

논문 2004-41SP-4-14

HAQ 알고리즘과 Moment 기반 특징을 이용한 내용 기반 영상 검색 알고리즘

(Content-Based Image Retrieval Algorithm Using HAQ Algorithm and
Moment-Based Feature)

김 대 일*, 강 대 성*

(Dae-il Kim and Dae-Seong Kang)

요 약

본 논문은 내용 기반 검색 기법에 의한 보다 효율적인 특징 추출 및 영상 검색 알고리즘을 제안하였다. 먼저, MPEG 비디오의 key frame을 입력 영상으로 하여 Gaussian edge detector를 이용하여 객체를 추출하고, 그에 따른 객체 특징들, location feature, distributed dimension feature와 invariant moments feature를 추출하였다. 다음, 제안하는 HAQ (Histogram Analysis and Quantization) 알고리즘으로 characteristic color feature를 추출하였다. 마지막으로 key frame이 아닌 shot frame을 질의 영상으로 하여 제안된 matching 기법에 따라 4가지 특징들의 단계별 검색을 수행하였다. 본 논문의 목적은 사용자가 요구하는 장면이 속한 비디오의 shot 경계 내의 key frame을 검색하는 새로운 내용 기반 검색 알고리즘을 제안함에 있다. 제안된 알고리즘을 바탕으로 10개의 뮤직비디오, 836개의 시험 영상으로 실험한 결과, 효과적인 검색 효율을 보였다.

Abstract

In this paper, we propose an efficient feature extraction and image retrieval algorithm for content-based retrieval method. First, we extract the object using Gaussian edge detector for input image which is key frames of MPEG video and extract the object features that are location feature, distributed dimension feature and invariant moments feature. Next, we extract the characteristic color feature using the proposed HAQ(Histogram Analysis and Quantization) algorithm. Finally, we implement an retrieval of four features in sequence with the proposed matching method for query image which is a shot frame except the key frames of MPEG video. The purpose of this paper is to propose the novel content-based image retrieval algorithm which retrieves the key frame in the shot boundary of MPEG video belonging to the scene requested by user. The experimental results show an efficient retrieval for 836 sample images in 10 music videos using the proposed algorithm.

Keywords : content-based retrieval, moments feature, HAQ algorithm, MPEG video

I. 서 론

최근 컴퓨터와 통신의 급속한 발달로 멀티미디어 데이터가 증가함에 따라서 이를 효율적으로 관리하는 기술에 관한 연구가 활발히 진행 중이다. 멀티미디어 데이터는 문서, 음성, 영상, 비디오 등으로 구성되는데, 그 중 영상이나 비디오 등 영상에 관련된 데이터가 많은

용량을 차지하며 중요한 정보를 가지고 있다. 다양한 종류의 영상 자료를 사용자의 입장에서 효율적으로 검색할 필요성이 증대하고, 대용량 멀티미디어 정보를 효과적으로 검색하기 위해서는 효율적인 색인 및 관리가 필수적이다. 따라서 이를 효율적으로 검색하기 위한 특징 추출 기법이 요구되며, 이는 영상 관련 데이터 뿐 아니라 전체 멀티미디어 데이터의 관리 및 활용에 직접 관련된다. 내용 기반 검색^[1]은 키워드와 관계없이 영상의 자체 정보인 시각적인 예(visual example)로써 사용자가 원하는 영상을 쉽게 browsing하고, 질의, 검색할

* 정회원, 동아대학교 전자공학과
(Dept. of Electronic Eng., Dong-A University)
접수일자: 2003년12월2일, 수정완료일: 2004년5월28일

수 있는 기법이다. 즉, 컬러(color), 모양(shape), 그리고 질감(texture) 등의 내용 표현 요소들을 통하여 얻어진 특징 값을 사용하여 영상들 간의 유사도를 계산하고 검색을 수행한다. 그러므로 효율적인 영상 검색을 위해서는 무엇보다도 각 영상을 대표하는 표현 요소들의 효과적인 특징 추출이 중요하다. 영상의 표현 요소들 중에서 컬러 특징(color feature)은 영상을 표현하는 주된 특징으로 내용 기반 영상 검색의 초기부터 사용되어 온 방법이다. 컬러 특징은 영상 내의 조명 변화나 관측 위치, 크기 변화 등에 어느 정도 무관하게 적용 가능하다는 장점을 가지고 있으며 방법이 간단하여 내용 기반 영상 검색에서 가장 많이 사용되는 방법 중의 하나이다. 컬러 특징을 추출하는 대표적인 방법이 컬러 히스토그램(color histogram)을 이용하는 것이다^[2]. 컬러 히스토그램을 특징 정보로 하면 컬러의 수 만큼에 해당하는 크기의 공간이 필요하므로 보통 유사한 컬러 별로 군집하는 컬러 집합을 사용하게 된다. 초기에는 히스토그램을 균등한 비율로 나누어 히스토그램 구간(bin)을 결정하는 방식으로 컬러를 군집했지만, 효율적인 검색 시스템을 위해서 유사한 컬러끼리 군집하는 방법들이 제안되었다. 단, 컬러 히스토그램이나 컬러 집합을 이용하면 영상 내 객체들의 배치 즉 컬러 분포의 공간적인 정보를 전혀 이용할 수 없다는 단점이 있다.

질감은 영상 내부의 동일한 패턴을 가지는 영역에 대한 특징으로 하나의 컬러 정보로만은 표현할 수 없다. 질감은 영상 내의 객체의 표면이 가지는 영상에게는 매우 중요한 특징이 된다. 특히 항공사진을 분석하는 경우에 질감 정보가 많이 사용되어 왔다. Gabor 필터, wavelet 변환 등을 통해서 질감 특징을 분석하는 방법이 연구되고 있는데^[3], 이러한 방법은 인공적인 질감 즉 건물의 벽돌 무늬나 의류에서 나타나는 무늬 등의 분석에 큰 효과를 보이고 있다.

영상의 내용을 파악하기 위해서는 무엇보다도 영상 내의 객체(object)의 형태(shape)에 관심을 가질 필요가 있다. 형태 분석 및 표현 방법은 크게 형태의 외곽선 정보를 이용한 것과 영역 내부를 이용한 것으로 분류가 가능하다. 외곽선 정보를 이용하는 방법은 영상에서 외곽선을 추출하여 이의 분석을 통해 그 변화를 특징 정보로 삼는 것이다. 영역 내부 정보를 이용하는 것은 영상 전체의 화소의 분포를 분석하는 것으로 영상에 대한 모멘트를 구하는 방법과 같이 형태의 분포 정도를 특징 정보로 삼는 것이다^[4]. 또한 분석 결과에 따라서 연산 결과가 일련의 수치 데이터로 나오는 경우와 영상으로

결과가 나오는 경우로 분류해 볼 수 있다. 전자를 스칼라(scalar) 변환 방법이라 하고, 후자를 공간상의 방법이라고 한다. 스칼라 변환에는 Fourier 기술 등이 있고, 공간상의 방법에는 중앙 축 변환(medial axis transform) 등이 있다.

위에서 설명한 여러 가지 특징들을 추출하여 효과적으로 사용함으로써 내용 기반 영상 검색의 효율성을 높일 수 있다. 본 논문에서는 Laplacian of Gaussian edge detector를 이용한 객체 추출 기법과 그에 따른 객체 특징들, location feature, distributed dimension feature와 invariant moments feature^[5]를 추출하고, HAQ 알고리즘으로 characteristic color feature를 추출하여, 제안하는 matching 기법에 따라 검색을 하는 새로운 내용 기반 검색 알고리즘을 제안한다. 또한 본 논문은 MPEG 동영상에서 shot 경계 검출 알고리즘^[6]을 적용하여, shot frame들을 추출하고, 각 shot 경계를 검출한 후, 각 shot 경계 내에서 key frame 추출 알고리즘^[7]을 이용하여 key frame들을 뽑아 database 내에 저장한다. 본 논문의 목적은 제안하는 내용 기반 검색 알고리즘을 이용하여 query 영상에 가장 잘 부합하는 key frame을 결과 영상으로 정확하게 검색함으로써 user가 검색하고자 하는 query 영상을 포함하는 비디오와 비디오 내의 key frame을 검색해, still 영상 하나라도 검색하고자 하는 비디오와 장면을 검색하는 데 있다.

이후 논문은 다음과 같이 구성되어있다. 제 II장에서 객체 추출과 객체 특징 추출 기법을 소개한다. 제 III장에서는 컬러 특징 추출 기법인 HAQ 알고리즘에 대해 논한다. 제 IV장에서는 전체적인 검색 알고리즘에 대해서 살펴본다. 제 V장에서는 실험을 통해 효과적인 검색 성능을 측정한다. 마지막으로 제 VI장에서는 간단한 결론과 추후 연구에 대하여 논한다.

II. 객체 특징 추출

제안한 객체 특징들은 location feature, distributed dimensions feature와 invariant moments feature이다. 이 특징들을 추출하기 위해 우선 database 내의 객체 정보를 얻는다. 객체 추출 알고리즘은 다음과 같다.

Step 1. image의 gray를 입력받는다.

Step 2. gaussian detector로 edge를 검출한다.(edge image I_e 생성)

Step 3. 검출된 edge의 선을 모두 연결하여, 폐곡선을

생성한다.

$$t_{k,j} = \sum_{k=0}^5 m_{(i-2)(j+k-2)} \quad \text{for } m_{ij}$$

$$t_{(i-1)(j-1)} = m_{ij} \quad \text{for } k < 2$$

$$t_{(i-1)(j)} = m_{ij} \quad \text{for } k = 2$$

$$t_{(i-1)(j+1)} = m_{ij} \quad \text{for } k > 2$$

Step 4. 폐곡선 안을 통합하고, 객체 정보로 추출한다.(object image I_o 생성)

$$O(x, y_k) = \text{the same object}$$

$$\text{for } Start_{edge} < k < End_{edge}$$

Step 5. 원 image(I)에 객체 정보를 mask한다.(masking image I_m)

$$I_m(n, m) = I_o(n, m)$$

여기서, (n, m) 은 객체 정보 픽셀이다.

추출된 객체에 대한 특징 중 첫 번째 invariant moments feature는 scale, translation 그리고 rotation에 불변한 비선형 함수를 영상에 적용하여, 영상의 기하학적 moments를 정의하는 것이다. 이 특징은 형태가 유사한 key frame들을 포괄적으로 검색하는데 쓰인다. 이 차원 $M \times M$ 영상 $g(x, y)$ 가 주어지면, 여기서 $g(x, y)$ 는 $\{g(x, y), x, y = 0, \dots, M-1\}$ 이고, $(p+q)$ 번째 기하학적 moment는 다음과 같이 정의된다.

$$m_{pq} = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} x^p y^q g(x, y) \quad \text{for } p, q = 0, 1, 2, \dots$$

서로 다른 영상에 대한 m_{pq} 의 동적 범위를 유지하기 위해 $M \times M$ 영상 평면은 $x \in [-1, +1], y \in [-1, +1]$ 범위의 평면으로 mapping된다.

$$m_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} x^p y^q g(x, y)$$

translation에 불변한 moments를 만들기 위해, central moment를 다음과 같이 정의한다.

$$\mu_{pq} = \sum_{x=-1}^{+1} \sum_{y=-1}^{+1} (x-x')(y-y')^q g(x, y)$$

여기서, $x' = \frac{m_{10}}{m_{00}}$ 이고 $y' = \frac{m_{01}}{m_{00}}$ 이다. central moments는 scale 변화에 불변하기 위해 다음과 같이

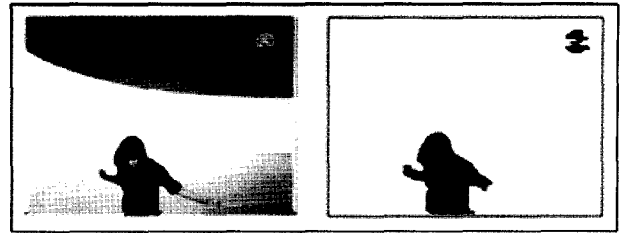


그림 1. 원영상과 객체 추출 영상

Fig. 1. Original image and Image of an extracted object.

normalization된다.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \quad \text{with } \gamma = \frac{p+q}{2} + 1$$

rotation, translation, scale 변화에 불변한 비선형 함수 집합의 첫 여섯 개의 함수들을 다음과 같이 정의한다.

$$\phi_1 = \eta_{20} + \eta_{02}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{00} + \eta_{20})[(\eta_{00} + \eta_{12})^2 - 3(\eta_{21} + \eta_{11})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{00} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} - \eta_{12})(\eta_{03} + \eta_{12})$$

ϕ_1 에서 ϕ_6 의 값은 매우 작다. 따라서 영상을 표현하는 특징으로 $\log \phi_i, i=1, \dots, 6$ 을 사용한다. 그림 1은 원 영상과 객체가 추출된 영상 그리고 표 1은 그에 따른 invariant moment feature를 나타내었다.

동일 shot 경계 내의 영상들의 객체는 일정한 범위와 크기를 가지게 되는데, 이에 제안하는 특징은 location feature와 distributed dimension feature이다. location feature $l_o(x, y)$ 는 다음과 같이 정의된다.

$$l_o(x, y) = \left(\frac{1}{n_{ox}} \sum_{k=0}^{n_x} O_{xk} \right) \left(\frac{1}{n_{oy}} \sum_{k=0}^{n_y} O_{yk} \right)$$

여기서, n_{ox} 와 n_{oy} 는 각각 전체 x 축과 y 축이며, O_x 와 O_y 는 객체의 x 축 좌표와 y 축 좌표이다. Distributed dimension feature D_o 는 다음과 같이 정의된다.

$$D_o = \frac{n_o}{n_u}$$

표 1. 불변 모멘트 특징 값
Table 1. Invariant moments feature.

-11.957397
-28.456831
-26.989985
-26.940418
-55.070702
-41.763634

여기서, n_0 는 객체가 차지하는 픽셀 수이며, n_u 는 전체 픽셀 수이다.

III. 컬러 특징 추출

사람의 시각에서는 인식하기 위해 유사한 컬러들을 서로 grouping한다. 여기에 착안하여, 영상의 유사한 컬러들을 grouping하는 것과 검색을 위한 feature로써 characteristic color들을 추출하는 기법을 소개한다. 제안하는 컬러 특징 추출 기법은 영상에 대한 컬러 palette^{[8][9]}의 디자인과 유사하다. 이것은 원 color image에서 제한된 컬러 영상의 quantization이 필요하다. HAQ 기법은 영상의 컬러 component의 상호 관계를 활용하기 위해 영상의 컬러 space를 분할한다.

HAQ algorithm은 크게 두 가지로 구성되는데, the bit allocation^[10]과 histogram analysis method가 바로 그것이다. 영상의 전체 좌표 space를 I 로 표시하면, I 의 컬러 value를 $I=(I_1, I_2, I_3)$ 로 표시한다. 여기서 I_i 는 컬러 component value이다. bit allocation의 목적은 영상의 각 컬러 component에 주어진 bit 몫을 자동적으로 할당하기 위함이다. 즉 bit B 의 고정된 수를 컬러 component (I_1, I_2, I_3) 사이에서 나누어 놓는다. 일반적으로, 변화가 심한 컬러 component에 더 많은 bit들을, 변화가 적은 쪽에 더 작은 bit들을 할당한다. 만약 데이터가 특정 컬러 component에 퍼져있으면, 그 데이터의 component는 다른 것보다 더 중요하기 때문에 더 조밀하게 grouping한다. 즉, 중요한 component는 quantization error를 줄이기 위해 더 많은 bit들이 주어질어야만 하는 것이다. Histogram analysis method는 quantization threshold level을 결정짓기 위한 기법이다. 각 color component에 대한 histogram을 분석해 보면, count가 크게 변화하는 구간이 존재하기 마련이다. 이 구간을 grouping하기 위해 threshold level이 필요하고,

histogram analysis method로 이 level을 결정한다. 다음은 HAQ 알고리즘을 이용하여 characteristic 컬러를 추출하는 과정이다.

Step 1. 입력 영상들을 RGB space로 표현한다.

Step 2. 각 color component(I_i)에 주어진 bits B 의 고정된 수를 할당한다.

$$I_i=B, \quad i=1,2,3$$

Step 3. I_i 의 bit allocation 근사 solution을 구한다.

$$B_i=B+\frac{1}{2}\log_2\frac{\sigma_i^2}{[\prod_{j=1}^3\sigma_j^2]^{\frac{1}{3}}}, \quad j=1,2,3, \dots$$

$$\text{여기서, } \sigma^2=\frac{1}{n}\sum_{i=0}^n(x_i-m)^2$$

Step 4. Bit allocation procedure로 I_i 의 histogram analysis procedure 반복 회수를 결정한다.

$$r_i=2^{B_i}-1, \quad i=1,2,3$$

Step 5. histogram H_{ij} 을 구한다.

$$H_{ij}=\frac{1}{n}\sum_{k=0}^{L-1}n_{ik}(z_{ik})^j, \quad j=1,2,3, \dots$$

여기서, n 은 총 입력 픽셀 수, L 은 histogram I_i 의 총 분포 level, n_{ik} 은 histogram I_i 내 z_{ik} 값을 가지는 픽셀 수이다.

Step 5. 다음 과정을 r_i 번 반복한다.

I_i 의 threshold level $t_i(m_{r_i})$ 을 찾는다.

$$t_i(m_{r_i})=\text{MIN}(n_{ik})$$

$$\text{for } S_i \leq n_{ik} \leq E_i, \quad k=0,1,2, \dots, L-1$$

여기서, S_i 는 histogram I_i 내의 첫 maximum n_{ik} 이고, E_i 는 histogram I_i 내의 마지막 maximum n_{ik} 이다.

Step 6. Step 6에서 구한 threshold level $t_i(m_{r_i})$ 에 따른 각 color point들의 중심을 color palette의 특징 color($C_{i,(m_{r_i})}$)로 지정한다.

$$C_{i,(m_{r_i})}=\frac{\sum_{k=0}^{t_i(m_{r_i})}n_{ik}z_{ik}}{\sum_{k=0}^{t_i(m_{r_i})}n_{ik}}, \quad i=1,2,3$$

그림 2는 각 컬러 component의 histogram에서 threshold level을 도시한 그림이다. 그림 3은 pepper 영상에 HAQ algorithm을 적용하여 characteristic color들로 이루어진 color palette와 복원된 영상을 나타낸 그

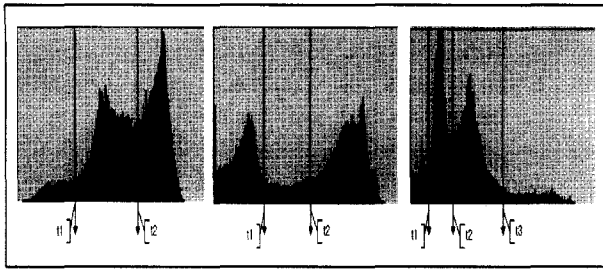


그림 2. 각 컬러 component의 threshold level
Fig. 2. threshold level of each color component.

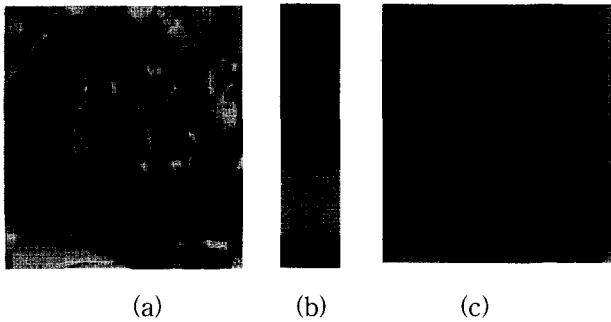
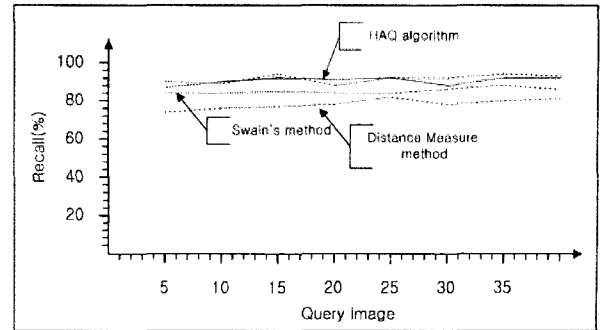


그림 3. HAQ 알고리즘을 이용하여 추출한 컬러 특징
(a) 원 영상 (b) 컬러 palette (c) 복원 영상
Fig. 3. The extracted color feature using HAQ algorithm
(a) original image (b) color palette (c) restoration image.

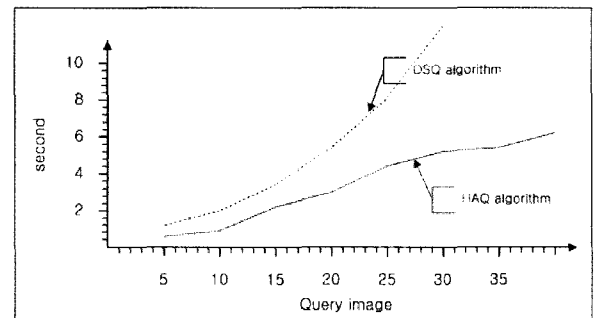
림이다. 본 논문에서 제안하는 HAQ 알고리즘은 Soo-Chang Pei와 Ching-Min Cheng이 제안한 DSQ(Dependent Scalar Quantization) 알고리즘^[11]과 같은 결과를 보인다. 그러나 DSQ 알고리즘보다 계산량을 줄여 경제적인 특징 추출을 하였다. DSQ 알고리즘은 이미 여러 다른 컬러 특징 추출 알고리즘 보다 우수함을 나타내었다. HAQ 알고리즘이 이러한 DSQ 알고리즘의 계산 시간을 줄인 결과를 그림 4에 나타내었다.

IV. 검색 알고리즘

본 영상 검색 알고리즘의 전반적인 구조는 그림 5와 같다. 이 구조는 크게 MPEG 비디오 분석 섹션, 원 영상들의 특징 추출 및 database 저장 섹션, query 영상 feature와 database 영상 feature와의 matching 섹션의 세 가지로 구성된다. 첫 번째로 비디오 분석 과정에서는 [6]과 [7]의 알고리즘을 거쳐 key frame이 추출되며, 이 key frame이 원 영상이 된다. 두 번째로 특징 추출 및 저장 과정에서는 원 영상들이 먼저 특징 추출 카테고리들을 거쳐 4개의 특징들이 추출되어 database에 저장된다. 마지막으로 query 영상과 database 영상의 matching 과정은 다음과 같다.



(a)



(b)

그림 4. HAQ 알고리즘의 성능 비교
(a) 검색 효율성 비교 (b) 검색 시간 비교
Fig. 4. The performance comparison of the HAQ Algorithm (a) the comparison of the retrieval efficiency (b) the comparison of the retrieval time.

Step 1. 질의 영상의 invariant moments feature(F_{mi})를 입력받는다.

Step 2. F_{mi} 의 invariant moments threshold(T_{mi}) 내 모든 영상을 database에서 검색한다.

$$I(F_{DBmi}), \quad -T_{mi} < F_{DBmi} < T_{mi}$$

여기서, $I(F_{DBmi})$ 는 검색된 영상이고, F_{DBmi} 는 database내의 invariant moments feature이다.

T_{mi} 는 실험값이다.

Step 3. 질의 영상의 location feature(F_{ql})를 입력받는다.

Step 4. F_{ql} 의 location threshold(T_l) 내 모든 영상을 $I(F_{DBmi})$ 에서 검색한다.

$$I(F_{DBl}), \quad -T_l < F_{DBl} < T_l$$

여기서, $I(F_{DBl})$ 는 검색된 image이고, F_{DBl} 은 $I(F_{DBmi})$ 내의 location feature이다. T_l 은 실험값이다.

Step 5. 질의 영상의 distributed dimensions feature(F_{qr})를 입력받는다.

Step 6. F_{qr} 의 distributed dimensions threshold(T_r)

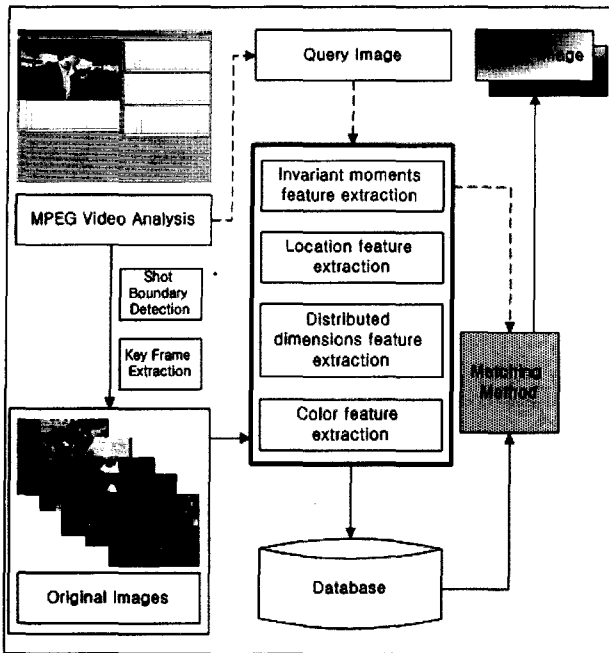


그림 5. 제안된 영상 검색 알고리즘의 전체적인 구조
 Fig. 5. The general structure of the proposed image retrieval algorithm.

내 모든 영상을 $I(F_{DBI})$ 에서 검색한다.

$$I(F_{DBr}), \quad -T_r < F_{DBr} < T_r$$

여기서, $I(F_{DBr})$ 는 검색된 영상이고, F_{DBr} 은 $I(F_{DBI})$ 내의 distributed dimensions feature이다. T_r 은 실험값이다.

Step 7. $I(F_{DBI})$ 내의 color feature를 histogram내 비중이 높은 color부터 차례로 가중치(W_n)를 두어 비교한다.

Step 7-1) 다음을 n번 반복하여 W_n 을 결정한다.

$$W_n = B_c - n,$$

$$B_c = \# \text{ of color palette's bits}$$

Step 7-2) W_n 가 가장 높은 순서대로 각 color threshold(T_c)내의 영상을 선택한다.

Step 8. 가장 근접한 color palette를 가진 영상들을 검색하여 출력한다.

V. 실험 결과 및 고찰

본 논문에서는 제안한 알고리즘의 검색 효율성을 평가한다. 또 객체 특징들에 의한 검색 성능과 컬러 특징들에 의한 검색 성능, 그리고 제안한 객체 정보와 컬러 정보를 이용한 검색 성능을 비교 평가하고자 한다. 제안한 알고리즘을 이용하여 10개의 비디오 데이터에서

표 2. 객체 특징과 컬러 특징을 이용한 검색 성능 분석 결과

Table 2. The results of comparison using the object features and the color features.

0.72	0.75	0.66	0.69	0.70	0.81	0.81	0.79	0.74	0.69	0.64
0.41	0.43	0.36	0.36	0.39	0.32	0.47	0.47	0.40	0.34	0.40
0.59	0.63	0.64	0.63	0.55	0.61	0.63	0.70	0.56	0.54	0.61
0.37	0.36	0.36	0.33	0.31	0.33	0.36	0.42	0.37	0.31	0.35
0.57	0.58	0.57	0.61	0.55	0.59	0.58	0.63	0.63	0.55	0.59
0.29	0.29	0.28	0.31	0.27	0.30	0.30	0.31	0.33	0.28	0.30
0.62	0.65	0.65	0.63	0.65	0.59	0.69	0.59	0.61	0.58	0.63
0.68	0.70	0.71	0.69	0.62	0.60	0.63	0.62	0.62	0.61	0.65

추출한 총 836개 key frame들의 특징들로 데이터베이스를 구성하고, 각 비디오 데이터에서 key frame이 아닌 10개의 frame들을 질의 영상으로 하여 총 100개의 질의 영상을 실험한다. 앞서 II장과 III장에서 객체 특징과 컬러 특징을 추출하는 방법에 대하여 살펴보았다.

표 2는 객체 특징들과 컬러 특징들을 이용하여 검색 성능을 분석한 것이며, 성능 평가 척도인 Recall과 Precision을 이용하여 검색의 효율성을 평가하였다. 객체 정보를 이용한 invariant moments feature, location feature, dimension feature, 그리고 HAQ 알고리즘을 이용한 color feature의 4가지 방법을 비교한 결과를 나타낸 것이다. 각각 10개의 query class에 대해 비교한 결과, recall 측면에서는 Invariant moments feature와 color feature가 우수했고, precision 측면에서는 color feature의 검색율이 우수하였다. 따라서 각각의 feature에 대한 검색 성능은 color feature가 다른 feature에 비해서 보다 나은 검색 효율을 나타내는 결과를 보였다.

본 논문에서 최종적으로 제안하는 객체 특징과 컬러 특징을 병합한 검색 기법에 대한 효율성을 비교, 분석한다. 표 3은 객체 특징을 이용한 검색 기법과 컬러 특징을 이용한 검색 기법, 그리고 제안하는 객체 특징과 컬러 특징을 병합한 검색 기법에 대한 성능을 비교, 분석한 것이다. recall과 precision 외에 accuracy는 질의를 포함하는 비디오 데이터 내의 질의와 동일한 shot 경계에 있는 key frame을 검색 결과의 3순위 내 포함

표 3. 각 특징 추출 방법을 이용한 검색 기법과 제안한 검색 기법과의 성능 분석
Table 3. Analysis between retrieval using each feature extraction and the proposed retrieval.

0.83	0.84	0.79	0.80	0.72	0.85	0.89	0.85	0.79	0.77	0.81
0.80	0.83	0.75	0.75	0.70	0.77	0.83	0.73	0.68	0.73	0.76
0.64	0.59	0.88	0.59	0.66	0.80	0.67	0.58	0.55	0.64	0.62
0.62	0.65	0.65	0.63	0.65	0.59	0.69	0.59	0.61	0.58	0.63
0.66	0.70	0.71	0.69	0.62	0.60	0.63	0.62	0.62	0.61	0.65
0.66	0.72	0.69	0.69	0.67	0.70	0.66	0.73	0.63	0.66	0.69
0.89	0.91	0.91	0.94	0.85	0.96	0.91	0.98	0.88	0.90	0.91
0.83	0.88	0.88	0.88	0.79	0.89	0.89	0.93	0.90	0.85	0.87
0.93	0.97	0.90	0.89	0.94	0.90	0.90	0.93	0.98	0.91	0.93

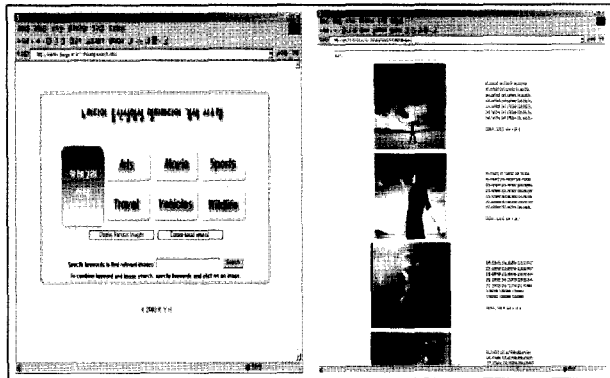


그림 6. 실시간 내용 기반 영상 검색 시스템
(a) 시작 화면 (b) 질의 화면
Fig. 6. Real-time content-based image retrieval system
(a) start page (b) query page.

여부를 판단하는 항목이다. 결국 본 논문의 목적인 하나의 영상을 질의로 비디오의 shot 부분을 찾아내는 검색 성능을 보여준다. 세 가지의 특징을 병합한 객체 특징 검색 기법은 recall과 precision이 컬러 특징 보다 월등히 높지만, accuracy 부분에서 컬러 특징보다 성능이 떨어짐을 보인다. 반면에 컬러 특징은 recall과 precision의 성능은 떨어지나, accuracy가 상대적으로 높다. 본 논문에서 제안하는 객체 특징과 컬러 특징을 병합한 검색 기법은 이 둘의 장점을 살려 평균 93%의 높은 accuracy를 보여주고 있다. 즉, still 질의 영상 하나로 질의가 포함되는 비디오는 물론, 비디오 데이터 내 질의가 포함되는 shot 경계까지 검색이 가능함을 본 실험

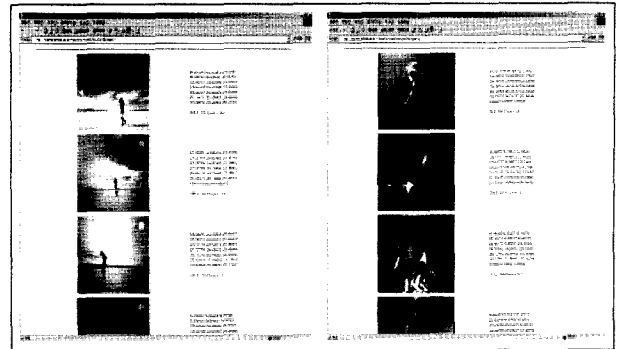


그림 7. 질의에 대한 결과 영상
Fig. 7. The result images about the query image.

을 통해 알 수 있다. 본 논문에서 제안한 특징 추출 및 검색 기법을 적용하여 실시간 데이터베이스 검색 시스템을 구축하였다. 데이터베이스 상에는 key frame 과 특징 값들이 저장되는 공간과 질의 영상들과 특징 값들이 저장되는 공간, 두 가지 저장 공간으로 분류된다. 이 시스템은 사용자가 선택한 질의 영상에 대하여 matching 과정을 통해 그에 따른 결과 영상들을 matching 순서대로 화면에 출력한다. 본 검색 시스템은 카테고리 선택 부분과 내용 기반 검색 부분으로 나뉘어져 있다. 그림 6은 본 영상 검색 시스템의 시작 화면과 질의 영상 화면을 보여주고 있고, 그림 7은 질의 영상 선택 후 화면에 출력된 matching 결과 영상들이다.

VI. 결론 및 추후 연구

본 논문에서는 사용자의 질의에 부합하는 MPEG 비디오의 장면을 찾는 새로운 내용 기반 영상 검색 알고리즘을 제안하였다. 객체의 3가지 특징들과 HAQ 알고리즘을 통해 추출된 컬러 특징을 제안한 matching 기법을 통해 결과 영상을 검색하였다. 실험의 결과 제안된 알고리즘은 질의 영상이 속한 shot 경계 내의 key frame을 효율적으로 검색할 수 있었다. 실험에서 836개의 영상들을 실험 대상으로 하여 검색 효율을 분석하였다. 100개의 질의 영상들에 대하여 recall과 precision으로 평가한 결과, 제안한 영상 검색 기법이 각각 91%와 87%의 검색 효율을 보였다. 또한 accuracy는 93%로, 제안된 알고리즘은 질의 영상이 속한 shot 경계 내의 key frame을 효율적으로 검색할 수 있었고, 이를 인터넷 실시간 데이터베이스로 구축하였다. 향후 좀 더 정확한 객체 추출 기법에 대한 연구와 많

은 실험을 통한 정밀한 threshold 값을 찾는 연구가 수행하여야 할 것이고, 나아가 동영상과 연결된 검색 시스템을 구현하고자 한다.

참고 문헌

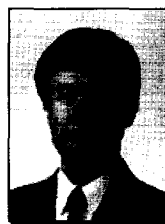
- [1] Y. H. Ang, Zhao Li and S. H. Ong, "Image Retrieval based on Multidimensional Feature Properties", The International Society for Optical Engineering(SPIE), vol. 2420, pp. 47-57, 1995.
- [2] M. J. Swain and D. H. Ballar, "Color index", Int. J. Computer Vision, vol. 7, no. 1, pp. 11-32, July 1991.
- [3] Gross M. H., Koch R., "Visualization of multidimensional shape and texture features in laser range data using complex-valued Gabor wavelets", IEEE Transactions on Visualization and Computer Graphics, vol. 1, pp. 44-59, March 1995.
- [4] Willian I. Grosky, Peter Neo, Rajiv Mehrotra, "A Pictorial Index Mechanism for Model-based Matching", Data and Knowledge Engineering, vol. 8, pp. 309-327, 1992.
- [5] A. Khotanzad and Y. H. Hong, "Rotation and Scale Invariant Features for Texture Classification", Proc. of IASTED Int. Symposium Robotics and Automation, Santa Barbara, CA, USA, pp. 16-17, May 1987.
- [6] 성장우, 강대성, "Multi-Pass 구조를 가지는 Shot 경계 검출 기법", 한국 신호처리 시스템 학회 논문지, 제 1권, 1-1호, pp. 58-63, 2000. 10.
- [7] 강대성, 김영호, "멀티미디어 검색을 위한 shot 경계 및 대표 프레임 추출", 한국 신호처리 시스템 학회 논문지, 제 2권, 1호, pp. 38-43, 2001. 1.
- [8] P. Heckbert, "Color image quantization for frame buffer display", Comput. Graph., vol 16, no. 3, pp. 297-307, July 1982.
- [9] M. T. Orchard and C. A. Bouman, "Color quantization of images", IEEE Transactions on Signal Processing, vol. 5, pp. 124-139, 1995.
- [10] A. Segall, "Bit allocation and encoding for vector sources", IEEE Transactions on Information Theory, vol. IT-22, pp. 162-169, Mar 1976.
- [11] S. C. Pei and C. M. Cheng, "Dependent scalar quantization of color images," IEEE Transactions on Circuits Systems and Video Technology, vol. 5, pp.124-139, April 1995.

저자 소개



김대일(정회원)
2002년 동아대학교 전자공학과
학사 졸업.
2004년 동아대학교 전자공학과
석사 졸업.
현재 동아대학교 정보기술연구소
전임연구원.

<주관심분야: 영상 검색, 영상 압축, 신호처리>



강대성(정회원)
1984년 경북대학교 전자공학과
학사 졸업.
1991년 Texas A&M 대학교
전자공학과 석사 졸업.
1994년 Texas A&M 대학교
전자공학과 박사 졸업.

현재 동아대학교 전기전자컴퓨터공학부 부교수.
<주관심분야: 영상처리, 비디오 색인 및 검색, 영상 압축>