

시간축과 공간축 화소 정보를 이용한 배경 생성

조 상 현[†] · 강 행 봉^{**}

요 약

비디오 감시 시스템에서 정확한 물체 추적을 위해서는 움직이는 물체가 없는 정적인 배경 영상이 필수적이다. 하지만 기존의 배경 생성 방법들은 주로 시간 축에 따른 화소 정보를 이용하여 오랫동안 정지해 있는 물체들이 존재하는 경우에는 적용하기 어려운 단점이 있다. 이러한 문제점을 해결하기 위해 본 논문에서는 mean-shift와 fast marching method(FMM)을 이용해 시간 축 화소 정보와 공간 축 화소 정보를 이용하여 배경을 생성하는 방법을 제안한다. mean-shift를 이용해 시간 축에 따른 화소 값의 최빈값을 추정하여 배경을 생성하고, FMM을 이용해 공간 축에 따른 화소 정보를 이용하여 일정 기간 동안 움직이지 않은 물체가 있는 환경에서 바람직한 배경을 생성한다. 실험 결과는 제안한 방법이 기존의 시간에 따른 빈도만을 이용하는 방법보다 더 효율적임을 보여준다.

키워드 : Mean-shift, Fast Marching Method, 배경 생성, 감시 시스템, 배경 모델

Background Generation using Temporal and Spatial Information of Pixels

Sang-Hyun Cho[†] · Hang-Bong Kang^{**}

ABSTRACT

Background generation is very important for accurate object tracking in video surveillance systems. Traditional background generation techniques have some problems with non-moving objects for longer periods. To overcome this problem, we propose a new background generation method using mean-shift and Fast Marching Method (FMM) to use pixel information along temporal and spatial dimensions. The mode of pixel value density along time axis is estimated by mean-shift algorithm and spatial information is evaluated by FMM, and then they are used together to generate a desirable background in the existence of non-moving objects during longer period. Experimental results show that our proposed method is more efficient than the traditional method.

Keywords : Mean-Shift, Fast Marching Method, Background Generation, Surveillance System, Background Model

1. 서 론

비디오 감시 시스템에서 물체 검출과 추적은 매우 중요한 요소로서 최근에도 활발한 연구가 이루어지고 있는 분야이다. 많은 물체 검출 및 추적 방법들이 제안되었지만 대부분의 방법들이 정확한 검출 및 추적을 위해 정확한 배경을 필수적으로 요구한다. 대표적인 예로 물체를 검출하고 추적하는 방법 중 가장 간단한 방법인 배경 차(background subtraction) 기반의 방법을 들 수 있다. 배경 차 기법은 미리 가지고 있는 배경 영상과 현재 프레임 영상과의 차이를 이용해 물체를 검출하고 추적하는 방법이다. 배경 차 기반의 방법을 이

용해 물체 검출 및 추적을 하는 경우 정확한 배경을 이용하지 않으면 유령(ghost) 현상과 같은 문제점이 발생할 수 있다. 배경 차 기반의 방법뿐만 아니라 기존의 많은 물체 검출 및 추적 방법에서도 정확한 결과를 얻기 위해서 움직임이 없는 정적인 배경을 가지고 있다고 가정하고 이러한 배경을 이용해 배경 모델을 세우거나 물체 위치를 추정한다. 하지만 실제로 이러한 움직임이 없는 정적인 배경을 가지는 경우는 매우 드물기 때문에 이러한 방법들을 실세계에 적용하는 것은 어렵다.

배경 영상을 만드는 데는 시간에 따라 획득되는 영상 정보를 이용하는 것이 일반적이다. 대표적인 예로 배경 영상의 화소값을 프레임 영상들의 평균이나 중간 값으로 취하는 방법을 들 수 있다. 하지만 이러한 방법들은 시간 축 방향에서의 영상 정보만을 이용하므로 장시간 정지해 있었거나 움직임이 적은 물체가 있는 경우에 그 물체까지 배경 영상에 포함되는 문제점이 있다.

예를 들면, 주정차 금지 구역에 차량이 잠시 주차되고 있

* 본 연구는 MKE/KETI의 IT R&D 프로그램 [2009-S-040-01, 개인신변 안전 보장을 위한 영상보안기술개발]의 지원 및 2009년도 가톨릭대학교 교비 연구비 지원으로 수행되었음.

† 준 회원 : 가톨릭대학교 컴퓨터 공학과 박사과정
** 종신회원 : 가톨릭대학교 디지털 미디어학부 교수(교신저자)
논문접수 : 2009년 8월 25일
수정일 : 1차 2009년 10월 15일, 2차 2009년 11월 24일
심사완료 : 2009년 11월 24일

다고 가정하자. 기존의 시간 축의 정보만을 이용해 배경을 생성하는 시스템의 경우 일정 시간을 넘기면 주차된 차량은 배경에 포함되어 진다.

배경 모델을 이용하는 시스템은 그 특성상 특정 주기로 배경 모델을 갱신해야 한다. 하지만 움직임이 적은 환경에서는 대상 물체가 배경과 같이 움직임이 없으므로 배경과 전경의 경계가 모호해지고 이로 인해 시스템의 검출 및 추적 성능의 저하를 일으키게 된다. 이러한 문제점을 해결하기 위한 방법으로 먼저 생각할 수 있는 것은 갱신 주기를 길게 하는 것이다. 하지만 갱신 주기가 길어지면 그만큼의 고용량 메모리가 요구되게 되고 또한 최근의 배경 정보를 배경 모델에 반영하기 어려운 단점이 있다. 따라서 단순히 갱신 주기를 길게 하기보다는 적절한 갱신 주기에서 불필요한 물체들이 제거된 정확한 배경을 생성해 이용하는 것이 보다 효율적이라 볼 수 있다.

그러므로 본 논문에서는 이러한 전경이 없는 정확한 배경을 생성하기 위해 시간과 공간에 따른 화소정보를 이용한 방법을 제안한다. 제안한 방법은 각 위치의 화소에 대해서 시간 축에 따른 화소값에 대한 분포의 최빈값을 mean-shift 알고리즘을 이용해 추정하고 이를 이용해 배경을 생성한다. 생성된 배경으로부터 잘못 분류된 화소들을 분류하고 분류된 화소나 영역에 대해서 FMM을 이용하여 잘못 분류된 영역을 복원한다. 그 후 복원된 영역과 다른 프레임에서의 동일 위치 영역에 대한 차이를 이용하여 복원된 영역을 보정하게 된다.

본 논문의 구성은 다음과 같다. II장에서는 배경 생성에 대한 관련 연구를 소개하고, III장에서는 본 논문에서 제안하는 mean-shift와 FMM을 이용한 배경 생성 방법을 설명한다. IV장에서는 제안한 방법과 기존의 방법을 비교하고 제안한 방법에 대한 실험결과를 설명한다. 마지막으로 V장에서는 결론 및 향후 연구과제를 제시한다.

2. 관련 연구

배경 차를 위해 배경 모델을 만드는 방법은 매개변수적(parametric) 방법과 비매개변수적(non-parametric) 방법으로 구분할 수 있다.

매개변수적 방법은 화소의 확률밀도함수가 가우시안 혹은 혼합 Gaussian(Mixture of Gaussian : MOG)과 같이 특정 모델이라는 가정을 하고 그에 따른 매개변수를 추정하는 방법이다[1-3]. 매개변수적 방법은 가우시안이나 혼합 가우시안과 같은 특정 분포를 가정하기 때문에 학습 단계가 요구되지만 모델이 간결하여 효율적인 평가 단계를 갖는다. Stauffer와 Grimson[4]은 적응적 배경 차에 기반한 모션 분할 문제를 다루었다. 그들은 각각의 화소를 혼합 가우시안으로 확률밀도함수를 모델링하고, 모델을 갱신하기 위해 온라인 근사화를 사용했다. Toyama[5] 등은 조명 스위칭과 같은 현존하는 많은 배경 유지에 대한 문제점을 해결하기 위해 Wallflower scheme를 제안했다. Haritaoglu[6]등은 각 화

소에 대해서 최대값(M), 최소값(N), 최대 inter-frame 절대차(D)값을 이용한 W4라는 방법을 제안했으며, W. Wren 등 [1]은 단일 가우시안 모델을 이용한 Pfind 라는 방법을 제안했다. 또한 Junxian Wang[7] 등은 SVR(Support Vector Regression)을 이용해 배경을 모델링하여 물체를 검출하는 방법을 제안했다.

비매개변수적 방법은 매개변수적 방법과는 달리 특정 모델이 필요치 않아서 임의의 데이터 분포에 적용할 수 있는 장점이 있다. 하지만, 높은 계산 복잡도로 인해 매개변수적 방법보다는 비효율적이라는 단점이 있다. Elgammal[8] 등은 커널 밀도 추정을 이용해 각각의 화소에 대한 화소값을 이용해 확률밀도 함수를 추정하고 이를 이용해 임의의 화소값에 대한 확률을 구하는 비매개변수적 배경 모델을 만드는 방법을 제안했다.

mean-shift는 원래 밀도함수의 최빈값을 찾는 효율적인 방법으로 Fukunaga[9]에 의해 제안된 후 Comaniciu와 Meer [10]에 의해 smooth trajectory 성질과 수렴성이 증명되어 컴퓨터 비전 응용분야에서 많이 이용되고 있다. Yazhou Liu [11] 등은 mean-shift를 이용한 비매개변수적 배경 생성 방법을 제안했다. 그들은 영상의 각 위치별로 시간에 따른 화소값 분포에 대해 mean-shift를 이용해 화소값의 최빈값을 추정함으로써 배경을 생성했다. 이들의 방법은 기존의 방법에 비해 간단하고 효율적이면서도 고품질의 배경 영상을 얻을 수 있는 장점을 가지고 있다. 하지만 이들의 방법은 시간에 따른 화소값만을 고려하여 장시간 정지해 있는 물체가 존재할 경우에는 물체까지도 배경에 포함되는 단점을 가지고 있다.

이미지 복원 기술은 영상의 손상된 부분을 복원하기 위해 복원 영역 외부의 화소값과 그라디언트를 이용해 복원 영역을 채우는 기술이다. Bertalmio[12] 등은 이미지 라플라시안에 의해 추정된 이미지 smoothness 정보를 등광도선(isophotes) 방향으로 전파시키는 방법을 제안했으며, Chan[13] 등이 제안한 TV(Total Variational) 모델은 등광도선 방향을 유지하기 위해 이방성 확산(anisotropic diffusion)과 결합된 오일러-라그랑지 방정식(Euler-Lagrange equation)을 이용했다. 또한 Chan[14]등은 두꺼운 영역을 복원할 수 있도록 TV모델을 개선한 CCD(Curvature-Driven Diffusion)모델을 제안했다.

FMM(Fast Marching Method)은 원래 경계값 문제를 풀기 위해 Sethian[14]에 의해 도입된 방법이다. 이러한 FMM을 이용한 이미지 복원 기법은 이미지 복원을 위해 화소의 주변(neighborhood)의 가중치 평균(weighted average)를 이용해 이미지 smoothness를 추정하고 이를 전달하는 방법으로 A. Telea[15]에 의해 제안되었다.

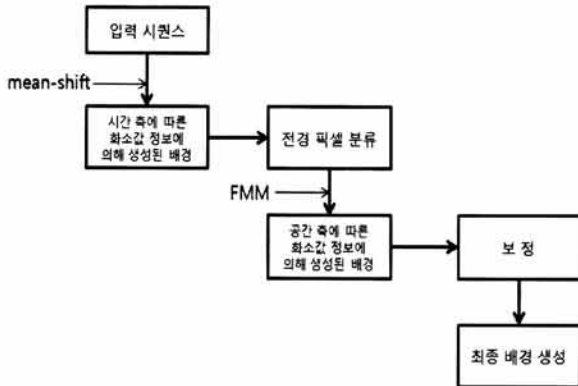
본 논문에서는 mean-shift와 FMM을 이용해 시간과 공간에 따른 화소 정보를 이용한 배경 생성 방법을 제안한다. [11]에서 제안한 방법과 같이 mean-shift를 이용해 시간 축에 따른 화소값에 대해 배경을 생성한 후, 각 화소에 대한 화소값의 변화 빈도를 이용해 잘못 생성된 화소들을 분류한다. 잘못 분류된 영역을 보정하기 위해 분류된 영역에 대해서 FMM을 이용해 영역을 복원한 후 복원 영역과 각 프레임 안에서 동일 위치의 영역간의 차이를 이용한 보정 과정

을 거쳐 배경을 생성하게 된다.

3. 시간 축과 공간 축 화소 정보를 이용한 배경 생성

3.1 개요

(그림 1)은 본 논문에서 제안한 배경 생성 방법의 개요를 나타낸 것이다. 먼저 비디오 시퀀스에 대해서 mean-shift를 이용해 시간에 따른 화소값 밀도의 최빈값을 추정하여 배경을 생성한다. 시간에 따른 화소값의 빈도만을 이용해 생성되므로 장시간 존재하는 물체가 존재할 경우 생성된 배경에는 이러한 물체가 포함된다. 따라서 이러한 물체를 제거하기 위해 화소값의 변화 빈도를 이용해 잘못 분류된 화소를 구별한 후 해당 화소나 영역들에 대해 FMM을 적용하여 주변 공간상의 화소값에 따른 이미지 복원을 수행한다. 마지막으로 복원된 영역들에 대해서 다른 프레임 이미지 안의



(그림 1) 제안한 배경 생성 방법의 개요

동일위치의 화소나 영역에 대한 차이를 이용해 복원된 배경의 왜곡을 보정하고 최종 배경을 생성한다.

3.2 시간에 따른 화소값을 이용한 배경 생성

먼저 각 위치에 대해서 가장 많이 등장하는 화소값이 배경 화소값이라고 가정하자. (그림 2) (a)부터 (그림 2) (c)까지는 PETS(Performance Evaluation of Tracking and Surveillance) 2000 Data set의 한 시퀀스를 나타낸 것이다. (그림 2) (d)는 (193, 135) 위치에서의 시간에 따른 화소값의 변화를 나타낸 것이다. (그림 2) (d)에서 배경 화소값은 많은 부분에서 안정적으로 나타나는 것을 볼 수 있다. (그림 2) (e)는 각 시간대로 나타난 화소들의 값을 나타낸 것으로 실제 배경 화소값에 해당하는 값들이 높은 밀도를 나타낸다는 것을 볼 수 있다.

시간에 따른 화소값의 최빈값을 효율적으로 찾기 위해 본 논문에서는 mean-shift를 이용한다.

mean-shift 을 이용한 배경 생성 방법은 다음과 같다.

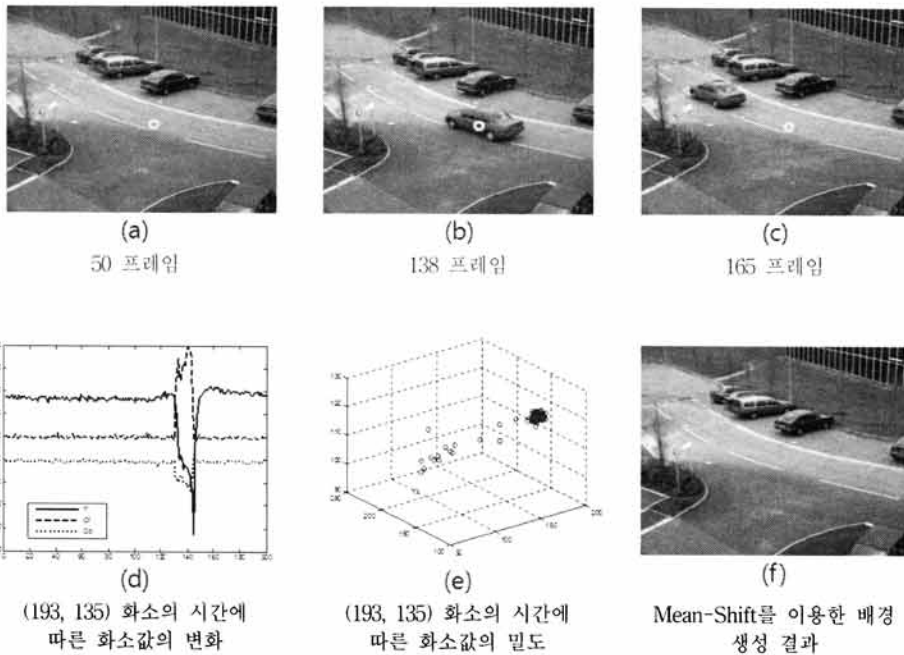
각 위치의 점 p 에 대하여

- (1) Mean-Shift의 시작 점의 초기화 : $y_1 = p_i$
- (2) 다음의 y_{t+1} 이 수렴할 때까지 Mean-Shift 과정 적용

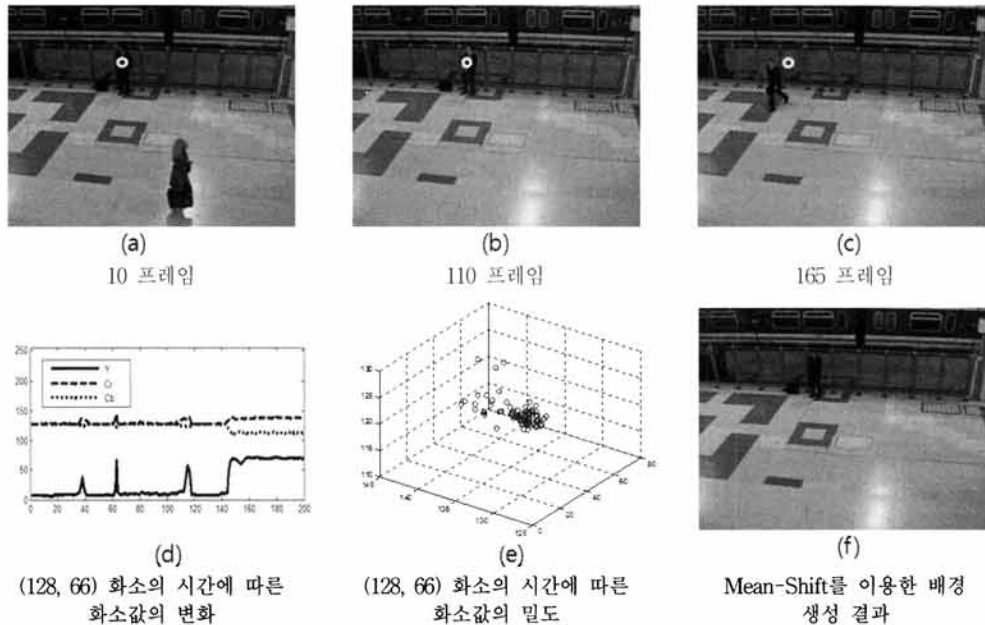
$$y_{t+1} = \frac{\sum_{i=1}^n x_i g\left(\left\|\frac{y_t - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n g\left(\left\|\frac{y_t - x_i}{h}\right\|^2\right)} \quad (1)$$

여기서 $g(x)$ 는 kernel 함수의 profile 함수이다.

- (3) 수렴점 y_{conv} 을 y 의 최빈값으로 취한다.



(그림 2) 시간에 따른 화소 값의 변화와 밀도



(그림 3) 장시간 정지해 있는 경우의 시간에 따른 화소값의 변화와 밀도

하지만, 단순히 시간에 따른 화소값 최빈값을 배경 화소값으로 취하는 것은 (그림 3)과 같이 장시간 정지해 있는 물체가 있는 경우와 같은 복잡한 영상에서는 적용하기 힘들다.

(그림 3) (a) 부터 (그림 3) (c)까지는 PETS 2006의 다른 데이터 시퀀스의 일부를 나타낸 것이다. (그림 3) (d)는 (128, 66) 위치에서의 시간에 따른 화소값의 변화를 나타낸 것이다. 영상에서 한 남성이 해당 위치에 장시간 서 있음에 따라 남성에 해당하는 색이 장시간 나타나는 것을 확인할 수 있다. (그림 3) (e)는 각 시간대로 나타난 화소들의 값을 나타낸 것으로 실제 남성의 화소값에 해당하는 값들이 높은 밀도를 나타낸다는 것을 볼 수 있다. (그림 3) (f)는 배경 생성 결과로 남성이 배경에 포함되어 있는 것을 확인할 수 있다.

본 논문에서는 이러한 문제점을 해결하기 위해 먼저 잘못 분류된 화소나 영역을 분류하기 위해 화소의 변화 빈도를 이용한다. 자세한 내용은 다음의 3절에서 다루었다.

3.3 공간에 따른 화소값을 이용한 배경 복원

오랫동안 정지해 있는 물체가 존재할 경우, 물체에 해당하는 영역이 그 위치에 머무는 시간이 길어지므로, 앞에서와 같이 시간에 따른 화소값 밀도의 최빈값을 배경으로 취할 경우, 정지해 있던 물체가 배경으로 될 가능성이 높다.

따라서 시간에 따른 화소값을 이용해 생성한 배경에 포함되어 있는 물체를 제거하기 위해서 먼저 영상에서 그 물체 영역을 검출하는 기술이 필요하다.

이러한 영역을 검출하기 위해 본 논문에서는 각 위치의 화소 p 에 대해서 화소의 시간에 따른 변화 수를 측정하여 p 의 $flag$ 를 다음과 같이 정의한 후 그 값에 따라 화소를 표 1과 같이 분류한다.

모든 $t=1, \dots, T$ 에 대해서

$$flag(p) = \begin{cases} 1 & \text{where } \|C(p)\| = 0 \\ 2 & \text{where } \|C(p)\| < T_r \\ 3 & \text{where } \|C(p)\| > T_p \end{cases} \quad (2)$$

여기서, $C(p) = \{t | |I_{t+1}(p) - I_t(p)| > T_p\}$
 T_r, T_p : threshold
 $\| \cdot \|$: number of elements of the set

$C(p)$ 의 값에 따라 본 논문에서는 <표 1>과 같이 화소들 화소값의 변화 빈도에 따라 안정적인 점, 덜 안정적인 점, 불안정한 점으로 분류한다. 만약 점 p 의 화소값이 시퀀스 전체에 대해서 안정적이라면, 점 p 의 화소값을 점 p 에서의 배경 화소값으로 한다. 또한 만약 점 p 의 화소값의 불안정적이라면 기존의 시간 축 화소 정보를 이용해 배경 화소값을 계산한다.

따라서 본 논문에서는 덜 안정적인 점에 대해서 이미지 복원을 수행한다. Threshold 값 T_r, T_p 는 정확한 화소 분류를 위해서 매우 중요하다.

T_p 는 시스템 잡음이나 조명 변화와 같이 시스템 특성이나 영상의 특성에 의해 결정되며 T_r 은 화소값의 변화 빈도 정도로서 화소의 안정성을 결정한다. 만약 T_r 또는 T_p 가 너무 작으면 분류 품질이 시스템 잡음이나 작은 조명 변화에 매우 민감하게 되고, 너무 크면 분류 품질이 매우 좋지 못하게 된다. 하지만, T_r, T_p 값은 시스템이나 영상에 따라 다르게 적용되므로 경험적 실험에 의해서 최적의 T_r, T_p

<표 1> 화소의 분류

$flag(p)$	의 미
1	안정적인 점(배경)
2	화소값이 변하는 빈도수가 적은 점
3	화소값이 변하는 빈도수가 많은 점

을 구한다.

앞에서 언급했듯이 본 논문에서는 각 위치에 대해서 가장 많이 나타난 화소값을 배경 화소값으로 가정한다. 하지만 이러한 가정은 움직임이 많은 환경에서는 매우 효율적이지만, 오랫동안 정지해 있는 물체가 존재하는 경우에는 비효율적이다.

따라서 이러한 문제점을 해결하기 위해 본 논문에서는 FMM을 이용한 복원 기법을 이용한다. FMM을 이용한 이미지 복원 방법은 화소의 주변 화소값을 이용해 복원하므로 간단하고 속도가 빠르며 좁은 영역을 복원하는데 효과적이지만 넓은 영역을 복원하는 데는 많은 왜곡이 생기는 단점이 있다. 따라서 이러한 왜곡을 개선하기 위해 많은 이미지 복원 방법이 제시되었지만 대부분 많은 계산량을 요구한다. 하지만 동영상의 경우 다수의 프레임이 존재하고 또한 복원해야 할 영역에 대한 ground truth를 다른 프레임이 포함하고 있는 경우가 많으므로 본 논문에서는 FMM을 이용해 간단하게 영역을 복원한 후 다른 프레임 안의 동일 위치 영역과의 유사도를 계산하여 가장 유사한 영역을 복원된 영역과 대체하는 방법을 이용한다.

영상이 그레이 이미지라고 가정하자. (그림 4)와 같이 영역 Ω 의 경계(boundary) $\partial\Omega$ 위의 점 p 를 복원할 점이라 하고, $B_\epsilon(p)$ 을 크기가 ϵ 인 p 의 주변이라 하자. 충분히 작은 ϵ 에 대해서 점 q 의 이미지 $I(q)$ 과 그라디언트 $\nabla I(q)$ 의 값에 대해 점 p 에 대해서 이미지의 1차 근사 $I_q(p)$ 는 다음과 같다.

$$I_q(p) = I(q) + \nabla I(q)(p - q) \quad (3)$$

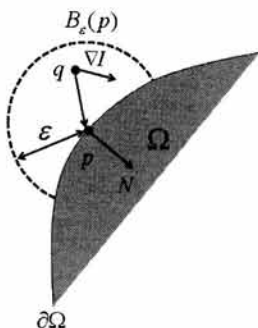
그러면 점 p 의 화소값을 $B_\epsilon(p)$ 안의 점 q 들이 함수로서 정규화된 가중치 함수 $w(p, q)$ 에 의해 다음과 같이 표현한다.

$$I(p) = \frac{\sum_{q \in B_\epsilon(p)} w(p, q) [I(q) + \nabla I(q)(p - q)]}{\sum_{q \in B_\epsilon(p)} w(p, q)} \quad (4)$$

여기서 $w(p, q) = \text{dir}(p, q) \cdot \text{dst}(p, q) \cdot \text{lev}(p, q)$

where

$$\text{dir}(p, q) = \frac{p - q}{\|p - q\|} \cdot N(p)$$



(그림 4) 복원 점과 주변

$$\text{dst}(p, q) = \frac{d_0^2}{\|p - q\|^2}$$

$$\text{lev}(p, q) = \frac{T_0}{1 + |T(p) - T(q)|}$$

영역 Ω 전체를 복원하기 위해서는 (4)식을 반복적으로 적용하는데, 초기 경계 $\partial\Omega_i$ 까지의 거리순으로 $\partial\Omega$ 의 화소들을 전진시킴으로서 $\partial\Omega$ 을 Ω 으로 전개시키는데 이를 위해 FMM을 이용하게 된다. 자세한 복원 방법은 [15]에 나타나 있다.

3.4 복원된 화소의 보정

복원된 이미지는 지정된 영역 경계 주변 정보를 이용하므로 복원된 영역은 왜곡이 발생하게 된다. 정확한 배경 생성을 위해서는 이를 보정하는 단계가 필요하다.

복원된 영역을 보정하기 위해 (5)식과 같이 복원된 영역과 가지고 있는 모든 프레임에서의 영역간의 차이를 계산하고 그 차이가 가장 적은 영역 복원된 영역으로 대체한다.

모든 $t = 1, \dots, T$ 에 대하여
(여기서 T 는 시퀀스의 길이)

$$R_B = \text{argmin}_{R_{F_t}} |R_I - R_{F_t}| \quad (5)$$

여기서

$$|R_I - R_{F_t}| = \sum_{(x, y) \in R_I} (I_{R_t}(x, y) - I_{F_t}(x, y))^2$$

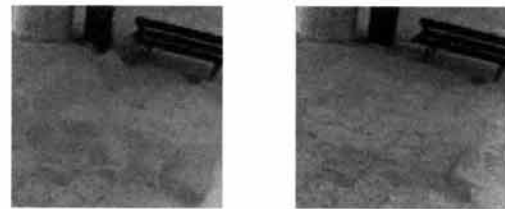
R_B : 배경 영역

R_I : 복원된 영역

R_{F_t} : t 번째 프레임에서의 R_I 와 동일한 위치의 영역

$I_R(x, y)$: 영역 R 에서의 (x, y) 의 화소값

(그림 5)는 복원된 배경 영상과 보정된 배경 영상의 비교를 보인 것이다. 보정된 영상이 복원된 영상보다 더 좋은 결과를 보이는 것을 확인할 수 있다.



(a) 이미지 복원 결과

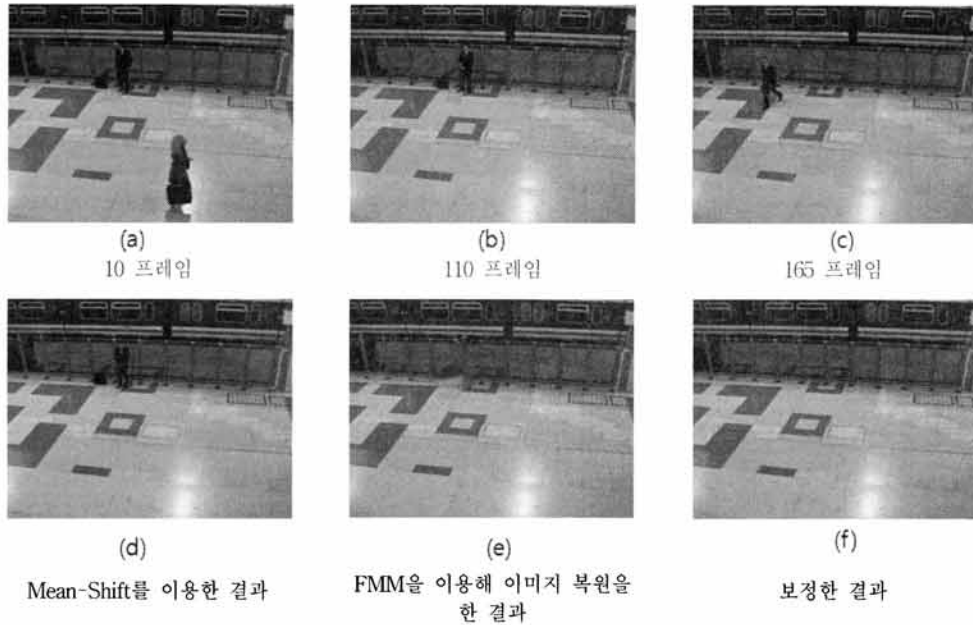
(b) 보정한 결과

(그림 5) 복원 이미지와 제안한 방법의 결과 비교

4. 실험

제안한 방법은 Microsoft Visual C++로 구현했다.

(그림 6)은 (그림 3)에서 보인 데이터 시퀀스에 대한 실험



(그림 6) 테스트 시퀀스 1 실험 결과

결과를 보인 것이다. (그림 6) (a)부터 (그림 6) (c)는 각 시퀀스를 보인 것이고, (그림 6) (d)는 기존의 mean-shift를 이용해 시간축에 따른 화소값의 최빈값을 추정하여 생성한 배경 결과이다. 사람의 위치가 장시간 변하지 않아 사람까지 배경으로 인식하는 것을 확인할 수 있다. (그림 6) (e)는 시간에 따른 화소 변화 빈도를 이용해 이미지를 복원한 결과이다. (그림 6) (f)는 복원된 이미지를 이용해 왜곡된 이미지를 다른 프레임 정보를 이용해 보정한 것이다.

(그림 7) 은 (그림 6) (e)와 (그림 6) (f)의 일부분을 확대해서 나타낸 것이다. (그림 7) (a)에서 볼 수 있듯이 원래 사람이 있던 영역은 주변 값 정보를 이용해 복원되어 영역에 왜곡이 발생했음을 확인할 수 있다. 하지만 제안한 방법의 경우에는 왜곡된 영역을 다른 프레임 정보를 이용해 보정함으로써 (그림 7)(b)와 같이 왜곡을 줄여 보다 정확하게 배경을 생성함을 확인할 수 있다.

(그림 8)은 다양한 환경에서의 실험 결과를 보인 것이다. (그림 8) (a)는 PETS 2001 데이터 시퀀스에 대한 실험결과를 보인 것이고 (그림 8) (b)는 <http://www.openvisor.org>에

서 얻은 테스트 시퀀스에 대한 실험 결과를 보인 것이다. (그림 8) (c)와 (그림 8) (d)는 AVSS(Advanced Video and Signal Based Surveillance) 2007에서 제공한 데이터 시퀀스의 결과를 보인 것으로 다양한 환경에서도 정확한 배경이 생성됨을 보여주고 있다.

<표 2>는 사용된 테스트 시퀀스의 에러율을 나타낸 것이다. 에러율은 다음의 식 (6)과 같이 계산했다.

I_g : groundtruth 이미지

I_b : 생성된 배경 이미지

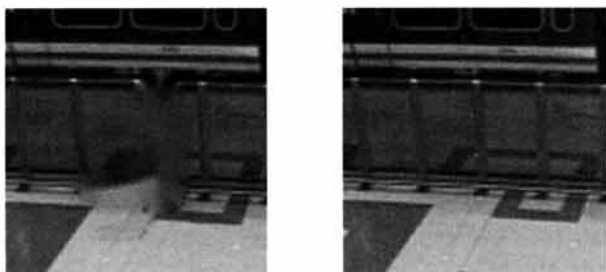
T_e : threshold

$$E = \{(x,y) \mid |I_g(x,y) - I_b(x,y)| < T_e\}$$

$$Err = \frac{\|E\|}{\text{이미지 사이즈}} * 100 \quad (6)$$

$\| \cdot \|$: 집합의 cardinality

표 2를 보면 그림 8(d)의 테스트 시퀀스를 제외하고는 대체로 낮은 에러율을 나타내고 있다. 그림 8(d)의 테스트 시퀀스의 경우에는 시퀀스 자체의 해상도와 품질이 고르지 못해 ground truth와 생성된 배경과의 차이가 많이 나타났다.

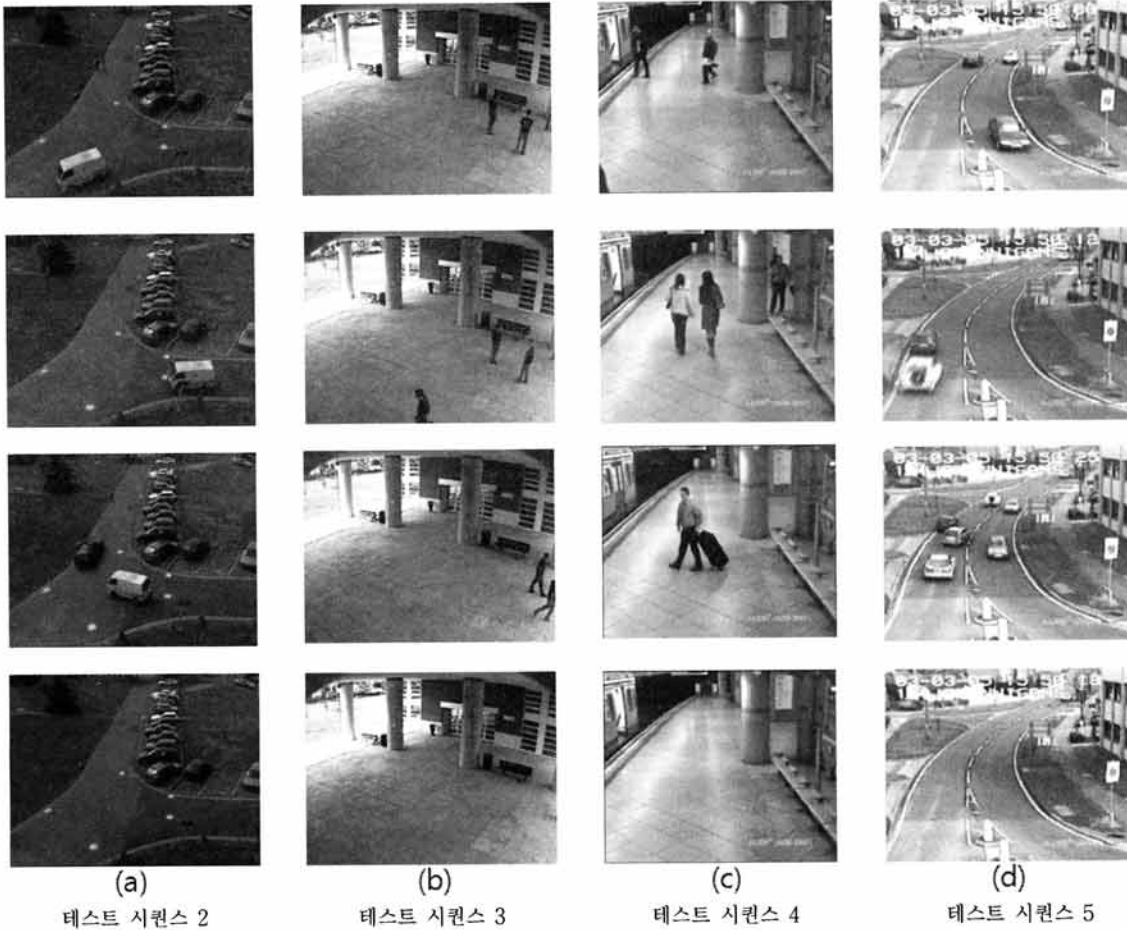


(a) 이미지 복원 결과
(b) 보정한 결과

(그림 7) 복원 이미지와 제안한 방법의 결과 비교

<표 2> 테스트 시퀀스의 에러율

	mean-shift		제안한 방법	
	$T_e = 5$	$T_e = 10$	$T_e = 5$	$T_e = 10$
테스트 시퀀스 1	8.53	6.12	0.11	0.01
테스트 시퀀스 2	14.20	6.44	6.31	1.13
테스트 시퀀스 3	4.11	3.68	0.25	0.03
테스트 시퀀스 4	4.60	2.29	3.17	0.68
테스트 시퀀스 5	71.92	45.99	35.65	19.5



(그림 8) 다양한 환경에서의 실험 결과

5. 결 론

시간 축에 따른 영상 정보는 배경을 생성하는데 매우 중요한 정보이다. 하지만 시간 축에 따른 정보만을 이용할 경우 움직임이 없거나 적은 물체가 존재할 때 정확한 배경을 생성하기가 어렵다. 이럴 경우 도로나 주차장, 터미널과 같이 경우에 따라 움직임이 없거나 적은 물체가 존재하는 환경에서는 생성한 배경영상에 물체가 포함될 가능성이 높기 때문에 시스템의 물체 검출이나 추적 품질이 저하될 수 있다. 이를 해결하기 위해 본 논문에서는 시간에 따른 화소값 정보와 공간 정보에 따른 화소값 정보를 이용하는 방법을 제안했다.

시간에 따른 화소값들에 대하여 mean-shift를 적용해 각 화소값의 최빈값을 추정하여 배경을 생성하고 잘못 생성된 배경 화소를 분류한 다음 FMM을 적용해 잘못된 영역을 복원 및 보정과정을 거쳐 최종 배경을 생성한다. 제안한 방법은 시간 축에 따른 화소 정보와 공간 축에 따른 화소 정보를 이용하여 배경을 생성했기 때문에 장시간 움직임이 없는 물체가 존재하는 복잡한 환경에서 시간에 따른 화소값 정보만을 이용했던 기존의 방법과 비교했을 때 보다 바람직한 결과를 얻었다.

하지만, 제안한 방법은 환경의 조도 변화나 물체들의 움

직임의 정도에 따라 T_r , T_p 값을 다르게 적용해야 한다. 따라서 물체 조도의 변화 정도와 물체들의 움직임의 정도를 측정하여 하므로 최적의 T_r , T_p 값을 갱신하는 별도의 과정이 필요하다. 따라서 차후연구과제로서 최적의 T_r , T_p 값을 자동으로 계산해 최적의 배경을 생성하는 방법을 연구할 예정이다.

참 고 문 헌

- [1] C. R. Wren, A. Azarbayejani, T. Darrell, A. P. Pentland, Pfunder: Real-time tracking of the human body, IEEE Transactions on Pattern Analysis and Machine Intelligence 19(7) (1998) 780-785.
- [2] C. Stauffer, W. Grimson, Adaptive background mixture models for real-time tracking, in: IEEE Conference on Computer Vision and Pattern Recognition, Fort Collins, USA, 1999, pp.246-252.
- [3] S. Rowe, A. Blake, Statistical background modelling for tracking with a virtual camera, in: British Machine Vision Conference, Birmingham, UK, 1995, pp.423-432.
- [4] C. Stauffer, W. E. L. Grimson, Learning patterns of activity

using real-time tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence 22(8) (2000) 747-757.

[5] K. Toyama, J. Krumm, B. Brumitt, B. Meyers., Wallflower: Principles and practice of background maintenance, in: IEEE International Conference on Computer Vision, Corfu, Greece, 1999, pp.255-261.

[6] I. Haritaoglu, D. Harwood, L. S. Davis, W4: real-time surveillance of people and their activities, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) (2000) 809-830.

[7] Junxian Wang, George Bebis and Ronald Miller, "Robust Video-Based Surveillance by Integrating Target Detection with Tracking," CVPR, 2006.

[8] A. Elgammal, D. Harwood, L. Davis, Non-parametric model for background subtraction, in: European Conference on Computer Vision, Dublin, Ireland, 2000, pp.751-767.

[9] K. Fukunaga, L. Hostetler, The estimation of the gradient of a density function, with applications in pattern recognition, IEEE Transactions on Information Theory 21 (1975) 32-40.

[10] D. Comaniciu, P. Meer, Mean shift: a robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence 24(5) (2002) 603-619.

[11] Yazhou Liu, Hongxun Yao, Wen Gao, Xilin Chen, Debin Zhao, Nonparametric Background Generation, Pattern Recognition, 2006. ICPR 2006. 18th International Conference on Volume 4, 0-0 0 Page(s):916-919.

[12] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. Image Inpainting, In Proceedings SIGGRAPH 2000, Computer Graphics Proceedings. Annual Conference Series, edited by Kurt Akeley, pp.417-424, Reading, MA: Addison-Wesley, 2000.

[13] M. Bertalmio, A.L. Bertozzi, and G. Sapiro. Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting, In Proc. ICCV 2001, pp. 1335-1362, IEEE CS Press 1. [CITY]:[PUB], 2001.

[14] J. A. Sethian. Level Set Methods and Fast Marching Methods, Second edition. Cambridge, UK: Cambridge Univ. Press, 1999.

[15] A. Telea, "An image inpainting technique based on the fast marching method," J. Graphics Tools, Vol.9, No.1, pp.25-36, 2004.



조 상 현

e-mail : cshgreat@catholic.ac.kr
 2003년 가톨릭대학교 수학과 졸업
 2005년 가톨릭대학교 컴퓨터공학과(석사)
 2005년~현 재 가톨릭대학교 컴퓨터공학과
 박사과정
 관심분야: 컴퓨터 비전, 패턴인식, 컴퓨터
 그래픽스



강 행 봉

e-mail : hbkang@catholic.ac.kr
 1980년 한양 대학교 전자공학과 졸업
 1986년 한양대학교 전자공학과(석사)
 1989년 Ohio State Univ. 컴퓨터공학(석사)
 1993년 Rensselaer Polytechnic Institute
 컴퓨터공학(박사)
 1994년~1997년 삼성종합기술원 수석연구원
 1997년~현 재 가톨릭대학교 디지털미디어학부 교수
 2005년 UC Santa Barbara Visiting Professor
 관심분야: 컴퓨터비전, 컴퓨터그래픽스, HCI, 인공지능, 기계학습