

베이시안 모델링 물체 검출에 관한 초당 프레임 처리량 유지 기법

신수광*, 윤희용*

*성균관대학교 정보통신공학부

e-mail : shinsukwang@naver.com, youn@ece.skku.ac.kr

A Method Sustaining Frame Process Rate on Object Detection of Bayesian Modeling

Su-Kwang Shin*, Hee-Yong Youn*

*School of Information and Communication Engineering, SungKyunKwan University

요 약

사생활 보호에 대한 인식이 커지고, 인터넷 시대에 접어들면서 네트워크 기반의 보안시스템의 개발이 활발하다. 실시간 비디오 카메라를 통한 움직이는 물체를 검출하기 위해서는 불필요한 잡음이나 조명의 변화에 대처해야 한다. 이러한 많은 요소들을 고려하여 움직이는 물체를 검출하려면 많은 계산 복잡도를 가지게 된다. 또한, 카메라의 영상크기가 증가함에 따라 움직이는 물체를 검출하기 위해서 더 많은 계산 복잡도를 가지게 된다. 본 논문에서는 기존의 통상적인 움직임 검출방법과 적응적 배경방식인 ‘물체 검출을 위한 동적인 장면의 베이시안 모델링 기반 물체 검출 방법’을 분석하고, 실시간으로 처리되는 동적 비디오 영상에서 이동 물체를 검출하는 과정에서의 영상의 크기가 커지고, 이동하는 물체의 개수가 많아짐에 따라 발생하는 계산의 복잡도를 ‘CPU 성능과 영상 resize 를 이용한 계산 복잡도 감소 방법’을 통해 초당 프레임 처리속도를 유지시키는 방법을 제시한다.

1. 서론

사생활 보호에 대한 인식이 커지고, 인터넷 시대에 접어들면서 네트워크 기반의 보안시스템의 개발이 활발하다. 원격지의 카메라로부터 전송된 영상을 통하여 현재의 상황을 파악할 수 있으며, 적절한 조치를 웹을 통해 취할 수 있다. 이러한 웹 멀티미디어 보안 시스템은 활용영역이 점점 확대되면서 영상의 해상도와 전송속도, 그리고 보안시스템의 핵심인 객체영역 인식, 영상 정보의 처리, 저장, 검색 기술 등의 연구가 추가적으로 요구되고 있다 [1].

이동 물체의 검출은 배경으로부터 움직임 물체를 분할함을 의미한다. 일반적으로 움직임 분할에는 두 가지 방식이 있다. 첫 번째 방식은 연속되는 두 프레임간의 차분을 취함으로써 물체를 분할하는 프레임간 차분 방법이다. 두 번째 방식은 움직이는 물체가 없는 배경 영상과의 차분을 취함으로써 이동 물체를 분할하는 배경 영상 방법이다. 카메라가 고정된 조건에서의 영상 감시 응용으로는 후자의 배경 영상 방식이 보다 효율적이다.

베이시안 모델링 기반 물체 검출 방법은 적응적 배경 방식으로 이러한 문제를 해결하기 위한 수시로 상황에 맞게 배경이 변화되기 때문에, 그만큼 계산량이 늘어나게 된다. 일반적인 QVGA(320 * 240), VGA(640 * 480)에서는 이러한 계산량이 실시간 처리를 할 수 있을 정도가 되지만, 영상의 크기가 두 배로 늘어 날수

록 계산량은 4배 이상 복잡하게 되어 실시간 처리라고 여겨지지 않을 초당 5 프레임 이하의 처리량을 보일 것이다.

본 논문은 이러한 문제점에 대해서 CPU 성능과 영상 resize 를 이용한 계산 복잡도 감소 방법’을 통해 초당 프레임 처리속도를 유지시키는 방법을 제시한다. 영상의 크기가 커짐에 따라 그리고 영상 안의 이동하는 물체가 많이 짐에 따라 계산 복잡도는 증가 하게 된다. 따라서 이 두 가지 경우가 발생 될 때 CPU 자원소모량을 분석하여, 일정 임계값보다 CPU 자원소모량이 커지게 될 때 프레임의 크기를 작게 Resize 하는 방식으로 계산복잡도가 증가되는 것을 막을 수 있다. 이 방법은 초당 프레임 처리량이 일정하게 유지 하기 때문에 좀 더 정확한 물체 감지를 할 수 있고, 실시간 처리에 알맞다.

본 논문의 구성은 다음과 같다. 2 장에서 기본 움직임 물체에 검출에 관한 차영상을 이용하는 방법과 제안 방법인 물체 검출을 위한 베이시안 모델링을 분석하고, 3 장에서는 베이시안 모델링 기반한 물체 검출 방법의 문제점을 제시하면서 초당 프레임 처리량을 유지시키는 방법을 제안한다. 끝으로 4 장에서 논문의 결론을 맺는다.

2. 움직임 정보 검출 방법

움직임 정보는 프레임 단위의 움직임과 각 픽셀에

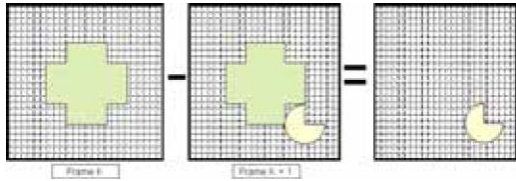
서의 움직임을 추정하여 움직임 벡터를 전송한 다음 프레임의 움직임 보상을 행한다. 그러나 각 픽셀에서의 움직임 정보는 프레임 움직임만큼 보상된 이전 프레임과 현재 프레임과의 차이가 큰지 작은지를 나타내는 움직임 검출 정보로서 표현된다.

움직임을 검출하기 위한 방법에는 차영상을 이용한 방법, 적응적 배경 방법 등이 있다.

2.1 차영상을 이용한 방법

움직임이 있는 후보영역을 검출하기 위해 연속된 두 프레임간의 차영상 분석 방법을 사용한다. 그림 1[2]의 차영상 분석 방법은 연속된 두 프레임간의 밝기 차이를 구한 후, 임계값을 사용하여 임계값보다 낮은 밝기 차이를 가진 부분은 움직임이 없는 배경으로 구별하고 임계값보다 큰 밝기 차이를 가진 부분은 움직임이 있는 물체로써 구별한다[3].

이러한 임계값의 선택은 획득한 영상 안의 잡음뿐만 아니라 시간에 따라 변하는 조명에 상당히 의존적이다. 따라서 움직임이 있는 후보영역들을 검출하기 위한 임계값은 유동적으로 선택되어야 한다. 또한 물체가 정지해 있는 상황에서는 배경으로 인식될 수 있다.



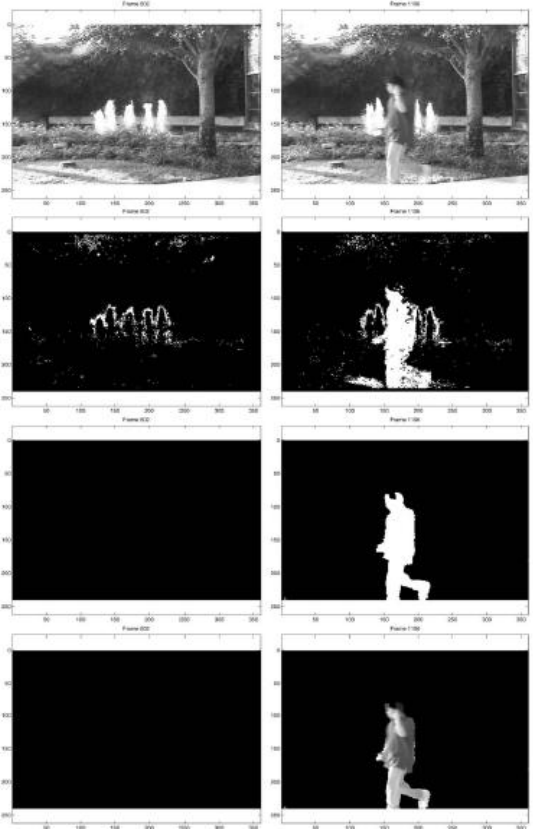
(그림 1) 차영상을 이용한 방법

2.2 물체 검출을 위한 동적인 장면의 베이시안 모델링

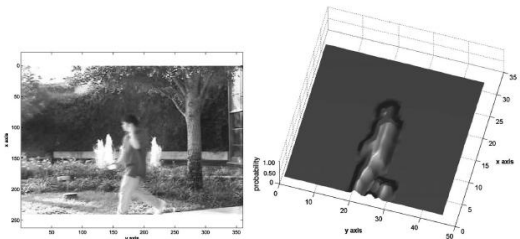
지난 1970년 이래로, 고정식 카메라에서 인접한 프레임의 차영상을 이용하여 물체 검출 방법으로 사용되어왔다[4]. 그러나 차영상을 이용한 방법이 실시간 처리에서는 부적절하여, 통계적인 기술이 배경 픽셀 값을 모델링하기 위해서 도입되었다.

물체 검출을 위한 동적인 장면의 베이시안 모델링은 적응적 배경 영상기법을 전제로 3 가지의 참신한 방법을 가지고 있다. 첫 번째 이 방법은 관찰된 밀도의 공간적인 종속성을 모델링하는 원칙적인 수단을 공급한다. 독립적인 임의의 변수를 사용한 이미지 픽셀들의 모델은 변하게 되고 공간적으로 인접한 픽셀들에는 유용한 구조가 존재하게 된다. 이 구조는 인접 지역 범위 표현보다도 비모수 밀도 측정(nonparametric density estimation - 데이터에 근거하여 Parametric assumptions 의 가정을 하지 않고 밀도를 직접적으로 추정하는 방법, Histogram 을 사용하여 측정하는 것이 가장 단순)을 사용함으로써, 배경 데이터가 단일 분포로서 모델링 되고, 다양한 공간적인 불확실

성들을 직접 다루게 될 수 있다. 두 번째, 이전의 접근과 달리 전경 부분에서 추적 정보를 사용할 필요 없이 물체 검출을 증대시키기 위해 명확히 설계된다. 일시적인 지속성의 기준은 배경차이의 전통적인 기준과 동시에 사용되는 것으로 제안된다. 세 번째 집합이 될 가능성에 직접 임계값을 적용시키는 대신, 물체 검출을 위해서 전경과 배경 설계화에 경쟁적으로 사용하는 MAP-MRF framework 를 제안한다[5].



(그림 2) 분수가 일정시간 후에 배경이 되는 현상



(그림 3) 분수가 보이는 영상에 사람만 검출

그림2와 그림 3은 물체 검출을 위한 동적인 장면의

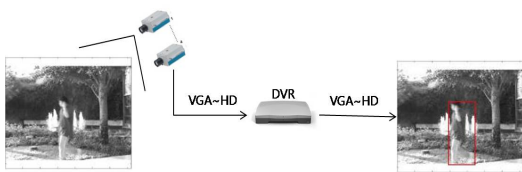
베이지안 모델링의 장점을 보여준다. 적응적 배경방법 중에 MAP-MRF Framework를 사용하여 반복적으로 움직이는 작은 요소들은 배경으로 저장되고, 임계값보다 큰 값들은 전경으로 설계하여 실제 움직이는 대상만 검출할 수 있게 되었다.

3. 고정 프레임 처리량 유지 기법

본문에서 소개한 차영상을 이용하는 방법은 영상안의 잡음, 시간에 따른 조명 변화, 외부환경(나무, 분수 등)에 대처하지 못한다. 현재는 적응적 배경 방법으로 많이 사용되고 있고, 따라서 본문에서 제시한 베이지안 모델링을 사용하여 3.06 GHz Intel Pentium 4 Processor, 1GB RAM, 360*240(QVGA)의 영상크기 환경에서 테스트 하였다[5]. 이 때 초당 11프레임을 처리하였고, 이때 움직이는 물체는 하나로 한정되어 있다. (카메라의 영상 초당 30 프레임으로 동작).

따라서 360*240(QVGA) 영상크기보다 표준 카메라 크기인 640*480(VGA)으로 테스트 한다면 계산 복잡도는 4배 이상 증가하여 초당 30 프레임으로 영상을 받는 카메라에서 영상을 받아, 초당 5~6 프레임의 영상만 베이지안 모델링 하는 것에 쓰이게 될 것이다. 또한 테스트 시에 움직이는 물체는 하나로 한정되어 있지만, 실제 실시간 처리를 요구하는 환경에서는 무수히 많은 물체를 화면에서 검출해 내어야 한다. 많은 물체를 각각 검출하기 위해서는 그만큼 계산 복잡도가 증가 될 것이다.

초당 5~6 프레임만을 처리하게 되면, 빠른 속도로 화면에서 이동하는 물체에 관해서는 물체가 t 프레임과 t+1 프레임에서의 이동간격이 커질 것이고, 추후에 물체의 추적을 하기가 어려워 지게 된다. 또한 화면에 잠깐 나타났다가 사라지는 물체에 관해서는 물체 검출을 못 할 경우도 생기게 된다.



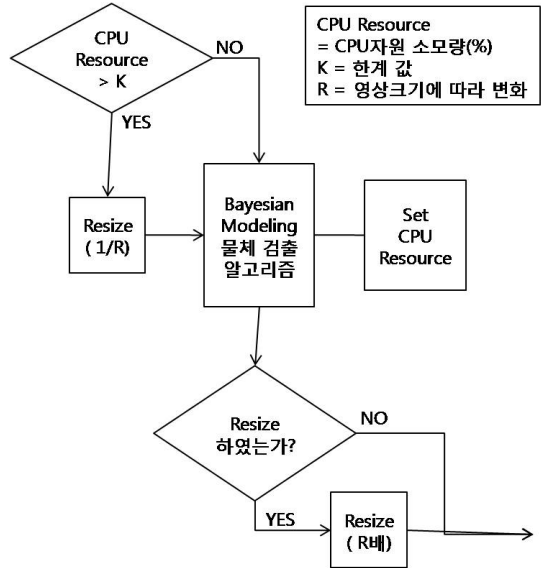
(그림 4) 카메라의 처리 과정

그림 4는 일반 카메라 감시에 관련된 처리 흐름이다. 카메라에서 나오는 영상 (VGA-HD)을 DVR(Digital Video Recorder)에서 기록하고, 움직이는 물체를 검출하여 검출된 물체에 대한 정보와 카메라에서 받은 영상을 모니터에 출력하게 된다.

이러한 시스템에서 영상사이즈가 커지고, 움직이는 물체가 많아짐에 따른 계산 복잡도의 증가는 실시간 처리에서 초당 프레임 처리량을 감소하게 만드는 가장 중요한 요소가 된다.

본 논문에서는 복잡도에 상관없는 고정 프레임 처리량을 유지하기 위해 CPU성능과 영상 resize를 이용하여 계산 복잡도를 감소시키는 방법을 제안한다.

그림 4에서 DVR에서 움직이는 물체를 검출하는 알고리즘을 사용한다는 가정하에 본 논문에서 제시하는 방법에 대한 자세한 처리과정을 그림 5에 표현하였다.



(그림 5) DVR에서의 처리과정

그림 5의 처리 과정은 다음과 같다. 움직이는 물체를 검출하기 위해서, 실시간으로 카메라로부터 영상을 받아 온다. 적응적 배경 영상에서는 보통 초기 배경 영상을 만들기 위한 Learning Time이 존재한다[5]. 이러한 Learning Time을 처리할 때 초기 CPU Resource를 설정한다. 설정된 CPU Resource는 Learning Time이 끝나고 초기 배경 영상이 만들어 지면, Learning Time 동안에 만들어진 배경 영상과 베이지안 모델링 물체 검출 알고리즘을 이용하여 실시간으로 들어오는 프레임에 대해서 비교 분석과정을 한다.

CPU Resource의 값은 크게 두 가지 요소에 의해 바뀌게 된다. 하나의 프레임의 크기이고, 또 다른 하나는 프레임 안에 변화되는 픽셀의 수이다. 프레임의 크기가 320*240 영상에서 640*480영상으로 2배 커지게 됨에 따라 CPU Resource는 4배 이상의 자원을 소모하게 될 것이다. 또한 프레임 안에서 움직이는 물체가 많아 변화된 픽셀들의 수가 많아지면 베이지안 모델링 알고리즘의 복잡도가 커지기 때문에 CPU Resource가 커지게 될 것이다.

기준 값 K를 중심으로 CPU Resource가 K보다 크면 물체검출 알고리즘 수행 시 한계값 보다 CPU 자원 소모량이 많다는 것을 의미하고, 이것은 효율적인 초당 프레임 처리가 이루어 질 수 없다는 것을 뜻한다. 반대로 CPU Resource가 K값 보다 작으면, 이것은 충분한 CPU 자원이 있기 때문에 효율적인 초당 프레임 처리가 이루어 질 수 있다는 것을 의미한다.

CPU Resource가 K값보다 크면, 카메라에서 받은 프레임을 $1/R$ Resize한다. 사이즈를 $1/R$ 만큼 작게 하여, 베이시안 모델링 기반한 알고리즘에 원하는 초당 프레임 처리량이 이루어 질수록 한다. R값은 영상 크기에 따라 달라 질 수 있다. 640*480(VGA)의 영상에서는 일반적으로 R값을 2로 정하여 320*240의 영상으로 변환하고, 움직이는 물체를 검출하는 베이시안 모델링 기반 알고리즘을 동작 시킨다. 영상의 크기가 커지면 커질수록 R값도 비례하여 증가 한다. 또한 영상 크기가 아닌 베이시안 모델링 기반한 알고리즘 동작 중 움직이는 물체가 다수 발생되어 CPU Resource가 증가한 경우에도, 다음 프레임부터 R값을 증가하여 원하는 초당 프레임 처리량이 이루어 질수록 한다.

적절한 R 값을 이용하여 프레임을 resize 하여 베이시안 모델링 기반한 알고리즘을 통해서 움직이는 물체를 검출하였다면, 검출된 물체에 관한 정보를 다시 원래 프레임에 맞게 resize 시켜야 한다. 따라서 R 값을 사용하여 resize 하였다면 다시 물체에 관련된 픽셀 정보에 관해서는 다시 resize 하여 원래 프레임에 맞게 화면에 출력 할 수 있도록 정보를 변환 시켜준다.

4. 결론

본 논문에서는 물체 검출에 관한 방법에서 문제점에 관한 해결책을 다루었다. 베이시안 모델링에 관한 적응적 배경 방법은 많은 알고리즘들이 응용되어 사용되어 있기에, 영상크기와 움직이는 물체의 수가 증가함에 따라 계산 복잡도가 증가하여 초당 프레임 처리량이 감소 하게 된다. 초당 프레임 처리량이 감소하게 되면, 물체 검출 이후의 추적이나 관련된 이벤트에 대한 신뢰성을 잃게 만들 수 있고, 정확한 물체 검출을 하지 못하게 된다.

이런 문제점을 해결하기 위해 본 논문에서는 CPU 자원 소모량을 확인하여, 자원 소모량이 일정 한계값보다 크면 영상 프레임을 Resize 하는 방식으로 영상크기와 움직이는 물체의 수의 증가에 적절히 대처하여, 초당 프레임 처리량을 유지 시키는 해결 방안을 제시하였다.

이 방법의 장점은 초당 프레임 처리량이 일정하게 유지 되기 때문에, 실시간 처리에 알맞다. 또한 추후에 하드웨어가 바뀌어도 CPU 에 따라 능동적으로 대처하기 때문에 전체 알고리즘을 바꾸지 않아도 된다.

향후 연구과제로 CPU 자원소모량과 프레임 처리량과의 상관관계에 대한 정확한 임계값을 분석하고, 움직이는 물체가 많아짐에 따른 CPU 자원소모량과의 상관관계 또한 분석할 필요가 있다.

참고문헌

- [1] H.Zang et al., "Automatic Partitioning of Full-motion Video", Multimedia System, Vol. 1, No. 1, pp. 10-28, 1993.
- [2] 이창수 외, "적응적 배경영상과 그물형 픽셀 간격의 윤곽점 검출을 이용한 객체의 움직임 검출", 한국통신학회논문지 '05-3 Vol.30 No.3C, 2005

- [3] J.L. Starck, F. Murtagh, E.J. Candes, D.L.Dnoho, "Gray and Color image contrast enhancement by curvelet transform", IEEE Transactions on Image Processing, Vol. 12, No. 06. Pp. 0706-0717, JUNE 2003
- [4] R. Jain and H. Nagel, "On the Analysis of Accumulative Difference Pictures from Image Sequences of Real World Scenes", IEEE Trans. Pattern Analysis and Machine Intelligence, 1979.
- [5] Yaser Sheikh et al., "Bayesian Modeling of Dynamic Scenes for Object Detection", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 27, No.11, November 2005