

컬러를 이용한 효과적인 비디오 검색 시스템 구현

이 효 종, *문 정 렬
전북대학교 전자공학과, 공업기술연구소
전화 : 063-277-7312

Implementation of Efficient Video Retrieval System using Color

Hyo Jong Lee, Jeong Ryeol Mun
Dept. of Electronics, Chonbuk University
E-mail : jlmun@sel.chonbuk.ac.kr

Abstract

In this paper, we propose an efficient database system for video retrieval. Using the color spaces, it shows results of user's request. Each color space used following user's selection. We suggest adaptive three color systems for database. Experimental results based on a video database containing 331 shots are included.

I. 서론

최근 인터넷이나 디지털 도서관에서 오디오와 함께 비디오 정보가 전달되어야 하므로 이에 디지털 데이터 사용이 필수적이다. 또한 빠르게 증가하는 디지털 데이터를 효과적으로 사용하기 위해서는 그에 맞는 검색 방법이 제안되어야 한다. 이에 따라 최근에 컬러 히스토그램에서 추출한 여러 가지 특징을 이용한 비디오 데이터베이스 시스템들이 연구되고 있다[2].

비디오 데이터베이스 시스템은 일반적으로 다음과 같은 과정으로 이루어진다[2]. 먼저 비디오를 동일한 카메라에서 연속적으로 촬영된 프레임들, 즉 화면 단위로 나눈다. 비디오를 화면 단위로 나누는 방법으로 RGB, HSV, XYZ, LAB, LUV, Munsell 등의 여러 컬러 공간 등을 통해서 구해진 히스토그램을 많이 이용하고 있고 히스토그램간의 차를 구하는 데는 Bin-to-bin

difference(B2B), Chi-square test histogram difference(CHI)[3], Histogram intersection

(INT) [4], Average color(AVG)[5] 같은 방법들이 쓰이고 있다[2]. 두 번째로, 각 화면에서 화면의 특성을 대표하는 대표 프레임(key frame)을 고른다. 마지막으로, 대표 프레임에서 추출해낸 특징 정보를 여러 트리를 통해 계층화해서 저장하고 검색에 이를 이용함으로써 비디오 데이터 베이스를 구성할 수 있다.

그러나, 많은 비디오 데이터베이스 시스템이 사용자의 입장에서 요구되는 검색의 정확성과 검색 속도를 충분히 만족시키지 못하고 있다. 이에 본 논문에서 구현된 시스템은 컬러 정보를 이용해 사용자의 요구에 따라 검색 분야를 한정하고 추가함으로써 효과적인 검색과 함께 사용자의 요구를 충족시킬 수 있다.

본 논문의 구성은 다음과 같다. 제 2장에서는 화면 분류와 대표 프레임 추출에 관해 서술하고, 제 3장에서는 프레임에서 추출된 정보를 저장하고 이를 효과적으로 활용하는 방법에 관해 서술한다. 제 4장에서는 실험 결과를 제시하고 마지막으로 제 5장에서는 결론을 맺는다.

II. 화면 분류 및 대표 프레임선출

2.1 화면 분류

화면 분류와 대표프레임 선출 과정은 그림 1과 같다. 비디오 화면 분류에서 히스토그램을 이용하는 방법

중 가장 효과적인 Munsell 컬러 공간에서 계산된 히스토그램 값을 계산해내는 방법을 통해 비디오의 화면을 분리해 냈고[1] 이 때 각 프레임간의 차를 구하는 방법으로는 히스토그램 차를 계산해내는 가장 효과적인 Bin-to-bin (B2B) 방법을 사용했다[1].

Munsell 컬러 공간(S_1, S_2, L)은 다음과 같은 식으로 유도된다.

$$M_1 = V(X_c) - V(Y), X_c = 1.020X$$

$$M_2 = 0.4 \cdot (V(Z_c) - V(Y)), Z_c = 0.847Z$$

$$M_3 = 0.23 \cdot V(Y)$$

여기서 X, Y, Z는 각각 R, G, B값을 의미한다.

$$S_1 = (8.880 + 0.966 \cdot \cos \theta) \cdot M_1$$

$$S_2 = (8.025 + 2.558 \cdot \cos \theta) \cdot M_2$$

$$L = M_3 \quad \text{with } \theta = \text{atan}(M_1/M_2)$$

Bin-to-bin 방법으로 차를 구하는 방법은 다음과 같다.

$$fd_{b2b}(h_1, h_2) = \frac{1}{2N} \sum_i |h_1[i] - h_2[i]|$$

h_i : 히스토그램 N : 픽셀수

h : 히스토그램값 N : 픽셀수

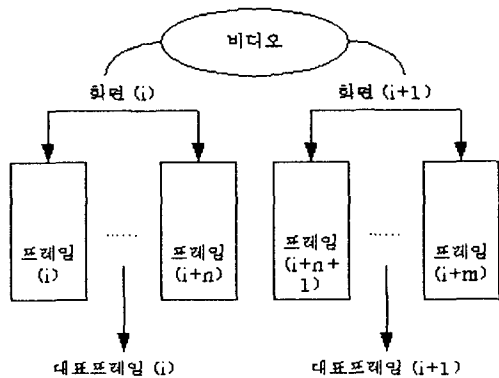


그림 1 화면 분류와 대표프레임 선출 과정

2.2 키 프레임 결정

키프레임은 각 화면에서 3가지 히스토그램방식으로 각각 구할 수 있었고 이를 통해 화면마다 프레임간의 순도를 측정함으로써 각각의 대표 프레임을 뽑아낼 수 있었다. 순도 측정방법으로는 root-mean-square error 방법을 이용했다.

$$E = \sum_{i,j} \sqrt{\Delta D_{ij}^2}$$

D_{ij} 값은 연속 프레임에서 계산된 값들의 차를 나타내고, 이때 대표 프레임은 $\min(E)$ 값을 갖는 프레임으로 정한다.

(1) RGB값을 이용한 키 프레임 결정

RGB 컬러 모델은 우리 눈의 망막에서 세 가지 시색소의 흥분 정도에 따라 색을 감지하는 시각의 3차원 원리를 이용했으며 디스플레이 장치에 중점을 두는 검색을 하기 위해 이용된다. 키프레임 결정에 사용되는 값을 구하는 과정은 아래와 같다.

1) 한 화면에 있는 프레임 각각의 gray 값의 평균을 구한다.

$$2) E_{ij} = \sum \sqrt{(\text{mean}(rgb_{ij}))^2 - (\text{mean}(rgb_{j0}))^2}$$

의 수식을 이용해서 $\min(E)$ 를 갖는 프레임을 가장 순도가 높은 프레임으로 정하고 이를 대표 프레임으로 선출한다.

(2) HSV값을 이용한 키 프레임 결정

HSV 컬러 모델은 각각 색상, 채도, 명도를 나타내고 사용자가 직관적으로 인식하는데 바탕을 두는 목적으로 시스템에 이용된다. 검색에 사용되는 값은 위의 RGB에서 값을 구하는 방식과 같이 구한다.

(3) YIQ값을 이용한 키 프레임 결정

YIQ 컬러 모델은 그 특성 방식에 TV 포맷에 중점적인 검색을 하는데 이용되고 검색시 사용하는 값은 역시 RGB에서 값을 구하는 방식과 같이 구한다.

III. 정보 저장과 효과적인 데이터 활용

계산된 정보를 효과적인 활용하기 위해서는 이를 사용자의 목적에 맞게 색인화해야 한다. 다차원 색인 구조를 갖는 R-Tree를 사용함으로써 여러 정보를 손쉽게 저장할 수 있다[6]. R-Tree의 구조는 다음과 같다.

컬러에 대한 이미지 키 X_m 를 R-Tree로 검색하기 위해 본 논문에서 다음과 같은 과정을 거쳐서 R-Tree에 저장될 키로 만든다. τ 는 문턱 값으로써 색인 할 키 값에 유연성을 부여하기 위해서 사용된다.

1) 컬러에 대한 프레임 키 X_m 에 대해 $X_m - \tau$ 의 좌표가 되는 X_m' 를 구한다.

2) 다시 X_m 에 대해 $X_m + \tau$ 가 되는 좌표 X_m'' 를 구한다.

3) $X_m'()$ 과 (X_m'')로 구성되는 사각형 영역을

R-Tree에 삽입 할 키로 사용한다.

4)위와 같은 과정을 각각의 히스토그램 적용한다.

2차원 공간상의 하나의 점으로 표현되는 컬러 특성에 대한 이미지 키는 위와 같은 과정을 거쳐 사각형 영역으로 바뀌게 된다. 만약 서로 유사한 이미지 데이터의 컬러 특성에 의한 이미지 키 K_c 와 K_c' 의 좌표가 특정 범위 내에 있다면 해당 이미지 데이터들의 컬러 분포가 서로 유사한 것임을 알 수 있게 된다. 따라서 R-Tree의 인덱스에 저장된 K_c 의 사각형 영역들은 R-Tree 탐색 시 중첩되어 컬러 특성에 대해 서로 유사한 이미지로 판단하게 된다. 본 논문에서는 컬러 히스토그램을 사용하여 RGB, HSV, YIQ에서 계산된 세 가지의 값을 이미지 키로 구성하였다.

검색 시스템에 유연성을 부여하기 위해 컬러 특징에 대한 이미지 데이터 검색에서는 r 를 조절하는 방법을 사용한다. r 는 최대 값과 계수 값으로부터 R-Tree에 저장 될 사각형 영역을 만들 때, 사용되는 값으로 r 값의 증가/감소는 해당 사각형 영역이 확장 또는 축소하게 되어 히스토그램상의 최대 값과 계수 값에 어느 정도의 유사성이 나타날 때 결정하게 된다.

사용자는 자신의 검색 조건에 맞게 세 가지 컬러에 맞는 방법을 하나에서 세 개까지 조합하여 선택함으로써 이에 맞게 데이터베이스에서 값들을 호출하여 계산함으로써 효과적으로 계산 량을 줄일 수 있고 그에 따른 검색 효과를 높일 수 있다. 이는 그림2에 나타나 있다.

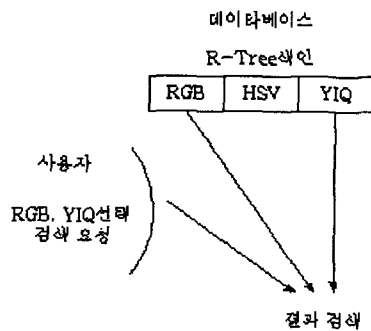


그림 2 효과적인 데이터 검색 구조

IV 실험 결과

4.1 데이터베이스 생성

비디오 데이터베이스의 성능을 측정하기 위해 비디오 정보들을 저장해야 한다. 저장된 비디오는 초당 30프레임을 갖고 320×240 해상도를 가진 MPEG 비디오 포맷이다. 비디오 데이터베이스에 저장된 데이터는 총 MPEG길이 808초로 화면은 331개로 구성되어있으며 각각 꽃 영상이 36개, 뉴스는 2개, 스포츠는 24개이며 영화는 271개의 화면을 가지고 있다.

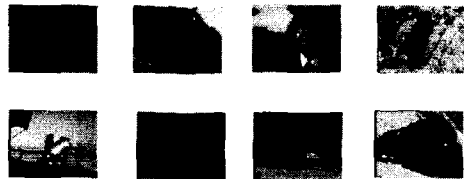
4.2 검색 실험

제시된 질의 이미지에 따른 세 가지 히스토그램의 결과 값을 나타냈다. 결과 이미지는 각 히스토그램마다 가장 질의 이미지와 비슷하다고 계산된 8장의 이미지들을 추출했다. 실험에 사용된 질의 이미지는 그림4이다.



그림 4 질의 이미지

(1) RGB 히스토그램 값을 이용한 결과 이미지



(2) HSV 히스토그램 값을 이용한 결과 이미지



(3) YIQ 히스토그램 값을 이용한 결과 이미지





4.3 검색 결과

검색 결과 세 가지 히스토그램 모두 미미한 검색물을 보이고 있다. 사용자가 요구하는 검색의 다양성과 정확성, 속도를 모두 만족시키기 위해서는 더 많은 컬러 특징뿐만 아니라 모양과 질감 특징 등 여러 값을 이용한 즉, 사용자의 요구에 맞출 수 있는 더 많은 검색 옵션이 있는 시스템이 필요하겠다.

IV. 결론

비디오에서 추출한 정보를 이용해 사용자가 요구하는 비디오를 검색하는 것은 여러 분야에 활용될 수 있다.

본 논문에서는 컬러 값을 이용해 사용자의 요구에 적합한 시스템을 구현해보았다. 그러나 사용자의 요구에 더 부합하고 정확성을 높이기 위해서는 컬러뿐만 아니라 다른 여러 비디오 정보의 특징들을 이용해야 할 필요성이 있었다.

이번 논문은 사용자의 요구에 부합하려는 시도로 인터넷의 발전에 따른 정보를 효과적으로 제공하는 분야에 대한 연구로 멀티미디어분야에 있어서 중요한 과제이며, 효과적인 데이터 베이스 검색시스템을 구현하는 자료로 이용될 수 있을 것이다.

참고문헌(또는 Reference)

- [1] Ullas Gargi, Rangacbat Kasturi, and Susan H. Strayer, "Performance Characterization of Video Shot Change Detection Methods", IEEE Trans. on circuit and systems for video tech., vol 10, no.10, Feb,2000
- [2] Hyun Sung Chang, Sanghoon Sull, Sang Uk Lee, "Efficient Video Indexing Scheme for Content-based Retrieval", IEEE Trans. on circuit and systems for video tech, vol.9, no.8 ,Dec,1999
- [3] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-video Search for Object Appearances", in Proc. IFIP 2nd Working Conf. Visual Database Systems, pp.113-127, 1992
- [4] M. J. Swain, D. H. Ballard, "Color Indexing", Int.J.Comput. Vis., vol.26, no.4, pp.461-470, 1993
- [5] J. Hafner, H. S. Sawhney, W. Equitz, M.

Flickner, and W. Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions", IEEE Trans. Pattern Anal. Machine Intell, vol.17, 1995

- [6] A. Guttman, "R-Trees:A Dynamic Index Structure for Spatial Searching", Proc. ACM SIGMOD, 47-57, 1984.