

컬러 레이아웃을 이용한 키 프레임 추출 기법

김 소 희*, 김 형 준*, 지 수 영**, 김 회 율*
*한양대학교 영상공학연구소, **한국전자통신연구원
전화 : 02-2290-0561 / 팩스 : 02-2292-6316

The Extracting Method of Key-frame Using Color Layout Descriptor

So-Hee Kim*, Hyoung-Joon Kim*, Su-Young Ji**, and Whoi-Yul Kim*
*Image Engineering Lab, Hanyang University, **ETRI
E-mail : sohee@vision.hanyang.ac.kr

Abstract

Key frame extraction is an important method of summarizing a long video. This paper propose a technique to automatically extract several key frames representative of its content from video.

We use the color layout descriptor to select key frames from video. For selection of key frames, we calculate similarity of color layout features extracted from video, and extract key frames using similarity. An important aspect of our algorithm is that does not assume a fixed number of key frames per video; instead, it selects the number of appropriate key frames of summarizing a long video.

Experimental results show that our method using color layout descriptor can successfully select several key frames from a video, and we confirmed that the processing speed for extracting key frames from video is considerably fast.

I. 서론

최근 압축 기술과 통신 기술이 빠르게 발달함에 따라 멀티미디어 데이터의 양은 급속도로 증가하지만 정

보량이 늘어날수록 원하는 정보를 검색하기는 더욱 어려워지고 있는 상황이다. 대용량 비디오 데이터베이스에서 사용자가 원하는 정보를 빠른 시간 내에 검색하기 위해서는 비디오의 내용을 자동으로 요약하고, 효과적으로 검색해주는 기술이 중요하다[1][2]. 대용량 비디오 데이터베이스에서 비디오를 검색하는데 많은 시간이 소요되기 때문에 비디오에서 중요 장면을 검출한 후 장면을 대표할 수 있는 키 프레임을 추출하는 것은 비디오의 내용을 효과적으로 요약하는데 필요하다[3][4].

본 논문에서는 비디오 상에서의 컬러 레이아웃 기술을 이용한 효과적인 키 프레임 추출 알고리즘을 제안한다. 컬러 레이아웃 기술자는 비디오에서 키 프레임을 추출하기 위한 색상 정보를 제공하는 기술자로 색상의 공간적인 분포를 표현할 수 있으며 유사도 계산도 단순하여 매우 빠른 속도로 검색과 브라우징을 가능하게 해준다. 컬러 레이아웃 기술자의 효율성은 멀티미디어 데이터의 국제 표준화인 MPEG-7에서 이미 검증되었다[5][6][7].

본 논문에서는 컬러 레이아웃 기술자를 이용하여 비디오의 색상 특징 정보를 추출하여 비디오의 내용을 효과적으로 요약할 수 있는 한 개 이상의 키 프레임을 추출한다. 비디오의 전체 프레임에 대해 컬러 레이아웃 특징 값을 추출하고, 현재 프레임과 이전 프레임이 얼마나 유사한가를 판별하는 유사도를 컬러 레이아웃

특징 값들을 이용하여 계산한 후, 유사도가 가장 높고 오랜 시간동안 비디오에서 지속되는 프레임들 중 첫 번째 프레임을 그 비디오를 대표하는 키 프레임으로 선정한다.

본 논문의 구성은 다음과 같다. 2장에서는 컬러 레이아웃 기술자에 대한 설명과 컬러 레이아웃을 이용한 색상 정보 특징 값 추출 알고리즘을 기술하고, 3장에서는 2장에서 추출한 컬러 레이아웃 특징 값들을 이용하여 유사도 구하는 방법과 유사도를 이용하여 키 프레임을 추출하는 알고리즘에 대해 기술하고, 4장에서는 실험결과를 보이고, 마지막으로 5장에서는 결론을 보인다.

II. 컬러 레이아웃 특징 값 추출

비디오 상에서 효과적인 키 프레임을 추출하기 위해서는 비디오의 특성 정보를 추출해야 하는데 본 논문에서는 컬러 레이아웃 기술자를 이용하여 비디오의 색상 정보 특징 값을 추출한다. 컬러 레이아웃 기술자는 영상의 색상 정보를 기술해 주는 기술자로서 색상의 공간적인 분포를 잘 나타내 주며 유사도 계산도 단순하여 매우 빠른 속도로 검색과 브라우징을 가능하게 해주고, 이미지 대 이미지 검색, 비디오 대 비디오 검색뿐만 아니라 다른 색상 기술자들에서는 지원하지 않는 스케치 대 이미지 검색도 가능하게 해주는 기술자이다[5][6].

컬러 레이아웃 특징 값을 추출하는 과정은 다음과 같다. 우선 R, G, B 컬러 모델을 Y, Cb, Cr 컬러 모델로 변환한 후, 이미지 분할, DCT(discrete cosine transform)변환, 지그재그 스캔을 적용한 특징 값 추출, 추출한 특징 값의 정규화 4단계의 과정을 거친다. 원 이미지의 각 픽셀 값들의 평균값을 구하는 방법은 간단하지만, 이미지의 명확한 색상 정보를 추출하기에는 부적절하다. 그래서 좀 더 명확한 색상 정보를 얻기 위해 컬러 레이아웃 기술자에서는 전체 이미지의 크기에 상관없이 가로와 세로의 길이를 각각 8로 나눈다. 즉, 원 이미지가 8x8 크기의 블록으로 나누어지고 블록 내의 모든 픽셀 값들의 평균값이 그 블록을 대표하는 픽셀 값이 된다. 그리고 DCT 변환을 하면 저주파 성분인 상위 몇 개의 DCT 계수만으로 효율적인 색상 정보를 추출하게 되므로 Y, Cb, Cr 이미지에 대해 각각 DCT 변환을 하여 DCT 계수를 산출한다. 계수의 추출은 지그재그 스캔을 적용하여 Y계수는 상위 6개, Cb와 Cr계수는 상위 3개를 컬러 레이아웃 특징 값으로 추출한다. 마지막으로 추출된 Y계수 6개, Cb와 Cr 계수 3개의 컬러 레이아웃 특징 값 각각에 대해 정규화 과정을 수행한다. 그림1은 컬러 레이아웃 특징 값을 추출하는 과정을 나타낸 것이다.

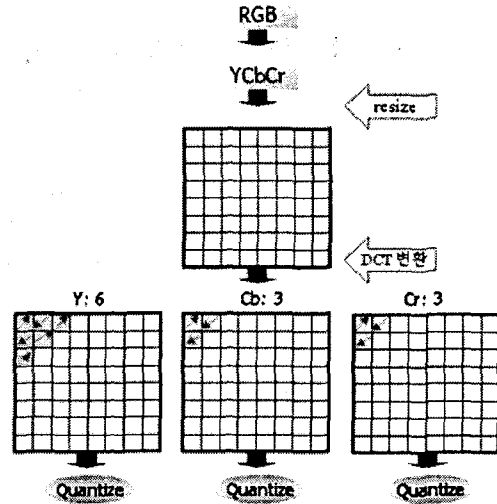


그림 1. 컬러 레이아웃 특징 값 추출 알고리즘

III. 키 프레임 추출

본 논문에서는 비디오 전체 프레임들의 컬러 레이아웃 색상 정보 특징 값을 이용하여 현재 프레임과 이전 프레임의 유사도를 구해 비디오의 내용을 효과적으로 요약하는데 필요한 키 프레임들을 추출한다. 키 프레임은 비디오 중에서 변화가 가장 적고 오랜 시간동안 비디오에서 지속되는 프레임들 중 첫 번째 프레임으로 선정한다. 유사도는 추출된 특징 값들을 이용하여 현재 프레임과 이전 프레임간의 유사한 정도를 나타내며 현재 프레임의 컬러 레이아웃 특징 값을 $YCoeff[6]$, $CbCoeff[3]$, $CrCoeff[3]$ 라 하고 이전 프레임의 컬러 레이아웃 특징 값을 $YCoeff'[6]$, $CbCoeff'[3]$, $CrCoeff'[3]$ 이라 했을 때, 프레임간의 유사도는 식(1)과 같고, λ 계수는 표1과 같다.

$$D = \sqrt{\sum_{i=0}^6 \lambda_{Yi} (YCoeff[i] - YCoeff'[i])^2} + \sqrt{\sum_{i=0}^3 \lambda_{Cbi} (CbCoeff[i] - CbCoeff'[i])^2} + \sqrt{\sum_{i=0}^3 \lambda_{Cri} (CrCoeff[i] - CrCoeff'[i])^2} \quad (1)$$

표 1. λ 계수

	Coefficient Order					
	0	1	2	3	4	5
$YCoeff$	2	2	2	1	1	1
$CbCoeff$	2	1	1			
$CrCoeff$	4	2	2			

컬러 레이아웃을 이용한 키 프레임 추출 기법

식(1)을 이용하여 현재 프레임과 이전 프레임간의 유사도가 구해지면 유사도가 가장 높은 즉 비디오 중에서 변화가 가장 적고 오랫동안 비디오에서 지속되는 프레임들 중 첫 번째 프레임을 키 프레임으로 선정한다. 키 프레임은 비디오에 따라 각각 내용이나 길이 등이 다르므로 적어도 1개 이상, 서로 다른 개수가 추출된다. 그림2는 특징 값들의 유사도를 계산하여 키 프레임을 추출하는 전체 알고리즘 과정을 나타낸다.

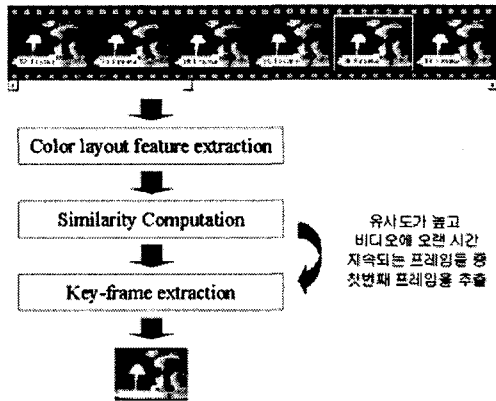


그림 2. 전체 알고리즘

IV. 실험 및 결과

본 논문의 실험은 MPEG-1으로 인코딩 된 2~6초 길이의 MPEG-7 표준 실험 영상, 영화 비디오 등을 대상으로 하였다.

그림3(a)~그림3(e)은 표2의 실험영상을 대상으로 본 논문에서 제안한 컬러 레이아웃 기술자를 이용한 키 프레임 추출을 실험한 결과를 나타낸다.

그림3(a)는 실험 비디오의 대표적인 내용을 효과적으로 나타낼 수 있는 영화 캐릭터의 가만히 있는 모습, 입을 벌린 모습, 눈이 튀어나온 모습이 키 프레임으로 추출되었음을 보여준다.

그림3(b)는 사람의 눈이 튀어나왔다가 다시 들어가는 모습이 키 프레임으로 추출되었음을 보여준다.

그림3(c)는 풀을 베기 시작하려는 모습과 풀을 베고 있는 모습이 키 프레임으로 추출되었음을 보여준다.

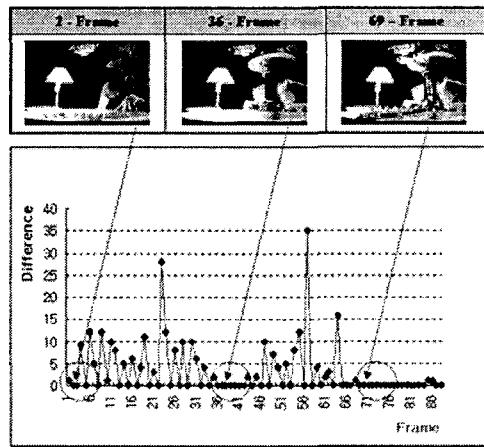
그림3(d)는 사람의 일련의 동작을 잘 나타낼 수 있는 박 점프를 시작한 모습, 점프의 최고점에 이르는 모습, 점프가 끝나는 모습이 키 프레임으로 추출되었음을 보여준다.

그림3(e)는 사슴들이 오른쪽 방향으로 서 있는 모습과 사슴들이 왼쪽 방향으로 달리는 모습이 키 프레임으로 추출되었음을 보여준다.

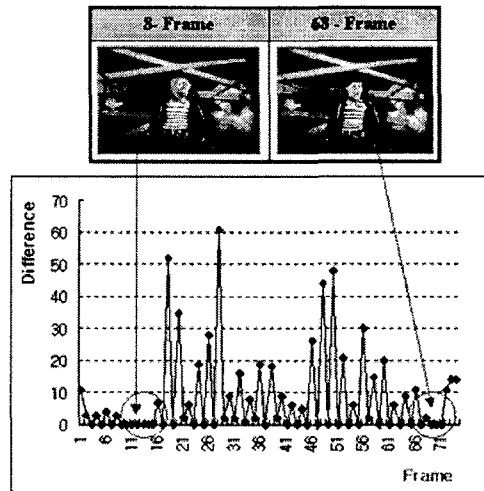
실험결과 비디오의 전반적인 내용을 효과적으로 요약하는데 적절한 키 프레임들이 추출됨을 알 수 있다.

표 2. 실험에 사용된 비디오

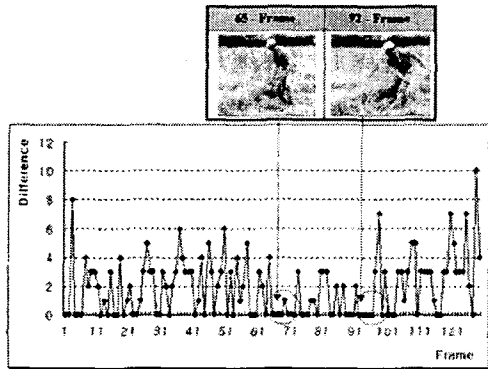
	사이즈	프레임 수	길이
실험 1	352x240	76	2초
실험 2	352x288	146	6초
실험 3	352x240	90	3초
실험 4	352x288	60	2초
실험 5	352x288	131	6초



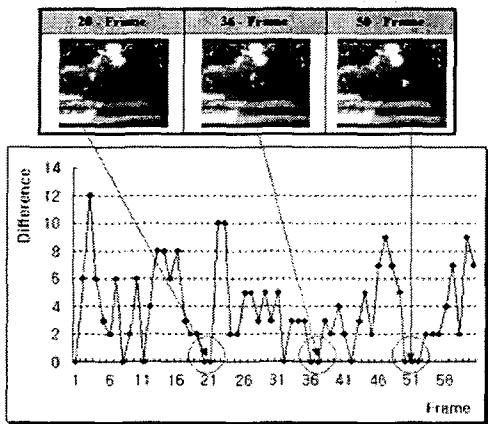
(a) 실험 1



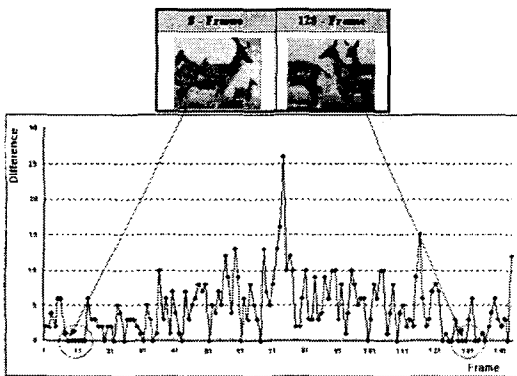
(b) 실험 2



(c) 실험 3



(d) 실험 4



(e) 실험 5

그림 3. 키 프레임 추출 결과

V. 결론

본 논문에서는 비디오의 내용을 요약 할 수 있는 키 프레임을 추출하기 위한 방법을 제안하였다.

키 프레임을 추출하기 위해서 컬러 레이아웃 기술을 이용하여 색상 정보가 담긴 특징 값들을 추출하고, 추출된 특징 값들의 유사도를 계산하여 비디오 중에서 변화가 가장 적고, 오랫동안 비디오에서 지속되는 프레임들 중 첫 번째 프레임을 키 프레임으로 추출한다.

키 프레임 추출 알고리즘은 비디오의 내용을 요약할 수 있는 효과적인 키 프레임들이 단순한 유사도 계산식에 의해 빠른 속도로 추출됨을 실험으로 확인 할 수 있었다. 그리고 MPEG으로 인코딩된 비디오에서는 Y, Cb, Cr의 DCT 계수를 직접 이용할 수 있으므로 더 빠르게 추출이 가능하다.

앞으로 연구되어야 될 내용으로는 색상 정보와 더불어 움직임 정보, 모양 정보, 질감 정보 등을 사용하여 프레임 내 전체적인 내용을 분석하고 그 정보를 키 프레임 추출 특징 값으로 이용하는 것이다.

참고문헌

- [1] B.Furht, S.W Smoliar and H.J. Zhang, *Video and Image Processing in Multimedia System*, Kluwer Academic Publishers, 1995.
- [2] Ahmed K.Elmagarmid, *Video Database Systems : Issues, Products, and Applications*, Kluwer Academic Publishers, 1997.
- [3] W. Wolf, "Key Frame Selection by Motion Analysis," ICASSP-96, Vol.2, pp.1228-1231, 1996.
- [4] "MPEG-7 Multimedia Description Schemes XM Version 7.0," ISO/IEC JTC1 /SC29/ WG11, N3964, Mar. 2001.
- [5] "Text of ISO/IEC 15938-3/FCD Information Technology - Multimedia Content Description Interface - Part 3 Visual," ISO/IEC JTC1/SC29/-WG11, N4062, Mar. 2001.
- [6] "CD 15938-3 MPEG-7 Multimedia Content Description Interface - Part3 Visual", ISO/IEC JTC1/SC29/WG11, W3703, Oct. 2000.
- [7] "MPEG-7 Visual part of eXperimentation Model Version 8.0", ISO/IEC JTC1/SC29/WG11, N3673, Oct. 2000.