

내용 기반 MPEG-4 를 위한 비디오 분할 기법 연구

김준기, 이호석
호서대학교 컴퓨터 공학과

A novel video segmentation approach for content-based MPEG-4

Jun-Ki Kim, Ho Suk Lee
Department of Computer Engineering, Hoseo University

요 약

본 논문은 MPEG-4 내용 기반 비디오 코딩을 위한 객체 추출기법을 소개한다. MPEG-4 표준화는 비디오 객체에 대한 접근성, 사용자와 객체의 상호작용, 높은 압축률을 위한 비디오 코딩 알고리즘등을 요구한다. 비디오 장면에서 "객체"에 대한 접근과 조작성 지원하기 위해서는 우선적으로 배경화면으로부터 움직이는 실제 객체를 분리하여야 한다. 즉 비디오는 부호화 처리 이전에 VOP(Video Object Plane) 로 구분되어야 한다. 본 논문은 이미지 시퀀스로부터 움직이는 객체를 분리하기 위하여 제안된 여러 방법들과 알고리즘을 소개하며 새롭게 제안한 방법을 사용하여 효율적인 VOP 추출 과정을 기술한다.

1. 서론

새로운 비디오 코딩 표준인 MPEG-4 는 내용 기반 부호화(content-based video coding)기능을 사용한다. 내용 기반 부호화 기능은 이전 표준인 MPEG-1, MPEG-2 와 뚜렷이 구별되는 코딩 기법중에 하나이다. MPEG-4 는 한 장면에서 영상과 오디오 데이터를 고정된 배경, 비디오 객체, 오디오 객체, 음악 등으로 분리한다. 즉 객체 기반 부호화를 사용함으로써 각 객체의 비트스트림(bitstream)에 대한 용이한 접근과 객체 기반 비트스트림의 조작, 객체와의 상호작용, 장면 안에 포함된 객체의 재사용성등의 많은 기능을 지원한다.

내용 기반 상호작용 기능을 사용하기 위해 MPEG-4 는 입력된 비디오 시퀀스의 각 프레임을 부호화전에 몇몇 객체로 분리하여 각 장면을 다룬다. 즉 VOP(Video Object Plane)로 불러오는 의미있는 객체를 우선적으로 분리하여야 한다. 비디오 객체 분리의 주요 목적은 각 지역안에 명암도나 색, 구성, 움직임의 동일한 특징을 갖는 영상 객체를 집단화 하는데 있다.

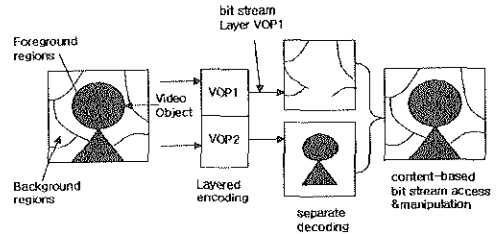
MPEG-4 비디오 부분의 처리에 대하여서는 여러 기관과 연구소에서 비디오 시퀀스에서 움직이는 객체에 대한 자동 분리 알고리즘을 개발하였다.[1][2][3] MPEG-4 표준 버전 1. 에서 권장하는 객체 분할 구조는 시간적 분할, 공간적 분할, 시-공간적 분할 결합의 세 부분으로 나누어져 있다. 시간적 분할이라는 것은 영상안에 있는 객체의 움직임 위치를

추출하는 작업이고 공간적 분할은 주어진 기준에 따라서 정확한 객체의 경계와 영상안에 있는 의미있는 객체의 지역을 추출하는 작업이다. 이 두 가지의 작업을 통합하여 마지막으로 영상으로부터 각 객체를 추출하는 것이다.

본 논문은 비디오 객체 추출을 위한 통합된 접근 방법을 제안하기 위하여 이미 제안된 알고리즘을 소개하고 공간적, 시간적 특성을 이용하여 정확한 영상의 경계를 찾고 객체를 추출하는 접근 방법을 설계하였다.

2. 관련 연구

영상에서 움직임의 특징과 시간의 흐름에 따른 방법을 사용하여 객체를 추출하는 다양한 접근 방법이 개발되었다. 객체를 추출을 위한 영상 구조와 흐름은 <그림 1> 과 같다.



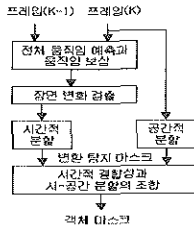
<그림 1> VOP 추출 구조

2.1 현재의 VOP 추출 방법

최근 VOP 추출 방법에 대하여 많은 제안이 발표되고 있

다.[3][4][5][6]

MPEG-4 비디오 부분에 적용된 시-공간 분할 단계는 <그림 2>와 같다.



<그림 2> 시-공간 분할 단계

참고문헌[4]은 시-공간적 정보를 기반으로 객체를 추출한다.

1) 이미지 시퀀스에서 움직이는 객체의 위치를 측정하기 위하여, 시간적 방향으로 연속된 두 이미지 차이의 조사와 가정 테스트인 두 연속적인 다른 이미지로부터 두 이미지 편차를 측정 비교 함으로서 이루어진다. 2) 공간적 분할은 움직이는 객체의 정확한 경계를 발견하고 의미있는 지역으로 각각 이미지를 나누어 수행되어진다. 시간적 분할은 움직임이 없는 지역으로부터 움직인 객체를 변환 탐지 마스크로 산출한다. 마지막으로 시-공간적 조합을 통하여 VOP 추출을 수행한다. 시간적 분할 방법은 명암도 차이와 편차 비교를 기반으로 통계적 가정 테스트를 포함하여 객체를 추출하고 공간적 분할은 watershed 검출 기반 지역 성장 알고리즘(region growing algorithm) 과 수리 형태학 필터(morphological filter)를 사용한다. Watershed 알고리즘은 분할 중복의 문제점이 발생하므로 지역 병합 연산을 필요로 하는 단점이 있다. 움직인 객체의 위치측정을 위해 전처리 과정을 수행하는데 제안된 논문에서는 1) 전체적 움직임 추정과 보상 2) 장면 변화 검출 알고리즘을 사용한다. 움직임 추정 기법은 6 파라미터 affine 움직임 모델을 사용한다. 장면 변화 검출은 장면 변화를 추출하기 위해 MAD(Mean Absolute Difference)와 SDMAD(Sign Difference MAD) 방법을 사용한다. 공간적 분할단계는 이미지 표준화, 기원기 근사치, watershed 검출, 지역 병합으로 이루어진다. 이미지 표준화는 수리 형태학 필터의 기능을 사용한다. 이것은 주어진 영상 요소에 기하학적인 특징은 보존하면서 불필요한 요소들을 제거 함으로서 영상을 단순화 시키고 동시에 유용한 기하학 정보를 추출하는 뛰어난 기능을 가지고 있기 때문이다.

참고문헌[5]은 움직임 정보를 기반으로 배경(background)으로부터 실제 객체(foreground)를 분리한다. 전체적인 알고리즘은 움직임 객체 추출, 모델 초기화, 모델 업데이트, VOP 추출의 4 단계 함수로 구성되어있다. 움직임 추출을 위한 방법으로는 수리 형태학 움직임 필터(morphological motion filter)와 변환 탐지 마스크(change

detection mask) 기법을 사용한다[5]. 객체 추적을 위한 매칭 표준은 Hausdorff distance 방법을 사용한다. 이것은 계산적 효율과 모양 변화와 잡음에 강력한 특징을 가진다. 움직임 추출을 위해 두 가지 방법을 사용한다. 수리 형태학 움직임 필터는 이미지의 다른 부분은 정확히 보존하면서 전체 움직임에 따르지 않는 부분을 제거하는 기능을 가지고 있다. 이 움직임 추출 기법은 작은 움직임이나 정확한 움직임에 사용되어지고 다른 방법으로 변화 탐지 마스크는 보다 빠른 움직임에 적용한다. 변화 탐지 마스크의 사용은 배경이 움직일 경우 문제가 발생한다. 그 이유는 배경의 움직임에 대하여 정확한 표시가 불가능 하기 때문이다. 그래서 변환 탐지 마스크는 고정된 배경을 기반으로 제한 되어 사용된다. 모델 초기화는 처음 객체의 위치와 모델의 존재를 발견하기 위하여 사용되어진다. 모델 업데이트는 다음 프레임의 존재에 따라서 수행되어진다. VOP 추출은 객체 경계를 결정하기 위하여 2 단계로 이루어진다. 처음 단계는 filling-in technique 을 사용하고 차후 boundary post-processor 사용하여 VOP 를 추출 한다.

참고문헌[6]에서는 joint marker 추출, 경계 결정, 움직임 기반 지역 융합 방법을 사용하였다. Joint marker 추출은 휘도와 움직임 사이에서 동일한 지역의 존재를 나타낸다. 다음으로 시-공간적 경계는 watershed 검출을 사용하여 결정되어진다. 마지막 단계로 움직임 기반 지역 융합을 사용하여 중복된 지역을 제거 함으로서 객체를 추출한다.

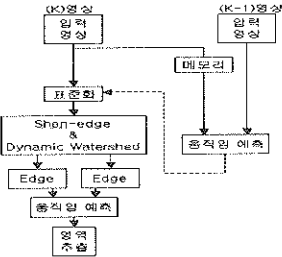
3. 제안된 방법

대부분의 비디오 객체 추출 알고리즘은 배경으로부터 객체를 분리하기 위하여 비디오 시퀀스에서 시간과 공간의 결합성을 이용하였다[1][2]. 본 논문에서는 새로운 제안을 위해 객체 추출에 필요한 일반적인 사항을 기술한다. 움직임은 객체 추출의 핵심이다. 영상에서 동적인 객체를 추출하기 위해서는 먼저 영상의 움직임 특성, 명암도, 물체, 구조, 크기, 형태의 변화를 발견하고 움직임의 속도에 따라 움직임 특성을 결정하여야 한다. 객체를 추출하기 위한 고려사항으로는 객체와 배경이 단순히 고정되어 있는 경우, 배경은 고정되어 있고 객체가 움직이는 경우, 배경에는 움직임이 있고 객체가 고정되어 있는 경우, 가장 일반적인 영상으로 배경과 객체가 움직이는 경우를 고려 하여야 한다.

제안된 방법은 2 단계로 구성되어 있다. 처음 단계는 이미지 표준화, Shen 에지 검출과 동적 watershed 알고리즘을 사용하여 객체의 윤곽을 설정한다. 두 번째 단계로는 객체 추출을 위하여 전 탐색 블록매칭 알고리즘을 사용하여 객체의 움직임 범위를 측정한다. 움직임 범위는 MSD(Mean Squared Difference)나 MAD(Mean Absolute Difference)를 사용하여 객체의 벡터를 구함과 동시에 다음 프레임의 객체를 비교하여 실질적인 객체를 추출 하는 단계로 진행되어진다.

3.1 분할 블록도

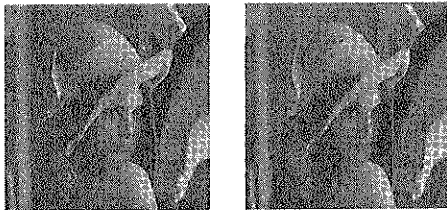
영상 분할의 전체구조는 <그림 3> 와 같다.



<그림 3> 분할 알고리즘 구조

3.2 공간적 분할

영상 이미지에서 정확한 객체의 경계를 찾기 위하여 우선 영상의 잡음을 제거하여야 한다. 잡음 제거 알고리즘은 median 필터를 사용하여 객체 이외의 잡음을 제거 한다. 다음으로 객체의 정확한 윤곽과 지역을 설정하기 위하여 Shen 에지 검출과 동시에 watershed 알고리즘을 사용하여 객체의 윤곽을 결정한다. Median 필터의 처리 결과는 <그림 4>와 같다.



(a) 원 영상 (b) 필터 적용 영상

이후 적용된 영상의 객체 윤곽 설정을 위해 Shen 에지 알고리즘을 적용한다.



(c) Shen 에지 검출결과 (d) watershed 검출결과

<그림 4> 공간적 영상 분할

watershed 알고리즘은 gradient 영상 픽셀의 휘도값을 기술기 정보로 생각하고 gradient 영상에서 위치 최소치에 연관된 동일(homogeneous)영역을 분할하는 알고리즘이다.

3.3 시간적 분할

영상의 시간적 분할이란 입력되는 프레임 사이에 객체의 위치를 측정하데 가장 중요한 움직임이 포함되어 있다. 시간적 프레임에서 객체를 추출하기 위하여 프레임 사이의 픽셀, 예지, 지역, 명암도 사이의 변화를 탐지한다. 즉 두 픽처사이의 차이를 구하여 객체의 움직임을 파악한다. 두 객체사이의 움직임 측정을 위해서 변환 탐지 마스크 방법을 사용한다. 이 방법에서는 두 영상 사이에 휘도 구성요소의 MAD 값을 구하여 움직임 객체의 특성을 파악한다. 두 영상사이의 MAD 값이 주어진 임계값보다 크면 그 영상의 장면 변화가 있다고 가정한다. 이후 현재 영상의 객체와 다음 영상의 객체에 대한 변화 추적을 통하여 정확한 객체를 추출한다. 마지막으로 공간적 경계 검출과 시간적 영상 검출에 따른 영상 추적을 통합하여 객체의 정확한 경계 추출을 설계한다.

4. 결론과 향후 계획

본 논문은 내용 기반 표준 MPEG-4 비디오 코딩을 위한 자동 VOP 객체추출 방법을 설계하였다. 설계된 방법은 정확한 에지를 검출할 수 있는 Shen 과 watershed 알고리즘을 사용하였고 정확한 객체이외의 잡음을 제거하기 위한 방법으로는 median 필터를 사용하였다. 또한 시간적 방향에서 객체의 추적을 위하여 전 탐색 블록 매칭을 이용한 변환 탐지 마스크를 사용하여 정확한 벡터의 움직임을 설정할 것이다. 현재 공간적 분할을 위한 기본 단계가 완료되었으며 차후 연속되는 비디오 프레임 사이의 객체를 찾기 위한 작업이 진행될 것이다.

5. 참고 문헌

- [1] Thomas Sikora, " The MPEG-4 Video Standard Verification Model", IEEE Transactions on Circuits and Systems for Video Technology, Vol.7, No.1, February 1997.
- [2] Touradj Ebrahimi, " MPEG-4 video verification model: A video encoding/decoding algorithm based on content representation, Signal Processing :Image Communication, pp. 367 - 384, 1997.
- [3] P.Salembier and F.Marques et al, " Region-Based Representations of Image and Video: Segmentation Tools for Multimedia Services " IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No.8, December 1999.
- [4] M.C.Kim, J.G.Choi, D.H.kim, H.Lee, M.H.Lee, C.T.Ahn and Y.S.Ho, " A VOP Generation Tool: Automatic Segmentation of Moving Objects in Image Sequences Based on Spatio -Temporal Information, IEEE Transactions on Circuits and Systems for Video Technology, Vol.9, No.8, December 1999.
- [5] Thomas Meier and King N.Ngan, " Video Segmentation for Content-Based Coding ", IEEE Transactions on Circuits and Video Technology, Vol. 9, No.8, December 1999.
- [6] J.G.Choi, S.W.Lee and S.D.Kim, " Spatio-temporal video segmentation using a joint similarity measure, " IEEE Trans, Image Processing, Vol 7, pp.279-286, April 1997.