

고속의 세미오토매틱 비디오객체 추적 알고리즘

(A Fast Semiautomatic Video Object Tracking Algorithm)

이 종 원*, 김 진 상**, 조 원 경***

Lee Jong Won, Kim Jinsang, Cho Won-Kyung

Abstract - Semantic video object extraction is important for tracking meaningful objects in video and object-based video coding. We propose a fast semiautomatic video object extraction algorithm which combines a watershed segmentation schemes and chamfer distance transform. Initial object boundaries in the first frame are defined by a human before the tracking, and fast video object tracking can be achieved by tracking only motion-detected regions in a video frame. Experimental results shows that the boundaries of tracking video object are close to real video object boundaries and the proposed algorithm is promising in terms of speed.

Key Words : object tracking, bounding box, watershed, chamfer distance transform

1. 서론

의미있는 비디오객체 추출은 비디오 객체추적과 객체기반 동영상 압축 응용분야에 필수적이다. 객체 단위로 영상을 압축하는 MPEG-4의 연구가 활발해지면서, MPEG-4의 입력으로 쓰일 객체를 추출하는 연구가 시작되었다. MPEG-4는 블록단위의 압축 방식인 MPEG-1, 2와 달리 객체 기반 압축 기법을 사용하는데 객체 기반 압축이란 2차원 평면상에 표현되어 있는 영상에서 사람이 영상을 인식 하듯이 전경과 배경을 구분하고, 각 영역에 맞게 적절한 압축방식을 도입하여, 블록단위 압축의 블록킹 효과를 없애고 압축효과와 화질 향상을 높인다는 취지에서 개발되었다 [1].

객체 기반 압축 방식을 사용하면 동영상 서비스를 제공하는 입장에서는 전경과 배경에 각각 다른 압축 기법을 적용하여, 주로 관심의 대상이 되는 전경에서는 압축률을 낮춰서 선명한 영상을 얻고, 관심 밖의 배경에는 압축률을 높여서 압축 효율을 증가시킬 수가 있다. 반면 서비스를 제공받는 사용자의 경우 자신이 원하는 물체만을 얻어서 사용할 수 있고, 합성이나 편집에 물체를 자유롭게 이용할 수 있다는 이점이 있다. 따라서 객체 기반 압축에서 무엇보다 중요한 기법은 자동적으로 얼마나 인간이 느끼는 것과 같이 물체를 찾아내는가 하는 것이다. 일반적인 자연 영상의 경우 똑같은 영상이 거의 없고, 또한 객체와 배경을 완벽하게 따로 분리할 수 있는 특징을 찾기 어렵기 때문에 의미 있는 객체만을 자동화된 방법으로 구분하기란 쉽지 않다. 특별히 지식 기반 방식을 사용하지 않는 한 2차원 영상 자체의 특성을 최대한 이용해

야 하는데, 이때 사용되는 특징으로는 명도(luminance), 컬러(color), 에지(edge), 질감(texture), 모션(motion) 등 여러 가지가 있다. 하지만 픽셀단위로 이루어진 한 프레임의 영상에 위와 같은 특징들을 이용해서, 객체와 배경을 분리하는 과정에는 상당히 많은 시간과 계산이 필요하다. 동영상의 경우 많은 프레임으로 이루어져 있는데 객체를 추출하기 위해 모든 프레임에 이와 같이 복잡한 과정을 똑같이 적용하는 것은 비효율적이라고 할 수 있다. 따라서, 첫 프레임에서 객체를 추출하고, 그 추출된 객체에 대한 정보를 이용해서 그 이후 프레임에 대해서는 좀더 간단한 방법들을 사용함으로써 추출된 객체를 계속해서 추적하는 방법에 관한 연구가 필요해졌다 [2-5]. 객체 추적 응용분야로서는 감시카메라, 화상회의, 화상 전화기, 휴대 단말기 등을 들 수 있다.

본 논문에서는 워터셰드 알고리즘(watershed algorithm)과 챔퍼변환(chamfer distance transform)을 적용한 고속 비디오 객체추적 알고리즘을 제안하며, 고속의 객체추적을 위하여 움직임이 있는 영역에 대해서만 객체추적을 수행한다. 첫 번째 프레임의 비디오 객체는 사람에게 의하여 초기화되며 이후 프레임에 대해서는 자동적인 객체추적이 수행된다.

본 논문의 2장에서는 고속의 객체 추적 알고리즘을 제안한다. 3장에서 실험결과를 논의하며 4장에서는 결론을 맺는다.

2. 고속의 객체 추적 알고리즘

2.1 움직임이 감지된 영역에 대한 Bounding Box 설정

그림 1은 제안된 알고리즘의 흐름도이다. 연속된 두 프레임 $I(n)$, $I(n+1)$ 이 들어 올 경우 두 프레임의 차영상을 구함으로써 두 프레임 사이에서 객체의 움직임에 의해 변화된 부분을 찾을 수 있다. 그러나 단순한 차영상의 경우, 윤곽선의 변화 외에도 객체 내부의 변화 또한 나타나게 된다. 하지만 우

저자 소개

* 이 종 원 : 경희대학교 전자공학과 석사과정

** 김 진 상 : 경희대학교 전자공학과 교수

***조 원 경 : 경희대학교 전자공학과 교수

리가 관심이 있는 부분은 객체의 내부에서의 변화가 아니라, 객체의 윤곽선에서의 변화가 있는지 없는지에 초점을 맞추고 있기 때문에 탐지된 객체 내부의 변화는 제거해 주어야 한다. 그러기 위해서는 앞에서 변화가 있다고 탐지된 픽셀들 중에 어떤 픽셀들이 객체 윤곽선 변화에 의해 탐지된 픽셀들인지, 어떤 픽셀들이 객체 내부 변화에 의해 탐지된 픽셀들인지를 판단해야 한다. 그것을 판단하기 위해 이전 프레임으로부터 추출된 객체의 윤곽선정보 ($O(n-1)$)를 이용하여 Chamfer Distance Transform 을 적용하였다 [3].

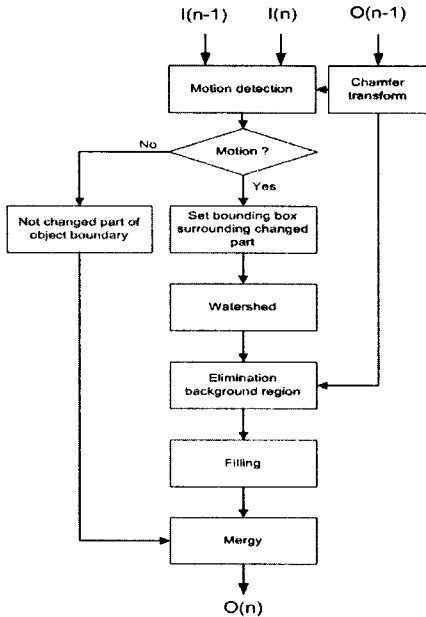


그림 1. 객체 추적 알고리즘

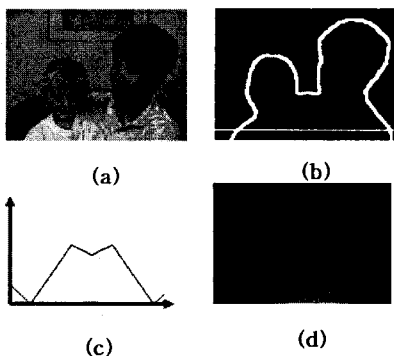


그림 2. (a) Mother and daughter 이미지 (b) 초기 객체의 윤곽선 (c) (b)에 가로로 횡단하는 직선에 대한 Chamfer distance transform의 그래프 (d) (b)에 Chamfer distance transform을 적용한 이미지

Chamfer Distance transform의 결과는 이미지내의 픽셀들과 이전 프레임 객체 윤곽선사이의 거리를 나타낸다. 그러므로 변화가 있었다고 판단되는 픽셀들 중에 윤곽선으로부터 일정거리 이상의 픽셀들은 내부변화에 의해 탐지된 픽셀로 간주하고 제거해 준다. Mother&daughter 동영상에 대한 초기 윤곽선은 그림 2(b)와 같으며 이를 Chamfer Distance transform 한 결과는 그림 2(d)이다. 그림 2(c)는 그림 2(b)의 한 수평방향에 대한 Chamfer Distance transform의 결과를 그래프로 나타낸 것이다. 그림 2(c)와 같이 Chamfer Distance transform을 수행하면 윤곽선이 있는 부분의 결과값은 0에 접근함을 알 수 있다.

2.2 Bounding Box 내부의 객체 추출

앞 절에서 두 프레임 사이에서 객체의 윤곽선 변화가 있다고 추정되는 영역을 설정하였다. 제안된 알고리즘은 배경과 객체를 분리하기 위해 이 영역에 대해서만 Watershed 알고리즘을 적용한다.

Watershed 알고리즘[5]은 영상처리 분야에서 인접한 영역을 효과적으로 군집화 하는 과정을 통해 영상을 분할하는 목적으로 사용되어지는 기법이다. 처음 Lantuejoul 와 Beucher에 의해 소개되었고, Vincent 와 Soille에 의한 방법과 Mayer에 의한 방법으로 대표된다 [5].

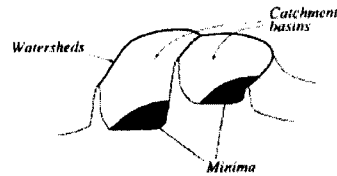


그림 3. Minima, Catchment basin and Watershed

Watershed 알고리즘은 4단계로 구성된다.

- 1단계, 원 영상으로부터 고도를 가진 경사 영상(gradient image)으로 변환한다.
- 2단계, 경사영상으로부터 지역적 낮은고도 (local minima)를 가진 픽셀들을 추출한다 (그림 3 참조).
- 3단계, 지역적으로 낮은 고도로부터 높은 고도로 범람과정(flooding)을 수행하여 유사한 영역을 분할한다.
- 4단계, 유사한 영역에 대한 영역병합(region merging)을 수행한다.

여기서 Bounding box 내에서만 Watershed 를 적용한 후에 Bounding box 내부의 배경에 대한 Watershed region 만을 추출하기 위해, 앞에서 추출했던 Chamfer distance transform 의 정보를 이용하여 이전 프레임 객체 윤곽선으로부터 일정 거리 이상의 떨어져 있으며, 이전 객체의 외부에 있는 Watershed region(배경의 Watershed region) 들을 선택하여 제거해 준다. 이를 통하여 남아있는 영역들은 Bounding box 내에 존재하는 현재 프레임의 객체를 추출할 수 있다.

2.3 Merging & Filling

지금까지 두 프레임 사이에서 변화가 있다고 추정되는 영역(Bounding box)을 설정하고, 그 영역 내부에 속한 객체의 일부분을 추출하였다. 여기서 추출된 Bounding box 내부의 객체의 윤곽선은 이전 프레임의 윤곽선 정보를 이용하여 갱신한 현재 프레임에서의 객체의 윤곽선이다. 이 갱신된 윤곽선과 이전 프레임에서 변화가 없었던 객체의 윤곽선을 통합해서 현재 프레임의 객체의 윤곽선을 만든다. 이로써 우리는 완전한 객체를 추출할 수 있다.

여기서 추출된 객체는 이전 단계의 배경 영역(Watershed region)만을 추출하는 과정과 갱신된 객체의 윤곽선과 변화가 없었던 이전 프레임의 객체 윤곽선을 합치는 과정에서 잘못 실제 객체의 일부가 함께 제거 될 수 있고, 현재 프레임의 객체 내부에 빈 공간이 존재 할 수 있기 때문에 이 부분을 채워줌으로써 우리가 목표한 현재 프레임의 최종 객체를 추출 할 수 있다.

3. 실험 결과

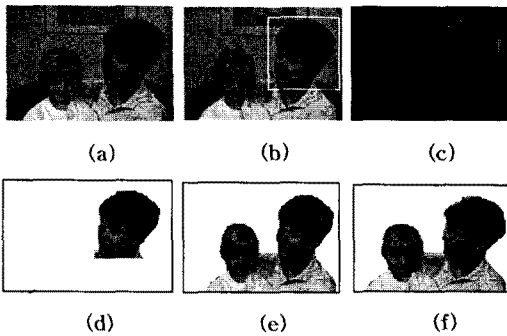


그림 4. (a) Mother and daughter 이미지(frame 26) (b) Bounding box 설정 (c) Bounding box 내부에 Watershed 적용 (d) Bounding box 내부에 객체추출 (e) Merging & Filling(frame 27) (f) 객체추적(frame 45)

표 1. 처리시간 비교(PentiumIII 1 GHz 사용)

	mother & daughter (sec)	salseman (sec)	image size
bounding box 적용	0.217	0.151	176*144 (qcif)
bounding box 미적용	0.770	0.843	176*144 (qcif)

그림 4. (a)는 Mother and daughter의 26번째 프레임의 나 타낸다. Mother and daughter의 26번째 프레임과 27번째 프레임사이에서 변화가 있다고 추정되는 영역을 설정하고 그 영역을 27번째 프레임에 설정한 모습을 그림 4.(b)에서 볼 수

있다. 영역 설정 후에 설정된 영역 내에서만 Watershed를 적용 한 영상을 그림4.(c)에서 볼 수 있으며, 그림4.(d)는 영역 내에서의 추출된 객체이다. 영역 내에서 추출된 객체와 이전 프레임의 변화가 없는 객체 윤곽선을 합친 후 Filling 작업을 거친 후 최종적으로 추출된 27번째의 객체를 (e)에 나타냈으며, (f)는 45 프레임까지 객체의 윤곽선이 성공적으로 추적된 모습을 나타낸다. 또한 표1에서 bounding box를 적용한 것과 적용하지 않은 경우 한 프레임을 처리하는데 소요된 시간을 나타냈다. 실험결과와 같이 제안된 알고리즘을 이용하면 추출된 객체의 윤곽이 실제 윤곽에 가깝고 bounding box를 적용하면 처리속도가 최대 6배 정도 향상됨을 알 수 있다.

4. 결론

본 논문에서는 워터셰드 알고리즘(watershed algorithm)과 챔퍼변환(chamfer distance transform)을 적용한 고속 비디오 객체추적 알고리즘을 제안하며, 고속의 객체추적을 위하여 움직임이 있는 영역에 대해서만 객체추적을 수행한다. 첫 번째 프레임의 비디오 객체는 사람에 의하여 초기화되며 이후 프레임에 대해서는 자동적인 객체추적이 수행된다. 실험 결과, 제안된 알고리즘은 실제 윤곽에 가까운 객체를 추출 할 수 있으며 bounding box를 적용하면 처리속도가 최대 6배 정도 향상됨을 알 수 있다.

감사의 글

본 연구는 경희대학교 지원(20040077)으로 수행되었습니다.

참 고 문 헌

- [1] 후지와라 히로시, "그림으로 보는 최신 MPEG", (주) 교보문고, 1995.
- [2] Jinsang Kim and Tom Chen, "A VLSI Architecture for Video Object Segmentation," IEEE Trans. on Circuits and Systems for Video Technology, vol. 13, no.1, pp. 83-96, Jan. 2003.
- [3] G. Borgfors, "Distance transformations in digital images," *Comput.Vis., Graph, Image Processing*, vol. 34, pp. 344-371, 1986.
- [4] A. Murat Tekalp, "Video object tracking with feedback of performance measures", IEEE Transactions on, Vol.13, pp. 310-324, 2003
- [5] L.E. Band. " Topographic partition of watersheds with digital elevation model," *Water Resources Res.*, Vol. 22, no. 1, pp. 15-24, 1986

저 자 소 개



이 중 원(李鍾源)

2003년 2월 : 경희대학교 전자공학과 (공학사)

2003년 2월~현재 : 경희대학교 전자공학과 석사과정

관심분야: 영상처리, VLSI 시스템 설계



김 진 상(金鎭想)

1985년 2월 : 경희대학교 전자공학과(공학사)

1987년 2월 : 경희대학교 전자공과(공학석사)

2000년 2월 : 콜로라도 주립대 전기 및 컴퓨터 공학과 (공학박사)

1990년 5월 ~ 2001년 8월: 한국통신연구소 선임연구원 1996년 7월 ~ 2000년 7월: 콜로라도주립대 및 휴랙팩 카드

2001년 9월~ 현재: 경희대학교 전자정보학부 조교수

관심분야: multimedia signal processing and VLSI system design for arithmetic units, and wireless and consumer electronics applications



조 원 경(趙源敬)

1971년 2월 : 경희대학교 전자공학과(공학사)

1973년 2월 : 한양대학교 전자공학과 (공학석사)

1986년 8월 : 한양대학교 전자공학과 (공학박사)

1978년 ~ 1980년:경남대학교 전자공학과 전임강사

1980년 ~ 현재 : 경희대학교 전자공학과 교수

1988년 ~ 1989 : 미국 오레곤 주립대학 교환교수

관심분야: computer system architecture and VLSI system design for arithmetic units, computer vision, and wireless communications