

색상과 움직임 정보를 이용한 내용기반 동영상 검색 시스템

김 소 희*, 김 형 준*, 정 연 구**, 김 회 율*
*한양대학교 영상공학연구소, **한국전자통신연구원
전화 : 02-2290-0561 / 팩스 : 02-2292-6316

Content-Based Video Retrieval System Using Color and Motion Features

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

Abstract

Numerous challenges have been made to retrieve video using the contents. Recently MPEG-7 had set up a set of visual descriptors for such purpose of searching and retrieving multimedia data. Among them, color and motion descriptors are employed to develop a content-based video retrieval system to search for videos that have similar characteristics in terms of color and motion features of the video sequence.

In this paper, the performance of the proposed system is analyzed and evaluated. Experimental results indicate that the processing time required for a retrieval using MPEG-7 descriptors is relatively short at the expense of the retrieval accuracy.

I. 서론

최근 컴퓨터와 통신 기술의 급속한 발달로 인해 인터넷을 통한 멀티미디어 데이터의 전송이 보편화되어 정보 통신망 사용자가 시간과 장소에 상관없이 다양한 정보에 접근할 수 있게 되었으며 이용할 수 있는

멀티미디어 정보의 양도 폭발적으로 증가하게 되었다. 하지만 정보량이 늘어날수록 원하는 정보를 찾기는 더욱 어려워지고 멀티미디어 정보의 검색, 저장, 관리 기술에 대한 요구가 늘어나면서 대용량 동영상 데이터베이스에서 원하는 정보를 빠른 시간 내에 검색할 수 있는 검색 기법이 필요하게 되었고 현재 활발한 연구가 진행 중이다[1][2][3].

본 논문에서는 동영상 검색에 적합한 색상 정보와 움직임 정보를 표현하는 복합 특징 값을 쿼리 동영상과 데이터베이스내의 동영상들에 대해 추출하고 유사도를 판별하여 사용자의 쿼리에 가장 적합한 동영상을 검색하는 내용 기반 동영상 검색 시스템을 구현한다. 검색을 위해 동영상의 색상 정보와 움직임 정보를 이용하고, 이러한 정보를 기술하기 위해 MPEG-7에 정의되어 있는 비주얼 기술자(visual descriptor) 중 에서 컬러 레이아웃 기술자(color layout descriptor), 대표 컬러 기술자(dominant color descriptor), 움직임 활동성 기술자(motion activity descriptor) 3개를 이용한다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 특징 값 추출 알고리즘을 기술하고, 3장에서는 추출한 특징 값들의 유사도를 측정하는 방법을 기술하고, 4장에서는 구현한 검색 시스템에 대하여 기술하고, 5장에서는 실험결과를 보이고, 마지막으로 6장에서는 결론을 제시한다.

II. 특징 값 추출 알고리즘

컬러 레이아웃 기술자는 색상의 공간적인 분포를 잘 나타내주는 기술자로 매우 빠른 속도로 검색과 브라우징을 가능하게 한다. 대표 컬러 기술자는 영상 전체 또는 일부 영역에서 두드러지게 나타나는 색상을 기술하는 기술자이다. 그리고 움직임 활동성 기술자는 움직임의 격렬함을 표현하는 것으로, 이는 사람이 동영상 보면서 느리게 움직이는 장면 혹은 빠르게 움직이는 장면 등으로 구분하는 시각적 특성을 반영한다 [2][3][4]. 컬러 레이아웃과 대표 컬러 특징 값은 동영상의 전체 프레임 중 그 동영상을 대표할 수 있는 키 프레임을 추출하여 그 키 프레임에 대해 추출하였고, 움직임 활동성 특징 값은 동영상의 전체 프레임의 평균을 구해서 추출하였다.

2.1 컬러 레이아웃 기술자

컬러 레이아웃 특징 값 추출은 R, G, B 컬러 모델을 Y, Cb, Cr 컬러 모델로 변환한 후, 전체 이미지의 크기에 상관없이 가로와 세로의 길이를 각각 8로 나눈다. 즉, 원 이미지가 8×8 크기의 블록으로 나누어지고 블록 내의 모든 픽셀 값들의 평균값이 그 블록을 대표하는 컬러 값이 된다. 그리고, DCT(discrete cosine transform) 변환 후 지그재그 스캔을 적용하여 Y 계수는 상위 6개, Cb와 Cr 계수는 상위 3개의 값을 컬러 레이아웃 특징 값으로 추출한다. 마지막으로 추출된 Y 계수 6개, Cb와 Cr 계수 각각 3개 특징 값들에 대해 정규화 과정을 수행한다. 추출된 특징 값은 Y_{Coeff} 6개, Cb_{Coeff} 3개, Cr_{Coeff} 3개로 기술된다.

2.2 대표 컬러 기술자

대표 컬러 특징 값 추출은 R, G, B 컬러 모델을 L, U, V 컬러 모델로 변환한 후, distortion과 split 과정을 반복하다가 대표 컬러의 개수가 8개가 넘으면 반복 수행을 멈추는 GLA(Generalized Lloyd Algorithm)를 적용한다. 그리고 merging 과정을 수행 후 다시 L, U, V 컬러 모델을 R, G, B 컬러 모델로 변환한다. 마지막으로 특징 값들에 대해 정규화 과정을 수행한다. 추출된 특징 값은 대표 컬러의 개수, 추출된 대표 컬러의 R, G, B 값, 추출된 대표 컬러가 차지하는 백분율로 기술된다.

2.3 움직임 활동성 기술자

움직임 활동성 기술자는 움직임 활동성의 정도를 나타내는 intensity와 방향, 움직임 활동의 공간적 특성으로 이루어지며, 각 특성을 추출하기 위해 16×16 크기의

블록에 대한 모션 벡터를 이용한다.

Intensity는 움직임 활동성의 높고 낮음을 표현하기 위해 각 블록에 대한 모션 벡터들의 크기의 분산으로 정의되고, 방향은 서로 다른 움직임들에 대한 대표 방향을 표현하기 위해 모션 벡터들의 방향에 대한 평균으로 정의된다. 공간적 특성은 움직임 활동성이 하나의 넓은 지역에서 나타나는지 아니면 여러 개의 좁은 지역에서 나타나는지를 표현하기 위해, 상대적으로 움직임이 없는 블록들의 크기를 이용한다. 움직임이 없는 블록의 판별 기준은 모션 벡터들의 크기의 평균값을 이용하고, 평균값보다 작은 크기의 모션 벡터를 갖는 연속된 블록들의 run-length를 short run, middle run, long run의 세 부류로 나누어서 각 부류의 개수로 공간적 특성을 정의한다. 이때, run-length를 세 부류로 나누는 기준은 run-length가 영상을 이루는 블록의 가로 개수에 대해 1/3 보다 작으면 short run, 1/3 보다 크고 2/3 보다 작으면 middle run, 2/3 보다 크면 long run으로 정한다. 추출된 특징 값은 intensity를 표현하는 σ , 방향을 표현하는 A_{avn} , 그리고 공간적 특성을 표현하는 N_{sn} , N_{mr} , N_{lr} 들로 기술된다.

III. 유사도 측정

사용자가 쿼리로 준 비디오와 데이터베이스에 저장된 비디오와의 유사도를 측정하기 위해 2장에서 설명한 각 특징 값에 대한 distance를 구하고, 식(1)과 같이 구해진 각 distance들의 가중치 합을 이용한다.

$$Distance = w_1 D_L + w_2 D_D + w_3 D_A \quad (1)$$

$$\text{단, } w_1 + w_2 + w_3 = 1$$

3.1 컬러 레이아웃 기술자

쿼리 비디오의 컬러 레이아웃 특징 값을 $Y_{Coeff}[i]$, $Cb_{Coeff}[i]$, $Cr_{Coeff}[i]$, 데이터베이스에 들어있는 비디오의 특징 값을 $Y_{Coeff}'[i]$, $Cb_{Coeff}'[i]$, $Cr_{Coeff}'[i]$ 라 할 때, 컬러 레이아웃에 대한 distance는 식(2)과 같이 정의된다. 이 때, 표1의 λ 는 가중치를 나타내며 식(2)에 대입하여 distance를 계산한다.

$$D_L = \sqrt{\sum_{i=1}^6 \lambda_{Y_i} (Y_{Coeff}[i] - Y_{Coeff}'[i])^2 + \sum_{i=0}^3 \lambda_{Cb_i} (Cb_{Coeff}[i] - Cb_{Coeff}'[i])^2 + \sum_{i=0}^3 \lambda_{Cr_i} (Cr_{Coeff}[i] - Cr_{Coeff}'[i])^2} \quad (2)$$

색상과 움직임 정보를 이용한 내용기반 동영상 검색 시스템

표 1. λ 계수

	Coefficient Order					
	0	1	2	3	4	5
YCcoeff	2	2	2	1	1	1
CbCoeff	2	1	1			
CrCoeff	4	2	2			

3.2 대표 컬러 기술자

식(3)에서 $a_{k,l}$ 는 두 컬러 c_k, c_l 사이의 similarity coefficient를 나타내고, 식(4)의 $d_{k,l}$ 는 두 컬러 c_k, c_l 사이의 Euclidean distance를 나타낸다. 이때, $d_{max} = \alpha T_d$ 인데, T_d 는 두 컬러의 maximum distance 로써 기본값은 LUV 컬러 모델에서 10~20이며 α 는 1.0~1.5이다. 식(5)에서 $D(F_1, F_2)$ 는 두 특징 값 사이의 similarity를 나타내며 F_1 은 쿼리 대표 컬러 특징 값, F_2 는 레퍼런스 대표 컬러의 특징 값을 의미한다. 최종적으로 대표 컬러에 대한 distance는 두 대표 컬러의 difference를 나타내는 DC_Diff 에 W (가중치 : 0.7)를 곱한 것으로 식(6)과 같이 정의된다.

$$a_{k,l} = \begin{cases} 1 - d_{k,l} / d_{max} & d_{k,l} \leq T_d \\ 0 & d_{k,l} > T_d \end{cases} \quad (3)$$

$$d_{k,l} = \| c_k - c_l \| \quad (4)$$

$$D(F_1, F_2) = \sum_{i=1}^{N_1} P_{1i}^2 + \sum_{j=1}^{N_2} P_{2j}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{1i,2j} P_{1i} P_{2j} \quad (5)$$

$$D_D(D_1, D_2) = W \times DC_Diff \quad (6)$$

- F : 대표 컬러
- N : 대표 컬러의 개수
- c : 색상 값
- P : 퍼센트
- W : 가중치 set to 0.7

3. 움직임 활동성 기술자

쿼리 비디오를 S_q , 데이터베이스에 들어있는 비디오를 S 라 할 때, 움직임 활동성에 대한 distance는 식(7)과 같이 정의된다. 이때, 분모가 0인 경우 그 항은 계산에 포함하지 않는다.

$$D_A = \frac{W_{int}}{\sigma(S_q)} |\alpha(S_q) - \alpha(S)| + \frac{W_{avr}}{A_{avr}(S_q)} |A_{avr}(S_q) - A_{avr}(S)| + \frac{W_{sr}}{N_{sr}(S_q)} |N_{sr}(S_q) - N_{sr}(S)| + \frac{W_{mr}}{N_{mr}(S_q)} |N_{mr}(S_q) - N_{mr}(S)| + \frac{W_{br}}{N_{br}(S_q)} |N_{br}(S_q) - N_{br}(S)| \quad (7)$$

$$\text{단, } W_{int} = \frac{\alpha(S_q) + A_{avr}(S_q) + N_{sr}(S_q) + N_{mr}(S_q) + N_{br}(S_q)}{5}$$

IV. 검색 시스템

본 논문에서 구현한 검색 시스템은 컬러 레이아웃 기술자, 대표 컬러 기술자, 움직임 활동성 기술자 3개를 이용하여 색상과 움직임 정보를 기술하고 그 기술된 정보를 이용해서 사용자가 원하는 비디오를 검색한다. 본 검색 시스템의 특징은 색상과 움직임 정보 특징 값 추출 및 검색 등을 압축영역에서 수행한다는 것이다. 전체 시스템은 그림1에서처럼 사용자의 쿼리 동영상의 특징 값을 추출하여 데이터베이스에 저장되어 있는 동영상들의 특징 값들과의 유사도를 비교한 후 사용자의 쿼리에 적합한 동영상을 데이터베이스에서 검색하여 유사도가 높은 순서로 보여준다. 쿼리란 어떤 종류의 특징벡터를 이용하여 검색할 것인지를 기술하는 것으로 본 논문에서 구현한 내용기반 동영상 검색 시스템에서는 색상 혹은 움직임 정보를 포함한 일차원, 다차원 쿼리가 모두 가능하다. 구현한 검색 시스템은 사용자가 색상과 움직임 특징 값에 직접 가중치를 두어 쿼리를 할 수 있게 구성되어 있으므로 사용자가 원하는 동영상을 정확히 검색할 수 있다.

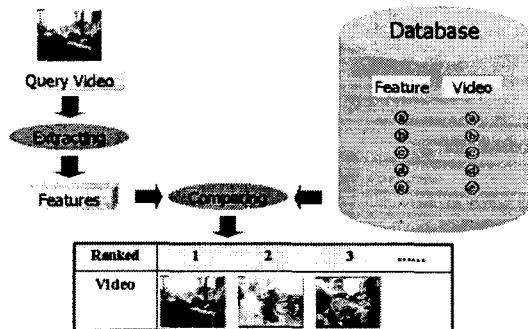


그림 1. 전체 시스템 구조도

V. 실험 및 결과

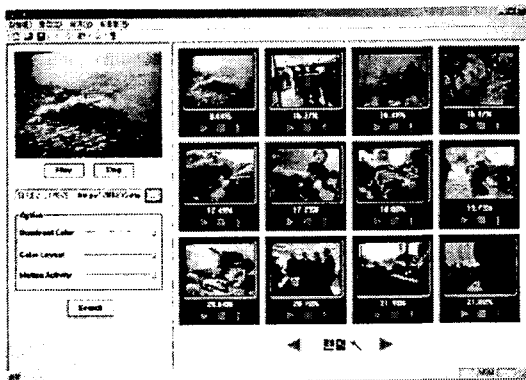
본 논문에서는 실험을 위해 MPEG 형식의 길이가 5~10초인 1000개의 비디오로 데이터베이스를 구성하였다. 그림2는 구축한 데이터베이스의 실험영상을 대상으로 본 논문에서 구현한 검색 시스템을 이용한 동영상 검색 결과를 나타낸다. 그림2(a)와 그림2(b)의 결과화면에서 왼쪽 부분은 쿼리 화면으로써 사용자가 검색하고자 하는 동영상을 색상과 움직임 특징 값에 가중치를 두어 쿼리를 할 수 있게 구성되어 있고, 오른쪽 부분은 검색 화면으로써 유사도가 높은 순서로 디스플레이 되도록 구성되어 있다.

실험결과 사용자의 쿼리 동영상과 같은 동영상은 물론 유사한 색상과 움직임 정보를 가지고 있는 동영상들이 유사도가 높은 순서로 잘 검색되었다.

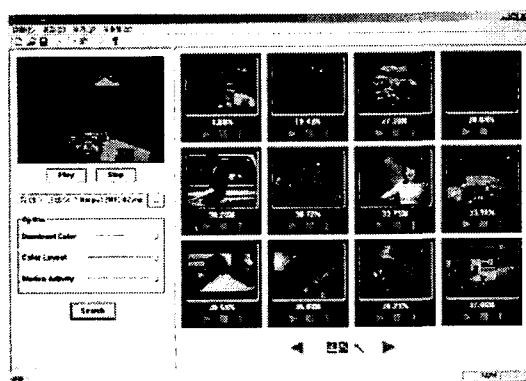
VI. 결론

본 논문에서 구현한 동영상 검색 시스템은 색상과 움직임 정보를 기술하기 위해 MPEG-7에서 정의한 컬러 레이아웃 기술자, 대표 컬러 기술자, 움직임 활동성 기술자를 이용해서 기술된 정보를 이용해서 동영상을 검색한다. 검색 시스템은 색상과 움직임 정보에 가중치를 주어 사용자가 쿼리한 동영상에 대해 특징 값들을 추출한 후, 데이터베이스에 저장되어 있는 동영상들의 특징 값들과 유사도를 비교해서 유사도가 높은 순서대로 동영상을 검색해준다.

향후 연구과제는 움직이는 물체의 컬러 영역 또는 움직임 궤적 등을 이용한 동영상 검색 방법과 다양한 쿼리 방식을 검색 시스템에 추가해서 검색 효율을 높이는 것이다.



(a)



(b)

그림 2. 검색 결과

참고문헌

- [1] 이동호, 유광석, 김희을, "컬러와 모양 정보를 이용한 캐릭터 이미지 검색," 방송공학회 논문지, 제5권, 제1호, pp. 50-60, 2000년 6월.
- [2] "MPEG-7 Multimedia Description Schemes XM," ISO/IEC JTC1/SC29/WG11, N3112, Dec. 1999.
- [3] "MPEG-7 Multimedia Description Schemes WD," ISO/IEC JTC1/SC29/WG11, N3113, Dec. 1999.
- [4] "Text of ISO/IEC 15938-3/FCD Information Technology - Multimedia Content Description Interface - Part 3 Visual," ISO/IEC JTC1/SC29/WG11, N4062, Mar. 2001.
- [5] "CD 15938-3 MPEG-7 Multimedia Content Description Interface - Part3 Visual," ISO/IEC JTC1/SC29/WG11, W3703, Oct. 2000.
- [6] "MPEG-7 Visual part of eXperimentation Model Version 8.0," ISO/IEC JTC1/SC29/WG11, N3673, Oct. 2000.