

논문 2007-44CI-5-1

영상 분할을 이용한 영역기반 내용 검색 알고리즘

(Region-based Content Retrieval Algorithm Using Image Segmentation)

이 강 현*

(Kang Hyeon RHEE)

요 약

영상 정보의 이용이 증가함에 따라 영상을 효율적으로 관리할 수 있는 시스템의 필요성이 증가하고 있다. 이에 따라, 본 논문에서는 영상 분할 알고리즘, 색상 특성, 질감, 그리고 영상의 형태와 위치 정보의 효율적인 결합에 근거한 영역기반 내용 검색 알고리즘을 제안한다. 색상 특징으로는 색상의 공간적인 상관관계를 잘 나타내는 HSI 색상 히스토그램을 선택하였고, 영상의 분할과 질감특성은 각각 Active contour와 CWT(Complex wavelet transform)를 사용하였다. 그리고 형태와 위치 특징들은 HSI의 휘도 성분에서 불변 모멘트를 이용하여 추출하였다. 효율적인 유사도 측정을 위해 추출된 특징(색상 히스토그램, Hu 불변 모멘트, CWT)을 결합하여 정확도와 재현율을 측정하였다. www.freefoto.com에서 제공하는 DB를 사용하여 실험한 결과, 제안된 검색엔진은 94.8%의 정확도와 82.7%의 재현율을 가지며 성공적으로 영상 검색 시스템에 응용할 수 있다.

Abstract

As the availability of an image information has been significantly increasing, necessity of system that can manage an image information is increasing. Accordingly, we proposed the region-based content retrieval(CBIR) algorithm based on an efficient combination of an image segmentation, an image texture, a color feature and an image's shape and position information. As a color feature, a HSI color histogram is chosen which is known to measure spatial of colors well. We used active contour and CWT(complex wavelet transform) to perform an image segmentation and extracting an image texture. And shape and position information are obtained using Hu invariant moments in the luminance of HSI model. For efficient similarity computation, the extracted features(color histogram, Hu invariant moments, and complex wavelet transform) are combined and then precision and recall are measured. As a experimental result using DB that was supported by www.freefoto.com, the proposed image retrieval engine have 94.8% precision, 82.7% recall and can apply successfully image retrieval system.

Keywords : Image Retrieval, Color Histogram, Hu Invariant Moments, Active Contour, CWT

1. 서 론

최근 수년 동안 영상정보의 크기와 양의 폭발적인 증가에 따라 디지털 도서관, 영상처리 및 데이터베이스 시스템분야 등에서 영상검색 기술은 많은 관심을 받고 있다. 더욱이, 지난 몇 년 동안에 하드웨어기술의 놀라운 발전은 막대한 멀티미디어 데이터의 저장과 검색을 가능하게 하였다.

연구 초기에는 검색의 대상이 되는 멀티미디어 데이

터에 사람이 직접 색인을 첨가하는 방법이 연구되었으나, 이러한 방법은 시간과 비용이 많이 들며, 색인을 첨가하는 사람과 검색하는 사용자의 관점이 불일치 할 경우 검색의 효율성이 크게 떨어지게 된다. 이러한 단점을 극복하기 위해 1990년대 초에 질의어 대 이미지(query-by-image)^[1~3] 기법이 제안되었다.

일반적으로 내용기반 검색을 위해서는 영상 데이터의 '색상', '질감', '형태'의 세 가지 특징 요소를 주로 활용한다. 이 중에서도 특히 색상은 영상을 구성하는 가장 지배적인 요소로써 추출과 활용이 쉽고 빠르며, 영상의 크기와 방향에 독립적인 특성을 가지고 있어서 다른 두 가지에 비해 훨씬 더 많이 사용한다. 일반적으로 사용하는 색상 특징으로는 색상 히스토그램 (Color

* 평생회원, 조선대학교 전자공학과
(Dept. of Electronic Engineering, Chosun University)
접수일자: 2007년5월27일, 수정완료일: 2007년9월4일

Histogram)^[4], 칼라 코허런스 벡터(Color Coherence Vector)^[5], 바이너리 칼라 세트(Binary Color Set)^[6~7], MPEG-7 기술자(Descriptor)인 SCD(Scalable Color Descriptor), CLD(Color Layout Descriptor) 및 CSD(Color Structure Descriptor)^[8], 그리고 칼라 오토 코렐로그램(Color Autocorrelagram)^[5] 등이 있다.

질감은 구름, 나무, 벽돌, 직물 등과 같은 실세계 객체들의 표면특성과 관련된 중요한 특징으로 표면의 구조적인 배열과 주위 화소와의 관계에 대한 정보를 나타낸다^[9]. 내용기반 영상검색에 사용되는 질감 특징으로는 GLCM(Gray-level Co-occurrence Matrix)을 이용한 통계적 질감특징^[10], MPEG-7의 기술자인 EHD(Edge Histogram Descriptor)^[8], 웨이블릿 변환을 이용한 다해상도 질감 특징^[6] 등이 있다. 최근에는 영상내의 밸리(Valley)와 에지(Edge)를 잘 표현하는 BVLC(Block Variance of Local correlation Coefficient)를 결합하여 사용하는 검색기법이 제안된 바 있으며, 이는 비교적 낮은 특징의 차원을 가지면서도 우수한 검색성능을 나타내는 효율적인 기법으로 알려져 있다^[7].

모양은 객체의 윤곽선을 나타내는 특징으로 인간이 시각적으로 이미지를 분류하거나 인식하는 데 있어서 가장 중요한 정보이다. 그러나 모양 특징은 배경이 복잡하거나 객체와 배경의 구분이 명확하지 않은 영상에서는 객관적이고 정확한 추출이 어렵다는 단점이 있다^[11]. 모양 특징에는 Fourier 서술자, Chain 코드, Zernike 모멘트 등이 있다^[9,12]. 그러나 색상, 질감 및 형태 기반의 CBIR(content-based image retrieval)은 일반적으로 영상의 전체적인 특징들을 기반으로 검색이 이루어지기 때문에 영상의 부분적인 특징들을 표현하지는 못한다. 이에 따라, 최근 많은 CBIR 시스템은 이미지의 영역기반(region-based) 영상검색(QBIC^[13], Netra^[14], VisualSeek^[15], Blobworld^[16], IRM^[17], FRIP^[18], and ROI^[19~20])에 초점을 맞춰 연구가 이루어지고 있다.

IBM에서 제안된 QBIC^[13]는 정지영상과 동영상 모두에서 내용기반 검색을 제공해 주지만 사용가능한 영상 분류(class)와 분할(segmentation)에 제약조건이 있다. VisualSEEK^[15]는 영역 추출과 색상 정보를 이용한 검색 시스템이다. 하지만 정확한 검색을 위해서 여분의 스케일과 형태 정보를 필요로 한다. Blobworld^[16]는 색상, 질감, 그리고 형태정보를 이용하여 웹(world wide web)상에서 영상을 검색할 수 있으나 영상분할을 위해 몇 가지 파라메타 조정과 절단(trimming) 작업이

필요하다. IRM^[17]은 영역정합(region-matching) 알고리즘을 사용한 검색 방법으로 영상분할의 단점을 줄이기 위해 분할된 영상과 전체 영상과의 영역을 정합하여 검사하지만 특성이 비슷한 영역이 존재할 경우 전체 이미지의 유사도가 감소하는 단점이 있다.

본 논문에서는 영상의 부분적인 특징을 표현하지 못하는 내용기반 영상 검색의 단점과 영상의 전체적인 특징을 표현하는데 한계점을 가지는 영역기반 영상 검색의 단점을 보완하기 위하여 색상, 질감 그리고 형태정보를 이용한 영역기반 내용 검색 알고리즘을 제안하였다. 제안된 알고리즘은 영상분할 알고리즘^[21]을 적용하여 입력영상이 패턴기반의 영상인지 아니면 목적 이미지를 가지는 영상인지를 판단한 다음 판단 결과에 따라 영상의 색상과 형태정보(색상 히스토그램, 불변모멘트)를 이용할 것인지 색상과 질감(색상 히스토그램, CWT(complex wavelet transform)특성을 이용할 것인지를 결정하여 영상을 검색한다.

II. 영상분할과 특징 추출

영역기반 내용 영상 검색을 위해서는 영상데이터에서 특징 영역을 추출해야 하며, 이를 기반으로 검색을 수행해야 한다. 일반적으로 사용되고 있는 영상의 주요 특징으로는 색상, 질감, 형태 및 영상을 구성하고 있는 객체들의 공간적 위치 등이 있다. 본 장에서는 영상 검색에 사용될 영역 추출 알고리즘, 추출된 영역의 색상 특징(Hue Histogram), 질감 특징, 그리고 형태 및 공간적 위치 특징에 사용된 알고리즘에 대해 살펴본다.

1. 영역 추출

본 논문에서는 [21]에서 제안된 변형된 기하학적 Active Contour 알고리즘을 사용하여 영상을 분할하였다. Active Contour는 내부 에너지와 영상 에너지, 그리고 외부 제약 에너지에 따라 형태가 변화하는 스플라인 곡선이다. Active Contour는 매개변수를 써서 $v(s) = (x(s), y(s))$ 로 정의할 수 있는데 여기서 $x(s)$ 와 $y(s)$ 는 각각 $s \in [0, 1]$ 에 따른 곡선의 x 와 y 의 좌표 값이다. 이렇게 표현된 Active Contour의 에너지 함수는 식 (1)과 같이 정의된다.

$$E_s = \int_0^1 E_s(v(s)) ds = \int_0^1 E_{int}(v(s)) + E_{ing}(v(s)) + E_{con}(v(s)) ds \quad (1)$$

여기서 내부 에너지 $E_{int}(v(s))$ 는 Active Contour의 내적인 힘으로서 Active Contour의 연속성(continuity)과 부드러움(smoothness)에 의해 계산되는 값이고, 영상 에너지 $E_{img}(v(s))$ 는 Active Contour가 영상 내에서 물체의 특징(feature)을 얼마나 잘 나타내는가에 의해 결정되는 값이며, 외부 제약 에너지 $E_{con}(v(s))$ 는 물체의 형태에 대한 사전지식을 이용하기 위한 것이다. 이 세 에너지의 합이 최소화된 Active Contour는 물체에 대한 지식 및 특징을 잘 표현하면서 부드럽고 연속적인 곡선의 모양을 가지게 된다.

[21]에서는 식 (2)를 Active Contour의 경계함수(Edge Function)로 제안하였다.

$$g = \frac{1}{1 + |\nabla G_\sigma \times I|^2} \quad (2)$$

여기서, I 는 영상데이터, g 는 경계함수(Edge Indicator Function), G_σ 는 표준편차 σ 의 가우시안 커널(Gaussian Kernel)이다. 식 (2)를 이용하여 $\phi(x, y)$ 의 외부 에너지는 식 (3)과 같이 정의된다.

$$\epsilon_{g,\lambda,\nu}(\phi) = \lambda L_g(\phi) + \nu A_g(\phi) \quad (3)$$

여기서, $\lambda > 0$, ν 는 상수이고, $L_g(\phi)$ 와 $A_g(\phi)$ 는 각각 식 (4)와 (5)와 같이 정의된다.

$$L_g(\phi) = \int_{\Omega} g\delta(\phi)|\nabla\phi| dx dy \quad (4)$$

$$A_g(\phi) = \int_{\Omega} gH(-\phi) dx dy \quad (5)$$

여기서, δ 는 디랙 함수(Dirac Function)이며, H 는 헤비사이드 함수(Heaviside Function)이다. 결과적으로 전체 에너지 함수는 식 (6)과 같이 정의된다.

$$\epsilon(\phi) = \mu P(\phi) + \epsilon_{g,\lambda,\nu}(\phi) \quad (6)$$

여기서, $\mu P(\phi)$ 는 내부 에너지이다.

2. 질감 특징 추출

본 논문에서는 영상의 질감 특징을 추출하기 위하여 DT-CWT^[22]를 사용하였다. DT-CWT는 그림 1과 같이 영상을 6개의 방향 성분으로 분해하여 그림 2와 같이 영상의 방향 특징을 표시해주는 기법으로 현재 질

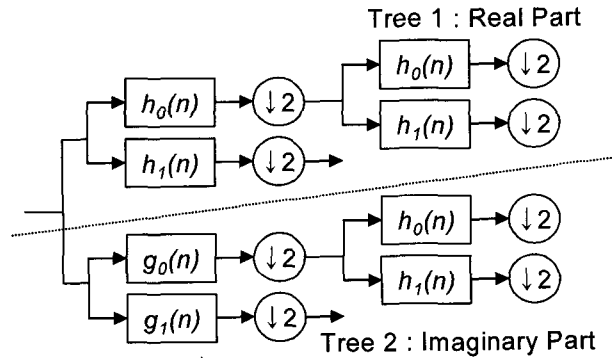


그림 1. 1차원 이진트리 CWT

Fig. 1. The 1-D dual-tree complex wavelets transform.

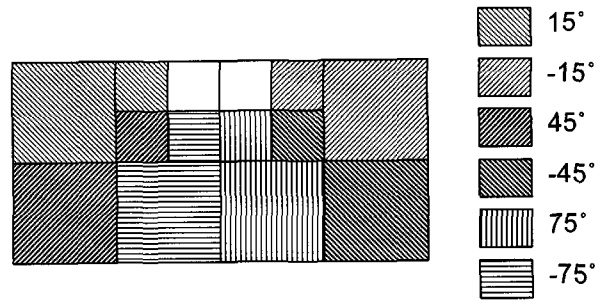


그림 2. CWT의 주파수 영역 분할

Fig. 2. Frequency-domain partition in complex wavelet transform.

감 특징을 이용한 영상 검색에서 많은 연구가 이루어지고 있다.

3. 색상 특징 추출

내용 기반 영상 검색의 초기부터 사용되어온 영상의 색상 정보는 영상의 내용을 가장 잘 대표하는 특징이라 할 수 있으며, 영상내의 조명 변화나 관측 위치, 크기 변화 등에 무관하게 적용 가능하다는 장점을 갖고 있다. 본 논문에서는 인간의 인지도와 유사한 HSI 모델의 H 값을 30° 간격의 12가지 색상으로 양자화(Quantization)하여 $H_n(n=1\sim 12)$ 으로 부호화하고, 이 값을 이용하여 전체 영상의 색상 히스토그램을 추출하

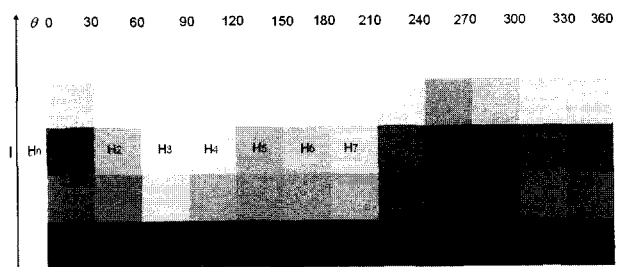


그림 3. 색상 양자화 테이블

Fig. 3. Color quantization table.

여 영상의 색상 정보를 기술하였다. 그림 3은 색상 히스토그램에 사용된 양자화 테이블이다.

4. 형태 및 공간적 위치 특징 추출

본 논문에서는 영상 내에서 물체의 형태 및 공간적 위치 특징을 추출하기 위하여 불변모멘트(Hu Invariant Moments)를 사용하였다.

Uniqueness Theorem^[23]에 의하면, 2차원 연속 함수 $f(x,y)$ 를 갖는 영상에서, $f(x,y)$ 가 구간에서 연속이고 x - y 평면의 일정 부분에서 영이 아닌 값을 갖는다면 모든 차수의 모멘트가 존재하며 $f(x,y)$ 에 의해 $(p+q)$ 차수의 모멘트 m_{pq} 가 유일하게 결정되고 또 m_{pq} 로부터 $f(x,y)$ 가 유일하게 결정된다. 디지털 영상 $f(i,j)$ 가 이진 영상이라면 모멘트는 식 (7)과 같다.

$$m_{pq} = \sum_i \sum_j i^p j^q, (p,q=0,1,2,\dots) \quad (7)$$

p, i : 이미지 픽셀의 x 축 좌표

q, j : 이미지 픽셀의 y 축 좌표

중심 모멘트는 식 (8)로 정의되며 이동에 불변인 특성을 갖게 된다^[24].

$$\mu_{pq} = \sum_i \sum_j (i - \bar{x})^p (j - \bar{y})^q \quad (8)$$

식 (8)에서 인덱스 i, j 는 각각 수평 x 축, 수직 y 축의 위치를 나타낸다.

모멘트 m_{00} 는 이진 객체 영역을 나타내는데 3차 모멘트까지의 중심 모멘트는 식 (9)와 같은 관계를 갖는다.

$$\begin{aligned} \mu_{00} &= m_{00} = \mu \\ \mu_{10} &= \mu_{01} = 0 \\ \mu_{20} &= m_{20} - \mu \bar{x}^2 \\ \mu_{11} &= m_{11} - \mu \bar{x} \bar{y} \\ \mu_{02} &= m_{02} - \mu \bar{y}^2 \\ \mu_{03} &= m_{03} - 3m_{02} \bar{y} + 2\mu \bar{y}^3 \\ \mu_{30} &= m_{30} - 3m_{20} \bar{x} + 2\mu \bar{x}^3 \\ \mu_{21} &= m_{21} - m_{20} \bar{y} - 2m_{11} \bar{x} + 2\mu \bar{x}^2 \bar{y} \\ \mu_{12} &= m_{12} - m_{02} \bar{x} - 2m_{11} \bar{y} + 2\mu \bar{x} \bar{y} \end{aligned} \quad (9)$$

중심 모멘트의 해석은 다음과 같다.

μ_{20} : 수평축 분산

μ_{02} : 수직축 분산

μ_{11} : 수평축과 수직축의 상호분산(covariance)

μ_{12} : 수평축에서 오른쪽에 비해 왼쪽으로 퍼져 있는 정도

μ_{21} : 수직축에서 위쪽에 비해 아래쪽으로 퍼져 있는 정도

μ_{30} : 수평축의 비대칭 정도(skew)

μ_{03} : 수직축의 비대칭 정도(skew)

또한 정규 모멘트는 중심 모멘트를 일정한 크기의 값으로 나눠주는 것으로 식 (10)과 같이 정의되며, 이는 크기에 불변인 특성을 부여해 준다^[25].

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma}, \quad \gamma = \frac{p+q}{2} + 1 \quad (10)$$

본 논문에서는 식 (9)와 (10)을 이용하여 Hu 불변 모멘트^[26]를 추출한 후, 영상의 형태 및 공간적 위치 특징을 추출하는 알고리즘에 사용하였다. Hu 불변 모멘트는 식 (11)과 같이 2차와 3차 중심 모멘트로 구성된다.

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} + \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ &\quad - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ &\quad - 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03}) \\ &\quad \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (11)$$

식 (11)에서 정의된 Hu 모멘트의 해석은 다음과 같다.

ϕ_1 : 수평과 수직 방향의 분산의 합, 수직과 수평축으로 많이 퍼져 있을수록 값이 커짐.

- ϕ_2 : 수직축과 수평축의 분산의 정도가 비슷할 경우 수직과 수평축에 대한 상호분산 값.
- ϕ_3 : 좌우, 상하로 치우친 값을 강조하는 결과
- ϕ_4 : 좌우, 상하로 치우친 값을 상쇄하는 결과
- ϕ_5, ϕ_6, ϕ_7 : 크기, 회전, 위치에 불변인 특징값을 추출

III. 제안된 영상검색 엔진

본 논문에서는 Active Contour 알고리즘을 사용하여 질의 영상의 일정 영역을 선택한 다음 선택된 영역에 따라 색상 특성, 질감 특성, 형태 및 공간적 위치 특성을 추출하여 영상을 검색한다. 그림 4는 본 논문에서 제안된 영상 검색 엔진의 전체 블록도이다.

질의 영상이 검색 시스템으로 입력되면, HIS 색상으로 변환된 다음, 영상의 I 성분에 Active Contour 알고리즘을 적용하여 영상을 분할한다. 영상 검색에 사용될 알고리즘은 분할된 영역의 픽셀수에 따라 즉, 질의 영상 내에 목적 물체가 존재하면 불변모멘트와 색상 히스토그램을 적용하며, 질의 영상이 패턴으로 이루어져 정확한 목적 물체가 존재하지 않을 경우에는 CWT와 색상 히스토그램으로 나누어 영상 특성을 추출한다. 제안된 알고리즘에서는 H 성분으로부터 색상 히스토그램을

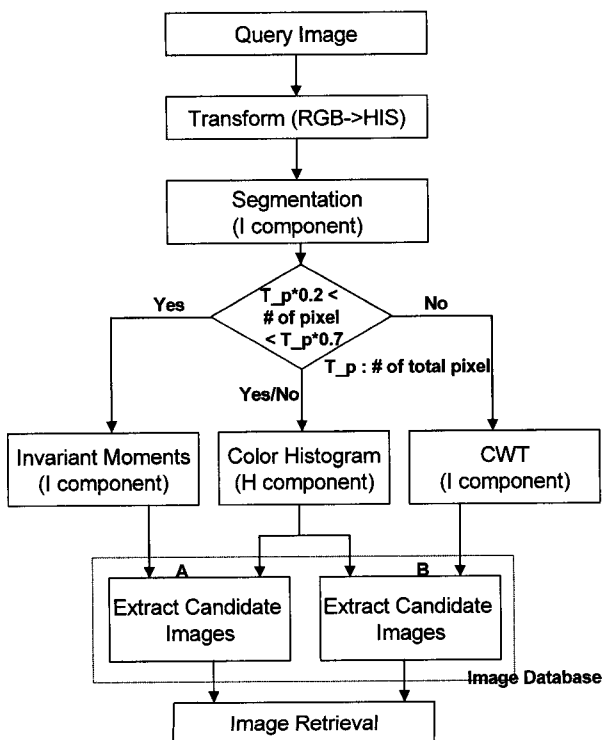


그림 4. 제안된 알고리즘의 전체 블록도
Fig. 4. Whole block diagram of the proposed algorithm.

계산하고 I 성분으로부터 불변 모멘트와 CWT를 계산하였다.

1. 영상 분할

질의 영상에서 목적영상을 추출하기 위하여 Active Contour 알고리즘을 적용하였다. 그림 5는 Active Contour 알고리즘이 적용된 영상을 보여주고 있다.

그림 5에서 (a)와 (c)는 질의 영상, (b)와 (d)는 Active Contour 알고리즘이 적용되어 일정 영역이 추출된 영상이다. 그림 5에서 알 수 있듯이 일정 크기 이상의 목적물이 질의 영상 내에 포함되어 있으면 (b)와 같이 일정 영역을 분할하여 나타낼 수 있으나 (c)와 같이 임의의 패턴으로 구성되어 있는 영상의 경우 일정 영역을 추출해내도 의미 없는 영역이 되기 쉽다. 그러므로 영상 분할 알고리즘을 이용하여 영상을 두 가지 종류 즉, 목적영역이 존재하는 영상과, 패턴으로 이루어지는 영상으로 나누어 영상 특성을 추출하는 알고리즘을 이원화하여 적용시킨다.

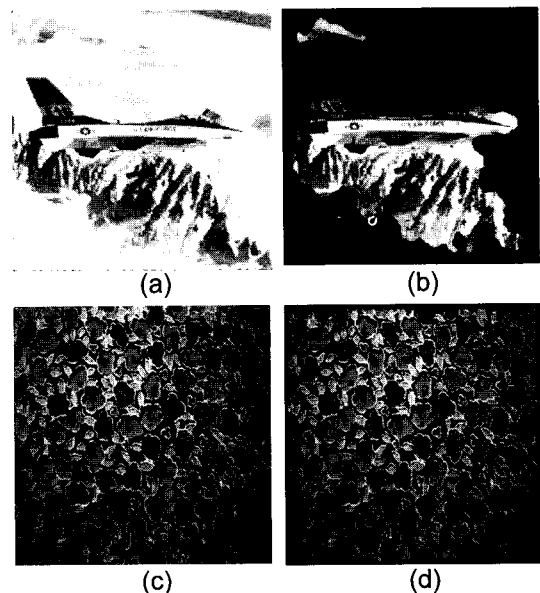


그림 5. 영상 분할
Fig. 5. Image segmentation.

2. 색상 특징 추출

질의 영상으로부터 색상 성분(Hue)을 추출하여 그림 3의 색상 양자화 테이블에 매핑시켜 색상 데이터를 추출하였다. 추출된 색상 데이터는 그림 6과 같이 12개의 색상으로 구분되어 히스토그램 형태로 저장되며 DB(Database)에 저장된 색상 히스토그램과의 상관관계 계산에 사용된다.

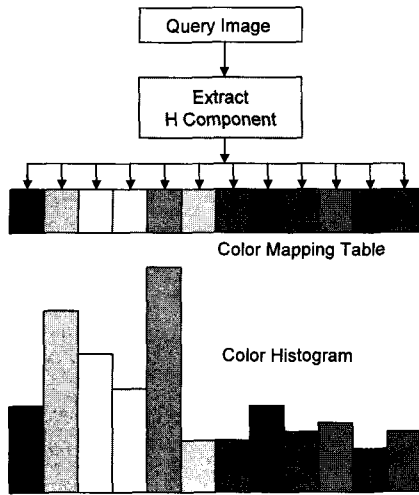


그림 6. 색상 특징 추출 과정
Fig. 6. The procedure of the extracted color feature.

3. 형태 및 공간적 위치 특징 추출

본 논문에서 사용한 Hu 불변 모멘트는 총 7개의 모멘트 성분에 영상의 형태와 공간적 위치를 담고 있으나 5번째 ϕ_5 와 7번째 ϕ_7 모멘트는 값의 변화가 미세하여 영상 검색에 큰 영향을 미치지 않으며, 또한 계산량의 증대를 가져오므로 검색 알고리즘에서 제외하였다. 영상 분할 알고리즘에서 목적영역이 존재하는 질의 영상으로 분류되면 영상의 질감 특성보다 목적영역의 형태 및 위치 특성을 이용한다. 그림 7은 형태 및 공간적 위치 특징에 사용되는 DB의 기본 형태로 5개의 Hu 불변 모멘트로 구성 되어 있다.

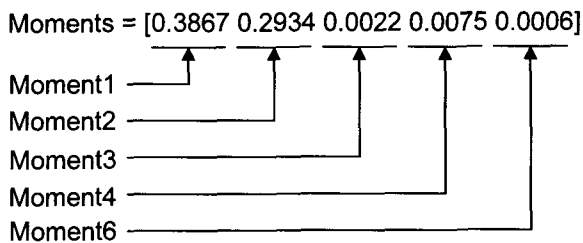


그림 7. 형태 및 공간적 위치 특징의 데이터베이스
Fig. 7. Database of the shape and spatial location feature.

4. 질감 특징 추출

영상 분할 알고리즘에서 패턴으로 이루어진 영상으로 구분된 영상은 특정 목적영역이 없으므로 질감 특징을 추출하여 영상을 검색한다. 그림 8은 본 논문에서 사용한 DT-CWT의 주파수 응답 특성을 나타낸다.



그림 8. DT-CWT의 주파수 응답 특성
Fig. 8. Frequency response of the DT-CWT.

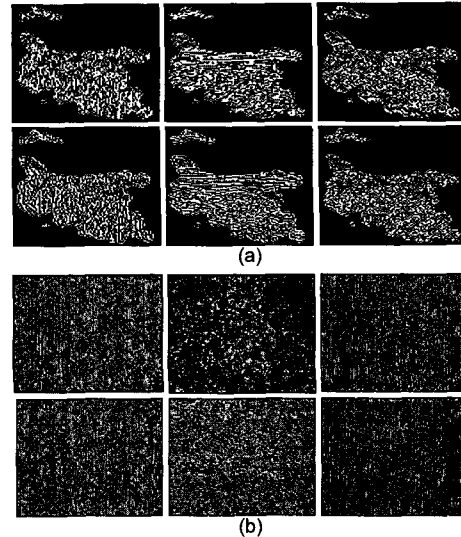


그림 9. 그림 5의 (b)와 (d)에서 추출된 질감 특성
Fig. 9. The extracted texture feature from (b) and (d) of the Fig 5.

그림 8에서 영상의 질감 특성은 {15, -15, 45, -45, 75, -75}로 이루어져 있으며 그림 9는 그림 5의 (b)와 (d)의 질감 특성을 나타낸다.

본 논문에서는 식(12)와 식(13)을 사용하여 추출된 질감의 전체 에너지와 표준편차를 계산하여 영상 검색에 사용하였다.

$$E_k = \frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N |W_k(i, j)| \quad (12)$$

$$\sigma_k = \left(\frac{1}{M \times N} \sum_{i=1}^M \sum_{j=1}^N (W_k(i, j) - \mu_k)^2 \right)^{\frac{1}{2}} \quad (13)$$

5. 제안된 영상검색 알고리즘의 데이터베이스

영상 검색을 효과적으로 하기위하여 그림 10과 같이 데이터베이스를 구현하였다.

데이터베이스는 총 5개의 항목으로 이루어져있다.

- Image Num. : 영상의 이름을 담고 있는 부분으로써 본 논문에서는 총 2000장의 이미지를 사용하므로 1부터 2000 사이의 값을 가진다.

Image Num.	D_region	C_Hist	Moments	CWT
Image Num. : 1~2000				
D_region : 0 or 1				
C_Hist = [1, 2, 3, ..., 12]				
Moments = [Moment1, Moment2, ..., Moment6];				
CWT = [std, mean]				

그림 10. 제안된 알고리즘의 데이터베이스 구조
Fig. 10. Database structure of the proposed algorithm.

- D_region : 영상분할 알고리즘을 통해 목적 영역이 있으면 0, 패턴 영상의 경우에는 1로 저장된다.
- C_Hist : 12개로 분할된 색상 히스토그램의 값을 가지고 있는 부분.
- Moments : 총 5개의 Hu 불변 모멘트가 저장되어 있다.
- CWT : CWT에 의해 추출된 질감 정보의 전체 에너지와 표준편차 값을 가지고 있다.

본 논문에서는 D_region이 '0' 일 때는 데이터베이스에 있는 C_Hist와 Moments 데이터를 이용하여 질의 영상과의 상관관계를 계산하고, D_region이 '1'일 경우에는 C_Hist와 CWT에 저장되어있는 전체 에너지와 표준편차 값을 이용하여 영상을 검색한다.

IV. 실험 및 결과 검토

본 논문에서 제안된 영상 검색 알고리즘의 성능평가를 위한 PC 환경은 펜티엄 3.2GHz Dual Core CPU, 4GB

메모리이며, 테스트 영상 DB는 <http://www.freefoto.com>에서 제공하는 2,000장(125X83)을 200개의 장르 별로 사용하였다. 그림 11은 실험에서 사용한 영상의 일부이다.

식 (14)와 (15)는 정확성(Precision)과 회수율(Recall)로 제안된 영상 검색 알고리즘의 성능을 평가에 사용되었다.

$$P = \frac{N_c}{N_c + N_f} \tag{14}$$

N_c : 검색된 영상 중, 질의 영상과 유사한 영상의 수

N_f : 검색된 영상 중, 질의 영상과 유사하지 않은 영상의 수

$$R = \frac{N_c}{N_t}, \quad N_t = N_c + N_m \tag{15}$$

N_c : 검색된 영상 중, 질의 영상과 유사한 영상의 수

N_t : 데이터베이스 안에 있는 질의 영상과 유사한 영상의 수

N_m : 유사한 영상임에도 불구하고 검색되지 못한 채 데이터베이스 안에 남아있는 영상의 수

본 논문에서는 영상분할 알고리즘을 사용하여 영상을 목적 영역이 있는 영상과 패턴 영상으로 분리한 다음 목적 영역이 있는 영상에는 색상 히스토그램과 Hu 불변 모멘트를, 패턴 영상에는 색상 히스토그램과

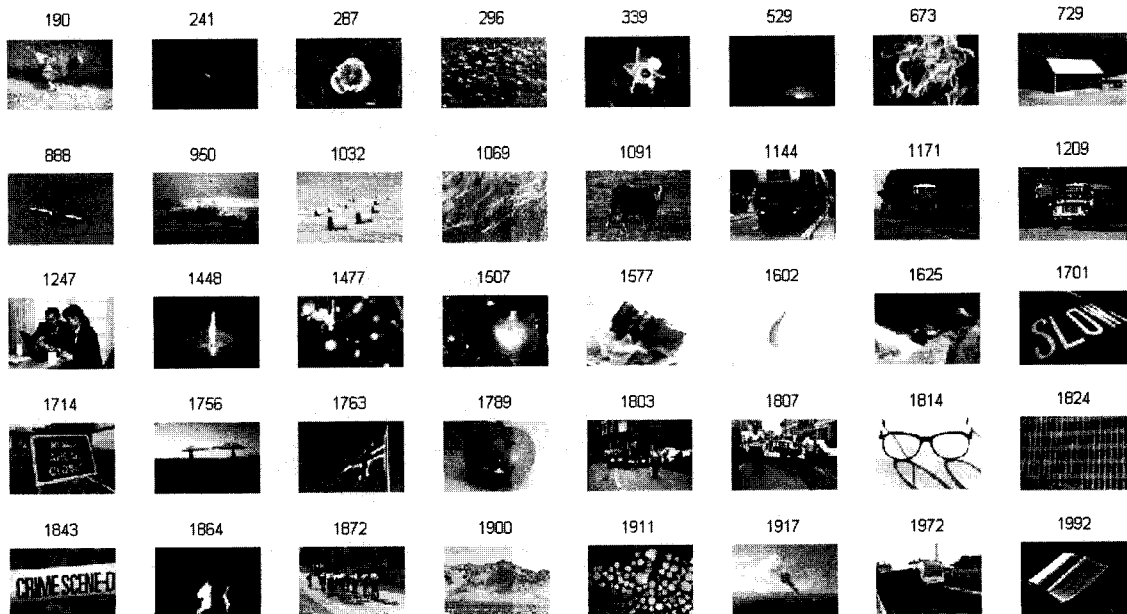


그림 11. 영상 데이터베이스
Fig. 11. Image database.

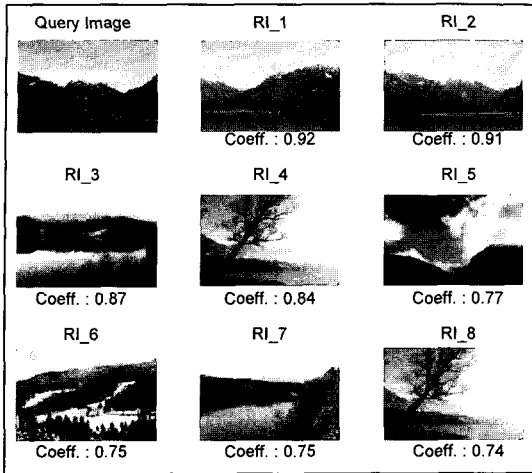


그림 12. 색상 히스토그램 적용
Fig. 12. Applying color histogram.

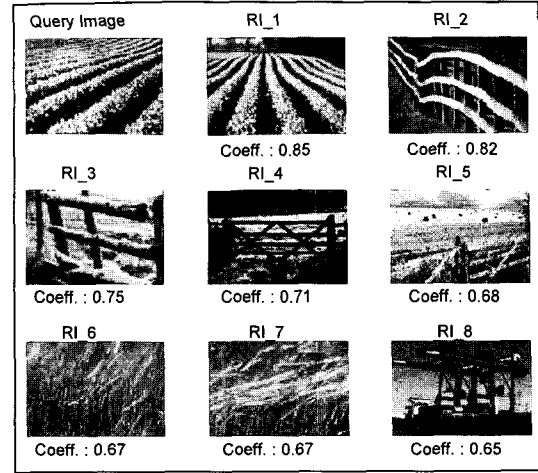


그림 14. CWT 적용
Fig. 14. Applying complex wavelet transform.

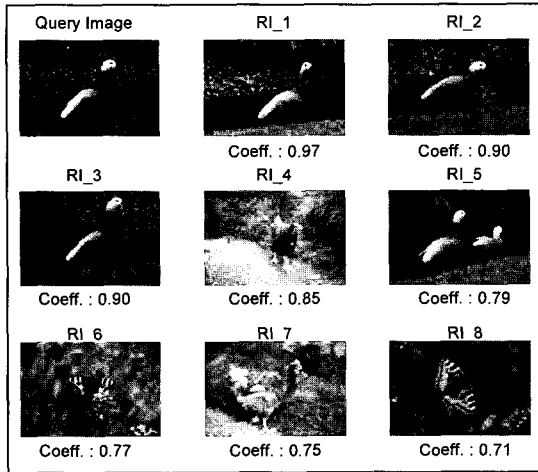


그림 13. Hu 불변 모멘트 적용
Fig. 13. Applying Hu invariant moment.

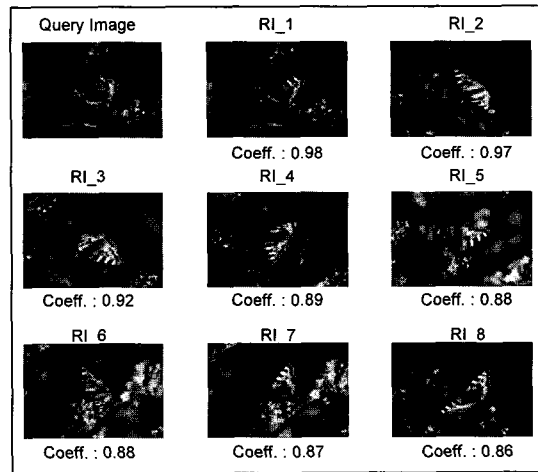


그림 15. 제안된 알고리즘
Fig. 15. The proposed algorithm.

표 1. 색상 히스토그램, Hu 불변 모멘트, CWT의 영상 검색 정확도와 재현율

Table 1. Image search precision and recall of the color histogram, Hu invariant moment, and CWT

Method	Color Histogram	Hu Invariant Moments	CWT	
			Energy	Standard Deviation
Precision	87.42%	84.18%	69.12%	74.73%
Recall	83.92%	84.47%	87.07%	86.09%

표 2. 제안된 영상 검색 알고리즘의 정확도와 재현율

Table 2. Precision and recall of the proposed image search algorithm.

D_region	0	1	-
Method	Color Histogram + Hu Invariant Moments	Color Histogram +CWT	Average
Precision	95.84%	93.71%	94.78%
Recall	82.51%	82.86%	82.69%

표 3. 성능 비교

Table 3. Performance compare table.

Method	Precision	Recall
BDIPBVLC[11]	90.5%	83.4%
SCD[4]	84.5%	84.4%
CSD[4]	77.0%	85.7%
EHD[4]	74.0%	86.2%
Wavelet Moments[7]	84.0%	84.5%
Color Autocorrelogram[8]	91.5%	83.2%
WM&BVLC[27]	94.2%	82.8%
Proposed	94.8%	82.7%

CWT를 적용시켜 각각의 후보 이미지를 추출하였다.

그림 12, 13, 14는 각각 색상 히스토그램, Hu 불변 모멘트, CWT만을 사용하여 영상을 검색한 결과를 나타내며 그림 15는 본 논문에서 제안된 영상 검색 알고리즘이 적용된 결과이다.

색상 히스토그램을 사용한 경우에는 영상의 형태와 상관없이 전체적인 색상 위주의 영상이 검색됨을 알 수 있으며, 불변 모멘트는 기본적인 물체의 형태와 위치 정보가 비슷한 영상을 검색함을 알 수 있다. 그리고 CWT는 영상의 전반적인 질감 즉, 뚜렷한 방향 성분을 가지는 영상을 검색함에 효율적임을 알 수 있다.

표 1은 영역분할 알고리즘을 적용하기전 색상 히스토그램, Hu 불변 모멘트, 그리고 CWT를 데이터베이스에 각각 적용했을 때의 영상 검색 정확도를 나타낸다.

표 2는 제안된 영상 검색엔진의 정확도와 재현율을 나타낸다. 영상 분할 알고리즘에 의해 색상 히스토그램과 Hu 불변 모멘트가 적용된 경우, 그리고 색상 히스토그램과 CWT가 적용된 경우의 정확도와 재현율이다.

본 논문에서 제안된 검색 엔진의 객관적 성능 평가를 위하여 SCD, CSD, EHD^[4], 웨이블릿 모멘트^[7], 칼라 오토코렐로그래^[8], BDIPBVLC^[11]를 이용한 방법들과의 검색 성능을 비교하였다. 표 3은 제안된 검색 알고리즘과 기존연구와의 정확도를 비교한 값이다.

결과적으로 제안된 검색 엔진은 웨이블릿 변환 영역의 색상 및 질감 특징을 이용한 영상검색 알고리즘^[27]에 비해 정확도에서 0.6% 정도의 성능 향상을 가짐을 확인하였으며, 영상검색에 소요되는 시간은 2,000장에서 8장을 추출하는데 0.02초가 소요되었다.

V. 결 론

현재 대부분의 영상 검색 알고리즘은 내용, 질감, 그리고 위치나 형태기반의 영상 정보를 추출하여 영상을 검색하고 있다. 하지만 영상은 크게 영상 내에 특정 영역에 목적물을 가지는 영상과 특정 목적물이 없이 색상과 질감만을 가지는 영상으로 구분되어진다. 이렇게 목적물을 가지는 영상과 목적물 없이 색상과 질감만을 가지는 영상에 똑같은 영상 검색 알고리즘을 적용한다는 점은 많은 문제를 내포하고 있다. 본 논문에서는 이러한 문제점을 해결하기 위하여 Active Contour를 이용한 영상 분할 알고리즘을 사용하여 특정 영역에 목적물이 존재하는 영상과 목적물이 존재하지 않는 영상으로 구분하여 목적물이 존재하는 영상에서는 색상 히스토그램과 Hu 불변 모멘트를 적용하여 영상 정보를 추출하였고, 목적물이 존재하지 않는 영상에서는 색상 히스토그램과 CWT를 적용하여 영상 정보를 추출할 수 있는 알고리즘을 제안하였다.

실험 결과 제안된 검색 알고리즘의 성능은 정확도 94.8%, 재현율 82.7%를 나타내었다. 결과적으로 내용, 질감, 위치나 형태 특성만을 개별적으로 사용하는 알고리즘에 비해 영상을 두 가지 종류로 구분하여 사용할 경우 검색 성능의 향상을 가져옴을 확인할 수 있었다.

본 논문에서는 Active Contour를 사용하여 영상을 구분하였으나 실험 결과 영상 구분 성능에 따라 전체 알고리즘의 성능이 변화 되는 것을 확인하였다. 이에 따라, 영상 구분 알고리즘의 성능을 향상 시킬 수 있는 방향의 연구가 필요하겠으며 또한 CWT를 이용한 질감 특성 추출에 소모되는 하드웨어 자원의 소모 및 계산 시간을 최소화 시킬 수 있는 방향의 연구가 필요하다.

참고 문헌

- [1] M. J. Swain and D. H. Ballard, "Color indexing," *Int. J. Comput. Vis.*, vol. 7, no. 1, pp. 11-32, 1991.
- [2] M. Stricker and M. Orengo, "Similarity of color images," *SPIE: Storage Retrieval Image and Video Database III*, vol. 2420, pp. 381-392, Feb. 1995.
- [3] Y. Gong, H. Zhang, H. Chuant, and M. Skauuchi, "An image database system with content capturing and fast image indexing abilities," in *Proc. Int. Conf. Multimedia Computing and Systems*, May 1994, pp. 121-130.
- [4] "ISO/IEC 15938-3/FDIS Information Technology-Multimedia Content Description Interface-part 3 Visual," *ISO/IEC/JTC1/SC29/ WG11, Doc. N4358*, Sydney, Australia. July 2001.
- [5] Kian-Lee Tan, Beng Chin Ooi, Chia Yeow Yee, "An Evaluation of Color-Spatial Retrieval Techniques for large Databases," *Multimedia Tools and Applications*, vol. 14, pp. 55-78, 2001.
- [6] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Texture features for image classification," *IEEE Trans. Syst. Man Cybern.*, vol. 8, pp. 610-621, Nov. 1973.
- [7] K. S. Thyagarajan, T. Nguyen, and C. Persons, "A maximum likelihood approach to texture classification using wavelet transform," in *Proc. of IEEE Conf. on Image Processing*, pp. 640-644, Austin, USA, Nov. 1994.
- [8] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing Zhu, and Ramin Zabih, "Image indexing using color correlograms," in *Proc. of Recognition*, pp. 762-768, Virginia, USA, July 1997.
- [9] J. R. Smith, S. F. Chang, "Integrated Spatial and Feature Image Query," *Multimedia Systems*, vol. 7, pp. 129-140, March 1999.
- [10] Y. Rui and T. S. Huang, "Image retrieval: Current techniques, promising directions, and open issues," *J. Visual Communication and Image Representation*, vol. 10, no. 4, pp. 39-62, Oct. 1999.
- [11] L. Cinque, S. Levialdi, K.A. Olsen, A. Pellicano, "Color-Based Image Retrieval Using Spatial Chromatic Histograms," In *Proc. of the Multimedia Systems*, vol. 2, pp. 969-973, June 1999.
- [12] D. Feng, W. C. Siu, and H. J. Zhang, *Multimedia Information Retrieval and Management-Technological Fundamentals and Applications*, Springer, pp. 4-24, 2003.
- [13] M. Flickner, H. Sawhney, W. Niblack, and J. Ashley, "Query by image and video content: The QBIC system," *IEEE Computer*, vol. 28, no. 9, pp. 23-32, Sep. 1995.
- [14] W. Y. Ma and B. S. Manjunath, "Netra: A toolbox for navigating large image database," in *Proc. Int. Conf. Image Processing*, vol. 1, 1997, pp. 568-571.
- [15] J. R. Smith and S. F. Chang, "VisualSEEK: A fully automated contentbased image query system," *ACM Multimedia*, pp. 87-98, 1996.
- [16] 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*, 1999, pp. 509-516.
- [17] J. Li, J. Z. Wang, and G. Wiederhold, "IRM: Integrated region matching for image retrieval," *ACM Multimedia*, pp. 147-156, 2000.
- [18] ByoungChul Ko, Hyeran Byun, "FRIP: A Region-Based Image Retrieval Tool Using Automatic Image Segmentation and Stepwise Boolean AND Matching," *IEEE Multimedia*, Vol. 7, NO. 1, Feb. 2005.
- [19] B. Moghaddam, H. Biermann, and D. Margaritis, "Defining image content with multiple regions-of-interest," in *Proc. IEEE Workshop on CBAIVL*, 1999, pp. 89-93.
- [20] Q. Tian, Y. Wu, and S. Thomas, "Combine user defined region-of-interest and spatial layout for image retrieval," in *Proc. IEEE Int. Conf. Image Processing*, vol. 3, 2000, pp. 746-749.
- [21] Chunming Li, Chenyang Xu, Changfeng Gui, and Martin D. Fox, "Level Set Evolution Without Re-initialization: A New Variational Formulation," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition Proc., CVPR/05*, 2005.
- [22] N. G. Kingsbury, "Image processing with complex wavelet," *Phil. Trans. Roy. Soc. London A*, vol. 357, pp. 2543-2560, Sep. 1999.
- [23] Papoulis, "Probability, Random Variables, and Stochastic Processes," McGraw Hill, 1965.
- [24] Morton Nadler and Eric P. Smith, "Pattern Recognition Engineering," Wiley-Interscience, pp.197~199, 1993.
- [25] Cho-Huak Teh and Roland T. Chin, "On Digital Approximation of Moment invariants," *Computer Vision, Graphics, And Image Processing*, Vol. 33, pp. 318~326, 1986.
- [26] M. K. Hu, "Pattern recognition by moment invariants," *Proc. IEEE*, vol. 49, no. 9, pp. 1428,

Sept. 1961.

[27] Y. K. Chun, J. K. Sung and N. C. Kim, "Image Retrieval using Multiresolution Color and Texture Features in Wavelet Transform Domain," Journal of The Institute of Electronics Engineers of Korea, Vol. 43-SP, NO. 1, January 2006.

저 자 소 개



이 강 현(평생회원)-교신저자
 1979년, 1981년 조선대학교 전자
 공학과 공학사 및 석사
 1991년 아주대학교 대학원
 공학박사
 1977년~현재 조선대학교 교수
 1991년, 1994년 미 스탠포드대
 CRC 협동연구원.

1996년 호주시드니대 SEDAL 객원교수
 2000년~현재 한국 멀티미디어 기술사 협회 이사
 2002년 영국 런던대 객원교수
 2002년 대한전자공학회 멀티미디어연구회전문
 위원장
 2003년 한국 인터넷 방송/TV 학회 수석부회장
 2003년~현재 대한전자공학회 정보화위원회 이사
 2005년~현재 조선대학교 RIS지원 사업단장
 <주관심분야 : 멀티미디어 시스템 설계, Ubiquitous
 convergence, 디지털 시네마 DRM>