

영상 합성과 에지 영상을 이용한 다수의 보행자 검출 방법

최양진⁰ 강희중

한성대학교 컴퓨터공학과

{pointor⁰, hjkang}@hansung.ac.kr

Detection of Multiple Pedestrians Using Image Composition and Edge Image

YangJin Choi⁰ and Hee-Joong Kang

Department of Computer Engineering, Hansung University

요 약

동영상 내에서 이동하는 객체를 추적하기 위해서는 우수한 객체 검출 방법이 필요하다. 이를 위하여, 본 논문에서는 연속된 영상에서의 인접한 프레임들을 이용하여 객체의 형태를 검출하고자 한다. 인접한 프레임들의 합성과 차를 이용하여 움직이는 객체의 대략적인 형태를 알아내고, 대략적인 형태를 이진화시킨 영상과 현재 프레임의 에지 영상과의 AND 연산을 통하여 객체의 형태를 알아 낼 수 있다. 그리고, 이 과정에서 생성되는 노이즈를 채움 연산과 영역화 연산을 통하여 제거할 수 있으며, 얻어진 객체의 크기 비율을 고려한 수직 투영을 통하여 다중 객체를 잘 분리해 낼 수 있었다.

1. 서론

비전시스템의 목적은 컴퓨터를 기반으로 스캐너, 카메라 등의 영상 감지 장치를 이용해서 입력영상을 얻고, 얻은 입력으로부터 주어진 문제에 대한 지식과 영상처리에 관한 일반적인 정보를 활용하여 유용한 정보를 검출하는데 있다[1]. 오늘날 비전시스템은 하드웨어와 기술의 발전으로 인하여 여러 응용 분야에 적용되고 있다. 그 예로서 비전시스템으로 검출할 수 있는 정보를 이용하여 범죄의 예방에 필수적인 감시 시스템 및 보안 시스템이 사용되고 있으며, 영상 회의 시스템 및 유비쿼터스 환경에서의 중요한 요소로도 활용될 수 있다. 이러한 시스템에서 움직이는 객체의 검출은 필수 요소이며 중요한 문제가 될 수 있다. 본 논문에서는 움직이는 객체를 검출하기 위한 방법으로 이전 영상과 현재 영상의 차이를 이용할 때 발생할 수 있는 객체 형태의 불확실성 문제점을 제거하기 위한 비교적 간단한 방법을 제안하고자 한다.

본 논문의 구성은 2장에서는 관련 연구를 소개하고, 3장에서는 본 논문에서 제안하는 객체의 검출방법을 소개하며, 4장에서는 제안한 방법의 실험 및 기존 방법[4]과의 비교를 하고, 5장에서는 결론 및 향후 연구과제를 기술한다.

2. 관련 연구

객체를 검출 및 추적하기 위해서 여러 가지 방법이 연구되어 왔다. 이들은 블록 정합 방법[3], 광류를 이용한 방법[2], 영상의 차이를 이용하는 방법[4]등 크게 세 가지로 분류할 수 있다. 블록 정합 방법은 이미 알고 있는 객체와 같은 것을 찾는 것이므로 간단한 것 같지만, 연

속되는 자연계의 영상으로부터 일정한 특징을 얻기가 어렵고, 일정한 특징을 얻는다 하더라도, 정합을 위한 시간이 많이 걸리게 된다[5]. 또한 광류를 이용한 방법은 많은 계산량이 요구되므로 실시간으로 처리하기에는 부적합하다[5]. 영상의 차이를 이용하는 방법은 그 방법이 비교적 쉽고 계산량이 많지 않아, 실시간으로 처리해야 하는 많은 시스템에서 이용하고 있다. 영상 차이를 이용하는 방법은 크게 참조영상(배경영상)을 이용하는 방법과 연속되는 영상 속에서 인접한 영상을 이용하는 방법이 있다[6]. 참조영상을 이용하는 방법은 갑작스런 날씨의 변화나, 조도 등에 영향을 많이 받으므로 참조영상의 업데이트가 필요하며, 연속되는 영상 속에서 인접한 영상을 이용하는 방법은 이러한 업데이트는 필요 없지만, 움직이는 객체의 형태를 얻기가 힘든 단점이 있다. 연속된 영상에서 인접한 프레임들을 이용하는 기존의 방법 중에는 차영상과 수평, 수직 투영을 통해서 객체를 분리하는 방법[4]이 있었으나, 차영상에서 객체의 윤곽선이 정확히 검출이 안된다면 수평, 수직 투영을 했을 때 객체의 형태를 제대로 얻을 수 없는 단점이 있을 수 있다. 따라서, 본 논문에서는 이전 영상을 이용하는 방법에서 객체의 형태를 정확하게 검출하기 위한 방법을 제안하고자 한다.

3. 이동 객체의 검출

본 논문에서 이동하는 객체를 검출하는 과정은 그림 1과 같이 크게 두 단계의 절차를 거친다. 이들 절차를 개략적으로 소개하면 다음과 같다. 먼저 첫 번째로 인접한 두 컬러 영상으로부터 변환된 흑백 영상과 에지 영상을 이용하여 객체의 대략적인 형태를 알아내고, 두 번째로

는 채움 연산(closing), 영역화(labeling), 수직 투영(vertical projection) 등을 통하여 좀 더 세밀한 형태와 노이즈 제거 및 다중 객체의 분할을 하게 된다.

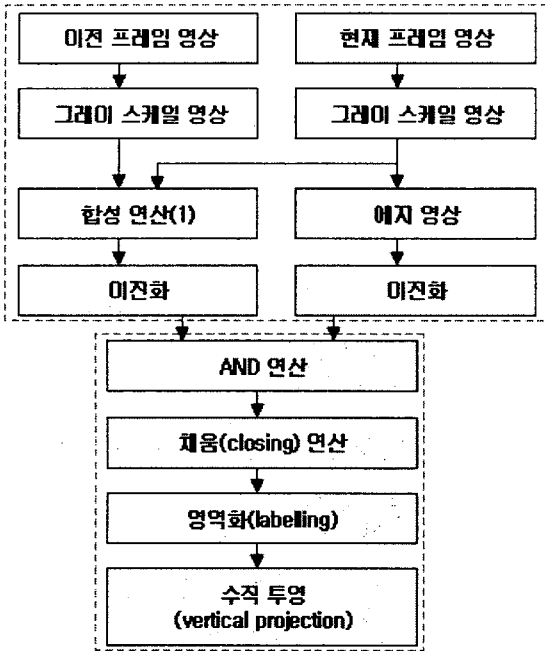


그림 1. 이동하는 객체의 검출 과정

계산량의 감소를 위해서 연속된 영상에서의 인접한 두 영상을 그레이 스케일 영상으로 변환한다. 이 결과로 나온 인접 영상을 식(1)과 같이 합성하여 처리하게 된다.

$$(F(x, y, t-1) + F(x, y, t)) / n - F(x, y, t-1) \quad (1)$$

여기서 $F(x, y, t-1)$, $F(x, y, t)$ 은 $(t-1)$ 과 t 시간에서의 프레임이고, x, y 는 프레임 안에서 픽셀의 위치 정보이고, n 은 합성한 프레임의 개수이다. 하지만 여기서는 두 프레임을 합성하므로 $n=2$ 가 된다. 두 프레임을 합성함과 동시에 n 으로 나누어 주고, 이를 다시 $(t-1)$ 시간의 영상으로 빼주는 것은 변화없는 부분을 제거하는 효과를 가질 수 있다. 이렇게 해서 얻은 프레임을 이진화시킨 영상과 에지 영상을 이진화시킨 후, AND 연산을 취하면 객체의 대략적인 형태를 얻을 수 있다. 이것을 식으로 표현하면 식(2)와 같다.

$$\text{Composed}F(x, y) \text{ AND } \text{Edge}F(x, y, t) \quad (2)$$

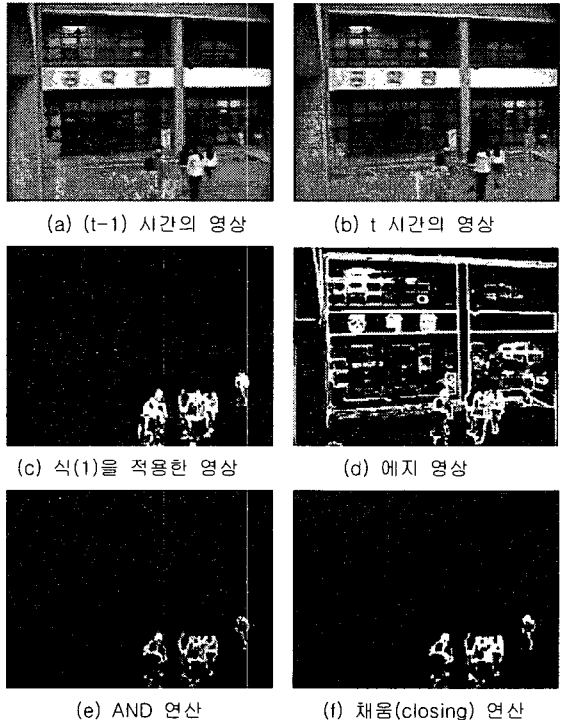
여기서, $\text{Composed}F(x, y)$ 는 식(1)을 이용해서 이진화시킨 프레임이고, $\text{Edge}F(x, y, t)$ 는 에지 영상이다. 그림 2는 제안한 방법을 수행했을 때의 결과를 보인 것으로 그림 2(a)와 (b)는 비교 대상이 되는 영상이다. 그리고 그림 2(c)는 두 영상을 식(1)과 같이 합성한 결과를 보여준다. 그림 2(d)는 t 시간의 에지 영상을 이진화시킨

영상이다. 에지 연산은 윤곽선이 강하게 나타나도록 하기 위해서 Canny 마스크를 사용하여 연산을 수행하였다. 대부분의 검출 마스크는 잡음에 대하여 매우 민감한 특성을 가지고 있어서 작은 잡음도 윤곽선으로 검출할 경우가 많이 있다. 그에 반해 Canny 마스크 방법은 영상의 잡음에 민감하지 않은 윤곽선 검출 방법이다[7].

그리고 그림 2(e)와 같이 AND 연산을 수행한다. AND 연산을 수행하는 이유는 합성된 영상에서 에지 영상을 이용하여 현재 영상의 객체를 분리하기 위함이다. 결과 영상에서 보듯이 합성과 에지 영상의 AND 연산으로 어느 정도의 노이즈는 제거가 된다. 하지만 완전한 노이즈 제거가 이루어 지지 않았기 때문에 그림 2(f)와 같이 채움(closing) 연산을 하게 된다. 채움 연산은 팽창 연산과 축소 연산을 수행하는 것으로 확장이 먼저 일어나므로 작은 구멍들은 메워 지고, 축소 작용 때에 원 영상의 크기로 복원된다[4]. 따라서, 채움 연산은 노이즈 제거와 더불어 하나의 객체임에도 불구하고 모양의 차이에 의해서 두 개로 분리된 객체를 하나로 연결할 수 있게 한다. 이렇게 얻어진 영상에 그림 2(g)와 같이 영역화(labeling)를 함으로써 객체에 대한 영역을 검출할 수 있게 된다. 이때 영역화는 라벨링 과정을 통하여 각각의 물체를 구성하는 개개의 픽셀들을 하나의 영역으로 묶어 주게 된다[8].

이렇게 얻어진 영상에서도 여러 명의 보행자가 이동할 때 하나의 객체로 보기 때문에 검출된 객체의 넓이와 높이의 비를 이용하여 단일 객체인지 또는 다중 객체인지를 분간할 수 있다[4]. 이렇게 함으로써 알아낸 다중 객체를 그림 2(h)와 같이 수직 투영을 통하여 분리해 낼 수 있다.

이와 같은 방법으로 연산을 하였을 경우, 그림 2(i)와 같이 최종적으로 객체를 검출해 내게 된다.



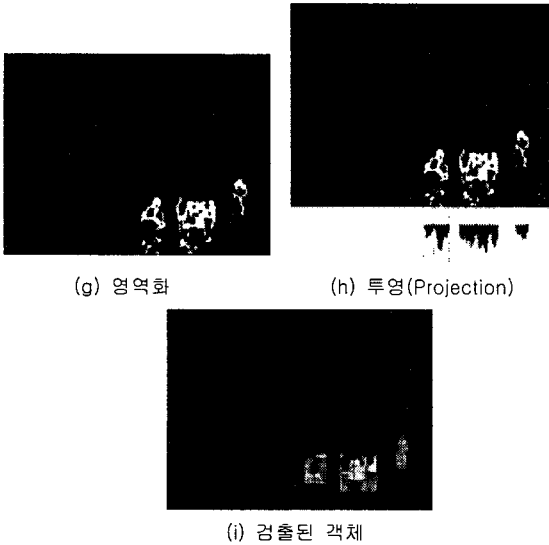


그림 2. 제안한 방법의 단계별 수행 예

4. 실험

학교 내의 건물 앞에서 이동하는 보행자를 대상으로 실험하였고, 네트워크로 연결되어 있는 비디오 서버로부터 320x240 크기의 영상을 초당 4 프레임씩 입력으로 받아들여서 펜티엄4 2.8G의 1.5G의 램이 장착된 컴퓨터에서 실험을 하였다.

그림 3은 기존의 방법[4]과의 차이를 보여주는 그림이다. 그림 3(a),(b),(e),(f)는 (t-1), t 시간에서의 영상이다. 그림 3(c)와 (g)는 두 영상의 차이를 보여주는 그림이다. 이와 같이 객체가 빨리 이동하게 되면 차영상만을 이용했을 때는 움직임의 거리가 크기 때문에 영상 간의 차이가 크게 나타난다. 따라서 이전 프레임에서 검출된 객체의 넓이까지 같이 검출되는 문제점이 있다. 하지만 그림 3(d)와 (h)와 같이 본 논문에서 제안한 방법으로 수행했을 때 객체를 잘 검출하는 것을 알 수 있다.

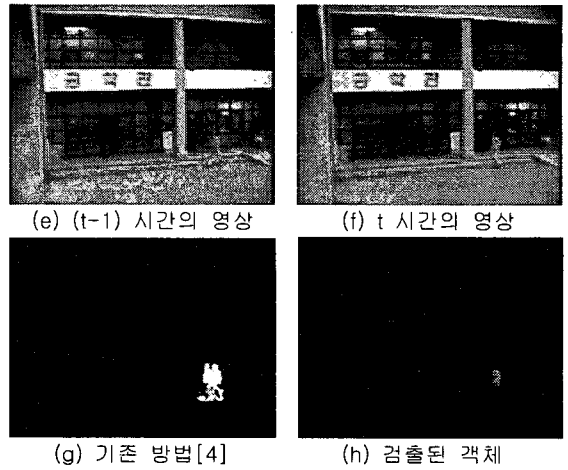
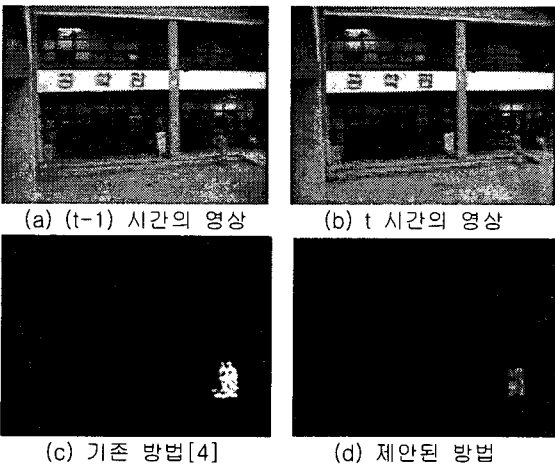


그림 3. 기존 방법과 제안 방법과의 비교

4. 결론 및 향후 연구과제

제안한 방법으로 실험한 결과, 움직임 검출이 잘되었지만 두 명 이상의 보행자가 밀접하게 붙어있을 경우 수평, 수직 투영만으로 보행자를 검출하여 분리하기 힘들었다. 이를 보완하기 위해서 투영과 동시에 객체의 색깔 정보를 이용해서 보행자의 분리가 필요하고, 또한 그림 2와 3에서 움직임이 없거나 배경색과 비슷한 옷을 입었을 경우 검출이 잘 안되는 결과를 보였다. 이는 식(1)의 현재 영상의 차와 예지 영상에서의 배경 영상과 영도의 차이가 거의 없기 때문인데, 이를 보완하는 방법이 필요하다.

5. 참고 문헌

[1] 김상훈, "다중색상 정규화와 움직임 색상정보를 이용한 객체 검출", 한국정보처리학회 논문지, Vol.12 No. 7 pp. 721- 728, 2005 . 12
 [2] S. M. Smith "ASSET-2: Real-Time Motion Segmentation and Object Tracking", Real-Time Imaging Vol.4 Issue 1 pp.295-304, 1999
 [3] Hariharakrishnan, D. Schonfeld, "Fast Object Tracking Using Adaptive Block Matching", IEEE Trans on Multimedia, Vol.7 No.5 pp.853-859, 2005 . 10
 [4] 임종석, 김옥현, "움직임 정보를 이용한 다수 보행자 추적", 한국정보처리학회 춘계학술대회논문집 제1권, 제1호, pp. 755-758, 2002
 [5] 김용현, "광류를 이용한 이동객체의 영역화", 광운대학교 대학원, 석사학위논문, 1994
 [6] 김진, "영역 기반의 이동 객체 검출 및 움직임 추정", 서강대학교 대학원, 석사학위논문, 1995
 [7] 장동혁, "디지털 영상처리의 구현", 정보게이트, 2002
 [8] 강동중, 하준은, "Visual C++ 을 이용한 디지털 영상처리", 사이텍미디어, 2003