

내용 기반 영상 검색 시스템을 위한 효율적인 특징 벡터 추출에 관한 연구

유 기 형[†] · 곽 훈 성^{**}

요 약

최근 다양하고 방대한 멀티미디어 데이터를 효율적으로 저장, 관리 및 검색할 수 있는 멀티미디어 데이터베이스 시스템이 정보화 사회의 중요한 핵심 기술로 대두되고 있다. 내용 기반 이미지 검색을 위해 본 논문에서는 웨이블릿 변환과 에너지 값을 사용하여 이미지 데이터로부터 특징 벡터를 완전 자동으로 추출하는 방법과 이를 이용한 효율적인 검색 기법을 제안한다. 웨이블릿 변환은 이미지 압축이나 신호 분석 등에서 많이 사용되며, 특히 웨이블릿 계수 값은 영상의 특성을 잘 반영하고 웨이블릿 영역에서 계산되는 예제영상(Query image)과 데이터베이스에 저장된 영상간의 유사성을 추정하는데 더 효율적이다. 영상 검색에 있어, 특징 벡터로 사용되는 표준편차와 평균 값을 에너지 값과 비교 분석하였다. 실험결과, 표준편차나 평균 값을 이용하는 것보다 에너지 값을 사용하는 것이 더 효과적이었다.

키워드 : 영상검색, 멀티미디어, 인터넷

A Study on Efficient Feature-Vector Extraction for Content-Based Image Retrieval System

Gi-Hyoung Yoo[†] · Hoon-Sung Kwak^{**}

ABSTRACT

Recently, multimedia DBMS is appeared to be the core technology of the information society to store, manage and retrieve multimedia data efficiently. In this paper, we propose a new method for content based-retrieval system using wavelet transform, energy value to extract automatically feature vector from image data, and suggest an effective retrieval technique through this method. Wavelet transform is widely used in image compression and digital signal analysis, and its coefficient values reflect image feature very well. The correlation in wavelet domain between query image data and the stored data in database is used to calculate similarity. In order to assess the image retrieval performance, a set of hundreds images are run. The method using standard derivation and mean value used for feature vector extraction are compared with that of our method based on energy value. For the simulation results, our energy value method was more effective than the one using standard derivation and mean value.

Key Words : Image Retrieval, Multimedia, Internet

1. 서 론

현재의 세계는 급격한 멀티미디어 정보의 증가와 활용을 경험하고 있다. 이는 다음과 같은 여러 가지 기술 발전에 기인하고 있다. 첫째로 디지털 오디오 비주얼(Audio-Visual) 데이터 처리 및 압축 기술의 발전과 이와 관련된 국제 표준화의 성공적인 도출이다. 둘째로 고성능 개인용 컴퓨터(PC), 대용량 저장 장치의 보편화 및 World Wide Web(WWW)으로 대변되는 컴퓨터 네트워크의 발전에 따라 디지털로 표현

된 멀티미디어 정보의 생성, 전송, 가공이 매우 용이해졌다. WWW은 e-mail, News group 등을 통한 정보 교환, 홈페이지를 통한 기업의 홍보, 상품의 광고 및 거래, Web방송 등 이미 그 활용도는 일상의 주요 활동에 없어서는 안될 수단으로 자리잡고 있다. 따라서 영상정보를 저장하고 검색하기 위한 구체적인 연구가 필요하며, 현재 활발하게 진행되고 있다. 그러나, 엄청난 속도로 증가하는 멀티미디어 정보 중에서 사용자가 필요로 하는 내용의 정보를 찾기 위해서는 기존의 키워드 기반의 검색은 한계에 도달한 상황이기 때문에 사용자가 원하는 정보를 내용에 기반하여 검색할 수 있는 방법이 요구되고 있다.

키워드 기반 검색 방법은 대상이 되는 각각의 영상에 사

[†] 정 회 원 : 전북대학교 컴퓨터공학과 박사과정

^{**} 정 회 원 : 전북대학교 컴퓨터공학과 주임교수

논문접수 : 2005년 7월 20일, 심사완료 : 2006년 5월 18일

람이 문자 주제어(keyword)를 일일이 부과하고 이를 이용하여 영상을 검색하는 방법이다. 이 방법은 제한된 범위 내에서는 정확도가 높다는 장점을 가지고 있지만 방대한 영상정보에 대해 사람이 일일이 주제어를 부여하기가 쉽지 않고, 같은 영상에 대해 다른 주제어가 주어져 있으면 검색이 불가능하다는 단점이 있다.

내용 기반 검색 방법은 영상 자체 내에서 대표할 수 있는 특징을 추출하여 이를 기반으로 색인 및 검색하는 기법이다 [1-3]. 이 방법은 텍스트 기반 검색기법의 단점을 극복하기 위하여 영상 내에서 고유한 특징을 자동적으로 추출하여 검색함으로써 시간 및 인력 소모를 줄일 수 있다는 장점을 가지고 있다. 즉, 칼라(color), 모양(shape), 그리고 질감(texture) 등의 내용 표현요소들을 통하여 얻어진 특징 값을 사용하여 영상들 간의 유사도를 계산하고 검색을 수행한다. 그러므로, 효율적인 영상 검색을 위해서는 각 영상을 대표하는 내용 표현 요소들의 효과적인 특징 추출이 중요하다. 따라서 최근 효율적인 특징벡터를 찾는 데 많은 연구가 진행되고 있다. 사용될 수 있는 특징벡터로는 색상, 무늬, 질감, 윤곽선, 형태, 공간 관계 등과 같은 각각의 영상에 고유한 특징들이 있다.

영상을 이용한 내용 기반 정보검색 시스템은 디지털 도서관에 필요한 멀티미디어 정보검색의 필요성에 따라 Columbia 대학의 VisualSeek[4] 시스템, MIT Media Lab의 Photo Book[5] 등 미국 및 캐나다[6] 등지에서 광범위하게 연구되고 있으며 그 응용성 역시 다양해지고 있는 추세이다. 최근 국내에서도 대학 및 연구소에서 대상체 이미지의 형태, 색, 질감 등의 특징을 기반으로 연구하고 있다.

내용 기반 검색 시스템 개발은 크게 두 가지 영역으로 나눌 수 있다. 첫째, 영상처리(Image Processing) 관점에서 영상(Image) 데이터의 특성과 변환(Transform) 방법을 연구하고, 이를 이미지 특징벡터(Feature Vector)를 추출하여 검색하는데 사용하는 것이다. 둘째, 데이터베이스 관점에서 영상으로부터 추출된 특징벡터를 사용하여 효과적으로 영상 데이터를 인덱싱(Indexing)하는 것이다. 본 논문에서는 전자인 이미지 프로세싱 관점에서 접근하여 영상 데이터의 특성과 변환(Transform) 방법을 연구하고, 영상 데이터의 효율적인 특징벡터를 추출하여 영상을 검색하는데 초점을 맞춘다.

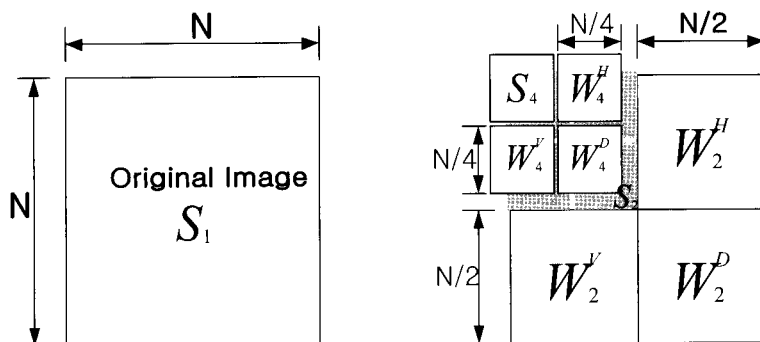
본 논문에서는 내용 기반 검색 시스템에서 사용할 수 있는 효율적인 특징벡터를 추출하기 위한 전처리 과정을 거친다. 먼저 RGB 칼라 영상을 Gray 레벨의 영상으로 변환시킨 다음 공간 정보를 잘 보존하면서 탁월한 압축 효과를 얻을 수 있는 웨이브렛 변환(Wavelet transform)을 이용한다. 유사도 측정 단계에서는 기존의 유사도 측정에 많이 사용하는 유클리디안 거리(Euclidean distance)를 사용한다.

현재 영상에 대한 정보를 얻기 위해서 많이 사용되는 특징벡터로는 평균과 표준편차가 있다[7]. 하지만, 이 값들은 영상들 사이의 차분 값이 작기 때문에 유사도 측정 시에 너무 많은 후보 영상을 추출함으로써 검색효율을 떨어뜨리기 쉽다. 따라서 본 논문에서는 영상의 특징을 잘 반영하면서 영상들 사이의 차분 값을 고려하여 비교적 구분이 명확한 에너지 값을 사용하는 방법을 제외하였다.

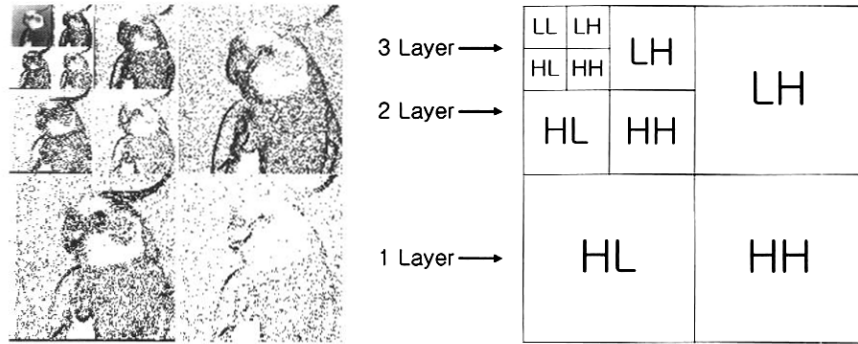
2. 웨이브렛 변환(Wavelet transform)

영상은 많은 압축 가능성을 내포하고 있는데, 이는 영상 내부의 화소(pixel)간에 높은 상관성이 있기 때문이다. 영상 내의 중복성을 효과적으로 제거하면 상대적으로 적은 양의 데이터를 가지고 효과적인 데이터 처리와 검색을 할 수가 있다[8]. 영상의 중복성 제거는 두 가지로 분리할 수 있는데 하나는 아무런 전처리 없이 공간 영역에서 수행하는 방법이며, 다른 하나는 변환영역에서 부호화 하는 것이다. 일반적으로 변환 기법을 이용하여 특정 대역으로 에너지를 집중시켜 압축하고 검색하는 것이 효과적이다.

(그림 1)는 웨이브렛 변환을 수행한 영상의 분할 형태를 보인 것으로, 크기가 $N \times N$ 인 원 영상 S_1 을 2단계 웨이브렛 분할한 것이다. 먼저 1단계 분할에서 하나의 저대역 부밴드 S_2 와 3개의 방향 성분을 가지는 고대역 부밴드 W_2^H, W_2^V, W_2^D 가 생성된다. 2단계 분할에서 저대역 부밴드 S_2 는 하나의 저대역 부밴드 S_4 와 3개의 고대역 부밴드 W_4^H, W_4^V, W_4^D 로 분해된다. 즉, 웨이브렛 변환은 가장 낮은 저대역 부밴드를 다음 단계의 부밴드들로 분할해 나가는 것으로, 다해상도의 피라미드 구조를 형성한다. 역 웨이브렛 변환은 이와 반대 방법으로 처리된다.



(그림 1) 2계층으로 분할한 웨이브렛 변환



(그림 2) 3계층(Layer) 웨이브렛 변환 영상과 부 밴드(Sub-Band) 표시

분할된 부밴드들은 하나의 저대역 부밴드와 수평, 수직, 대각선 방향 성분을 가지는 고대역 부밴드들로 구성된다. 저대역 부밴드는 저주파 성분으로 대부분의 에너지가 집중되어 있는 원 영상의 축소판(thumbnail version)이기 때문에 모든 웨이브렛 부밴드 중에서 가장 중요하다. 나머지 수평, 수직, 대각선 부밴드들은 각각 수평, 수직, 대각선 방향의 영상의 에지 정보를 제공한다[9, 10].

3. 특징벡터 추출과정

전처리 과정에서는 후보 영상들에 대해서 웨이브렛 변환을 수행한 후 정합하는데 필요한 특징벡터를 추출한다. 전처리 과정(Preprocessing)은 다음과 같다.

- ① 모든 영상을 128×128 크기로 표준화한다.
- ② 칼라(R,G,B) 영상을 8-비트 그레이(gray) 영상으로 변환한다.
- ③ 원 영상을 3계층(Layer) 웨이브렛 변환을 한다.

(그림 2)는 3계층 웨이브렛 변환한 결과와 각 대역을 표시한 것이다.

웨이브렛 변환의 특성상 3 계층 저대역 부밴드(LL)는 원 영상의 특징을 대부분 가지고 있으면서도, 크기는 64 : 1로 압축되어 있으므로 검색효율을 높일 수 있는 훌륭한 특징벡터가 된다.

4. 검색 과정

전처리 과정을 통해서 추출된 특징벡터를 다음과 같은 과정을 통하여 주어진 예제 영상과 후보 영상 간의 유사도를 계산한다.

- ① 각각의 후보 영상들에 대해 16×16 크기의 3계층 웨이브렛 저대역 부밴드(LL)대역의 에너지 값을 추출한다.
- ② 3계층 저대역 부밴드(LL)를 제외한 나머지 세 대역(LH, HL, HH)의 에지 성분을 비교한다.

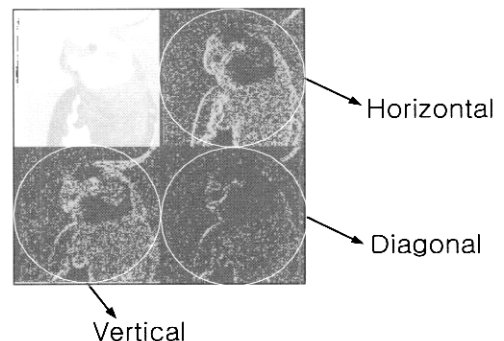
- ③ 3계층 저대역 부밴드(LL)를 이용하여 Euclidean Distance를 계산한다.

검색 효율을 높이기 위해서는 간단하면서도 영상의 특성을 반영하는 값을 사용하여 유사 정합되는 영상의 수를 줄일 필요가 있다. 그러한 목적으로 전처리 과정에서 웨이브렛 변환을 수행하였고 3단계까지 분해하여 16×16 크기의 3계층 LL대역의 에너지 값을 간단하면서도 효율적인 특징벡터로 사용한다. 3계층 LL대역의 에너지 값은 다음과 같으며, f_{LL}^* 는 LL대역의 변환계수이다.

$$E = \sum_i \sum_j |f_{LL}^*(i,j)|^2 \tag{4}$$

예제 영상(Query Image)의 에너지 값을 계산하여 그 에너지 값의 $\pm\beta\%$ 범주 안에 들어오는 영상들을 후보 영상으로 선택한다. 후보 영상을 많이 얻기 원하면 β 값을 조절하면 된다.

한편, 저주파 대역은 영상의 에지 정보 즉 윤곽선 정보는 거의 포함하고 있지 않으므로 저주파 대역에서 얻어진 에너지 값 외에 나머지 대역에서 윤곽선 정보를 추출하여 특징벡터로 사용한다. 윤곽선을 추출하는 방법은 다양하다. 웨이브렛 변환을 이용하는 경우에는 윤곽선 추출을 위한 별도의 작업이 필요치 않다. (그림 3)를 통해서 알 수 있는 바와 같이 저주파 대역을 제외한 나머지 대역이 가로와 세로, 대각선 방향의 윤곽선을 잘 나타내고 있다.

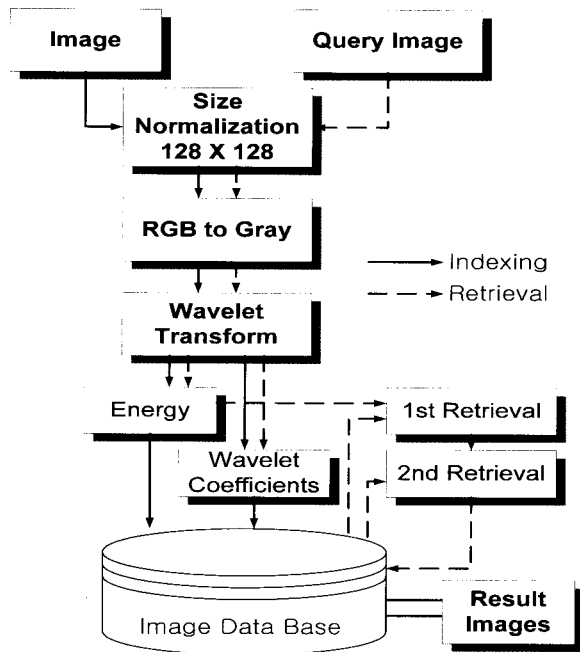


(그림 3) 웨이브렛 변환 영역에서 영상의 특성

마지막으로 예제 영상과 후보 영상의 각 pixel 간의 차분을 계산하는 유클리디안 거리(Euclidean distance)는 식 (5)와 같다.

$$D = \sum_x \sum_y |f(x,y) - w(x,y)| \quad (5)$$

식 (5)을 통해서 알 수 있듯이 비교하는 영상이 비슷할수록 작은 값을 나타내고, 만약 비교하는 두 영상이 같다면 값은 0이 될 것이다. 따라서 유클리디안 거리 값이 가장 작은 것부터 M 개만큼을 최종적인 검색 결과로 디스플레이 한다. (그림 4)는 제안한 시스템 구성도이다.



(그림 4) 제안한 영상 검색 시스템

5. 실험 및 결과

실험에서는 영상 검색 시스템의 1차 검색 과정에서 특징 벡터로 사용한 에너지 값의 효율성을 기존의 내용 기반 검색 시스템에서 주로 사용하는 표준편차와 평균 값과의 검색 결과를 비교하여 성능분석을 한 것이다.

제안한 알고리즘에 관한 실험 환경은 다음과 같다.

- IBM-PC/Pentium-4
- CPU : Intel Pentium-2.4GHz
- Main Memory : 1GBytes
- Compiler : C++

실험에 사용된 영상 데이터는 실험의 객관성을 위하여 Web 상의 Corel Draw Digitalriver에 분류되어 있는 200점의 영상을 사용하였다. 각 실험 영상과 사용자 질의 영상은 RGB 컬러 bmp 파일로 128×128 크기로 정규화된 256 컬러 영상이다.

실험에 사용된 영상 데이터를 유형별로 분류하면 <표 1>과 같다.

<표 1> 실험 영상의 종류

영상의 종류	개 수
자동차 영상	38
꽃 영상	60
인물 영상	39
풍경 영상	30
일몰 영상	33
합 계	200

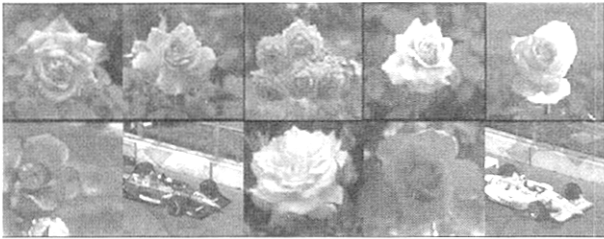
<표 1>에서 분류한 각각의 영상들을 예제 영상으로 주었을 때의 결과는 다음과 같다.

<표 2> 검색 결과

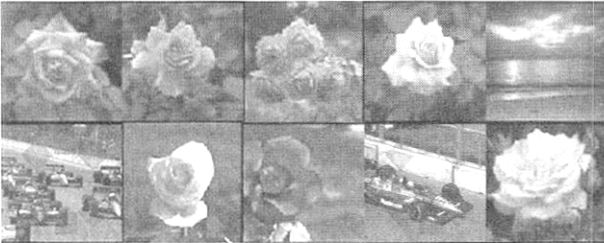
예제영상	검색 방법	1차 검색	2차 검색	
			상위 10개	상위 15개
꽃	에너지	80(24)	8	12
	평균	145(39)	7	10
	표준편차	95(31)	7	10
자동차	에너지	110(26)	6	7
	평균	176(38)	6	6
	표준편차	126(33)	5	7
인물	에너지	75(23)	6	11
	평균	115(27)	7	8
	표준편차	117(7)	2	3
일몰	에너지	100(23)	2	3
	평균	173(34)	2	3
	표준편차	123(25)	2	3

<표 2>의 1차 검색이란 각각 에너지, 평균, 표준편차 값을 사용하여 검색한 결과를 말한다. 2차 검색은 1차 검색된 후보 영상들 중 유클리디안 거리를 사용하여 검색된 결과이다. 괄호 안의 숫자는 검색된 영상 중 관련된 영상, 즉 예제 영상이 꽃 영상이라면 꽃 영상의 개수를 말한다. 1차 검색과 2차 검색은 주어진 예제 영상의 유사도가 ±25% 범주 안의 후보 영상들을 선택하므로 검색된 영상의 개수는 서로 다르다. 보통 영상에 대한 변화도가 평균이나 표준편차에 비해 에너지 값이 크기 때문에 유사도 ±25% 범주의 영상을 선택하였을 때 더 적은 수의 영상들이 검색된다. 즉, 검색효율이 증가하게 된다. <표 2>의 최종 검색에서 알 수 있듯이 상위 10개와 15개를 비교했을 때, 에너지 값을 특징벡터로 사용하는 방법이 비교된 방법들 보다 비슷하거나 더 좋은 결과를 얻었다.

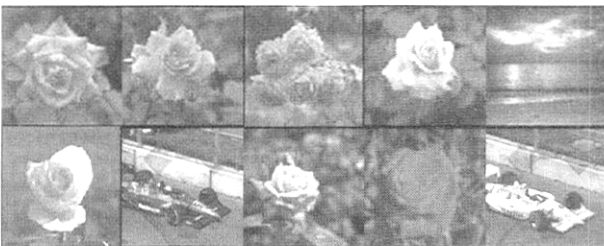
(그림 5)~(그림 7)은 꽃 영상을 예제영상으로 하였으며, 에너지 값, 평균, 표준편차에 의한 각각의 최종 검색 상위 10개의 영상을 보여주고 있다.



(그림 5) 에너지를 특징벡터로 사용하여 검색된 영상들



(그림 6) 평균을 특징벡터로 사용하여 검색된 영상들



(그림 7) 표준편차를 특징벡터로 사용하여 검색된 영상들

표와 그림을 통해서 제시된 바와 같이 에너지 값이 평균, 표준편차 등과 비교했을 때 영상의 특성을 정확하게 반영하는 효율적인 특징벡터임을 알 수 있다. 위에서 수행한 모든 실험의 결과를 Precision과 AVRR로 <표 3>에 나타내었다.

<표 3> <표 2>에 대한 Precision 및 AVRR 비교

검색 방법	상위 10개		상위 15개	
	Precision	AVRR	Precision	AVRR
에너지	0.8	3.8	0.8	5.87
평균	0.7	3.5	0.67	4.8
표준편차	0.7	3.3	0.67	4.8

<표 3>을 통해서 상위 10개를 비교하면 특징벡터로 에너지 값을 사용한 경우가 약간의 향상된 검색 효율을 얻었지만, 상위 15개를 비교한 경우 검색 효율에 있어서 보다 많은 차이를 보였다.

6. 결 론

본 논문에서는 간단하면서도 영상의 특성을 잘 반영하는

효율적인 특징벡터 추출 기법을 제안하였다. 영상 데이터은 압축 효과를 얻으면서 공간 영역의 정보가 잘 보존되는 웨이브렛 변환을 적용하였고, 웨이브렛 변환 영역에서 계산된 저대역의 에너지 값을 사용하여 유사한 영상을 검색하였다.

실험을 통하여 본 논문에서 제안한 저대역 에너지 값이 평균이나 표준편차를 이용하는 기존 방법에 비하여 우수한 검색효율을 가지고 있음을 실험을 통하여 확인하였다. Precision 관점에서 상위 10개를 검색하였을 경우 10% 이상 더 효율적임을 알 수 있었고, 상위 15개를 비교하였을 경우에는 13% 이상의 검색 효율 향상이 있음을 확인할 수 있었다. AVRR 관점에서는 상위 10개와 15개의 경우 30%이상 검색 효율이 증가함을 알 수 있었다.

향후 연구 방향으로는 색상정보를 이용한 특징벡터나 배경에 무관하게 객체를 검색할 수 있는 방법들이 연구되어야 할 것이다. 또한 사용자의 임의의 그림에 대한 검색 방법 역시 연구되어야 할 것이다.

참 고 문 헌

- [1] Gi-Hyoung Yoo, Hoon-Sung Kwak "Content-based Retrieval Method using Wavelet Transform and Correlation", ICEIC pp.II-132-II-135, Yambian, China, Aug., 1998.
- [2] 유기형, 곽훈성, "웨이브렛 변환을 이용한 내용기반 검색 시스템", 전자공학회 하계학술대회 논문집, Vol.21, No.1, pp.733-736, 1998. 6.
- [3] 유기형, 곽훈성, "Content-Based Image Retrieval using Adaptive Color Histogram", 한국통신학회논문지 '05-9 Vol.30 No.9C
- [4] John R. Smith and Shih-Fu Chang, "VisualSEEK: a fully automated content-based image query system," <http://www.ctr.columbia.edu/VisualSEEK>.
- [5] R.W. Picard, "A Society of Models for Video and Image Libraries", <http://www.media.mit.edu/~picard/>.
- [6] Zw-Nian Li and Bing Yan, "Recognition Kernel for Content-based Search," IEEE International Conf. on SMC, Vol.1, pp. 472-477, Beijing, China, Oct., 1996.
- [7] Yixin Chen and James Ze Wang "Content-based image retrieval by clustering", Multimedia Information Retrieval pp.193-200, Berkeley, CA, USA, Nov., 2003.
- [8] P. W. Jones, "Digital Image Compression Techniques", SPIE OPTICAL Engineering press.
- [9] M. K. Mandal, T. Aboulnasr, "Fast wavelet histogram techniques for image indexing", Computer Vision and Image Understanding Vol.75, Nos.1/2, pp.99-110, July/August, 1999.
- [10] Yongqing Sun and Shinji Ozawa, "Semantic-meaningful content-based image retrieval in wavelet domain", Multimedia Information Retrieval pp.122-129, Berkeley, CA, USA, Nov., 2003

유 기 형



e-mail : ghyoo@chonbuk.ac.kr
 1989년 전북대학교 컴퓨터공학과(공학사)
 1991년 전북대학교 전자공학과(공학석사)
 1991년 (주)아트로닉스 연구원
 1991년 전북대학교 반도체 물성연구소
 연구원

1997년~현재 전북대학교 컴퓨터공학과 박사과정
 2003년~현재 비전정보 대표
 관심분야: 영상검색, 영상신호처리, 멀티미디어

곽 훈 성



e-mail : hskwak@chonbuk.ac.kr
 1970년 전북대학교 전기공학과(공학사)
 1979년 전북대학교 전자공학박사
 1981년~1982년 미국 텍사스주립대학
 연구교수
 1994년~1995년 국가교육연구전산망
 추진위원

1997년~1998년 전주영상축전조직 위원장 및 전북대학교
 영상산업특성화사업단장
 1998년 과학기술법령정비정책위원
 1999년~현재 조달청우수제품(정보통신)심사위원
 1997년~현재 (사)영상산업연구센터대표
 현 재 전북대학교 전자정보공학부 컴퓨터공학 교수 및 영상공
 학과(대학원) 주임교수
 관심분야: 영상신호처리, 인공지능, 컴퓨터비전, 멀티미디어 등