

# 영역 기반의 영상 질의를 이용한 내용 기반 영상 검색

정회원 김낙우\*, 송호영\*, 종신회원 김봉태\*\*

## Content-based image retrieval using region-based image querying

Nac-woo Kim\*, Ho-young Song\* *Regular Members*, Bong-tae Kim\*\* *Lifelong Member*

### 요 약

본 논문에서는 효과적인 영상 검색을 위한 방법으로서 JSEG 영상 분할 기법을 통한 영역 기반의 영상 인덱싱 및 검색 기법을 제안한다. JSEG은 영상을 색상 분류에 따라 양자화하고 이에 영역 윈도우를 적용시켜 J-image를 만든 다음, 세부 분할된 영역의 성장과 병합을 통하여 영상을 효과적으로 분할하는 방법이다. 제안하는 영상 검색 시스템은 JSEG에 의해 분할된 영상을 사용자에게 질의 영상으로 주고, 사용자로부터 하여금 분할 영상에서 관심 영역군(群)을 선택하게 한다. 그리고 나서, 사용자 질의에 의해 선택된 영역의 MBR을 구하고 이 영역의 중심을 기준으로 다중 윈도우 마스크를 생성하여 적용시킴으로써 특정 관심 영역을 중심으로 한 영상의 전역적인 특징을 추출한다. 최종적으로 추출된 특징의 성능 비교를 위한 기술자로는 누적 히스토그램을 이용하였다. 제안된 방법은 특정 영역에서의 특징과 전역 특징을 동시에 추출하여 검색에 이용함으로써 보다 빠르고 정확하게 사용자가 원하는 영상을 제공할 수 있다. 실험 결과는 영상 색인 및 검색에 있어서 제안된 방법이 영상 기반의 검색 기법과 비교하여 더 효과적임을 보여준다.

**Key Words** : Region-based Image Retrieval, JSEG Segmentation, Minimum Boundary Rectangle.

### ABSTRACT

In this paper, we propose the region-based image retrieval method using JSEG which is a method for unsupervised segmentation of color-texture regions. JSEG is an algorithm that discretizes an image by color classification, makes the J-image by applying a region to window mask, and then segments the image by using a region growing and merging. The segmented image from JSEG is given to a user as the query image, and a user can select a few segmented regions as the query region. After finding the MBR of regions selected by user query and generating the multiple window masks based on the center point of MBR, we extract the feature vectors from selected regions. We use the accumulated histogram as the global descriptor for performance comparison of extracted feature vectors in each method. Our approach fast and accurately supplies the relevant images for the given query, as the feature vectors extracted from specific regions and global regions are simultaneously applied to image retrieval. Experimental evidence suggests that our algorithm outperforms the recent image-based methods for image indexing and retrieval.

### I. 서 론

내용 기반 영상 검색에 대한 연구를 지속하고 있

는 많은 수의 연구자들은 ‘어떻게 하면 영상에서의 전역적인 시각 특징을 보다 효과적으로 추출할 수 있을까’에 대한 의문을 해결하기 위해 많은 노력을

\* 한국전자통신연구원 광대역통합망연구단 광통신연구센터 FTTH서비스팀(nwkim, hsong)@etri.re.kr

\*\* 한국전자통신연구원 광대역통합망연구단 광통신연구센터 (bkim@etri.re.kr)

논문번호 : KICS2007-04-166, 접수일자 : 2007년 4월 9일, 최종논문접수일자 : 2007년 9월 7일

기술어왔다<sup>[1-3]</sup>. 이러한 연구는 내용 기반 영상 검색의 기초를 다지는 데에 큰 기여를 하였으나, 실제적인 영상의 검색 성능은 사용자가 원하는 기대 수준과는 여전히 거리가 있었다. 문제의 원인은 영상이 본래 가지는 의미와 추출된 시각 특징과의 차이에서 기인한다. 최근 이러한 차이를 좁히기 위해서 영역 기반의 영상 검색 기법에 대한 연구가 활발히 진행되고 있다<sup>[4-7]</sup>.

전체 영상이나 단순한 공간 재분할 영상에서의 특징 추출이 아닌 주어진 영상에서의 특정 관심 영역에 대해 사용자가 부분 질의를 할 수 있도록 함으로써 좀 더 영상의 의미에 기반을 둔 영상 검색이 이루어질 수 있도록 하는 것이다. 영역 기반의 영상 검색 기법을 적용함으로써 사용자가 배경은 고려하지 않고 '초록색 나비'를 질의하였을 때, 검색 시스템이 '들판'이나 '나무'가 포함된 잘못된 영상을 결과물로 내놓지 않도록 할 수 있다. 대표적인 영역 기반의 검색 시스템으로써는 Netra<sup>[6]</sup>, Blobworld<sup>[7]</sup> 등을 들 수 있다.

초기의 영역 기반 시스템은 대부분 영역 대 영역 간의 유사도를 비교하고 있다. Netra<sup>[6]</sup>는 웹상에서 구현된 시스템으로서 사용자의 파라미터 조정과 분할된 영상의 영역 보정을 지원함으로써 영상 내의 객체나 영역에 대한 검색을 지원하고 있다. Blobworld<sup>[7]</sup>는 기대치-최대화 기법을 이용하여 화소 기반으로 영상을 분할하고, 분할된 각 영역을 주요 색상과 질감, 모양 등으로 표현함으로써 이를 검색에 이용하는 기법을 말한다. 이러한 방법은 사용자가 하나의 영역 혹은 여러 개의 영역을 선택하여 질의하고, DB

(DataBase:데이터베이스)로부터 이에 대응하는 영역을 일대 일로 검색하는 영역 대 영역의 방식이나, 최신의 컴퓨터 비전 기술로서도 객체 영역을 정확히 분할할 수 없다는 점과 사용자가 이러한 검색 도구를 쉽게 이용하기도 어렵다는 점에서 그 한계가 있다. 또한, 질의에 있어서 단순히 중심 객체뿐 아니라 주변 배경 역시 일정 수준의 의미를 가질 경우 이러한 영역 대 영역 방식의 한계는 더욱 분명해진다.

우리는 본 논문에서 JSEG 알고리즘을 이용하여 각 영상을 효과적으로 객체 분할하고 분할된 객체 영역을 바탕으로 다중 윈도우 마스크를 생성함으로써 부분적이면서 전체적인 영상의 특징을 추출하였다. Deng & Manjunath<sup>[8]</sup>에 의해 제안된 JSEG 알고리즘은 영상을 몇 개의 대표 색상으로 양자화하고 이를 이용하여 영상 주파수 지도인 J-image를 구성함으로써 영상을 공간 분할하는 기법이다. JSEG 기법에 의해 객체 분할된 영상 중 일부가 사용자의 선택에 의해 질의되고, 이렇게 질의된 영역의 MBR(Minimum Boundary Rectangle:최소 영역 경계 상자)을 기준으로 다중 윈도우 마스크를 생성함으로써 전체 질의 영상의 특징을 마스크에 의한 가중치를 부여하면서 추출하였다. 최종적으로, 추출된 특징 벡터 간의 유사도 비교를 통해 효과적인 영상 검색을 제공하기 위한 영역 기반의 영상 검색 시스템을 구현하였다.

이 논문의 구성은 다음과 같다. 먼저, II장에서는 Deng & Manjunath가 제안한 JSEG 알고리즘을 소개하고, III장에서는 제안하는 영역 기반의 특징 추출 기법과 유사도 비교 방법에 대해 논의한다. IV장과 V장에서는 이를 통한 실험 결과와 결론을 도출하고 있다. 그림 1은 본 연구의 대략적인 블록도를 나타낸 것이다.

## II. JSEG을 이용한 영상 분할 기법

Deng & Manjunath는 [8]에서 영상을 동질적인 색상-질감 영역의 묶음으로 분할하는 방법인 JSEG 알고리즘을 제안하였다. JSEG은 동질 질감 영역의 추출을 위해 특정 모델을 추정하는 기존의 방식과 달리 단지 주어진 색상-질감 패턴의 동질성을 검사함으로써 계산량을 크게 줄일 수 있다. JSEG 기법은 크게 색상 양자화와 공간 분할의 두 단계로 나뉜다. 먼저, 색상의 공간적 분포에 대한 고려 없이 영상을 몇 개의 대표 색상으로 양자화한다. 양자화

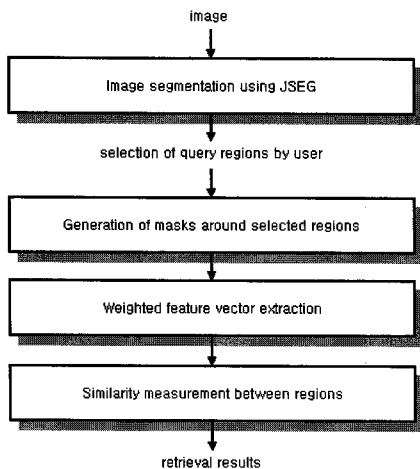


그림 1. 제안하는 알고리즘의 블록도  
Fig. 1 Block diagram of the proposed algorithm.

된 영상의 화소값을 색상 영역마다 순서값으로 대체하여 CCM(Color Class-Map: 색상 군집 지도)을 만든다. 다음으로, CCM을 이용하여 영상의 공간 분할 과정을 수행한다. 이 때, CCM의 화소와 대응하는 본 영상의 화소는 고려하지 않는다. 색상 간의 유사성과 이의 분포를 동시에 분석하는 것은 매우 어려운 작업이기 때문에, JSEG 알고리즘의 이러한 방법은 각 단계를 손쉽게 처리할 수 있도록 해준다.

### 2.1 CCM의 생성

일반적인 색상 영상의 처리 시 양자화를 통하지 않는 직접적인 색상 처리 방법은 많은 시간을 소요하기 때문에, 대부분의 영상 처리 방법들은 색상의 질을 크게 떨어뜨리지 않는 수준에서 색상 양자화를 전처리 기법으로 사용한다. JSEG에는 peer group filtering<sup>[9]</sup>을 통한 색상 양자화 기법이 적용되었다. 그리고 나서, 양자화 영상의 각 양자화 색상마다 고유의 레이블을 부여함으로써 CCM을 작성한다. CCM은 영상의 질감에 대한 일종의 조직도로 볼 수 있다. 이러한 CCM의 공간 데이터 점을 이용한 효과적인 영상 분할이 다음의 과정을 통해 이루어진다.

CCM에서 모든 데이터 점들의 집합을 Z라 하고, 그 총 수를 N이라 할 때, 평균 m은  $m = \frac{1}{N} \sum_{z \in Z} z$  이고, Z가 C개의 부류로 분류되었다고 가정하면,  $Z_i, i = 1, \dots, C$ 로 표현할 수 있다. 또한, 부류  $Z_i$ 의 총 데이터 수  $N_i$ 의 평균값  $m_i$ 는  $m_i = \frac{1}{N_i} \sum_{z \in Z_i} z$  이다. 이로부터 수식 (1)과 수식 (2)의  $S_T$ 와  $S_W$ 를 구하고, 최종적으로 수식 (3)에서 J를 구한다.

$$S_T = \sum_{z \in Z} \|z - m\|^2, \tag{1}$$

$$S_W = \sum_{i=1}^C S_i = \sum_{i=1}^C \sum_{z \in Z_i} \|z - m_i\|^2, \tag{2}$$

$$J = (S_T - S_W) / S_W. \tag{3}$$

여기에서  $S_W$ 는 같은 부류에 속해있는 점들의 분산을 의미한다. 만약 어느 한 영상이 동질의 색상 영역을 갖는 몇 개의 독립된 부류로 나뉘어 구성되어 있을 경우는 큰 J값을 갖고, 반대로 모든 색상 부류가 전체 영상에 걸쳐 고르게 분포되어 있다면 작은 J값을 갖는다는 것을 수식 (3)를 통해 알 수 있다.

### 2.2 JSEG 공간 분할 방법

JSEG 알고리즘은 CCM 전체 화소의 각각에 일 정크기의 윈도우 마스크를 씌운 후 해당 중심 화소에서의 J값을 구함으로써 전체 영상에 대한 J-image를 얻는다. 이러한 J값은 CCM에서의 각 화소가 영역 내부에 속하는지, 아닌지를 구분하는 표시자로 활용될 수 있다.

영상 분할은 큰 크기의 윈도우 마스크를 원 영상에 적용시킨 후 얻어진 듬성한 J-image에서 시작한다. 이 J-image를 이용하여 영역의 시작점을 설정한 후, 영역 성장법을 시행한다. 영역은 화소 단위로 성장하는 것이 아니라 윈도우의 크기에 따라 성장 영역의 크기가 따로 정해지고, 이 크기에 맞는 영역이 주위에 있을 경우에만 해당 영역을 합치면서 영역이 성장한다. 즉, 큰 윈도우를 이용할 경우 주변의 작은 영역을 다룰 수는 없다. 알고리즘은 이러한 작은 영역을 다루기 위해 윈도우 마스크 크기를 줄이면서 더 세밀한 J-image를 추출하고 이로부터 섬세한 영상 분할을 하는 방향으로 진행된다. 이러한 과정은 최소 크기의 윈도우 마스크에 도달할 때까지 반복된다. 영역 성장법을 통해 분할된 영상은 과-분할(over-segmented)되었을 수도 있기 때문에 최종적으로 agglomerative 방법<sup>[10]</sup>을 이용한 영역 병합법을 수행한다. 영역 병합은 인접 영역과의 색상 유사도의 비교를 통해 이루어지며, 색상의 기술자로서는 색상 히스토그램을 이용한다. 그림 2는 JSEG 공간 분할 기법의 적용 예를 보여준다.

Manjunath가 제시한 다른 영상 분할 기법으로써 EdgeFlow 알고리즘<sup>[11]</sup>이 있다. 이 알고리즘은 다양한 영상 특성에서의 색상 및 질감 변화 방향을 식별

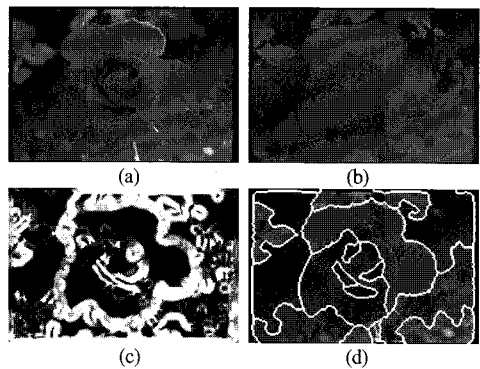


그림 2. JSEG을 통한 영상 분할의 예 (a) 원 영상 (b) 양자화된 영상 (c) J-image (d) 분할된 영상  
Fig. 2 The example of image segmentation using JSEG method (a) original image (b) quantized image (c) J-image (d) segmented image.

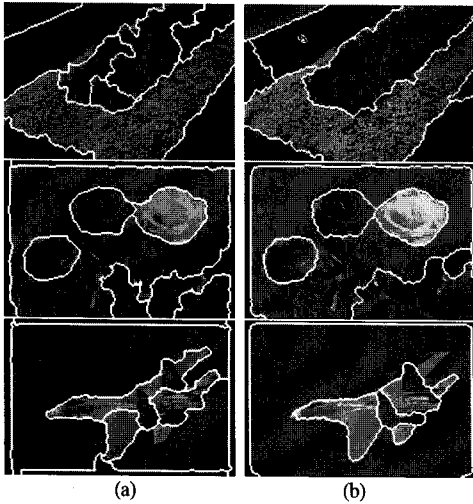


그림 3. 영상 분할의 예 (a) EdgeFlow 기법 (b) JSEG 기법  
Fig. 3 Examples of image segmentation (a) by EdgeFlow method (b) by JSEG method.

하기 위하여 예측 코딩 구조를 이용하여 에지 흐름 구조를 구축한 후, 반복적으로 이러한 에지 흐름을 주변 영역으로 전파시킴으로써 에지 흐름이 서로 상충되는 지역에서의 경계를 추출하는 방법이다. JSEG 기법과 비교해 볼 때 정확한 영역 경계를 검출하기 어렵다는 것과 영역 분할 후 후처리 과정이 충분하지 않아 과분할의 우려가 있다는 점에서 차이가 있어 본 논문에서는 JSEG 기법을 사용자 질의를 위한 영역 분할 기법으로 적용하였다. 그림 3은 EdgeFlow 기법과 JSEG 기법과의 영역 분할에 대한 비교 결과를 보여주고 있다.

### III. 선택 영역에서의 특징 추출 기법

영역 기반 검색 시스템은 사용자에게 의해 질의된 선택 영역으로부터 특징 벡터를 추출하고, DB에 있는 영상 특징과의 유사도 비교를 통해 사용자가 원하는 관련 영상을 찾아 그 검색 결과를 내보인다. 초기의 영역 기반의 검색 시스템은 질의 영상과 DB 내 영상 간의 영역 대 영역의 비교가 주(主)를 이루었으나, 컴퓨터 비전 분야의 영역 분할 기술의 한계와 단순한 영역 간 비교가 아닌 영역 이외의 배경에 대한 의미적 접근의 필요성 등으로 인하여 최근에는 사용자에게 의해 질의된 특징 영역뿐 아니라 그 배경에 대한 내용까지 취득할 수 있는 특징 기술자가 이용되고 있으며 본 논문에서 또한 이러한 특징 기술자에 대하여 제안하고 있다. 이에 따라, 본 장에서는 영상에서의 복합적인 특징 추출 기

법과 추출된 특징 간의 유사도 측정 기법에 대하여 설명한다.

#### 3.1 사용자 질의 영역 선택

영역 기반의 영상 검색을 위해 사용자는 먼저 JSEG 기법에 의해 기(既) 영역 분할 처리된 질의 영상을 선택하고, 분할된 영역 중 원하는 위치를 임의대로 선택하여 검색에 이용할 수 있다. 선택 영역은 하나 혹은 그 이상일 수 있으며 본 논문에서의 검색 시스템은 다수의 검색 질의 영역 선택 및 제거 등을 위한 효과적인 사용자 인터페이스를 제공하고 있다. 그림 4는 사용자에게 의한 질의 영역 선택의 다양한 예를 보이고 있다.

#### 3.2 영역 경계 상자의 추출 및 거리에 따른 윈도우 마스크의 생성

질의된 분할 영상의 선택 영역을 기반으로 하는 특징 벡터를 추출하기 위해 다음의 단계를 거친다. 먼저 그림 5-(a)에서와 같이 선택 영역에 대한

MBR, 즉,  $R^0 = \left( r_x^0, r_y^0 \right)^T$  를 생성한다.

여기에서,  $R$ 은 영역 경계 상자를 나타내고  $r_x$ 와  $r_y$ 는 각각 영역 경계 상자의  $x, y$  축 길이를 나타낸다. 생성된 MBR로부터 그림 5-(b)와 같이 수식 (4), (5)를 이용하여 영역을 점차로 증가시키면서 선택 영역의 중심으로부터 거리에 따른 외곽 경계 상자들을 만든다. 외곽 경계 상자의 최대 크기는 영상의 크기와 같고,  $I \supseteq \max \{ R^0, R^1, \dots, R^k, \dots, R^{n-1} \}$ ,  $k=0, 1, \dots, n-1$ 으로 표현된다.

$$\begin{pmatrix} r_x^{k+1} \\ r_y^{k+1} \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} \begin{pmatrix} r_x^k \\ r_y^k \end{pmatrix}, \quad (4)$$

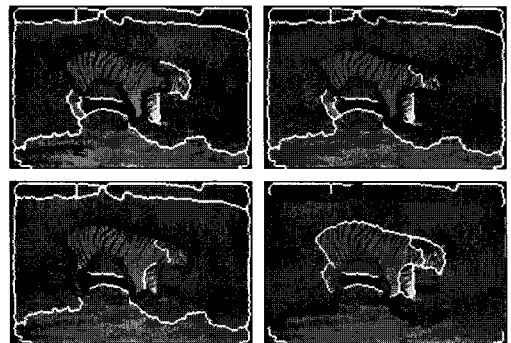


그림 4. 다양한 질의 영역 선택의 예  
Fig. 4 Examples of user selection for query regions.

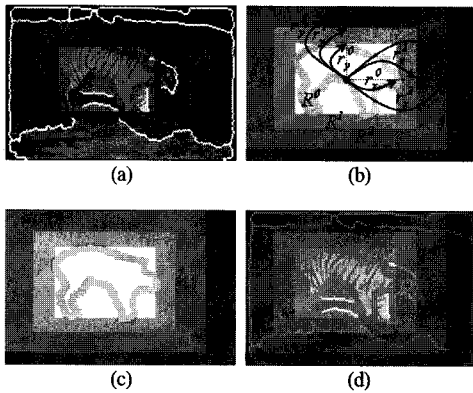


그림 5. 질의 영역의 설정 및 윈도우 마스크의 적용 (a) 질의 영역의 MBR (b) 윈도우 마스크의 생성 (c) 윈도우 마스크 (d) 윈도우 마스크의 적용

Fig. 5 Selection of query region and application of window mask (a) MBR of query region (b) generation of window mask (c) window mask (d) application of window mask.

$$R^{k+1} = \begin{pmatrix} r_x^{k+1} \\ r_y^{k+1} \end{pmatrix} - R^k \quad (5)$$

이러한 과정을 거친 후, 그림 5-(c)와 같은 최종 윈도우 마스크를 생성하며, 이를 영상에 적용시킴으로써(그림 5-(d)) 수식 (6)을 통해  $i^{\text{th}}$  영상에서의 특징 벡터,  $v_{k,l}^i$ , 를 얻는다.

$$v_{k,l}^i = \sum_{t=0}^{m-1} \sum_{k=0}^{n-1} w_{R^k} \cdot \tilde{H}_{R^k} \quad (6)$$

여기에서  $w_{R^k}$ 는 영역 가중치(weighting factor)로서  $1/\sqrt{2}$ 의 값을 갖는다.  $\tilde{H}_{R^k}$ 는 [12]에서 Stricker가 제안한 것으로 본 논문에서는 영역  $R^k$ 에서  $m$ 개의 빈을 갖는 누적 히스토그램을 나타낸다.

### 3.3 유사도 측정

영상 검색 시스템은 주어진 질의 영상과 색인된 자료 사이의 유사도 검색에 기반을 둔다. 영상 간 특징값 비교를 통하여 두 영상의 유사도를 계산한 후 유사도가 높은 영상부터 차례대로 사용자에게 결과 영상을 보여준다. 질의 영상을 Q라 하고, 색인 영상을 dB라고 할 때, L1 거리 측정법을 이용하여 Q와 dB 영상 간 유사도를 다음과 같이 비교한다.

$$\Delta simR = \sum_l \sum_k |v_{k,l}^Q - v_{k,l}^{dB}| \quad (7)$$

임의의 두 영상 쌍 사이의 비교에서 색상  $l$ 에서의 히스토그램 차분이 비록 같은 값을 갖더라도, 그 색상 빈에서의 히스토그램 값의 크기에 의해 그 의미가 크게 달라지므로 일반적으로 다음과 같은 표준화된 거리를 이용한다. Q와 dB 사이의 표준화된 거리는 다음과 같이 표현된다.

$$simR = \sum_l \sum_k \frac{|v_{k,l}^Q - v_{k,l}^{dB}|}{1 + v_{k,l}^Q + v_{k,l}^{dB}} \quad (8)$$

본 논문은 R\*-tree 색인 기법<sup>[13,14]</sup>을 이용해서 제안하는 알고리즘의 특징 기술자를 색인 한다. R\*-tree 기법은 현재 주로 사용되고 있는 여러 색인 기법의 기본이 되는 방법으로서 다차원 공간상에 위치나 크기에 관련된 정보를 원활히 관리하기 위한 동적 색인 구조이다.

## IV. 실험 결과

본 논문에서는 JPEG 압축된 2,920여장의 영상으로 DB를 구성하고, 제안된 검색 기법을 평가 하였다. 실험 DB는 실내영상, 식물, 동물, 풍경, 드라마, 뉴스 등의 다양한 영상으로 구성하였다. 그림 6 과 그림 7은 실험에 사용된 기본 영상들과 검색 시스템을 보여주고 있다.

본 논문에서 제안한 영상 검색의 정확성은 recall, precision과 ANMRR(Average of the Normalized Modified Retrieval Rank)을 통해 이루어진다. 먼저, 각 영상 질의 q에 대한 DB에서의 유사 영상의 개수를 NG(q)라 한다. 질의에 대한 검색 시 DB에서 찾은 최초 M개의 검색 결과 중에 유사 영상으로 기(既) 지정된 표준 기저 영상(ground truth images)과

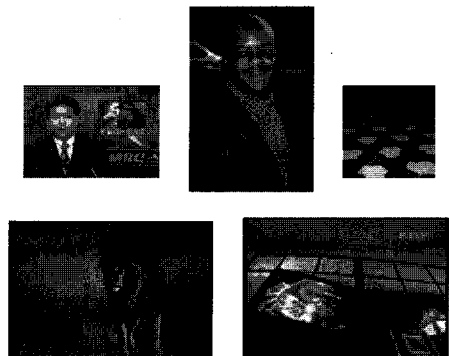


그림 6. DB 영상  
Fig. 6 DB images.

비교하여 정확하게 찾은 영상의 수를  $n_c$ , 놓친 영상의 수를  $n_m$ , 잘못 찾은 영상의 수를  $n_f$ 라 정의할 때, 질의 영상  $q$ 에 대한 precision과 recall 은 각각 다음과 같은 수식으로 얻어진다.

$$Precision = \frac{n_c}{n_c + n_f} = \frac{n_c}{M}, \quad (9)$$

$$Recall = \frac{n_c}{n_c + n_m} = \frac{n_c}{NG(q)}. \quad (10)$$

수식에서 알 수 있듯이 precision은 사용자의 질의를 통해 실제로 검색된 결과 영상과 사용자의 질의와 관련된 DB에서의 표준 기저 영상과의 비율을 뜻하며, recall은 표준 기저 영상과 실제 검색을 통해 사용자에게 보인 영상과의 비율을 나타내고 있다. 즉, precision은 검색 정확성을 측정하는 기준이며, recall은 적합한 자료를 얼마나 많이 검색했느냐를 보여주는 지표이다. Precision과 recall은 모두 [0.0~1.0]의 값을 가지며 높은 값을 가질수록 좋은 검색 성능을 나타낸다.

또한, 다른 성능 평가 기준으로서 MPEG-7에 정의된 ANMRR 을 이용한다. ANMRR는 모든 질의에 대하여 표준화된 수정 검색 순위의 평균을 나타낸다<sup>[15]</sup>. 먼저, 질의  $q$ 에 대한 평균 순위를 수식 (11)로 표현한다.

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}, \quad (11)$$

여기서 Rank(k)는 k번째 질의에 대한 검색 순위를 의미한다. 이러한 평균 순위 AVR은 다음의 수식

(12)로 수정된다.

$$MRR(q) = AVR - 0.5 - \frac{NG(q)}{2} \quad (12)$$

표준화된 수정 검색 순위는 수식 (13)으로 정의한다.

$$NMRR(q) = \frac{MRR(q)}{\max\{4 \times NG(q), \max_{q \in Q} NG(q)\} + 0.5 - 0.5 * NG(q)} \quad (13)$$

여기에서, Q는 모든 질의 영상의 수를 표현한다. 표준화된 수정 검색 순위는 [0.0~1.0]의 값을 가지며 recall, precision과는 반대로 낮은 값을 가질수록 높은 검색 성능을 나타낸다. 최종적으로, ANMRR은 모든 질의 영상에 대한 NMRR의 평균값을 의미하며 다음과 같이 계산한다.

$$ANMRR(q) = \frac{1}{Q} \sum_{q=1}^Q NMRR(q) \quad (14)$$

표 1 은 제안된 영역 기반의 영상 검색 기법에 대한 실험 결과를 보여준다. 본 논문에서는 제안하는 알고리즘과의 비교 성능 평가를 위하여 전역 영상 간(FtoF: Full image to Full image)의 특징 비교 방법과 선택 영역으로부터의 영역 대 영역 간(RtoR: Region area to Region area) 특징 비교 방법을 이용하였다. 추출된 특징에 대한 기술자로는 객관적 비교를 위해 누적 히스토그램 기법을 공통적으로 적용하였다. 표 1에 보이는 바와 같이 비슷한 recall 값을 가졌을 경우 제안하는 방법이 가장 좋은 precision 성능을 보여주고 있다. RtoR 방법이 가장 성능이 떨어지는 이유는 크게 두 가지로 볼 수 있는데, 우선 영상 간 영역의 비교 시에 서로 같은 위치의 대응 영역을 비교하기 때문에 두 영상 각각에 같은 모양과 색상을 갖는 객체가 존재한다고 하더라도 영상 내의 객체의 위치가 서로 같지 않다면 두 영역의 특징 벡터 간 차분이 큰 값을 가질 가능성은 항상 존재한다는 것이다. 또한, 사용자가 단순히 영상 내의 관심 객체 영역만을 대상으로 검색하는 것이 아닌 영상의 배경까지 고려할 경우에도 영역 대 영역 비교 방법이 전체 영상을 대상으로 비교하는 방법보다 오히려 성능이 떨어지는 요인이 될 수가 있다. 이와 반대로, FtoF방법의 경우는 사용자가 영상의 특정 영역에 중요 의미를 두고 질의할 경우, 배경을 고려하는 FtoF방법의 특성상 원치 않는 검색 결과가 자주 도출되기 때문에 제안하는 방법보다 성능이 떨어지는 것을 알 수 있다.

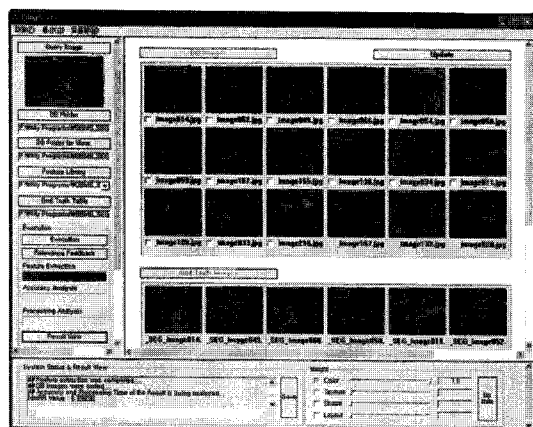


그림 7. 검색 시스템.  
Fig. 7 Retrieval system.

표 1에서는 각 방법에 의한 특징 추출 시간 (feature extraction time: FET)의 차이를 또한 비교하고 있다. 제안하는 방법은 선택된 영역을 기반으로 한 윈도우 마스크의 생성 등으로 인하여 다른 방법보다 조금 긴 특징 추출 시간이 소요된다. 그러나, 처음 추출된 특징 정보가 오프라인 상에서 이미 R\*-tree에 의해 정보 색인이 되기 때문에, 온라인상에서의 결과 검색된 최초의 M개의 영상에 대한 특징 추출 시간은 저장된 자료를 읽기 위한 아주 작은 시간만이 필요하게 되므로, 오프라인 상에서의 특징 추출 시간은 사실상 큰 의미를 갖지 않게 된다. 영상은 영역 분할이 이루어진 후에 검색 DB로 삽입되기 때문에 영상 간 비교 시의 특징 추출 시간에 영역 분할을 위한 시간은 계산하지 않았다.

그림 8은 다양한 시점 변화를 포함하고 있는 관련 영상에 대한 검색 결과를 보여준다. 그림 8-(a)는 질의 영상이고, 그림 8-(b)~그림 8-(d)는 검색된 결과이다. 제안된 방법과 여러 다른 검색 기법들과의 비교를 위해 각 영상마다 검색된 순위를 숫자로 표기하였다. 질의 영상에 대한 검색 결과에서 보이는 바와 같이, 제안된 방법은 영상 및 객체의 시점 변화에 대하여 더 강건한 결과를 보여주는 반면에, 다른 기법들은 이러한 변화에 대하여 상대적인 민감함을 가지고 있다. 질의 영상에 대한 검색 결과에

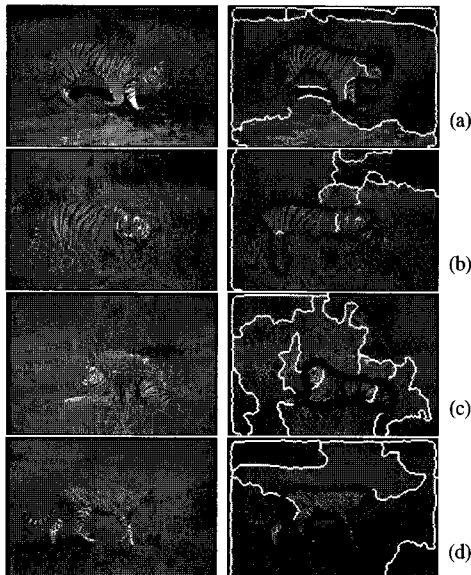


그림 8. 검색 결과 (a) FtoF: rank 1, RtoR: rank 1, Proposed (P): rank 1 (b) FtoF: rank 3, RtoR: rank 12, P: rank 2 (c) FtoF: rank 2, RtoR: rank > 36, P: rank 7 (d) FtoF: rank > 36, RtoR: rank 7, P: rank 10.  
Fig. 8 Retrieval results with rank.

서 그림 8-(c)는 FtoF 기법을 이용하였을 때 빠른 검색 순위를 가졌으며, RtoR 기법의 이용 시는 좋지 않은 결과를 나타내었다. 이는, 그림 8-(a)와 그림 8-(c)가 영상 전체로서는 매우 유사하나, 대응하는 위치 영역 간 비교 시에는 서로 다른 색상 구조를 가진다는 것을 의미한다.

또한, 그림 8-(d)는 RtoR 기법을 이용하였을 때에 비교적 좋은 검색 순위를 가지나, FtoF 기법의 이용 시는 결과가 좋지 않았는데, 이는, 그림 8-(a)와 그림 8-(d)의 대응 영역 간의 영상은 비슷한 색상 분포를 가지나 영상 전체로 보았을 경우 서로 다른 색상 조합을 가지고 있다는 것을 나타낸다. 결국, 두 방법의 장점을 고루 취합한 제안된 방법을 검색에 적용할 경우, 두 영상 모두 상위 검색 순위 내에서 올바르게 검색되는 것을 볼 수 있다. 그림 9는 카메라의 시점 변화를 가지는 다른 질의 영상에 대한 결과를 또한 제시하고 있다.

그림 10과 그림 11은 recall과 precision에 기초한 영상 검색 성능 비교 그래프를 표현하고 있다. 그림 10에서의 각각의 값은 전체적인 성능 비교를 위한 평균값을 나타낸다. 결과 영상이나 그래프에서 보이는 바와 같이, 제안한 검색 기법이 단순히 대응

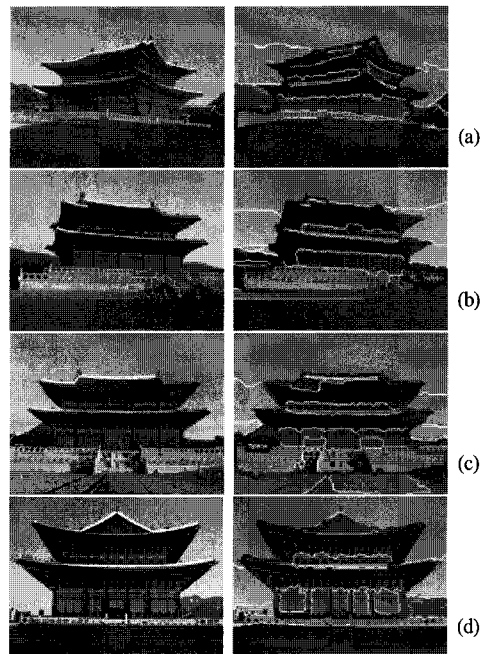


그림 9. 검색 결과 (a) FtoF: rank 1, RtoR: rank 1, Proposed (P): rank 1 (b) FtoF: rank 3, RtoR: rank 5, P: rank 3 (c) FtoF: rank 18, RtoR: rank 6, P: rank 5 (d) FtoF: rank 9, RtoR: rank 15, P: rank 7.  
Fig. 9 Retrieval results with rank.

표 1. 내용 기반 영상 검색 기법의 비교  
Table 1 Comparison for content-based image retrieval methods.

method	precision	recall	ANMRR	FET(ms)
FtoF	0.36	0.41	0.495	16
RtoR	0.32	0.40	0.561	22
Proposed	0.44	0.40	0.417	25

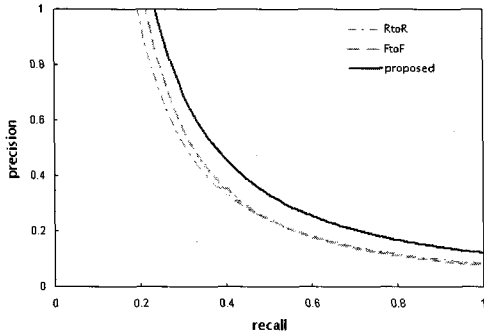
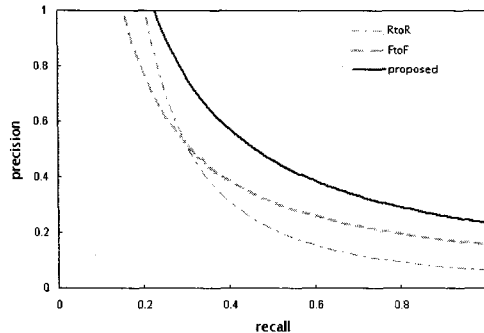
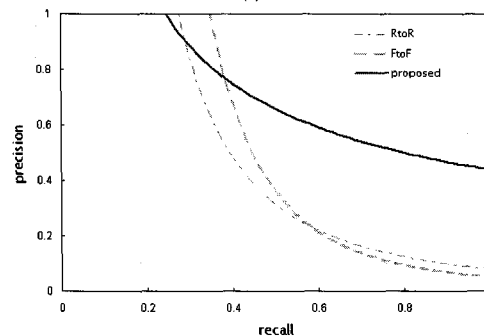


그림 10. Recall과 precision  
Fig. 10 Recall and precision.



(a)



(b)

그림 11. Recall과 precision (a) for 그림 8 (b) for 그림 9  
Fig. 11 Recall and precision (a) for Fig. 8 (b) for Fig. 9.

영역 간을 비교하는 RtoR 방법이나 전체 영상 간의 색상 비교 방법인 FtoF 기법 등과 비교하여 좀 더 우월한 성능을 가진다는 것을 보여주고 있다. 그림 11-(a)와 그림 11-(b)는 각각 그림 8과 그림 9에 대

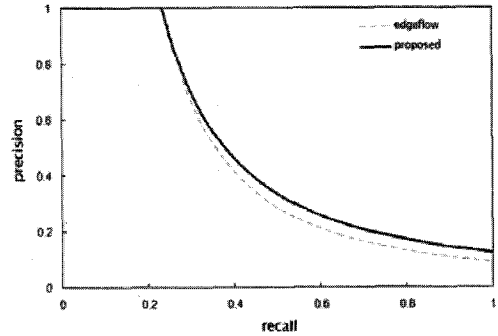


그림 12. Recall과 Precision (EdgeFlow vs. Proposed method)  
Fig. 12 Recall and Precision (EdgeFlow vs. Proposed method)

한 recall과 precision 그래프를 보여주고 있다.

그림 12는 2장에서 설명한 EdgeFlow 영상 분할 기법을 이용한 영상 검색 기법<sup>[16]</sup>과 JSEG기법을 이용한 제안한 알고리즘과의 비교 결과를 나타낸다. [16]에서의 영상 기법은 EdgeFlow를 통해 분할된 각 영역 간 색상의 거리 및 확률로 전체 영상에서의 특징자를 얻는 방법이다. 본 논문의 실험에서는 영역 간 검색에 초점을 두기 위하여, [16]에서의 영상 특징 추출 기법에서 전역 특징이 아닌 영역 내 특징만을 추출하여 영역 간 검색을 진행하여 이를 제안된 방법과 비교하였다. 그림에서 보는 바와 같이 그림 10에서의 결과와 마찬가지로 RtoR 기법에서의 문제가 이 결과에서도 드러나는 것을 알 수 있다.

## V. 결론

본 논문에서는 JSEG 알고리즘을 이용한 영역 기반의 영상 검색 기법에 대하여 제안하였다. 우리는 JSEG 기법을 이용하여 효과적으로 영상을 분할하고, 분할된 영역 중 일부를 사용자가 선택하여 질의 하게 함으로써 특정 관심 영역을 기반으로 영상 전체의 특징을 추출하는 기술자를 제안하고 이를 이용한 효율적인 영상 검색 시스템을 구현하였다. 제안하는 방법은 단순히 선택된 영역 만으로부터의 특징이나 질의된 영상 전체에서 추출된 특징을 이용하는 것이 아닌 영역과 배경의 특징을 가중치를 통해 복합적으로 이용함으로써 내용 기반의 영상 검색 성능을 향상시키고 있다. 최종적인 실험 결과는 영상 검색 결과에 있어서 기존의 방법에 비하여 제안하는 방법이 매우 효과적임을 잘 보여주고 있다. 현재 우리는 영역 기반의 마스크 생성으로부터 추출된 특징의 기술자로서 누적 히스토그램을 이용



하고 있으나, 영역 기반의 검색에 적합한 더 나은 성능의 기술자를 제안하는 것을 목표로 연구를 진행 중에 있다.

### 참 고 문 헌

[1] M. Flickner et al., "Query by image and video content: The QBIC system," *IEEE computer*, vol. 28, no. 9, pp. 23-32, 1995.

[2] J. Huang, S. R. Kumar, M. Mitra, W. J. Zhu, and R. Zabih, "Image indexing using color correlograms," *CVPR*, pp. 762-768, 1997.

[3] G. Pass and R. Zabih, "Histogram refinement for content-based image retrieval," *IEEE WACV*, pp. 96-102, 1996.

[4] J.R. Smith and C.-S. Li, "Image classification and querying using composite region templates," *J. Comput. Vis. Image Understand.*, vol. 75, no. 1, pp. 165-174, 1999.

[5] M.E. Wood, N.W. Campbell, and B.T. Thomas, "Iterative refinement by relevance feedback in content based digital image retrieval," in *Proc. 5th ACM International Multimedia Conference (ACM Multimedia 98)*, pp. 13-20, 1998.

[6] W.Y. Ma and B.S. Manjunath, "NETRA: A toolbox for navigating large image databases," in *Proc. IEEE Int. Conf. Image Processing*, vol. I, pp. 568-571, 1997.

[7] M. Carson, S. Thomas, J.M. Belongie, and J. Malik, "Blobworld: A system for region-based image indexing and retrieval," in *Proc. Int. Conf. Visual Information Systems*, pp. 509-516, 1999.

[8] Y. Deng and B.S. Manjunath, "Unsupervised segmentation of color-texture regions in images and video," *PAMI*, vol. 8, pp. 800-810, 2001.

[9] Y. Deng, C. Kenney, M.S. Moore, and B.S. Manjunath, "Peer group filtering and perceptual color image quantization," *Proc. of IEEE Intl. Symposium on Circuits and Systems*, vol. 4, pp. 21-24. 1999.

[10] R.O. Duda and P.E. Hart, *Pattern Classification and Scene Analysis*, John Wiley & Sons, 1970.

[11] W.Y. Ma and B.S. Manjunath, "Edge flow: a framework of boundary detection and image segmentation," *Proc. Of CVPR*, pp. 744-749, 1997.

[12] M. Stricker and M. Orengo, "Similarity of color images," *Proc. SPIE 2420*, pp. 381-392, 1995.

[13] N. Beckmann, H.-P. Kriegel, R. Schneider, and B. Seeger, "The R\*-tree: An efficient and robust access method for points and rectangles," *Proc. ACM SIGMOD*, pp. 322-331, 1990.

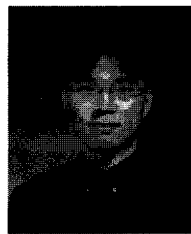
[14] A. Guttman, "R-trees: a dynamic index structure for spatial searching," *Proc. ACM SIGMOD*, pp. 47-57, 1984.

[15] "MPEG vancouver meeting," ISO/IEC JTC1/SC29/WG11, *Experimentation Model Ver. 2.0*, Doc. N2822, 1999.

[16] Y. Deng and B.S. Manjunath, "An efficient low-dimensional color indexing scheme for region based image retrieval," *Proc. IEEE Intl. Conference on Acoustics, Speech and Signal Processing (ICASSP-99)*, pp. 3017-3020, 1999.

김 낙 우 (Nac-woo Kim)

정희원

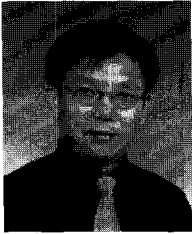


1997년 2월 중앙대학교 제어 계측공학과 졸업  
 2002년 2월 중앙대학교 첨단 영상대학원 영상공학과 석사  
 2006년 2월 중앙대학교 첨단 영상대학원 영상공학과 박사  
 2006년 3월~현재 한국전자통신

연구원 광대역통합망연구단 FTTH서비스팀  
 <관심분야> 영상 압축, 영상 통신, 영상 정보기술

**송 호 영 (Ho-young Song)**

정회원



1983년 2월 홍익대학교 전자 계  
산학과 졸업  
1996년 2월 충북대학교 전자계  
산학과 석사  
2003년 2월 충북대학교 정보 통  
신공학과 박사수료  
1983년 3월~현재 한국전자통신

연구원 광대역통합망연구단 FTTH서비스팀장

<관심분야> 통신, FTTH, 통방융합서비스

**김 봉 태 (Bong-tae Kim)**

종신회원



1983년 2월 서울대학교 전자 공  
학과 졸업  
1991년 12월 노스캐롤라이나 주  
립대학교 컴퓨터공학과 석사  
1995년 12월 노스캐롤라이나 주  
립대학교 컴퓨터공학과 박사  
1983년 3월~현재 한국전자통신

연구원 광대역통합망연구단 광통신연구센터장

2004년~현재 광인터넷포럼 운영위원장

2005년~현재 FTTH산업협의회 기술/표준화 분과위원장

<관심분야> 통신, FTTH, BcN, 네트워크