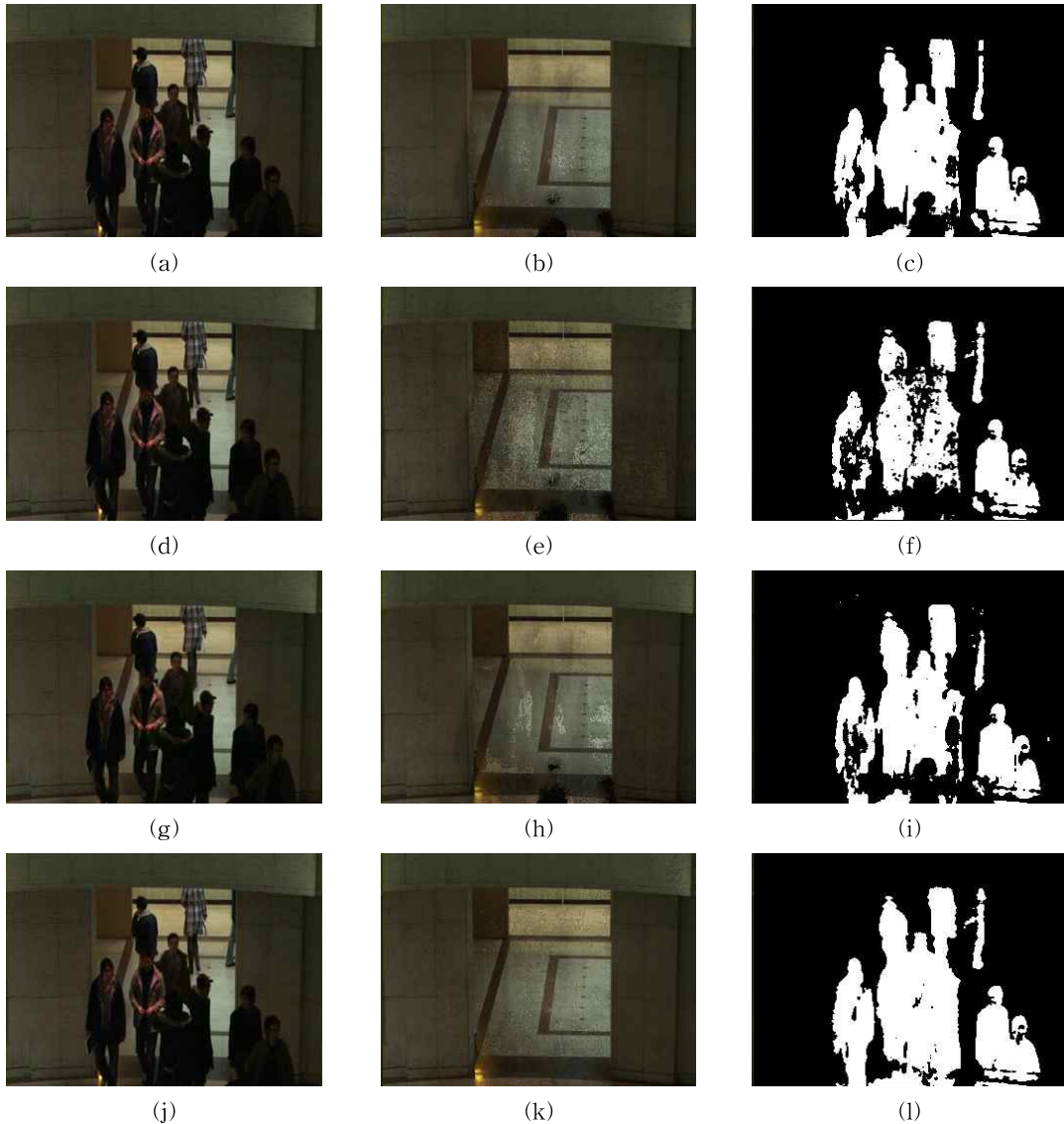


그림 9. Library 영상에 대한 실험 결과의 예: (a), (b), (c) - Grimson의 방법, (d), (e), (f) - Lee의 방법, (g), (h), (i) - Tang의 방법, (j), (k), (l) - 제안한 방법 (각각 (a) - 원영상, (b) - 배경모델, (c) - 전경분리 결과).

모델을 생성하기 위한 배경 모델링 방법을 제안하였다. 제안한 알고리즘은 Lee에 의해 제안된 배경 모델링 방법을 기반으로 하여, 영상의 혼잡도에 따라 배경 모델의 학습 속도를 조절하는 학습 비율을 조절하여 혼잡한 상황에서도 전경 영역이 배경 모델로 잘못 판단되는 것을 방지하였다. 또, 조명 등의 변화에 의해 영상에 전체적인 변화가 일어난 경우에도 이것이 혼잡한 상황으로 잘못 판단되어 배경 모델의 갱신이 이루어지지 않는 문제를 해결하기 위해 경계 정보를

이용하여, 영상에 전체적인 변화가 발생하였을 때에도 올바른 배경 모델을 생성할 수 있도록 하였다.

혼잡한 상황에서 촬영된 실제 비디오 데이터에 대한 실험을 통하여 제안한 방법의 유효성을 검증하였으며, 기존의 혼합 Gaussian 모델 기반의 배경 모델링 방법과 비교하여 혼잡한 상황에서 큰 성능 향상을 보이는 것을 확인하였다. 그러나, 제안한 방법은 혼잡도 계산을 위한 가중치 지도를 필요로 하며, 이는 별도의 충분한 초기 훈련 시간을 요구한다. 또한, 잘

못된 가중치 지도의 선택은 제안한 방법의 성능에 큰 영향을 끼칠 수 있으므로, 이러한 문제점을 개선하기 위한 연구가 향후 연구로써 요구된다.

참 고 문 헌

- [1] S. J. McKenna, S. Jabri, Z. Duric, A. Rosenfeld, and H. Wechsler, "Tracking Groups of People," *Computer Vision and Image Understanding* Vol.80, pp. 42-56, 2000.
- [2] S. A. Velastin, J. H. Yin, A. C. Davies, M. A. Vicencio-Silva, R. E. Allsop, and A. Penn, "Automated Measurement of Crowd Density and Motion using Image Processing," Seventh International Conference on Road Traffic Monitoring and Control, pp. 127-132, Apr, 1994.
- [3] H. Celik, A. Hanjalic, and E. A. Hendriks, "Towards a Robust Solution to People Counting," Proc. of International Conference of Image Processing, pp. 2401-2404, Oct. 2006.
- [4] G.-G. Lee, B. Kim, and W.-Y. Kim, "Automatic Estimation of Pedestrian Flow," First International Conference on Distributed Smart Cameras, pp. 291-296, Sept. 25-28, 2007.
- [5] G. Medioni, I. Cohen, F. Bremond, S. Hongeng, and R. Nevatia, "Event detection and analysis from video streams," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 23, pp. 873-889, August 2001.
- [6] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-Time Surveillance of People and Their Activities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.22, pp. 809-830, Aug. 2000.
- [7] D. Koller, J. Weber, T. Huang, J. Malik, G. Osagawara, B. Rao, and S. Russel, "Towards robust automatic traffic scene analysis in real-time," International Conference on Pattern Recognition, pages 1260-131, Nov. 1994.
- [8] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland. "Pfinder: Real-time tracking of the human body," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.19, No.7, pp. 780-785, 1997.
- [9] C. Stauffer and W.E.L. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," Proc. Conf. Computer Vision and Pattern Recognition, Vol.2, pp. 246-252, June 1999.
- [10] P. KaewTraKulPong and R. Bowden, "An Improved Adaptive Background Mixture Model for Real-time Tracking with Shadow Detection," chapter 11, pp. 135-144, *Video-Based Surveillance Systems*, Kluwer Academic Publishers, Boston, 2002.
- [11] D. Lee, "Effective Gaussian Mixture Learning for Video Background Subtraction," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.27, No.5, pp.827-832, May 2005.
- [12] Z. Tang and Z. Miao, "Fast Background Subtraction and Shadow Elimination Using Improved Gaussian Mixture Model," IEEE International Workshop on Haptic Audio Visual Environments and their Applications, pp. 38-41, Oct. 2007.
- [13] J. F. Canny, "A Computational Approach to Edge Detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 8, No.11, pp. 679-698, Nov. 1986.
- [14] "PETS 2006 Benchmark Data," Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, <http://www.cvg.rdg.ac.uk/PETS2006/data.html>, accessed June 2007.



이 광 국

2002년 한양대학교 전자전기공학부 학사
2004년 한양대학교 전자통신전파공학과 석사
2004년~현재 한양대학교 전자컴퓨터통신공학과 박사과정

관심분야 : 영상 감시, 패턴 인식, 내용기반 멀티미디어 분석



윤 자 영

2004년 한양대학교 건축공학부 학사
2005년~현재 한양대학교 건축환경공학과 석·박사통합과정
관심분야 : Building Information Modeling(BIM), 공간 정보관리, 경영시스템



송 수 한

2006년 한양대학교 전자전기컴퓨터공학부 학사
2008년 한양대학교 전자컴퓨터통신공학과 석사
2008년~현재 삼성전자 디지털미디어총괄 디지털프린팅사업부

관심분야 : 영상 감시, 패턴 인식, 영상 화질 개선



김 재 준

1983년 한양대학교 건축공학과 학사
1985년 Univ. of Illinois at Urbana-Champaign 건설관리 석사
1992년 Univ. of Illinois at Urbana-Champaign 건설경영/관리 박사

1993년~현재 한양대학교 건축공학부 교수
관심분야 : Building Information Modeling(BIM), 시스템경영, 건설자동화



가 기 환

2004년 명지대학교 전자공학과 학사
2007년~현재 한양대학교 전자컴퓨터통신공학과 석사과정

관심분야 : 영상 처리, 영상 감시, 패턴 인식



김 회 을

1980년 한양대학교 전자공학과 졸업 (공학사)
1983년 Pennsylvania State University 전기공학과 졸업 (공학석사)
1989년 Purdue University 전기공학과 졸업 (공학박사)

1989년 9월~1994년 2월 University of Texas 조교수
1994년~현재 한양대학교 전자통신컴퓨터공학부 정교수
관심분야 : 영상처리, 컴퓨터비전, 패턴인식, 머신비전, MPEG-7등