

논문 2008-08-02

웨이블릿 변환의 저주파수 부대역을 이용한 왜곡 영상 데이터베이스 검색

(Distorted Image Database Retrieval Using Low Frequency Sub-band of Wavelet Transform)

박하중, 김경진, 정호열*

(Ha-Joong Park, Kyeong-Jin Kim, Ho-Youl Jung)

Abstract : In this paper, we propose an efficient algorithm using wavelet transform for still image database retrieval. Especially, it uses only the lowest frequency sub-band in multi-level wavelet transform so that a retrieval system uses a smaller quantity of memory and takes a faster processing time. We extract different textured features, statistical information such as mean, variance and histogram, from low frequency sub-band. Then we measure the distances between the query image and the images in a database in terms of these features. To obtain good retrieval performance, we use the first feature (mean and variance of wavelet coefficients) to filter out most of the unlikely images. The rest of the images are considered to be candidate images. Then we apply the second feature (histogram of wavelet coefficient) to rank all the candidate images. To evaluate the algorithm, we create various distorted image databases using MIT VisTex texture images and PICS natural images. Through simulations we demonstrate that our method can achieve performance satisfactorily in terms of the retrieval accuracy as well as the both memory requirement and computational complexity. Therefore it is expected to provide good retrieval solution for JPEG-2000 using wavelet transform.

Keywords : Wavelet Transform, Lowest Resolution, Distorted Image Database, Content-based Image Retrieval(CBIR), Statistical Feature.

1. 서론

최근 들어, 컴퓨터, 모바일 단말 장치 및 인터넷 활용 기술의 발달로 멀티미디어 데이터는 기하급수적으로 증가하고 있다. 따라서 방대한 양의 멀티미디어 데이터베이스가 증가함에 따라 효율적인 멀티미디어 데이터 검색 시스템이 요구되고 있다[1]. 일반적으로 텍스트 기반 영상 검색기술은 질의어와 상이한 영상 데이터들이 찾아지는 경우가 많으므로 최근 들어 영상 데이터의 특징을 추출하는 내용 기반 영상 검색 시스템(Content-based Image Retrieval, CBIR) 알고리즘이 널리 사용되고 있다

[1-11]. 대부분의 내용기반 영상 검색 시스템은 영상이 포함하고 있는 색(color), 형상(shape), 질감(texture) 정보 등을 분석하여 특징 벡터로써 추출한 후 질의 영상에 가까운 영상을 다양한 유사도 측정 방법을 통해 검색을 수행한다. 이들은 효율적인 검색을 위해 선택적으로 사용될 수 있으며, 이들 중 질감 정보는 영상의 내용을 나타내는 중요한 시각적 특징 중의 하나로서 오랫동안 연구되어 왔으며 영상 또는 비디오 데이터를 내용기반 색인(index) 및 요약하는데 있어서 중요한 서술자(descriptor)로 이용된다[1]. 일반적으로 질감 영상의 유사도 측정을 위해 히스토그램과 같은 통계적인 분포를 이용한 방법을 많이 사용한다. 이는 매우 간단하지만 회전(rotation) 크기 변화(scaling) 천이(translation) 등과 같은 기하학적인 변형이 적용된 영상 검색에 강한 성능을 나타낸다[1].

대부분의 멀티미디어 데이터들은 저장 공간 및

* 교신저자(Corresponding Author)

논문 접수 : 2008. 3. 15 채택 확정 : 2008. 4. 15

박하중 : 삼성 테크윈 선임 연구원

김경진 : 영남대학교 교육대학원 석사 과정

정호열 : 영남대학교 정보통신공학과 교수

전송 대역폭(transmission bandwidth)을 줄이기 위해 JPEG(Joint Photographic Experts Group)[12], JPEG-2000[13], MPEG(Moving Picture Experts Group) H.26x 등으로 압축된다. 특히 최근에 JPEG 위원회는 "JPSearch"에서 정지 영상 검색을 위한 새로운 연구를 활발히 진행하고 있다[14]. 기존에는 이러한 압축된 영상들을 검색하기 방법으로 공간 영역 기반 검색 방법이 많이 제시되었으나[1] 최근 들어 데이터베이스 양이 기하급수적으로 증가함에 따라 채널 대역폭 요구량 및 검색 시간을 고속화하기 위한 방법으로 주파수 영역 또는 부분 복호화를 수행 후 특징을 추출하는 압축 영역 기반 검색 알고리즘이 많이 등장하고 있다[2-11]. 특히 "JPSearch" 그룹에서도 압축 영역에서 특징 벡터를 구성하여 검색을 수행하는 기술 개발이 중요한 관점으로 대두되고 있다[14]. 또한 압축 영상들은 인터넷 및 다양한 선로를 통해 많은 사용자들에게 전송되고 사용자 및 통신 네트워크 잡음에 의해 왜곡 및 변형이 수행된다. 따라서 이렇게 훼손된 압축 영상들을 효율적으로 검색하는 알고리즘 개발의 필요성이 대두되고 있다[3, 10, 11, 14]. 따라서 본 논문에서는 다양한 신호 처리 왜곡 및 기하학적으로 변형된 영상을 효율적으로 검색하는 알고리즘을 제안한다. 이를 위해 영상 처리 응용 분야에서 많이 이용되고 있는 웨이블릿 변환(wavelet transform)을 사용한다. 웨이블릿 변환은 공간 정보뿐만 아니라 다양한 주파수 특성을 나타내기 때문에 영상 검색 서술자를 기술하기 위해 많이 이용되고 있다[4-9]. 또한 웨이블릿 변환은 차세대 정지 영상 압축 표준인 JPEG-2000에서 기본 주파수 변환 기술로 채택되었기 때문에 본 논문에서 제안하는 검색 알고리즘의 활용성은 매우 방대하리라고 판단된다[13, 14]. 기존에 효율적으로 검색 서술자를 기술하기 위해 다양한 주파수 영역 기반 검색 알고리즘들이 많이 등장하였다. Feng 등은 효율적인 영상 검색을 위해 DCT(Discrete Cosine Transform) 계수의 평균(mean) 및 분산(variance)과 같은 통계적인 특성을 이용하였고[2], Cheng 등은 JPEG 복호화를 부분적으로 수행한 후 엔트로피 영역에서 DC(Direct Current) 계수의 상관관계와 AC(Alternate Current) 계수의 통계적 특성 분포를 특징 벡터로써 사용하여 압축 영상의 검색 시간을 많이 절약할 수 있었다[3]. 그러나 일반적인 신호처리 왜곡 영상 검색에는 검색 성능이 비교적 우수하지만 위치 변화에 민감하여 회전, 크기 변화, 천이 등과 같은 기하학적으로 변형된 영

상 데이터베이스 검색에는 성능이 떨어진다[3]. Xiong 등은 [4, 5]에서 웨이블릿 변환에서 모든 부대역(sub-band)의 통계적인 특성을 이용하여 메모리 관리 및 검색 시간 효율성에서 공간 영역 검색 방법보다 우수함을 증명하였다. JPEG-2000의 각 부대역 및 코드 블록(code-block)의 웨이블릿 계수의 분산을 특징 벡터로써 이용하여 질감 영상 검색을 수행하였으나 사람들의 지각적인 관점에서 패턴이 다른 영상이 검색되는 등의 단점을 나타낸다[4]. 또한 Arivazhagan 등은 모든 웨이블릿 부대역에서 다양한 통계적인 특성을 분석하여 질감 패턴을 효율적으로 정의하였다[6]. 그러나 [4-6]에서는 모든 웨이블릿의 부대역을 이용하기 때문에 특징 벡터를 구성하기 위해 요구되는 메모리 및 검색 시간에서 효율성은 떨어진다. 그리고 [8, 9]에서는 웨이블릿 계수의 중요 상태(significant state) 정보를 이용하였는데 일반적으로 중요 상태 정보는 영상의 주요 에너지 분포도를 나타낼 수 있기 때문에 이를 이용하여 질감 및 형상 특징 벡터를 효율적으로 서술할 수 있다. 또한 웨이블릿 기반 검색 방법에서 계수를 이용하는 방법과는 달리 비트 평면(bit-plane) 단위로 특징 추출이 가능하다는 장점이 있다. 그러나 많은 특징 벡터의 양을 요구함으로써 인해 모든 해상도의 각 비트에서 중요 상태 개수의 히스토그램 분포도를 추출해야하기 때문에 웨이블릿 전체 영역으로 복호화를 수행해야만 한다. 따라서 특징 벡터를 구성하기 위한 시간이 많이 걸리고 또한 국부적인 특징을 묘사하기 위해 저주파수 대역 크기를 일치해야하는 단점이 있다[8].

따라서 본 논문에서는 웨이블릿 변환의 모든 부대역을 이용하지 않고 오직 저주파수 부대역만을 이용하여 특징 벡터를 구성하는 영상 검색 알고리즘을 제안한다. 따라서 특징 벡터 구성을 위한 시간 및 메모리 요구량에서 많은 이득을 얻을 수 있다. 그리고 저주파수 부대역에서 통계적인 특성을 분석하여 2차에 걸친 검색 시스템을 구현한다. 1차 검색을 통해 질의 영상과 전반적으로 유사한 패턴을 나타내는 영상을 획득하고 재구성된 영상 데이터베이스에서 히스토그램 특성을 이용한 2차 검색을 수행하여 보다 세밀한 질감 패턴을 분석한다. 이후 최종 질의 영상과 유사한 영상으로 획득한다. 질의 영상과 데이터베이스 영상의 유사도 평가는 특징벡터 간 유클리디언 거리(Euclidean distance) 측정 방법을 이용한다.

본 논문의 구성은 다음과 같다. 2장에서는 웨이블릿 변환 특성에 대해 간략히 살펴보고 3장에서는

본 논문에서 제안하는 웨이블릿 기반 검색 알고리즘을 분석한다. 웨이블릿 저주파수 부대역(LL)에서 통계적인 패턴을 분석하는 2차 검색 과정을 제안한다. 4장에서는 다양한 왜곡 영상 데이터베이스를 이용하여 제안하는 검색 알고리즘의 성능을 기존의 웨이블릿 기반 검색 알고리즘과 비교 및 분석하고 마지막으로 본 논문의 결론을 맺는다.

II. 웨이블릿 변환

다양한 내용 기반 영상 검색 알고리즘들이 DCT 및 웨이블릿 변환 등과 같은 주파수 변환 기술의 특성을 분석하여 특징 벡터를 구성하였다[2-9]. 검색 성능뿐만 아니라 JPEG 및 JPEG-2000 등과 같은 멀티미디어 데이터 압축 영상들을 검색하기 위해 요구되는 검색 시간에서 많은 이득을 얻을 수 있기 때문이다. 특히 웨이블릿 변환은 영상 신호의 공간 및 주파수 해석이 가능하기 때문에 강인한 검색 성능을 나타내고 있다[4-9]. 웨이블릿은 음성 및 영상 등과 같은 신호 해석에 사용되는 기저 함수(basis function)의 집합(set)이다. 일반적으로 푸리에 해석(fourier analysis)은 진동(vibration) 또는 광학(optical)신호와 같은 주기적인(periodic) 특성을 나타내면서 무한한 에너지(infinite energy)를 가진 신호 해석에 탁월한 능력을 나타낸다. 그러나 비주기적인(non-periodic) 특성을 보이면서 유한한 에너지(finite energy or compact support)를 가진 시간(공간) 및 주파수 신호의 국부적인 해석(time-frequency localization analysis)에는 한계를 나타낸다. 이러한 단점을 극복하기 위한 대안이 웨이블릿이다. 웨이블릿은 지질 해석(geology analysis) FBI fingerprinting 데이터베이스 등과 같은 영상 압축 및 잡음 제거(denoising) 등과 같이 다양한 응용 분야에 적용되고 있다. 웨이블릿의 기저함수 중의 하나인 mother wavelet $\psi(x)$ 은 주기적으로 진동(oscillation)하며 시간(공간)적으로 유한한 특성을 나타낸다. 그림 1은 원본 영상 'building'에 차세대 정지 영상 압축 표준인 JPEG-2000에서 무손실 압축에 사용되는 5/3 필터[13]를 사용하여 3-레벨 웨이블릿 변환을 수행한 그림을 나타낸다.

그림 1에서 나타나듯이 웨이블릿 변환은 원본 영상의 다양한 주파수 성분 및 공간 성분을 효율적으로 표현할 수 있다. 웨이블릿 변환의 이러한 특성을 이용하여 EZW(Embedded Zerotree Wavelet) SPIHT(Set Partitioning In

Hierarchical Trees) 등과 같은 다양한 웨이블릿 기반 영상 압축 기술에 많이 이용되고 있다[15, 16]. 특히 웨이블릿 변환은 차세대 정지 영상 압축 표준인 JPEG-2000에서 입력 영상 신호의 기본 주파수 변환 기술로 채택되었다[13, 17].

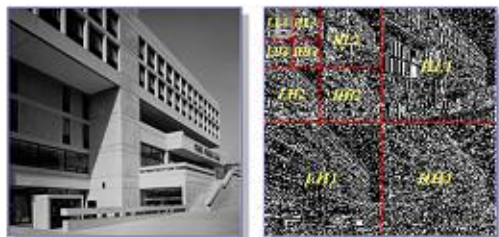


그림 1 원본 영상과 3-레벨 웨이블릿 변환을 수행한 영상(좌: 원본 영상 building 우: 3-레벨 웨이블릿 분해된 영상)

Fig 1. Original and wavelet decomposition image(left: original image, right: 3-level decomposition image).

JPEG-2000의 핵심 압축 기술인 Part 1 국제 표준안을 살펴보면 무손실(lossless) 압축을 위해 Le-Gall 5/3 필터가 사용되고 손실(lossy) 압축을 위해 Daubechies 9/7 필터가 기본 필터로 적용된다[13]. 앞의 숫자 5, 9는 저주파수 통과 필터 개수를 나타내고 뒤의 숫자 3, 7은 고주파수 통과 필터 개수를 의미한다.

III. 웨이블릿 기반 영상 검색

본 논문에서는 웨이블릿 변환 영역에서 특징 벡터를 구성하는 검색 알고리즘을 제안한다. 모든 웨이블릿 주파수 영역을 이용하지 않고 k -레벨로 분해된 다양한 부대역에서 오직 저주파수 부대역(LL_k)에서만 특징 벡터를 구성한다. 이를 위해 저주파수 부대역에서 평균, 분산 및 히스토그램 등과 같은 다양한 통계적인 특성을 분석한다. 이는 검색 시스템에서 특징 벡터 구성을 위해 요구되는 메모리 및 검색 시간에서 많은 이득을 얻을 수 있다. 본 논문에서 제안하는 웨이블릿 기반 영상 데이터베이스 검색 시스템은 그림 2에 나타난다. 그림 2에 나타나듯이 효율적인 검색을 수행하기 위해 2차에 걸친 검색 과정을 수행한다. 1차 검색 과정을 통해 질의 영상과 전반적으로 유사한 패턴을 나타내는 영상을 획득하고 재구성된 데이터베이스에서 히스토그램 특성을 이용한 2차 검색 과정을 수행하

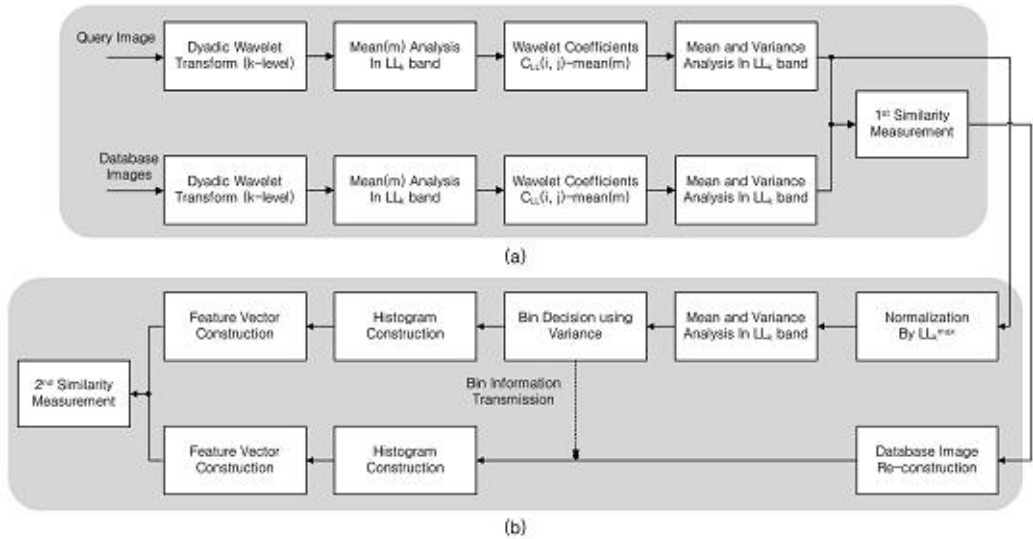


그림 2. 웨이블릿 기반 영상 데이터베이스 검색 시스템((a) 1차 영상 데이터베이스 검색, (b) 2차 영상 데이터베이스 검색)

Fig 2. Wavelet Transform based image database retrieval system((a) first image database retrieval (b) second image database retrieval).

여 보다 세밀한 질감 패턴을 분석한다. 먼저 질의 영상 및 데이터베이스 영상은 다해상도 웨이블릿 변환을 적용하여 그림 1에 나타나듯이 다양한 주파수 영역($LL_k, LH_k, HL_k, \dots, LL_k, LH_k, HL_k, HH_k, k=1, 2, 3, \dots$)으로 분해한다. 이 후 저주파수 부대역(LL_k)에서 1차 및 2차 검색을 위한 특징 벡터를 구성한 후 유사도 평가를 수행하여 데이터베이스 영상의 순위를 결정한다. 순위(Rank) R 이 내의 영상을 최종 질의 영상과 유사한 영상으로 획득한다. 1차 및 2차 검색 과정에서의 특징 벡터 구성을 위해 식 (1) 및 (2)와 같이 저주파수 부대역에서 평균 m 및 분산 σ^2 과 같은 통계적인 특성을 분석한다.

$$m = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} C_{LL_k}(i, j) \quad (1)$$

$$\sigma^2 = \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (C_{LL_k}(i, j) - m)^2 \quad (2)$$

여기서, $C_{LL_k}(i, j)$ 는 $N \times M$ 크기를 가지는 저주파수 부대역의 웨이블릿 계수를 나타낸다. 일반적으로 평균 및 분산 등과 같은 통계적인 분석은 기하

학적인 변형이 적용된 영상 검색에서 좋은 성능을 나타낼 수 있다[2].

1. 특징 추출

본 논문에서 제안하는 웨이블릿 기반 영상 검색 알고리즘의 특징 벡터 구성 과정은 다음과 같다. 본 논문에서는 모든 웨이블릿 부대역 중에서 오직 저주파수 부대역(LL_k)을 이용하여 통계적 특성을 분석한다. 저주파수 부대역에는 원 영상의 대부분의 에너지가 포함되어 있어 원 영상이 나타내고 있는 공간 및 주파수 성분의 대부분을 표현할 수 있다. 그림 2의 (a)에서 나타나는 1차 검색 과정을 살펴보면 먼저 질의 영상 및 데이터베이스 영상은 k -레벨 웨이블릿 변환이 수행된다. 이 후 저주파수 부대역(LL_k) 계수를 이용하여 저주파수 부대역의 평균값 m_{LL_k} 을 구한다. 다음 과정은 모든 저주파수 부대역 계수 $C_{LL_k}(i, j)$ 와 저주파수 부대역의 평균값 m_{LL_k} 과의 오차 값 $D_{LL_k}(i, j)$ 을 구한다. 이는 식 (3)을 이용하여 얻을 수 있다.

$$D_{LL_k}(i, j) = C_{LL_k}(i, j) - m_{LL_k} \quad (3)$$

식 (3)의 오차 값을 이용함으로써 밝기 조절

(brightening darkening) 등과 같은 일정한 크기에 의해 변화가 적용된 왜곡 영상을 효율적으로 검색할 수 있다. 자세한 내용은 부록 A를 참조 바란다. 이 후 수정된 저주파수 웨이블릿 계수 $D_{LL_k}(i, j)$ 를 이용하여 LL_k 부대역의 평균값과 분산 값을 다시 구한다. 이 후 데이터베이스 영상에서 질의 영상의 평균 및 분산 값의 오차 범위 $\pm 15\%$ 내에 공통적으로 포함되는 영상들을 1차 유사 영상으로 획득한다. 재구성된 데이터베이스 영상들은 2차 유사 영상 검색 과정을 수행하게 된다. 수정된 저주파수 부대역의 분산 값을 이용하여 질의 영상의 히스토그램 빈을 설정한다. 이를 위해 먼저 수정된 저주파수 부대역에서 웨이블릿 계수 $D_{LL_k}(i, j)$ 들을 동일한 조건의 히스토그램 빈 설정을 위해 계수들의 최대값을 구한다. 이 후 최대값을 이용하여 수정된 저주파수 부대역 계수를 -1 과 1 사이의 값으로 정규화를 수행한다. 자세한 과정은 식 (4)와 같다.

$$N_{LL_k}(i, j) = \frac{D_{LL_k}(i, j)}{\max}, \quad -1 \leq N_{LL_k}(i, j) \leq 1 \quad (4)$$

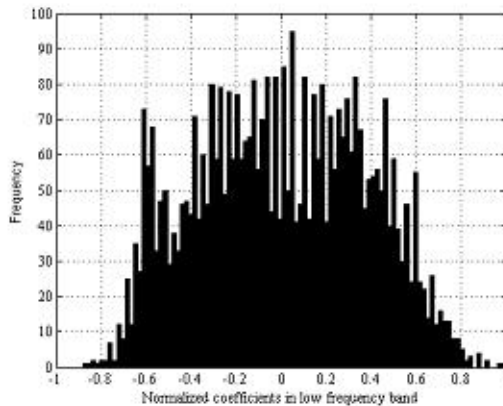


그림 3 정규화가 적용된 저주파수 부대역 계수의 히스토그램 (실험 영상: 'Flower02'(512x512))
수정된 저주파수 부대역 계수의 평균 $m: -3.6 \times 10^{-5}$,
분산 $\sigma^2: 0.14$.

Fig 3 The lowest sub-band histogram by normalization(test image: 'Flower02'(512x512))
mean $m: -3.6 \times 10^{-5}$, variance $\sigma^2: 0.14$.

이 후 정규화가 수행된 웨이블릿 저주파수 부대역 계수 $N_{LL_k}(i, j)$ 의 평균 m 및 분산 σ^2 을 구한다. 일반적으로 대부분의 웨이블릿 저주파수 부대역 계

수는 평균값을 중심으로 웨이블릿 계수들이 집중적으로 모여 있기 때문에 이러한 특성을 분석하여 히스토그램 빈 간격을 가변적으로 조정한다. 그림 3은 'Flower02' 질감 영상에서 정규화가 수행된 저주파수 부대역 계수의 히스토그램을 나타낸 것이다. 그림 3은 5/3 필터를 사용하여 3-레벨 웨이블릿 변환을 수행하였다.

그림 3에 나타나듯이 입력 영상의 저주파수 신호들은 그 부대역의 평균 값 근처로 집중되고 있음을 알 수 있다. 따라서 질의 영상에서 효율적으로 히스토그램 빈을 설정하기 위해 평균값을 기준으로 일정한 분산 범위 이내에는 히스토그램 빈을 촘촘한 간격으로 설정하여 세밀한 분석을 하고, 일정한 분산 값을 초과하는 경우에는 빈 간격을 넓게 하여 2차 검색을 위한 특징 벡터를 구성한다. 히스토그램 빈 간격을 다양하게 조절하여 실험을 수행한 결과 평균값을 기준으로 $\pm 2\sigma^2$ 범위에서 100개의 히스토그램 빈을 설정하여 검색을 수행한 실험에서 가장 높은 성능을 얻을 수 있었다. 자세한 히스토그램 빈 간격(Bin Interval) 설정은 식 (5)에 나타난다.

$$\text{Bin Interval} = \begin{cases} 2\sigma^2 / 100, & \text{if } C(i, j) \in m \pm 2\sigma^2, \\ \sigma^2, & \text{otherwise.} \end{cases} \quad (5)$$

질의 영상의 히스토그램 빈 정보는 데이터베이스 영상에서 히스토그램 특징 벡터를 구성하기 위해 이용된다. 식 (5)와 같이 히스토그램 빈을 설정함으로써 질의 영상과 데이터베이스 영상과의 차별화된 특징 벡터를 구성할 수 있다. 2차 유사 영상 획득 과정은 그림 2의 (b)와 같다. 1차 검색을 통해 질의 영상과 전반적으로 유사한 패턴을 나타내는 영상을 획득할 수 있으며 분산 값을 이용한 히스토그램 빈 설정을 구성하는 2차 검색은 1차 검색에서 획득된 영상 데이터베이스에서 보다 세밀한 검색을 수행할 수 있다. 또한 1차 검색 과정을 통해 대부분의 영상 데이터베이스를 걸러냄으로써 2차 영상 검색을 수행하기 위한 시간을 많이 단축할 수 있다. 앞으로 본 논문에서 제안하는 저주파수 부대역을 이용한 검색 알고리즘을 LRF(Lowest Resolution Feature)로써 정의한다.

2. 유사도 측정 및 순위 결정

질의 영상으로부터 2차 검색 과정을 위한 특징 벡터를 구성하여 데이터베이스 영상들과 유사도 평

가를 수행한다. 질의 영상 q 와 데이터베이스 영상 d 의 유사도 거리 $D(q, d)$ 를 측정하기 위해 식 (6)과 같이 특징벡터 간 유클리디언 거리(Euclidean distance) 측정 방법을 이용한다.

$$D(q, d) = \sum_{k=0}^{M-1} |E_k(q) - E_k(d)| \quad (6)$$

질의 영상에서 분산 값을 이용하여 구성된 히스토그램 빈 x 의 개수 M 개를 기반으로 질의 영상 q 및 데이터베이스 영상 d 의 각 히스토그램 빈의 확률 분포 $E_k(q)$, $E_k(d)$ 를 구성한다. 각 확률 정보의 오차 값의 절대치를 각 영상 데이터베이스의 색인 정보 값으로 저장한다. 이 후 그림 4와 같이 가장 유사한 순서대로 데이터베이스 영상을 정렬하여 최종 유사 영상으로 획득한다.

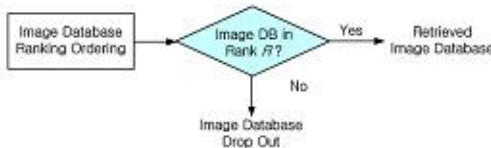


그림 4. 순위(Rank)를 이용한 영상 획득.
Fig. 4. Image acquisition by rank.

여기서, 순위(Rank) R 은 사용자에 의해 정의된 출력 영상 개수를 나타낸다. 순위 R 이내의 영상 데이터베이스는 질의 영상과 유사 영상으로 판단하여 출력한다. 그리고 본 논문에서 제안하는 알고리즘 및 기존 알고리즘의 검색 성능 평가를 위해 본 논문에서는 식 (7)과 같이 검색률 RR (Retrieval Rate)를 이용한다. 검색률 RR 은 사전에 질의 영상과 유사하다고 판단된 영상들이 순위 R 내에 획득되는 검색율을 의미한다.

$$RR(\%) = \frac{R_R}{T_R} \times 100(\%) \quad (7)$$

여기서, T_R 은 사전에 질의 영상과 유사하다고 정의된 영상 데이터베이스 개수를 나타내고 R_R 은 T_R 에서 획득된 영상의 개수를 의미한다.

IV. 실험 결과

본 절에서는 본 논문에서 제안된 웨이블릿 기반 영상 검색 시스템의 성능을 분석한다. 본 논문의

웨이블릿 기반 영상 검색 시스템은 Visual C++ 6.0으로 구현되었다. 펜티엄(pentium) 4 1.83GHz CPU, 2GB RAM 컴퓨터 환경에서 실험하였으며 MIT VisTex(MIT Vision Texture) 질감 영상 및 PICS(Psychological Image Collection at Stirling) 자연 영상 데이터베이스 검색 실험을 이용한다[18, 19]. 영상 데이터베이스에 다양한 신호 처리 왜곡 및 기하학적인 변형 방법 등 총 11가지 왜곡 사항을 적용하여 질감 영상 2880개, 자연 영상 1704개의 데이터베이스를 구축하였다. 영상의 왜곡 사항으로 밝기 조절(Brightening Darkening), JPEG (Q factor 30) 및 JPEG-2000 (Q factor 40) 압축, 회전(90°, 180°, 270°), 샤프닝(Sharpening), 스무딩(Smoothing), 천이(Translation, 우측방향 20%) 균일 잡음(Uniform noise) 20% 등을 적용하였다. 그림 5는 질감 영상 'Leaves00'에 다양한 신호 처리 및 기하학적 변형을 적용한 그림을 나타낸다.

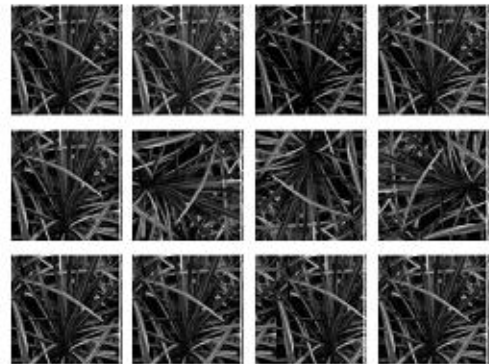


그림 5. 다양한 왜곡을 적용한 MIT VisTex(Vision Texture) 질감 영상 'Leaves00' (왼쪽 위부터 순서대로 원본 영상, 밝기 조절(brightening darkening), JPEG 압축(Q factor 30), JPEG-2000 압축(Q factor 40), 시계 방향 90° 회전, 180° 회전, 270° 회전, 샤프닝, 스무딩, 천이 20%, 균일 잡음 20%).

Fig. 5. MIT VisTex(Vision Texture) texture image 'Leaves00' with various distortions (from left top, brightening, darkening, JPEG compression(Q factor 30), JPEG-2000 compression(Q factor 40), clockwise 90, 180, 270 degree rotation, sharpening, smoothing, translation 20%, uniform noise 20%).

본 논문에서 제안하는 LRF 검색 알고리즘의 성

표 1. 질감 영상 및 자연 영상을 이용한 왜곡 영상 검색 실험 결과(LRF: 1차 검색($\pm 15\%$), RR(Retrieval Rate(%)))Table 1. Simulation results using texture and natural image database(LRF: first retrieval($\pm 15\%$), RR(Retrieval Rate(%)))

질의 영상	질감 영상			질의 영상	자연 영상		
	FSF(%)	SSF(%)	LRF(%)		FSF(%)	SSF(%)	LRF(%)
Bark05	83.33%	50.00%	100%	b2	83.33%	75.00%	100%
Bark05	75.00%	66.67%	83.33%	b2	75.00%	66.67%	83.33%
Brick05	83.33%	66.67%	83.33%	b6	41.67%	33.33%	66.67%
Brick07	75.00%	66.67%	83.33%	image1ifen	83.33%	83.33%	100%
BrickPaint01	50.00%	50.00%	83.33%	image10.i.indoor	41.67%	41.67%	83.33%
Building05	66.67%	41.67%	75.00%	image13.i.indoor	50.00%	41.67%	91.67%
Building04	83.33%	66.67%	100%	image15ifen	83.33%	75.00%	100%
Building09	91.67%	66.67%	91.67%	image6.i.indoor	75.00%	66.67%	91.67%
Clouds00	75.00%	41.67%	75.00%	image7.i.indoor	41.67%	33.33%	75.00%
clouds01	83.33%	41.67%	100%	image8.i.indoor	83.33%	66.67%	83.33%
Pabric05	50.00%	50.00%	83.33%	image9ifen	75.00%	66.67%	100%
Pabric09	75.00%	66.67%	83.33%	image9.i.indoor	91.67%	83.33%	91.67%
Pabric14	66.67%	66.67%	91.67%	image31	50.00%	50.00%	75.00%
Flower02	66.67%	41.67%	75.00%	image310	91.67%	75.00%	100%
Flowers05	50.00%	66.67%	100%	image314	50.00%	33.33%	83.33%
Food01	66.67%	41.67%	100%	image39	41.67%	41.67%	83.33%
Food04	66.67%	50.00%	75.00%	m061	91.67%	75.00%	91.67%
Grass00	83.33%	66.67%	100%	m5	75.00%	75.00%	75.00%
GrassLand04	91.67%	50.00%	75.00%	stereo10	66.67%	66.67%	100%
GrassPlan(Sky06	75.00%	41.67%	100%	stereo14	91.67%	91.67%	100%
GroundWaterCity00	91.67%	91.67%	100%	stereo2	83.33%	75.00%	100%
Leaves04	66.67%	41.67%	83.33%	stereo4	75.00%	66.67%	83.33%
Misc02	66.67%	66.67%	100%	tw01	83.33%	66.67%	83.33%
Paintings04	66.67%	50.00%	91.67%	tw05	83.33%	75.00%	75.00%
PrisonWindow01	75.00%	50.00%	100%	tw05	83.33%	75.00%	83.33%
PrisonWindow06	83.33%	75.00%	100%	tw09	41.67%	41.67%	75.00%
PrisonWindow09	91.67%	50.00%	91.67%	w12	66.67%	41.67%	75.00%
Sand05	66.67%	66.67%	100%	w2	91.67%	83.33%	83.33%
Stone01	66.67%	66.67%	75.00%	w4	91.67%	91.67%	91.67%
Stone02	83.33%	50.00%	100%	w7	66.67%	66.67%	83.33%
Tile02	75.00%	50.00%	100%	w8	83.33%	66.67%	91.67%
Tile05	83.33%	66.67%	91.67%	wb1	50.00%	33.33%	75.00%
Tile04	50.00%	50.00%	83.33%	wb2	50.00%	50.00%	75.00%
Tile07	75.00%	50.00%	91.67%	wb6	66.67%	41.67%	83.33%
ValleyWater00	83.33%	66.67%	75.00%	wb9	66.67%	50.00%	75.00%
ValleyWater01	75.00%	50.00%	83.33%	wb11	83.33%	66.67%	91.66%
Water00	66.67%	50.00%	91.67%	wb15	41.67%	33.33%	75.00%
Water04	75.00%	50.00%	91.67%	w11	83.33%	66.67%	75.00%
Water07	75.00%	50.00%	100%	w10	66.67%	50.00%	83.33%
Wood02	83.33%	66.67%	91.67%	w15	66.67%	50.00%	83.33%
Average	73.96%	56.46%	90.00%	Average	70.21%	60.83%	85.42%

능 비교를 위해 I 장에서 언급한 두 가지 웨이블릿 기반 검색 기술과 검색 결과를 비교 및 분석한다 [4-8], Xiong 등이 제안한 모든 부대역의 분산값을 이용하는 방법을 FSF(Full Sub-bands Feature)로 정의하고 [4], 모든 부대역의 비트 평면에서 에너지를 특징 벡터로 구성하여 검색 실험을 수행한 Mandal 등이 제안한 알고리즘을 SSF(Significant State Feature)로 정의한다 [8], 첫 번째 실험을 위해 모든 영상 데이터베이스는 JPEG-2000에서 무손실 압축을 위해 기본 필터로 사용되는 5/3 웨이블릿 필터를 사용하여 3-레벨로 분해하였다. 표 1은 질감 및 자연 영상 데이터베이스에서 임의로 선택한 40개의 원본 영상을 질의 영상으로 사용하여 검색 성능을 평가한 결과를 나타낸다. 표 1에 나타나듯이 본 논문에서 제안하는 방법 LRF 알고리즘이 기존의 다른 웨이블릿 기반 검색 알고리즘 FSF, SSF 보다 좋은 성능을 나타낸다. 본 논문에서 제안하는 LRF 검색 알고리즘이 2차에 걸친 검색 과정을 수행하기 때문에 단순한 특징 벡터만을 가지고 검색을 수행하는 기존 알고리즘 보다 효율적인 검색 결과를 얻을 수 있었다. 또한 LRF 알고리즘은 오직 저주파수 부대역만을 이용하기 때문에 특징 벡터를 구성하기 위해 요구되는 메모리 및 검색 시간에서도 많은 이득을 얻을 수 있다. 특징 벡터를 구성하기 위해 요구되는 메모리 요구량 M_R 은 웨이블릿 분해 레벨 k 에 따라 식 (8)과 같이 기하급수적으로 이득을 얻을 수 있다.

$$M_R = \frac{1}{4^k} \cdot M_T \quad (k=1, 2, 3, \dots) \quad (8)$$

여기서, M_T 는 웨이블릿 모든 부대역을 저장하기 위해 필요한 메모리 요구량이다. 다음 실험은 왜곡 영상 40개를 질의 영상으로 사용하여 제안하는 검색 알고리즘의 유연성을 확인한다. 표 2 및 표 3은 질의 영상으로 Darkening, 스무딩, 그리고 JPEG 압축(Q factor 30)이 적용된 영상을 사용하였을 경우 질감 및 자연 왜곡 영상 검색 결과를 나타낸다. 표 2에 나타나듯이 왜곡 영상을 질의 하였을 경우에도 기존 검색 알고리즘 보다 강인한 검색 성능을 얻을 수 있어 본 논문에서 제안하는 검색 알고리즘의 유연성을 확인할 수 있었다. 그림 6은 darkening이 적용된 질감 영상을 질의한 경우 검색이 수행된 결과를 나타낸다.

V. 결 론

본 논문에서는 웨이블릿 변환을 이용하여 효율적으로 특징 벡터를 구성할 수 있는 새로운 영상 검색 알고리즘을 제안하였다. 특징 벡터를 구성을 위해 다양한 웨이블릿 주파수 부대역에서 오직 저주파수 부대역만을 이용하였다. 웨이블릿 부대역 계수와 그 부대역 계수들의 평균값과의 오차 값을 이용하여 밝기 조절등과 같은 크기 변화가 적용된 왜곡 영상 데이터베이스를 획득할 수 있었다. 또한 2차에 걸친 검색 과정을 수행하여 전반적이면서도 세밀한 분석 과정을 통해 효율적으로 영상 검색을 수행할 수 있었다. 먼저 저주파수 부대역에서 평균과 분산 등과 같은 통계적인 특성을 분석하여 1차 검색을 수행하였다. 1차 검색을 통하여 획득된 영상 데이터베이스는 전반적으로 유사한 패턴을 나타내는 영상들로 구성될 수 있다. 1차 검색 결과를 대상으로 분산 정보를 이용한 히스토그램 분석은 영상에서 보다 세밀한 특징 벡터를 구성할 수 있기 때문에 최종 강인한 검색 결과를 얻을 수 있었다.

표 2. 질감 왜곡 영상 검색 결과 (RR: Retrieval Rate(%), 질의 영상: 왜곡 영상),
Table 2. Simulation results using texture image database (RR: Retrieval Rate(%), query image: distorted image)

질의 영상 (RR)	질감 영상		
	FSF(%)	SSF(%)	LRF(%)
Darkening	58,75%	45,88%	85,83%
Smoothing	70,13%	40,38%	87,29%
JPEG	64,62%	52,75%	88,54%
Average	64,50%	46,34%	87,22%

표 3. 자연 왜곡 영상 검색 결과 (RR: Retrieval Rate(%), 질의 영상: 왜곡 영상),
Table 3. Simulation results using natural image database (RR: Retrieval Rate(%), query image: distorted image)

질의 영상 (RR)	자연 영상		
	FSF(%)	SSF(%)	LRF(%)
Darkening	60,06%	42,66%	80,00%
Smoothing	69,83%	40,91%	84,17%
JPEG	61,44%	47,06%	77,08%
Average	63,78%	43,54%	80,42%

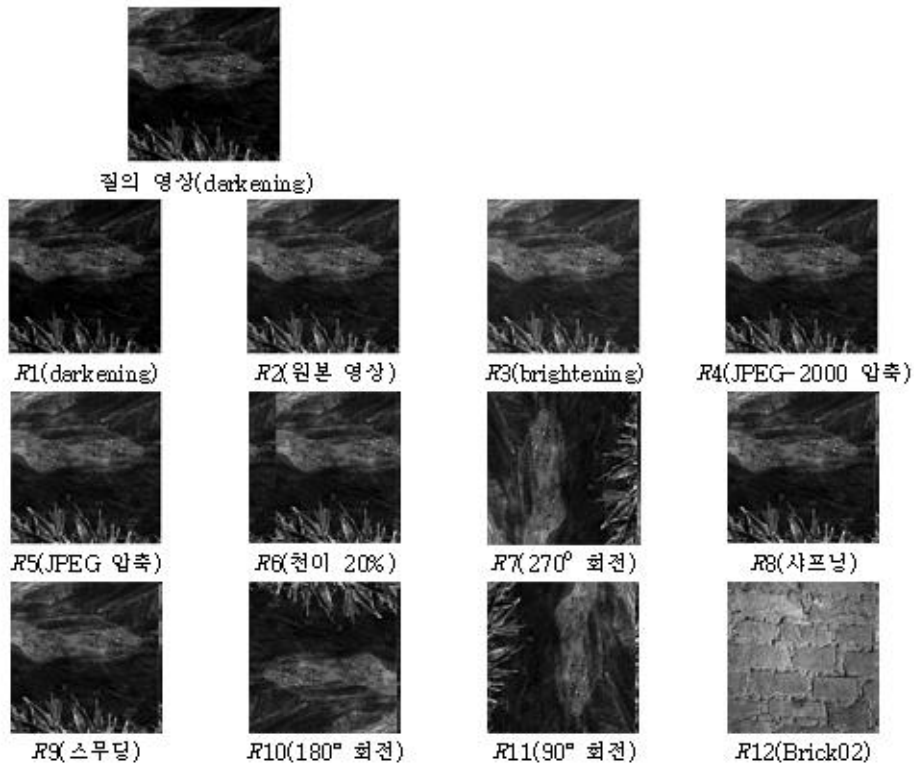


그림 6 왜곡 영상 검색 결과(질의 영상: Darkening이 적용된 'ValleyWater01', 1차 검색: $\pm 15\%$ $R=12$, $RR=91.67\%$).

또한 검색 영역으로 오직 웨이블릿 저주파수 부대역만을 이용하기 때문에 특징 벡터를 구성하기 위해 요구되는 메모리 및 검색 시간에서 많은 이득을 얻을 수 있었다. 따라서 본 논문에서 제안하는 검색 알고리즘은 다양한 웨이블릿 기반 영상 압축, 특히 차세대 정지 영상 압축 표준인 JPEG-2000에서 효율적인 검색 성능을 얻을 수 있을 것으로 판단된다. 본 논문에서는 효율적인 영상 검색을 위한 알고리즘 개발 및 실험 결과에 초점을 집중하였다. 따라서 향후 결과 증명을 위한 이론적인 분석이 요구되고 다양한 웨이블릿 필터를 이용하여 영상 검색 성능을 비교 및 분석하는 것이 필요할 것이다.

참고 문헌

[1] Arnold W. M, Smeulders, Marcel Worring, Simone Santini, Arnamath Gupta and Ramesh Jain, "Content-based image retrieval at the end of the early years", *IEEE Trans on*

Pattern Analysis and Machine Intelligence, Volume 22, Issue 12, Pages 1349-1380, Dec. 2000.

- [2] Guocan Feng and Jianmin Jiang, "JPEG compressed image retrieval via statistical features", *Pattern Recognition*, Volume 36, Issue 4, Pages 977-985, Apr. 2003.
- [3] Chin-Chen Chang, Jun-Chou Chuang and Yih-Shin Hu, "Retrieving digital images from a JPEG compressed image database", *Image and Vision Computing*, Volume 22, Issue 6, Pages 471-484, Jun. 2004.
- [4] Xiong Z. and Huang T. S., "Wavelet-based texture features can be extracted efficiently from compressed-domain for JPEG-2000 coded images", *IBBB International Conference on Image Processing (ICIP)*, Volume 1, Sep. 2002.

[5] Xiong, Z. and Huang, T. S., "Block-based, memory-efficient JPEG2000 images indexing in compressed-domain", *IBBB Southwest Symposium on Image Analysis and Interpretation(SSIAI)*, Apr. 2002.

[6] S. Arivashagan and L. Ganesan, "Texture classification using wavelet transform", *Pattern Recognition Letters*, Volume 24, Issues 9-10, Pages 1513-1521, Jun. 2003.

[7] Minh N. Do and Martin Vetterli, "Wavelet-based texture retrieval using generalized gaussian density and Kullback-Leibler distance", *IBBB Trans on Image Processing*, Volume 11, Pages 146-158, Feb. 2002.

[8] Mandal M. K. and Liu C., "Efficient image indexing techniques in the JPEG-2000 domain", *Journal of Electronic Imaging*, Pages 179-187, Jan. 2004.

[9] Jiang J., Guo B. and Li P., "Extracting shape features in JPEG-2000 images indexing in compressed domain", *ADVISED: Proceedings of the Second International Conference on Advances in Information Systems*, London, UK, Springer-Verlag Pages 123-132, Oct. 2002.

[10] Ha-Joong Park and Ho-Youl Jung, "JPEG-2000 Compressed Image Retrieval Using Partial Entropy Decoding", *International Workshop, MRCSLecture Notes in Computer Science*, 2006, Istanbul, Turkey, Pages 410-417, Sep. 2006.

[11] 박 하중, 정 호열, "문맥 정보를 이용한 JPEG 및 JPEG-2000 압축 영상 데이터베이스 통합 검색", 대한임베디드공학회 추계학술대회, 제주도, 2007년 11월.

[12] G. K. Wallace, "The JPEG still picture compression standard", *ACM Communication*, 34(4), 1991.

[13] Information technology, *JPEG-2000 image coding system, ISO/IBC International Standard 15444-1, ITU Recommendation T. 800*, Dec. 2000.

[14] JPSearch Call for Proposals(CFP), *ISO/IBC JTCl/SC29/WG1 N3792*, Nov. 2005.

[15] Shapiro, J. M., "Embedded image coding using zerotrees of wavelet coefficients", *IBBB*

Trans on Signal Processings, Volume 41, No. 12, Pages 3445-3462, Dec. 1993.

[16] A. Said and W. Pearlman, "A new, fast and efficient image codec based on set partitioning in hierarchical image coder using zerotrees of wavelet coefficients", *IBBB Trans Circuits Syst. Video Technol.*, Volume 6, Pages 243-250, Jun. 1996.

[17] D. Taubman, "High performance scalable image compression with ebcot", *IBBB Trans. on Image Processing*, Volume 9, Pages 1158 - 1170, Jul. 2000.

[18] Psychological Image Collection at Stirling, [online] Available: <http://pics.psych.stir.ac.uk>.

[19] VisTex(MIT Vision Texture), [Online]. Available: <http://vismod.www.media.mit.edu>.

부 록 A.

$N \times M$ 크기를 가지는 웨이블릿 부대역의 계수를 (α_i, j) 로 가정한다. 그리고 임의의 함수 $f(\alpha_i, j)$ 를 식 A-1과 같이 정의한다.

$$f(\alpha_i, j) = \alpha_i, j - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \alpha_i, j \quad A-1.$$

웨이블릿 부대역의 계수가 일정한 크기 α 만큼 증가 또는 감소될 경우 함수 $f(\alpha_i, j)$ 는 식 A-2를 만족한다.

$$f(\alpha_i, j) = f(\alpha_i, j) \pm \alpha \quad A-2.$$

식 A-2의 증명은 식 A-3과 같다.

$$\begin{aligned} f(\alpha_i, j) \pm \alpha &= \alpha_i, j \pm \alpha - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\alpha_i, j) \pm \alpha \\ &= \alpha_i, j \pm \alpha - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\alpha_i, j) - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\pm \alpha) \\ &= \alpha_i, j \pm \alpha - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\alpha_i, j) - \frac{(N \times M) \cdot (\pm \alpha)}{N \times M} \\ &= \alpha_i, j - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\alpha_i, j) \pm \alpha - (\pm \alpha) \\ &= \alpha_i, j - \frac{1}{N \times M} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} (\alpha_i, j) \\ &= f(\alpha_i, j) \end{aligned} \quad A-3.$$

따라서 웨이블릿 부대역 계수와 그 부대역 계수들의 평균값과의 오차 값은 일정한 크기 δ 만큼 변화한 부대역의 계수와 그 부대역 계수들의 평균값과의 오차 값과 같다.

저 자 소 개

박 하 중

2000년 영남대학교 전자공학과 학사, 2002년



영남대학교 정보통신공학과 석사, 2008년 영남대학교 정보통신공학과 박사. 현재, 삼성 테크윈 선임 연구원. 관심분야 멀티미디어 신호 처리 및 웨이블릿, JPEG/JPEG-2000, 웨이블릿 기반

영상 압축, 내용기반 영상검색, 영상 품질 측정.

Email: wavelet@yu.ac.kr

김 경 진

2005년 충북대 전기전자공학부 학사. 현재, 영



남대학교 교육대학원 전기전자통신교육학과 석사과정.

관심분야 멀티미디어 신호 처리 및 웨이블릿 내용 기반 영상 검색.

Email: susia1021@yu.ac.kr

정 호 열

1998년 프랑스 리옹 국립 응용과학원 전자공학과 박사. 현재, 영남대학교 공과대학 정보통신공학과 부교수.



관심분야 신호처리, 디지털 워터마킹, 손실/무손실압축, JPEG/JPEG-2000, 영상 품질 측정.

Email: hoyoul@yu.ac.kr