

Background Subtraction Algorithm Based on Multiple Interval Pixel Sampling

Dongeun Lee[†] · Young Kyu Choi^{††}

ABSTRACT

Background subtraction is one of the key techniques for automatic video content analysis, especially in the tasks of visual detection and tracking of moving object. In this paper, we present a new sample-based technique for background extraction that provides background image as well as background model. To handle both high-frequency and low-frequency events at the same time, multiple interval background models are adopted. The main innovation concerns the use of a confidence factor to select the best model from the *multiple interval background models*. To our knowledge, it is the first time that a confidence factor is used for merging several background models in the field of background extraction. Experimental results revealed that our approach based on multiple interval sampling works well in complicated situations containing various speed moving objects with environmental changes.

Keywords : Background Subtraction, Motion Detection, Sampling-Based Algorithm, Multiple Interval Sampling

다중 구간 샘플링에 기반한 배경제거 알고리즘

이 동 은[†] · 최 영 규^{††}

요 약

배경제거는 동영상의 내용을 자동으로 분석하기 위한 매우 중요한 기술의 하나로 움직이는 객체를 검출하고 추적하기 위한 핵심 기술이다. 본 논문에서는 배경 모델과 함께 배경 영상을 제공하는 새로운 샘플링 기반의 배경제거 알고리즘을 제안한다. 제안된 방법에서는 움직임이 빠른 객체와 느린 객체를 동시에 처리하기 위해 다중 구간 샘플링 기법을 이용하여 배경 모델을 생성한다. 이러한 다중 구간 배경 모델들로부터 최선의 배경 모델을 만들기 위해 “신뢰도”를 사용한 것이 본 논문의 특징이다. 배경 제거 분야에서 다양한 모델을 병합하여 하나의 모델을 만들기 위해 신뢰도를 정의하여 사용한 경우는 현재까지 보고되지 않았다. 실험을 통해 제안된 방법이 다양한 속도의 객체가 존재하고 시간에 따른 그림자의 이동과 같은 환경 변화가 있는 응용에서도 안정적인 결과를 나타내는 것을 알 수 있었다.

키워드 : 배경 제거, 움직임 검출, 샘플링 기반 알고리즘, 다중 구간 샘플링

1. 서 론

다양한 머신 비전 응용분야에서 지금까지 매우 중요하게 인식되고 있는 처리 과정의 하나인 영역 분할(segmentation)은 영상내의 관심 있는 물체 영역을 배경 영역에서 구분하는 것을 목표로 한다. 특히 움직이는 물체 영역을 실시간으로 배경에서 분리하는 것은 동영상 해석에서 매우 기본적이고 중요한 작업이다. 지능형 보안감시 분야의 경우 이와 같은 배경 분리를 통해 추출된 전경 객체들을 인식이나 추적, 행동 분석 등을 위한 기본 데이터로 사용하는

데, 추출되는 영역의 정확도가 전체 시스템의 성능에 크게 영향을 미친다. 본 논문에서는 동영상에서 움직이는 객체를 추출하기 위한 배경 제거(background subtraction) 방법에 대해 논하고자 한다.

가장 간단한 배경 제거 방법인 프레임 차이법은 인접한 프레임 사이의 차 영상을 만들고, 이를 분석하여 전경 영역을 추출한다. 이 방법은 간단하고 계산 량이 적다는 장점이 있으나, 잡음의 영향을 매우 많이 받고 영상에서 모션을 발생시키는 지역에 대한 정보를 알 수 없는 등의 문제점이 있으며, 따라서 복잡한 후처리 과정이 필요하다는 단점이 있다[1].

이에 비해 배경 차이법은 배경 모델(background model)을 먼저 만들고 갱신하면서 현재 프레임과 모델과의 일치도를 계산하여 전경 영역을 판단한다. 이는 인접한 두 프레임

[†] 정 회 원 : LG CNS 정보기술연구원 선임연구원

^{††} 종신회원 : 한국기술교육대학교 컴퓨터공학부 교수

논문접수 : 2012년 6월 26일

수정일 : 1차 2012년 8월 13일, 2차 2012년 9월 13일

심사완료 : 2012년 9월 26일

* Corresponding Author : Young Kyu Choi(ykchoi@kut.ac.kr)

의 화소 값을 직접 비교하는 것이 아니라 배경이라고 할 수 있는 모델을 이전 프레임들로부터 유추해 내고 현재 화소 값이 이 모델에 얼마나 잘 맞는지를 계산하여 그 화소가 배경인지 전경인지를 결정한다. 배경 모델의 갱신은 일반적으로 오래 전 프레임의 영향을 점진적으로 줄이면서 최근 프레임의 영향을 강조하는 방향으로 처리되는데, 시간적 평활법과 시간적 중간치법[2] 등 비교적 간단한 방법에서부터 복잡한 모델을 사용하는 다양한 방법들이 소개되고 있다. Wren 등은 배경 화소를 하나의 값이 아니라 가우시안 분포(Gaussian distribution)로 가정하고 모델링 하는 방법을 제안하였다[3]. 정적인 배경에서 이 방법은 기존의 방법들에 비해 유리하지만 나뭇가지가 움직이거나 잔잔히 물결이 일렁이는 것과 같이 동적인(dynamic) 배경에는 한계가 있었다. 이를 보완하기 위해 각 화소에 대한 모델로 여러 개의 가우시안 분포를 이용하는 적응적 가우시안 혼합모델(Mixture Of Gaussian model, MOG)이 제안되었는데[4], 하나의 확률 분포만을 사용하는 기존 방법에 비해 다수의 확률 분포로 한 화소의 배경을 모델링 할 수 있도록 허용함에 따라 동적인 배경을 더 잘 대응할 수 있다는 장점이 있다. 그러나 이 방법은 이론적으로 우수하지만 계산 량이 많아 실시간 응용에 적용하기가 어렵고, 배경 분포의 개수를 결정하는 방법이나 학습을 등의 내부 파라미터를 최적으로 결정하기도 어렵다는 단점이 있다. 또한 가우시안 분포를 화소에 대한 통계 모델로 사용하는 것이 항상 최적인가 하는 문제가 있으며, 동적인 배경에 대해 방법적으로는 개선되었으나 다양한 형태의 동적인 환경에 대한 완전한 해결책은 될 수 없다는 한계가 있다.

이후에 MOG를 보완하는 다양한 알고리즘이 소개되었다[5-8]. Elgammal 등은 확률 밀도 분포로 커널 기반 밀도 연산(kernel-based density estimation) 방법을 제시하고, 나뭇잎의 움직임과 같이 작은 움직임을 포함하는 배경에 효과적임을 제시하였다[5]. 그러나 이 방법은 각 화소에서의 커널 밀도 연산을 위한 계산 량이 많으며, 이를 줄이기 위해 다수의 미리 계산된 테이블을 사용하였다. 또한 각 화소에서 이전 밝기 값들의 미디언 절대 변위(median absolute deviation)를 이용하여 커널의 대역폭을 계산하는데, 배경 샘플의 분포가 다봉(multi-modal)인 경우 대역폭이 부정확하다는 단점이 있다.

최근에는 화소 값의 기반 확률 분포에 대해서 어떠한 가 정도 하지 않는 비매개변수형 알고리즘들이 제안되고 있는데, Pietikainen은 지역적 이진패턴(Local Binary Pattern, LBP)을 기반으로 한 텍스처 기반 알고리즘을 제안하였다[9]. 또한, 샘플링 기반 접근법으로, Wang 등은 이전 화소의 값들을 큐에 저장하고 현재 프레임의 화소 값과의 “의견 일치(cosensus)” 정도를 측정하여 전경 객체를 추출하는 방법을 소개하였으며[10], Barnich 등은 무작위로 샘플 큐 내의 샘플의 제거하는 방법을 제안하였다[11]. 연구 결과 이러한 샘플링 기반 접근이 처리 속도가 빠르고 알고리즘의 안정성

이 높은 것으로 나타났는데, 따라서 본 논문에서는 샘플링 기반의 새로운 배경제거 알고리즘을 제안한다.

본 논문의 구성은 다음과 같다. 먼저 2장에서 본 논문과 관련된 기존 연구들을 소개하고 제안한 알고리즘의 동기를 설명한다. 3장에서는 제안된 다중 구간 샘플링 기반 배경제거 알고리즘을 설명한다. 4장에서 실험결과를 보이고 5장에서 결론을 맺는다.

2. 관련 연구 및 제안된 방법의 동기

최근에 Wang 등은 화소의 배경 모델로 어떤 확률 분포를 사용하지 않고 이전 화소의 값을 직접 사용하는 “의견 일치(cosensus)” 기반의 배경제거 알고리즘인 SACON을 제안하고 복잡한 확률 분포를 사용하는 것에 비해 훨씬 단순하고 계산량이 적으면서도 안정적으로 동작함을 제시하였다[10]. 이 방법은 먼저 각각의 화소에서 나타난 이전 프레임의 값들을 캐쉬(history cache)에 저장하고 현재 프레임의 값을 이들과 비교하여 샘플(SAMPLE)의 의견 일치 정도(CONsensus)를 계산하고, 이 값이 문턱치 이하인 경우 전경 화소로 판단하는 방법이다. 이 방법은 복잡한 확률 밀도 함수를 사용하지 않으며 이전 화소 값을 그대로 배경 모델에 사용한다는 측면에서 알고리즘적으로 매우 간단하고, 처리시간 측면에서도 장점이 있다. 또한 컬러 영상에 채널 별로 알고리즘을 적용하는 데도 매우 효과적이다. 그러나 이전 값들의 저장을 위한 캐쉬의 크기가 커지면 메모리 사용량이 늘고, 비교할 샘플의 수가 증가함에 따라 처리시간이 느려지는 단점이 있다. 캐쉬의 크기가 너무 작으면 너무 짧은 시간에 대한 기록만이 배경 모델에 반영되므로 천천히 움직이는 객체에 대한 처리가 매우 어려워지는 단점이 있다. 본 논문에서는 SACON에서와 같이 샘플 기반의 알고리즘을 제안하는데, 특히 다중 구간 샘플링과 신뢰도를 바탕으로 한 다중 구간 모델의 병합 방법을 제시하여 움직임이 빠르고 느린 객체가 혼재하는 환경에서도 안정적으로 추출이 가능하도록 하는 것을 목표로 하였다.

현재까지 소개된 다양한 배경제거 방법들은 배경 영상의 제공 여부에 따라서 두 그룹으로 분류할 수 있는데, 각 화소에 대해 배경 값이 반복적으로 구해지는 그룹과, MOG와 같이 배경 값이 아니라 하나 또는 여러 개의 확률 분포로 그 화소의 통계적 특성이 모델링되고 이를 이용해 화소의 전경 여부를 결정하는, 즉 배경 영상이 만들어지지 않는 그룹이다. 배경 제거 알고리즘의 일부 응용분야에서는 배경 영상의 생성이 반드시 필요한 경우가 있는데, 예를 들어, 비디오 개요(video synopsis) 시스템의 경우 동영상의 이벤트 기반 압축과 검색을 위해 배경 제거와 함께 반드시 배경 영상이 필요하게 된다[12]. 따라서 본 연구에서는 전경 객체의 추출과 함께 기본적으로 배경 영상을 생성할 수 있는 알고리즘에 초점을 맞추었으며, 향후 이 알고리즘은 비디오의 효율적인 검색을 위한 압축 및 검색 시스템에 활용될 예정이다.

3. 샘플링 기반의 배경제거 알고리즘

제안된 방법은 화소의 배경 모델로 확률 밀도함수를 사용하는 것이 아니라 이전 프레임들로부터 추출된 값들을 사용하는 샘플링 기반 알고리즘으로, 특히 샘플링 구간(sampling interval)을 다중으로 적용하여 각 화소마다 다중 구간 샘플링 모델을 만들고, 이들을 통합하여 최종 배경 영상을 생성하고 전경 객체를 추출하는 것을 특징으로 한다. 이를 위해 먼저 단일 구간 배경 모델링 방법을 소개하고, 모델에서 배경 값과 신뢰도를 계산하는 방법을 제시하며, 마지막으로 다중 구간의 모델들을 통합하여 배경영상과 전경 객체를 추출하는 방법을 설명한다.

3.1 단일 구간 배경 모델링

일반적인 보안 감시 응용분야에서는 초당 수 프레임에서 수십 프레임의 매우 방대한 크기의 영상을 저장 및 분석하게 된다. 그러나 배경제거 알고리즘에서는 배경 갱신의 속도를 반드시 이와 동일하게 유지할 필요는 없는데, 이것은 동영상의 시간적 중복성(temporal redundancy)에 따른 것으로 특히 배경 제거와 같이 기본적으로 카메라가 고정되어 있는 경우 화소값에 많은 시간적 중복성이 발생하게 된다. 따라서 제안된 알고리즘에서는 먼저 입력되는 비디오 프레임들을 시간적으로 일정한 길이의 구간으로 나누고, 하나의 구간에서 하나의 대푯값을 추출하여 그 구간을 대표하도록 한다.

임의의 화소 위치 p 에서 t 번째 프레임에 측정된 화소 값을 $x_t(p)$ 라고 하고, 대푯값 추출 구간의 길이를 m 프레임이라 하면, i 번째 구간에서 연속적으로 측정되는 모든 화소 값 $S_i(p)$ 은 다음과 같다.

$$S_i(p) = \{x_{t_0+t}(p) | t = 1, 2, \dots, m, t_0 = (i-1) \times m\}$$

$S_i(p)$ 에서 이 구간을 대표하는 샘플 값 $s_i(p)$ 를 추출하기 위해서 다양한 방법을 적용할 수 있는데, 전체 화소 값의 평균이나 중앙값(median)을 이용하는 방법도 있으며, 무작위 추출이나 일정한 위치의 값을 선택하여 사용할 수도 있다. 연구 결과 다양한 연산이나 복잡한 추출 방법을 적용해도 처리시간에 비해 성능의 개선이 미미하다고 판단되었으며, 따라서 본 논문에서는 화소마다 일정한 위치에서 값을 추출하여 그 구간을 대표하는 샘플 값으로 이용한다.

샘플 기반 알고리즘에서는 추출되는 샘플 값들을 샘플 큐 또는 버퍼에 저장하게 되며, 버퍼에 저장된 샘플 값들이 해당 화소에 대한 배경 모델이 된다. 버퍼의 크기가 제한적이기 때문에 새로운 샘플들이 추출되었을 때 버퍼에 저장된 이전 샘플을 하나 버리고 새로운 샘플을 저장해야 한다. 보통의 경우 선입(first-in) 선출(first-out)의 큐 형태를 사용하지만, Barnich등은 버퍼에 저장된 순서와 상관없이 무작위로 샘플을 선택해 버리고 이곳에 새로운 샘플 값을 저장하는

방법을 사용하였다[11]. 본 연구에서 고찰 결과, 랜덤 잡음이 많은 영상에서는 이와 같은 방법이 의미가 있을 수 있지만, 영상 기술의 발전에 따른 최근 감시 카메라의 화질을 고려하면 큐 형태의 동작이 가능한 한 오래된 샘플의 영향을 줄이고 최근 샘플의 영향을 모델에 반영한다는 점에서 더 바람직하다고 판단하였다.

샘플 버퍼의 크기를 N 이라 하면, 처음 N 개의 값들이 입력되어 버퍼가 채워지고, 이후에 새로운 샘플이 입력되면 가장 오래전에 저장된 샘플이 삭제된다. 따라서 이 샘플 버퍼는 항상 가장 최근 N 개 구간에서 추출된 샘플들이 저장되게 되는데, p 화소에 대해 i 번째 구간에서의 샘플 버퍼 $Q_i(p)$ 에 저장되는 내용은 다음과 같다.

$$Q_i(p) = \{s_{i-N+1}(p), \dots, s_{i-1}(p), s_i(p)\}$$

입력 영상이 다수의 채널로 구성된 경우는 버퍼가 채널의 수만큼 생성되는데, 예를 들어 RGB 모델을 사용하는 경우 $Q_i^R(p), Q_i^G(p)$ 및 $Q_i^B(p)$ 의 샘플 버퍼가 생성되어 각 구간에서 갱신된다. 이러한 샘플 버퍼는 해당 화소의 배경 모델이 되는데, 확률 밀도 함수를 사용하지 않아 계산이 간단하고, 다수의 샘플 값들이 반영되므로 동적인 배경에도 대응할 수 있는 기본 자료를 제공한다.

3.2 배경 값 및 배경 신뢰도

앞에서 설명한 바와 같이 제안된 방법은 배경 분포만을 유지하는 것이 아니라 배경 영상을 제공하는 것을 목표로 하는데, 따라서 샘플 버퍼를 대표하는 값을 추출하여 이를 배경 화소 값으로 정의한다. 이를 위해 샘플 버퍼의 값들에 대한 다양한 연산을 적용할 수 있는데, 평균이나 중앙값을 선택할 수도 있다. 본 논문에서는 버퍼 내의 각 샘플 값들에 대해 다른 샘플들과의 일치도(agreement degree)를 구하고, 이 값이 최대인 샘플을 배경 화소 값으로 정의하였다. 먼저, 샘플 버퍼 내의 두 화소 $s_i(p)$ 와 $s_j(p)$ 의 일치(agreement) $A_{i,j}(p)$ 를 다음과 같이 정의한다.

$$A_{i,j}(p) = \begin{cases} 1 & \text{if } |s_i(p) - s_j(p)| < T_{agree} \\ 0 & \text{otherwise} \end{cases}$$

이때, T_{agree} 는 샘플간의 일치 여부를 결정하기 위한 문턱치로 카메라의 센서 특성이나 영상의 화질 등에 따라 적절히 설정할 수 있다. 이를 바탕으로 샘플 버퍼 내의 한 화소 $s_i(p)$ 의 일치도 $B_i(p)$ 를 다음과 같이 정의한다.

$$B_i(p) = \sum_{n=1}^{n=N} A_{i,n}(p)$$

일치도가 큰 샘플은 해당 화소에서 지금까지 측정된 다른 많은 샘플들과 유사한 샘플이며, 따라서 배경으로 선택될

가능성이 높다. 이러한 일치도를 바탕으로 i 번째 구간 샘플을 배경 값으로 선택할 때의 신뢰도(confidence) $C_i(p)$ 를 다음과 같이 정의한다.

$$C_i(p) = \frac{B_i(p)}{N}$$

$C_i(p)$ 는 최소 값이 $1/N$ 이고 최대 값이 1인데, 최대 값의 경우 $s_i(p)$ 가 샘플 버퍼 내의 모든 값들과 일치한다는 것을 나타낸다. 최종적으로 샘플 버퍼를 대표하는 배경 화소 값 $BG(p)$ 과 신뢰도 $CF(p)$ 는 최대의 신뢰도를 나타내는 구간의 샘플 값을 이용해 다음과 같이 정의된다.

$$i^* = \arg_i \max\{C_i(p) | i = 1, 2, \dots, N\}$$

$$BG(p) = s_{i^*}(p)$$

$$CF(p) = C_{i^*}(p)$$

이때, i^* 는 최대 신뢰도를 나타내는 구간을 나타내며, 영상 내의 모든 화소 p 에서 $BG(p)$ 가 추출되면 해당 샘플링 구간에 대한 배경 영상이 만들어진다. $CF(p)$ 는 배경 값의 신뢰도를 나타내지만, 마찬가지로 배경 모델을 얼마나 신뢰할 수 있는지를 나타내기도 하는데, 샘플 버퍼의 많은 샘플들이 비슷한 값으로 나타내면 $CF(p)$ 가 커지고, 버퍼 내의 값들이 제각기 다르게 나타나면 작아지기 때문이다. 따라서 $CF(p)$ 를 배경 화소 값뿐만 아니라 배경 모델의 신뢰도로 활용한다.

3.3 다중 구간 배경 모델의 통합 및 배경 제거

대부분의 배경제거 알고리즘은 영상 내의 객체의 움직임이나 환경 변화 특성과 배경 갱신의 주기에 따라 크게 영향을 받는데, 예를 들어, 객체의 움직임이 비교적 느린 환경에서 배경 갱신이 너무 빠르면 객체 영역의 정확한 추출이 어려워지며, 반대로 구름이나 해 등의 요인에 의해 영상 전체의 점진적인 변화가 있는 영상에서 배경 갱신 속도가 너무 느리면 실제 객체가 아닌 많은 영역들이 전경으로 잘못 검출될 가능성이 많아진다. 샘플 기반 알고리즘에서 이를 해결할 수 있는 쉬운 방법으로는 샘플 버퍼를 매우 크게 만들어 많은 샘플들을 사용하도록 하는 것이나, 이것은 시간과 메모리의 측면에서 바람직하지 않다. Elgammal 등은 단구간(short-term)과 장구간(long-term)의 두 가지 모델을 운영하면서 각 모델에서 추출된 결과 전경 영역의 교차(intersection) 영역을 취하는 등의 방법을 사용하여 이 문제를 해결하고자 하였는데[5], 다중 구간 모델로 일반화하지 못했으며 전경 영역의 교차로 최종 전경 영역을 구하는 것은 원하지 않는 미검출(suppress true positive)을 발생시킬 수 있다.

본 논문에서는 이를 위한 보다 체계적인 해결책을 제안하는데, 배경 모델의 신뢰도 값을 이용하여 다중 구간의 배경

모델 정보를 통합하는 것이 핵심이다. K 가지 길이의 구간에서 배경모델을 생성했다고 가정하면, 하나의 화소에 대해 K 가지의 배경 모델이 생성되며, 각 모델에 대해서 대표 화소 값과 배경 신뢰도가 구해진다. 이러한 다중 구간의 모델들 중에서 신뢰도가 가장 높은 모델을 최종 모델로 선택한다. 즉, $CF^k(p)$ 를 m_k 길이의 구간 모델에서 계산된 배경 신뢰도라 하면 다중 구간 배경 모델에서 최종적으로 선택되는 구간 k^* 는 다음과 같이 정의된다.

$$k^* = \arg_k \max\{CF^k(p) | k = 1, 2, \dots, K\}$$

이에 따라, p 화소에 대한 최종적인 배경 모델은 길이가 m_{k^*} 인 구간 모델이 되며, 이 화소의 최종 배경 값과 배경 신뢰도는 다음과 같이 정의된다.

$$BG(p) = BG^{k^*}(p)$$

$$CF(p) = CF^{k^*}(p)$$

배경 모델에 따른 전경 객체의 생성에는 다양한 기준이 적용될 수 있는데, SACON에서와 같이 현 프레임의 값을 배경 모델의 샘플 버퍼의 값들과 비교하여 일치 정도 구할 수도 있으며[10], VIBE에서와 같이 현재 프레임의 화소 값과 일정 거리 이내에 있는 샘플들의 개수가 기준 이하인 경우 전경 화소로 판단하는 방법을 채택할 수도 있다[11]. 이러한 방법들은 특정한 형태의 동적인 배경의 영상에 대해서는 보다 잘 동작할 수는 있지만, 계산량이 많고 모든 동적 배경에 대한 해결책이 될 수는 없다. 본 논문에서는 동적인 배경보다는 배경은 정적이지만 다양한 속도의 객체와 환경 변화가 있는 경우에 대해 초점을 맞추었으며, 최종 배경 영상을 만들고 현 프레임과의 차이를 이용하는 방법을 사용하였다. 제안된 방법은 기존의 방법들에 비해 짧은 구간 및 긴 구간의 데이터를 효과적으로 포함하고 있으므로 프레임 차이만으로도 좋은 결과를 나타냈으며, 다중 구간 모델을 활용하면 다양한 동적인 배경에서도 더 안정적인 결과를 나타낼 수 있는 여러 가지 방법을 채택할 수 있는 확장성이 있다.

3.4 제안된 알고리즘의 특징

제안된 알고리즘의 특징은 다음과 같다. 먼저 화소의 배경 모델로 확률 밀도함수를 사용하지 않고 이전 프레임에서 샘플링 된 화소 값을 사용하는 샘플링 기반 방식을 취하여 계산 속도 면에서 우월하며 실시간 다채널 응용에 적용하기에 적합하다.

이전 프레임의 모든 데이터를 사용하지 않고 구간을 정하고 대표 값을 샘플링하여 사용하므로 샘플 버퍼의 길이를 줄일 수 있는데, 특히 다양한 길이의 샘플링 구간을 지원할 수 있으며 이를 통해 빠른 객체와 느린 객체에 모두 안정적으로 동작할 수 있도록 하였다.

다중 구간의 모델을 운영할 때, 각 구간의 배경 값에 신뢰도를 부여하는 방법을 제안하였으며, 이를 바탕으로 다중 구간의 배경 값을 하나로 병합할 수 있으며, 최적의 배경 모델을 선택할 수 있다. 기존에 단구간 및 장구간의 두 가지 모델을 운영하는 방법이 제시된 바가 있었지만[5], 이를 다양한 구간의 모델 운영으로 확장한 사례는 보고되지 않았으며, 특히 두 모델을 이용한 결과를 병합하는데 신뢰도를 정의하고 이를 바탕으로 다양한 구간 모델의 결과를 병합하는 방법은 아직 보고된 바가 없다.

제안된 방법은 최종적으로 하나의 배경 영상과 이 영상에 대한 신뢰도 영상을 구할 수 있으며, 배경 영상은 객체 추출 뿐 아니라 다양한 용도로 활용할 수 있다. 최근에 발표되고 있는 동영상의 효율적인 압축과 검색을 위한 비디오 개요 시스템[12] 등에서는 배경 분포가 아니라 배경 영상이 필요한데, 제안된 알고리즘은 이러한 시스템을 위한 안정적인 배경 영상을 제공할 수 있다. 특히 구간을 단기, 중기, 장기 등 여러 구간으로 나누어 배경 모델을 만들고 신뢰도를 바탕으로 배경 영상을 생성하면 다양한 기간 정보가 반영된 매우 효과적인 배경 영상을 생성할 수 있다.

4. 실험

제안된 알고리즘은 Windows-7 운영체제하의 Pentium-PC에서 구현되고 실행되었는데, C와 C++를 이용하여 구현되었으며 Visual Studio 6.0 환경에서 Release 모드로 컴파일 되었다. 실험을 위한 데이터는 해상도가 640x360의 컬러 영상으로 다양한 실내외 환경에서 취득되었으며, 배경제거 알고리즘의 적용은 320x180으로 다운 샘플링하여 사용되었다. 모든 실험 데이터에는 동일한 파라미터가 적용되었는데, 샘플링은 3종류의 구간으로 하였으며($K=3$), 샘플링 주기는 각각 1초, 6초 및 20초로 설정하였다. 또, 구간별 배경 모델에서 샘플 버퍼의 크기 N 은 10으로 동일하게 설정하였다.

Fig. 1은 다양한 속도의 객체들을 포함하고 있는 “다산광

장” 데이터의 일부 프레임을 보여주고 있는데, 즉, 대부분의 사람들이 정상 속도로 걷고 있는데 비해 배낭을 멘 한 사람은 매우 느린 속도로 주위를 배회하고 있다(Loitering person 영역). 이 데이터에서 한 가지 더 주목할 부분은 아래쪽의 건물 그림자가 시간이 지남에 따라 점점 커진다는 것이다 (slow moving shadow 영역). 기존 대부분의 배경제거 기법들은 모델 갱신 (또는 샘플링) 속도가 일정하므로 단일 구간 알고리즘이라 볼 수 있고, 이와 같은 데이터에 대해 좋은 결과를 얻기가 어렵다. 예를 들어, SACON의 경우 많은 수의 샘플을 사용하더라도 샘플링 간격이 너무 좁으면 전체 샘플이 짧은 시간 범위의 정보만을 포함하게 되고, 느리게 배회하는 사람이 전경객체로 추출되지 못하는 경우가 발생한다[10]. 반대로 샘플링 간격이 너무 넓으면 너무 오랜 기간의 정보가 모델에 들어가서 해가 이동함에 따른 건물 그림자의 경계부분이 계속 전경 객체로 나타나는 문제점을 발생시킨다. 제안된 방법에서는 Fig. 2에서와 같이 단구간, 중구간 및 장구간 모델을 동시에 운영하여 이런 문제를 해결한다. 즉 (a)와 같이 각 구간별 배경 모델 및 배경 신뢰도를 독립적으로 운영하고, 이들을 종합하여 각 화소별로 최종 배경 화소 값과 신뢰도를 (b)와 같이 결정한다. 단구간 모델에서는 배회하는 사람 영역이 신뢰도가 낮게 나오고 건물 그림자 영역은 신뢰도가 높게 나타난다. 반대로 장구간 모델에서는 긴 시간 범위의 정보를 이용하므로 배회하는 사람 영역은 신뢰도가 높은 반면 건물 그림자 영역의 신뢰도가 낮게 나타난다. 따라서 최종 배경 모델과 신뢰도는 화소별로 가장 신뢰도가 높은 구간의 정보를 (b)와 같이 사용하게 되는데, 즉, 배회하는 사람의 영역에서는 긴 구간의 정보를 반영하고, 건물 그림자 영역에서는 보다 짧은 구간의 특성을 반영하였다. (d)는 현재 프레임 (c)와 배경 영상의 차이를 나타내며 (차이가 없는 경우 밝기가 RGB 각각 128입), (e)는 최종적으로 배경 영상에서 추출한 전경 객체 영상을 보여주고 있는데, 다양한 속도의 객체가 존재하는 경우에도 제안된 알고리즘이 안정적으로 객체를 추출하는 것을 알 수 있다.



Fig. 1. Some frames of "Dasan plaza" data

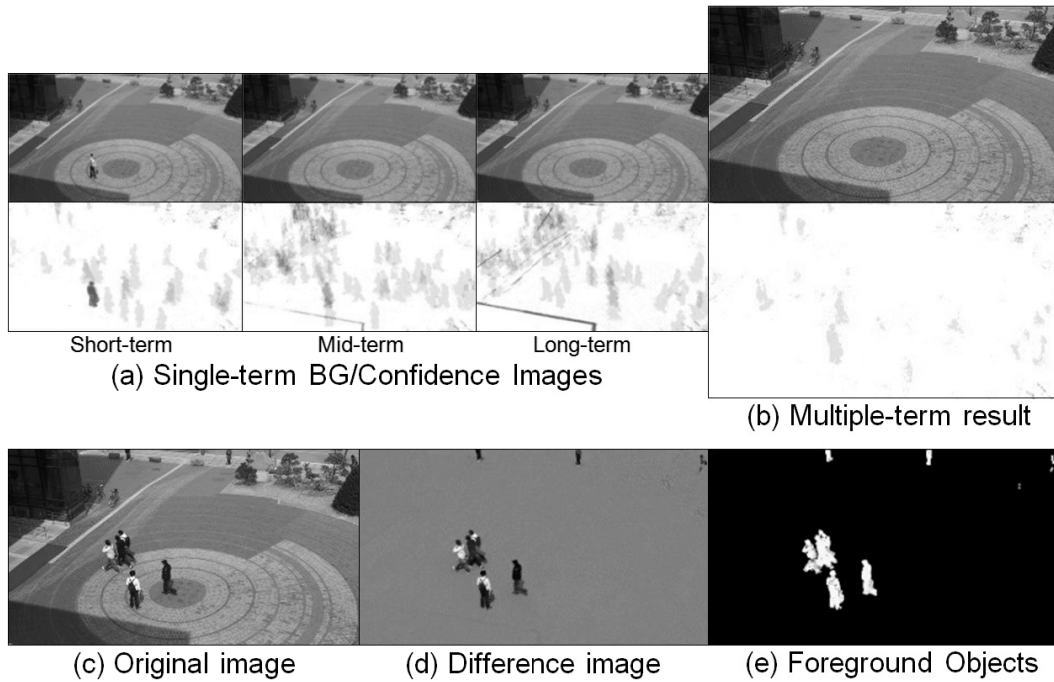


Fig. 2. Background subtraction result of "Dasan plaza" (8794th frame)

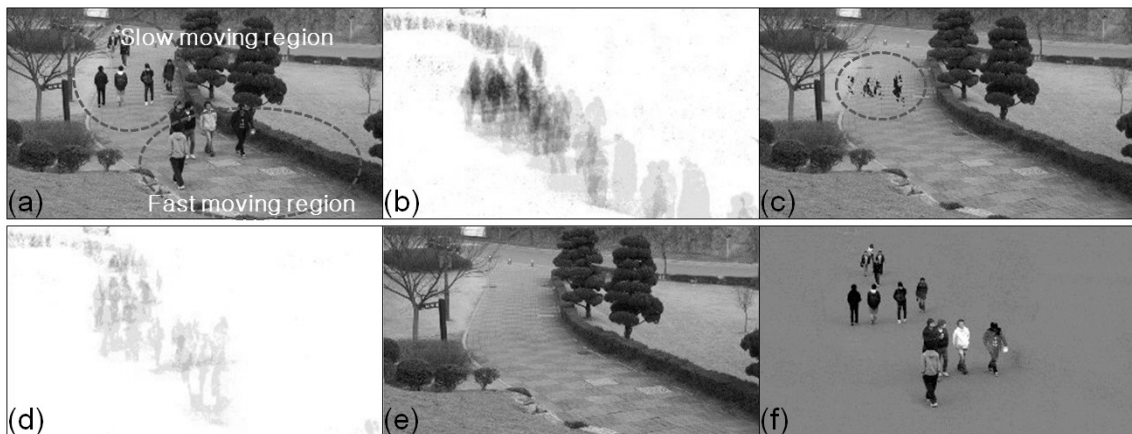


Fig. 3. "Soul-gwan road" data: (a) original image, (b) single-term confidence, (c) single-term BG, (d) multiple-term confidence, (e) multiple-term BG, and (f) multiple-term difference image

비록 객체들이 실제로는 비슷한 속도로 움직인다고 하더라도 카메라의 원근 효과와 지형적인 특성에 의해서 영상내에서 완전히 다른 형태로 나타나는 경우도 많다. Fig. 3은 이러한 전형적인 경우를 나타내고 있는데, 사람들이 걷는 속도가 일정하더라도 그림 (a)의 Slow moving region으로 표시한 부분과 같이 영상의 윗부분이나 카메라의 시선과 도로의 방향이 일치하는 경우 영상에서 화소단위의 이동 속도는 느려진다. 이에 비해 Fast moving region과 같이 영상의 아래쪽이나 굽어진 길에서는 상대적으로 객체의 움직임이 빠르게 나타난다. 이러한 설정의 영상들에서는 단일 구간을 사용하는 경우 (c)의 점선 부분과 같이 바람직하지 않은 배경 영상이 생성될 가능성이 매우 높으며, 이런 영역들은 (b)

의 검은 부분과 같이 배경 신뢰도가 낮다. 제안된 방법에서는 다중 구간 모델을 운영하여 (d)와 같이 화소별로 신뢰도가 가장 높은 모델을 선택하고, 이를 통해 안정적인 배경영상(e)을 생성할 수 있었으며, 결과적으로 안정적인 차영상(f)을 추출하고 배경을 제거할 수 있었다.

Fig. 4는 몇 가지 실내의 데이터에 대한 실험 결과를 보여주고 있는데, 위에서 아래로 원영상과 전경객체 추출 결과영상 및 배경영상을 보여주고 있다. "복지관삼거리"의 경우 다양한 움직임을 잘 검출하면서 건물 그림자의 영향이 객체 영역에 나타나지 않는 것을 볼 수 있다. "주차장"에서도 크기와 속도가 다른 객체들을 잘 추출할 수 있었으며, "졸업연구실 복도"와 같은 실내 환경에서도 제안된 알고리

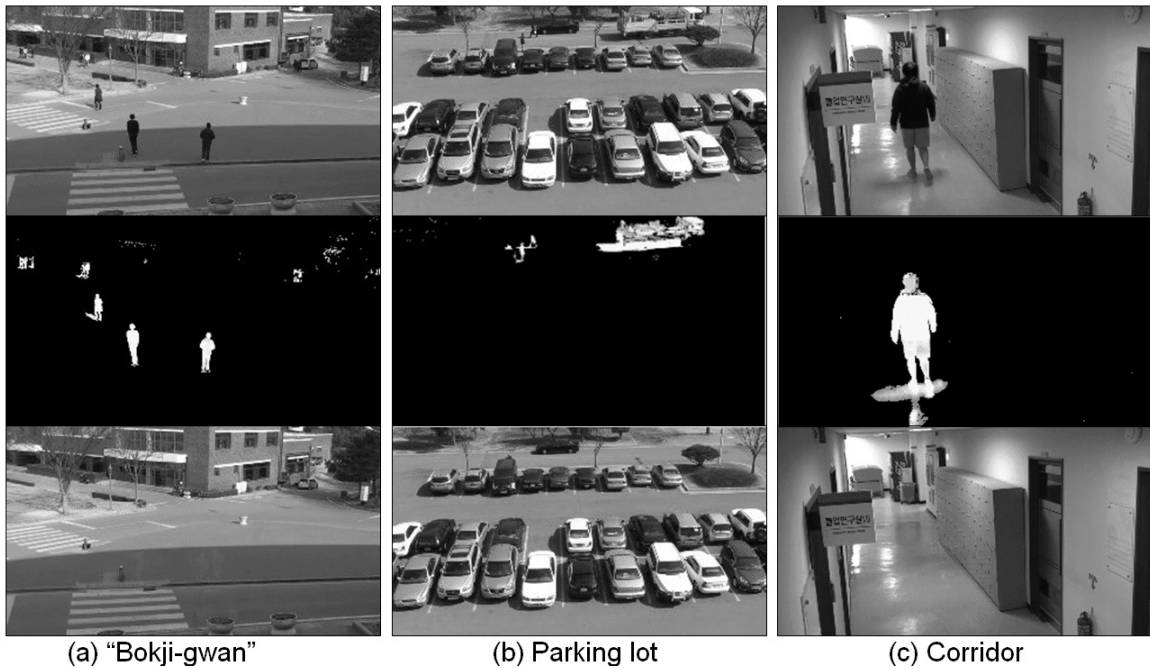


Fig. 4. Background subtraction results: original image, FG image and BG image (from top to bottom)

즘이 안정적으로 객체를 추출하는 것을 알 수 있다. 본 논문은 다양한 이동 속도에 대한 대응에 초점을 맞추었는데, 배경제거에서는 이 외에도 해결해야 할 많은 문제들이 있다. “복지관삼거리”에서와 같은 동적 배경(바람에 의한 나뭇가지 움직임)에 안정적으로 대응하는 것과, “주차장” 및 “졸업연구실 복도”에서와 같은 그림자 제거 문제 등에 대해서도 활발하게 연구되고 다양한 방법들이 보고되고 있다. 향후 연구에서는 제안된 알고리즘을 이러한 문제에 효과적으로 대응할 수 있도록 확장할 예정이다.

샘플링 기반 배경제거에서는 처리시간이 샘플의 개수에 가장 큰 영향을 받는다. SACON의 경우 200개 내외의 샘플을 사용하였는데, 이에 따라 매우 낮은 처리속도를 보고하고 있다 (1.6GHz Pentium PC, 160x120 color영상에 대해 6 frames/sec) [10]. ViBe에서는 새로운 샘플을 버퍼에 넣는 과정에서 가장 오래된 샘플을 없애는 것이 아니라 무작위로 기존 샘플을 선택해 제거하는 방법을 사용하는데, 이에 따라 잡음 영상에 강하고 보다 적은 수의 샘플 (20개 이상)로도 좋은 결과를 보인다고 보고하고 있다. 제안된 방법에서는 버퍼당 샘플의 수를 10으로 설정하고 3가지의 샘플링 구간을 사용하였으므로 전체 샘플의 개수는 30개이다. 그러나 처리시간에 주로 영향을 주는 것은 단구간 버퍼의 샘플 10개가 되는데, 이것은 중, 장구간의 경우 샘플 버퍼의 갱신이 단구간에 비해 월등히 적게 이루어지므로 처리시간에 큰 영향을 미치지 않기 때문이다. 실험을 통해 제안된 방법은 320x180 컬러 영상에 대해 초당 240 이상의 처리 성능을 보였다. ViBe의 경우 2.67GHz Core i7 CPU, 6GB RAM, C 언어 구현으로 640x480 영상에 대해 180장의 처리가 가능하다고 보고하고 있지만 [11], 이러한 실제 처리시간에 대한 주

장은 실험 시스템과 알고리즘 코드의 최적화 및 병렬화 등에 따라 크게 달라질 수 있다. 제안된 방법은 샘플의 수 ($N=10+a$)와 알고리즘에 따른 이론적인 측면에서 ViBe에 비해서도 처리시간 측면에서 우월하다고 볼 수 있다.

5. 결 론

본 논문에서는 샘플링 기반의 새로운 배경제거 알고리즘을 소개하였다. 제안된 방법은 화소의 배경 모델로 확률 밀도함수를 사용하지 않고 이전 프레임에서 샘플링 된 화소값을 사용하는데, 특히 샘플링 구간의 길이가 다른 다중 구간 모델을 지원할 수 있도록 하였으며, 신뢰도를 바탕으로 하여 다중 구간 모델에서 최적의 모델을 찾아 전경 객체 추출에 적용함으로써 움직임이 빠른 객체와 느린 객체에 모두 안정적으로 동작할 수 있도록 하였다. 제안된 방법에서는 배경 모델과 함께 배경 영상도 만들어지는데, 따라서 전경 객체의 추출 뿐 아니라 비디오 개요 시스템과 같이 배경 영상이 반드시 필요한 응용에도 활용될 수 있다. 실험을 통해 제안된 방법이 다양한 속도의 객체가 존재하고 시간에 따른 점진적인 환경 변화가 있는 다양한 데이터에서도 안정적인 결과를 나타내는 것을 알 수 있었다.

참 고 문 헌

[1] T. H. Cho and Y. K. Choi, “Motion Detection Using Multiple Distributions for Background,” The KIPS Transactions, Vol.8, No.4, pp.381-389, 2001.

[2] N. McFarlane and C. Schofield, "Segmentation and Tracking of Piglets in Images," Machine Vision Application, Vol.8, pp.187-193, 1995.

[3] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: Real-Time Tracking of the Human Body," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.19, No.7, pp.780-785, July, 1997.

[4] C. Stauffer and W. Grimson, "Adaptive Background Mixture Models for Real-Time Tracking," Proc. IEEE CS Conf. Computer Vision and Pattern Recognition, Vol.2, pp.246-252, 1999.

[5] A. Elgammal, R. Duraiswami, D. Harwood, and L.S. Davis, "Background and Foreground Modeling Using Nonparametric Kernel Density Estimation for Visual Surveillance," Proc. IEEE, Vol.90, No.7, pp.1151-1163, 2002.

[6] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in Proceedings of the International Conference on Pattern Recognition, pp.28-31, 2004.

[7] A. Mittal and D. Huttenlocher, "Scene modeling for wide area surveillance and image synthesis," In CVPR, 2000.

[8] O. Javed, K. Shafique, and M. Shah, "A hierarchical approach to robust background subtraction using color and gradient information," In in IEEE Workshop on Motion and Video Computing, pp.22-27, 2002.

[9] M. Heikkila and M. Pietikainen, "A Texture-Based Method for Modeling the Background and Detecting Moving Objects," IEEE Trans. on PAMI, Vol.28, No.4, pp.657-662, April, 2006.

[10] H. Wang and D. Suter, "A consensus-based method for tracking: Modelling background scenario and foreground appearance," Pattern Recognition, Vol.40, No.3, pp.1091-1105, 2007.

[11] O. Barnich and M. Van Droogenbroeck, "ViBe: A universal background subtraction algorithm for video sequences," In IEEE Transactions on Image Processing, 20(6):1709-1724, June, 2011.

[12] Y. Pritch, A. Rav-Acha and S. Peleg, "Non-Chronological Video Synopsis and Indexing," IEEE Trans. on PAMI, Vol.30, pp.1971-1984, 2008.



이 동 은

e-mail : delee@lgcns.com
 2005년 한서대학교 컴퓨터학과(학사)
 2007년 중앙대학교 첨단영상대학원
 영상공학과(공학석사)
 2008년~현 재 LG CNS 정보기술연구원
 선임연구원

관심분야: 영상처리, 객체추적, 워터마킹 등



최 영 규

e-mail : ykchoi@koreatech.ac.kr
 1989년 경북대학교 전자공학과(학사)
 1991년 KAIST 전기및전자공학과
 (공학석사)
 1995년 KAIST 전기및전자공학과
 (공학박사)

1995년~1999년 LG산전연구소 선임연구원
 1999년~현 재 한국기술교육대학교 컴퓨터공학부 교수
 관심분야: 영상처리, 머신비전, 컴퓨터그래픽스 등