

웨이브릿 계수의 퍼지 동질성과 고주파 에너지를 이용한 영상 검색용 특징벡터 추출

Visual Feature Extraction for Image Retrieval using Wavelet Coefficient's Fuzzy Homogeneity and High Frequency Energy

박원배
충북대학교 정보통신공학과

Won-Bae Park (nte92@dreamwiz.com)
Dept. of Computer & Communication Eng., Chungbuk National University

류은주
충북대학교 정보통신공학과

Eun-Ju Ryu (ejthebest4ever@hotmail.com)
Dept. of Computer & Communication Eng., Chungbuk National University

송영준
충북대학교 정보통신공학과

Young-Jun Song (songyjorg@dreamwiz.com)
Dept. of Computer & Communication Eng., Chungbuk National University

중심어 : 영상 검색, 웨이브릿, 퍼지, 특징 추출

Keyword : image retrieval, wavelet, fuzzy, feature extraction

요약

Abstract

본 논문에서는 공간-주파수 특성과 다중 해상도 특성을 모두 갖는 웨이브릿 변환을 이용하여 각 대역의 특성에 맞는 비주얼 특징을 추출하고 이를 내용기반 영상 검색에 이용하는 새로운 방법을 제시하였다. 웨이브릿 변환된 영상의 최저 주파 대역은 원 영상의 근사한 형태로 공간 정보를 충분히 활용할 수 있다. 이를 위해 웨이브릿 계수값과 각 계수간의 공간 정보를 모두 고려한 퍼지 동질성(FH : Fuzzy Homogeneity)를 이용하여 L개의 특징 벡터를 추출하였고, 나머지 고주파 대역의 에너지 값을 이용하여 3개의 특징 벡터를 추출하여 이를 영상 데이터베이스에 저장한다. 질의 시에는 L개의 FH 벡터 중 가장 크기가 큰 10개의 값과 3개의 고주파 대역의 에너지 값을 이용하여 가장 유사한 영상을 검색하였다. 90개의 텍스처 영상을 사용해 실험한 결과 좋은 정확성을 보였다.

In this paper, we propose a new visual feature extraction method for content-based image retrieval(CBIR) based on wavelet transform which has both spatial-frequency characteristic and multi-resolution characteristic. We extract visual features for each frequency band in wavelet transformation and use them to CBIR. The lowest frequency band involves spacial information of original image. We extract L feature vectors using fuzzy homogeneity in the wavelet domain, which consider both the wavelet coefficients and the spacial information of each coefficient. Also, we extract 3 feature vectors using the energy values of high frequency bands, and store those to image database. As a query, we retrieve the most similar image from image database according to the 10 largest homograms(normalized fuzzy homogeneity vectors) and 3 energy values. Simulation results show that the proposed method has good accuracy in image retrieval using 90 texture images.

1. 서론

의 보급, 그리고 WWW의 출현 및 보급 등으로 인해 엄청난 양의 디지털 영상이 시시각각 쏟아지고 있다. 이러한 엄청난 양의 디지털 영상을 효율적으로 관리하기 위해 영상 데이터

멀티미디어 시대의 도래와 디지털 카메라와 디지털 스캐너

접수번호 : #040113-001

*교신저자 : 류은주, e-mail : ejthebest4ever@hotmail.com

접수일자 : 2004년 1월 13일, 심사완료일 : 2004년 2월 23일

베이스가 필요하다. 기존의 영상 데이터베이스 검색 방법은 문자기반으로 사람이 입력한 키워드를 이용하여 원하는 영상을 검색하는 방법이다. 그러나 이 방법은 소규모 데이터베이스에서는 효율적이지만 데이터베이스가 커지면 효율이 떨어지고, 대용량의 데이터에 대하여 사람이 일일이 키워드를 부여해야 하기 때문에 많은 인력이 필요하다는 단점이 있다. 그리고 동일한 영상인 경우에도 주관적인 판단에 의해 키워드가 다르게 부여될 수 있어 정보의 일관성과 객관성이 보장되지 않는다[1]. 이러한 문제점을 해결하기 위해 영상 데이터를 나타내는 직접적인 표현 요소들을 통하여 원하는 정보를 쉽게 검색 할 수 있는 내용기반 영상검색(CBIR : Content Based Image Retrieval)에 대한 연구가 활발히 진행 중이다. CBIR은 영상의 색과 모양, 질감 등 영상의 내용적인 특징들을 이용하여 완전히 일치하는 영상을 데이터베이스로부터 검색하는 것이 아니라, 어느 정도의 오차를 허용하는 유사일치 검색을 사용한다. 특징 추출은 크게 공간 영역에서의 화소값이나 히스토그램 등을 이용하는 방법[2]과 변환 영역에서 주파수 계수를 이용하는 방법으로 분류될 수 있다[3]. 공간 영역에서의 특징 추출 방법은 알고리즘의 복잡도가 높고 특징 벡터의 차원이 커서 특징 추출 및 검색에 많은 시간이 소요되는 문제점이 있다. 도메인 변환으로 얻을 수 있는 가장 큰 이점은 영상의 지역적 요소를 효과적으로 특징지을 수 있다는 것이다. 그리고 웨이브릿 변환은 영상의 용량을 줄이는데 사용되는데 이것은 인덱싱의 저장 효율을 증가시킨다. 기존의 웨이브릿을 이용한 방법으로 Mandal[3]은 웨이브릿 계수의 히스토그램으로 인덱싱하는 방법을 제안했고 또한 Jacobs[4]는 계산이 빠른 Harr 웨이브릿 변환을 사용하고 특징 벡터로서 다중 해상도로부터 의미있는 계수값을 사용했다. Wang[5]은 최저주파 대역의 웨이브릿 계수 크기로 영상을 인덱싱하여 비교하는 방법을 제안했다. Gross[6]는 KL확장과 Kohonen 맵을 웨이브릿 변환과 함께 사용하여 질감을 분석하였다. Kundu[7]는 웨이브릿 변환과 co-occurrence 행렬을 사용하여 통계적 접근법과 변환기반 접근법을 동시에 사용하는 방법을 제시하였다.

본 논문에서는 공간주파수 특성과 다중 해상도 특성을 모두 갖는 웨이브릿을 이용하여 각 대역의 특성에 맞는 비주얼 특징을 추출하고 이를 영상 검색에 이용하는 새로운 방법을 제시한다. 웨이브릿 변환된 영상의 최저주파 대역은 원 영상의 근사한 형태로 공간 정보를 충분히 활용할 수 있다. 이를 위해 웨이브릿 계수값과 픽셀간의 공간 정보를 모두 고려한 퍼지 동질성[8]을 이용하여 L개의 특징 벡터를 추출하였고,

나머지 고주파 대역의 에너지 값을 이용하여 3개의 특징 벡터를 추출하여 이를 영상 데이터베이스에 저장한다. 질의 시에는 L개의 FH 벡터 중 가장 크기가 큰 10개의 값과 3개의 고주파 대역의 에너지 값을 이용하여 유사한 순서로 영상을 검색하였다.

본문의 구성은 본문에서 웨이브릿 도메인상에서 특징 벡터 추출에 대한 설명과 구체적인 실험 결과를 보여주고, 결론과 향후 과제에 대해 기술하였다.

II. 웨이브릿 도메인에서의 비주얼 특징 추출

1. 웨이브릿 도메인

웨이브릿 분해는 적은 계수를 가지고 매우 좋은 근사값을 만든다. 그리고 웨이브릿 분해는 영상의 크기에 선형적인 관계를 갖는 매우 적은 코드가 필요하기 때문에 계산이 빠르고 쉽다[4]. 본 논문에서는 에너지 응집이 좋은 9/7 탭 쌍직교 웨이브릿 필터를 사용하여 영상을 2-레벨로 분해하였다. 그림 1은 2-레벨 웨이브릿 분해 영상이다. LL 대역은 원 영상의 근사한 형태이고, LH 대역은 수평 방향, HL 대역은 수직 방향, HH 대역은 대각 방향의 에지 정보를 나타낸다.

2. LL 대역에서 FH 벡터 추출

히스토그램에 기반을 둔 영상 처리는 오직 그레이 레벨만을 고려하고 픽셀과의 관계를 나타낼 수 있는 공간 정보를 포함하지 않는다. 그래서 히스토그램을 통한 비교는 다른 모양의 영상이라도 색의 분포만 같으면 같은 영상으로 간주한다. Cheng[8]은 이런 결점을 보완하기 위해 영상 분할 응용에 퍼지 동질성 방법을 제안했다.

아래 그림 2는 같은 색 분포를 갖는 다른 모양의 영상이다. 이 두 영상을 비교했을 때 그림 2(c)의 경우 히스토그램을 보여주는데 공간 정보 없이 색 분포만을 고려하기 때문에 같은 영상으로 인식하지만 그림 2(d)의 경우 호모그램을 보여주는 데 공간 정보를 고려하기 때문에 픽셀의 분포가 서로 달라 다른 영상으로 간주한다. 따라서 기존의 히스토그램을 이용한 영상 검색 시스템에서 발생할 수 있는 문제점을 해결할 수 있다.

웨이브릿 LL 대역은 원 영상의 근사한 형태로 공간 정보를 충분히 활용할 수 있다. 따라서 본 논문에서는 LL 대역의 웨이브릿 계수값을 L 레벨로 정규화를 시킨 후 웨이브릿 계수

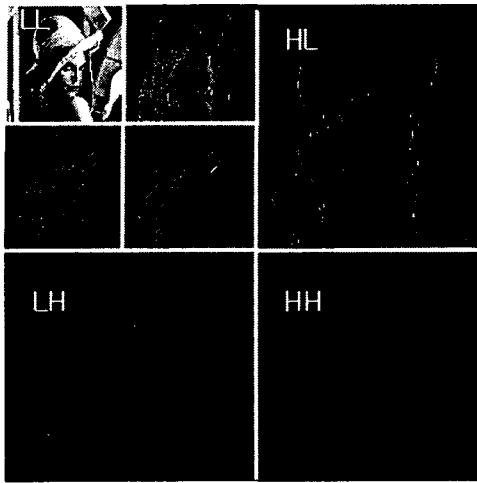


그림 1. 웨이블릿 변환 영상

여기서 i 와 j 는 웨이블릿 계수의 인덱스이다.

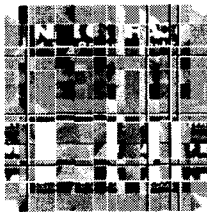
정규화된 LL 대역의 웨이블릿 계수 값들은 식 (2)~(5)에 의해 호모그램을 구한다. $\alpha(i, j)$ 와 $\alpha(k, l)$ 은 (i, j) 와 (k, l) 픽셀의 웨이블릿 계수이다. 식(2)는 퍼지 동질 벡터를 추출하기 전에 두 픽셀 사이의 동질한 정도를 나타내는 멤버십 함수를 나타낸다.

x 는 $\alpha(i, j)$ 와 $\alpha(k, l)$ 의 차이 값이고, a, b, c 는 일정한 상수 값이다.

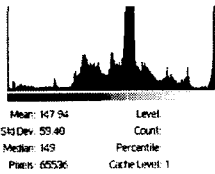
$$Z(x) \equiv \begin{cases} 1 & 0 \leq |x| \leq a \\ 1 - 2 \times \left(\frac{x-a}{c-a} \right)^2 & a \leq |x| \leq b \\ 2 \times \left(\frac{x-c}{c-a} \right)^2 & b \leq |x| \leq c \\ 0 & c \leq |x| \leq L-1 \end{cases} \quad (2)$$



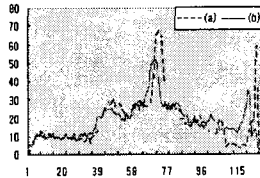
(a)영상 A



(b)영상 B



(c)영상 A, B 히스토그램



(d)영상 A, B 호모그램

그림 2. 히스토그램과 호모그램 비교

값과 픽셀간의 공간 정보를 모두 고려한 퍼지 동질성을 이용하여 8개의 특징 벡터를 추출하였다. 이러한 과정은 그림 3과 같다.

먼저 입력 영상을 $M \times N$ 크기로 정규화 한다. 정규화된 영상은 2-레벨 웨이블릿 변환을 한 후 L 레벨에서 LL 대역의 웨이블릿 계수를 정규화 한다.

정규화된 웨이블릿 계수인 $\alpha(i, j)$ 는 다음과 같다.

$$c(i, j) \in \{0, 1, 2, 3, \dots, L-1\}, 0 \leq i \leq M/4, 0 \leq j \leq N/4 \quad (1)$$

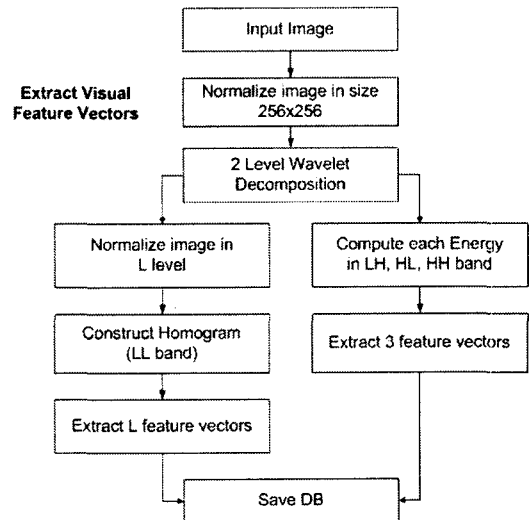


그림 3. 특징 벡터 데이터베이스 구성 흐름도

식 (3)은 퍼지 동질성 벡터 값으로 웨이블릿 계수 값이 t 레벨이고 계수사이 동질성과 θ 의 각을 갖는 주변 값을 합한 값이다. 즉, 8개 방향 θ 값의 동질성의 합을 구하였다.

$$h(t, \theta) = \left\{ \sum Z(|t-r|) \mid t = c(i, j), r = c(k, l), \right. \\ \left. 1 \leq i \leq M/4, 1 \leq j \leq N/4, \right. \\ \left. 0 < |(i, j) - (k, l)| \leq \sqrt{2}, \arg\{(i, j), (k, l)\} = \theta \right\} \quad (3)$$

$$\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ\}$$

호모그램은 정규화된 4개의 각 퍼지 동질성 벡터에 의해 아래 식 (4)와 같이, 호모그램 셋은 식 (5)와 같이 정의된다.

$$H(t) = \frac{1}{4} \left[\frac{h(t,0^\circ)}{(M-1)N} + \frac{h(t,45^\circ)}{(M-1)(N-1)} + \frac{h(t,90^\circ)}{M(N-1)} + \frac{h(t,135^\circ)}{(M-1)(N-1)} \right] \quad (4)$$

$$H = \{H(0), H(1), H(2), \dots, H(L-1)\} \quad (5)$$

위 식을 이용하여 웨이브릿 도메인 중 LL 대역에서 L개의 호모그램 값을 구해 영상 데이터베이스에 저장한다.

3. 고주파 웨이브릿 대역에서 에너지 벡터 추출

단일 특징이 아닌 다양한 특징들을 같이 사용함으로써 영상 검색의 효율을 높일 수 있다[1]. 본 논문에서는 또 다른 특징으로 LH, HL, HH 대역의 에너지 값을 이용하였다. 에너지는 명암도의 균일성을 재는 척도로서 화소간의 변화에 따른 특징 값을 산출하여 이를 CBIR을 위한 특징 벡터로서 이용할 수 있다. 이를 위해 웨이브릿 변환 영상에서 LH, HL, HH 대역들 각각에 대해 에너지를 구하여 역시 데이터베이스에 그 값을 저장한다. 각 대역의 에너지는 $B \in \{LH, HL, HH\}$ 일 때 식 (6)과 같다.

$$E(B) = \sum_{(i,j) \in B} c^2(i,j) \quad (6)$$

단, c는 각 대역에서의 변환계수이다.

4. 질의 결과 처리

입력 영상에 대한 질의 처리는 그림 4와 같다. 입력 영상에서 그림 3에서처럼 L개의 FHV 와 LH, HL, HH 대역들의 에너지를 구한다. 구해진 L개의 FHV를 모두 이용해 각각의 영상의 호모그램을 비교했을 때 많은 연산량이 필요하므로 본 논문에서는 호모그램을 크기 순으로 정렬한 후 상위 10개의 FHV만을 이용해 영상 검색에 이용하였다. 영상간의 유사도를 비교하기 위한 척도로서 일반적으로 특징 값들 사이의 거리를 구하여 이를 이용한다. 본 논문에서는 유사도 척도로 널리 사용되고 있는 유클리디안 디스턴스 함수를 이용하였다.

$$D_H(Q, I) = \sum_{i \in T_Q} |H_Q(t) - H_I(t)| \quad (7)$$

Q는 질의 영상을 나타내고, I는 데이터베이스내의 영상을 나타낸다. 같은 방법으로 질의 영상의 LH, HL, HH 대역의 에너지 값을 이용하여 데이터베이스내의 영상과의 유사성을 측정하였다.

$$D_E(Q, I) = \sum_{B \in \{LH, HL, HH\}} |E_Q(B) - E_I(B)| \quad (8)$$

입력 영상과 데이터베이스내 영상과의 특징 거리는 식 (9)와 같이 정의된다.

$$D(Q, I) = D_H(Q, I) / D_{H, \max} + D_E(Q, I) / D_{E, \max} \quad (9)$$

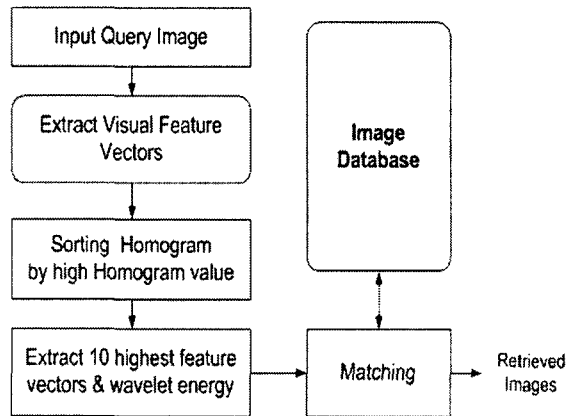


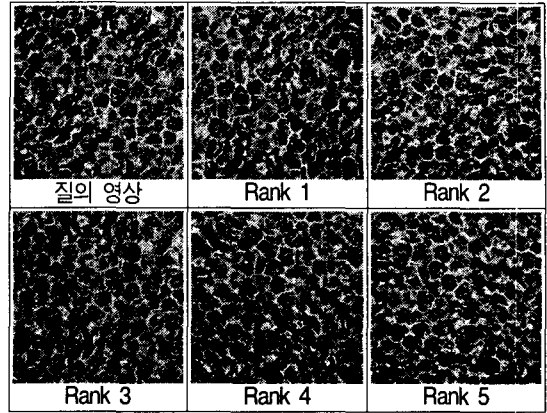
그림 4. 질의 처리

$D_E(Q, I)$ 는 LH, HL, HH 대역의 에너지 유사도를 나타내고 $D_{E, \max}$ 는 D_E 값의 최대 값을 나타낸다. 입력 영상과의 거리가 가장 작은 데이터베이스내 영상이 가장 비슷한 영상이다.

III. 실험 결과

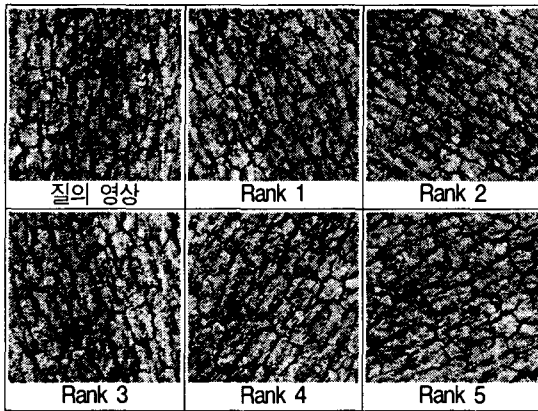
실험을 위해 본 논문에서는 총 90개의 텍스처 영상을 이용하였다. 그림 5는 질의 영상을 주었을 때 각각의 검색 결과를 나타낸다. 검색 결과는 위의 식 (9)을 이용하여 유사도를 측정 후 가장 작은 값을 갖는 영상을 Rank 1로 놓고 유사도 값을 올림차순으로 정렬하여 Rank 1~Rank 5까지 총 5개의 결과 영상을 나타낸다. 질의 영상이 그림 5 (a)인 경우 영상

의 텍스처가 성글어져 있어 공간 정보를 특징짓기 더 쉽기 때문에 상위 5개의 검색된 영상이 모두 같은 텍스처를 가진 영상을 검색해 내었다. 그림 5 (b)영상인 경우 Rank 1~4까지는 같은 텍스처를 가진 영상을 검색해 내고, Rank 5에서는 같은 텍스처는 아니지만 유사한 텍스처 영상을 검색해 내었다. 그림 5 (c) 영상인 경우 그림 5 (a)영상과 마찬가지로 Rank5까지 검색한 영상 모두 같은 텍스처를 갖는 영상이었다. 세 영상 모두 좋은 결과를 보여주고 있으나 모두 같은 텍스처 영상을 검색한 그림 5 (a)와 그림 5 (c)에 비해 다른 텍스처 영상을 검색한 그림 5 (b)는 조금 결과가 좋지 않은 것을 볼 수 있다. 텍스처 영상에서 검색이 좋은 결과를 보이는 데 텍스처 특징이 공간 정보를 충분히 특징짓게 성글어져 있는 영상이 더 정확한 검색을 할 수 있다.



(c)

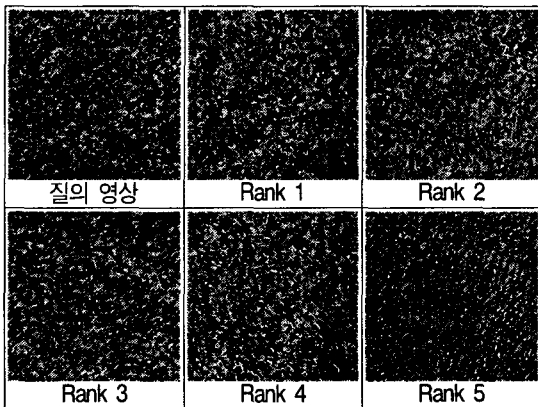
그림 5. 질의 영상과 검색된 영상



(a)

IV. 결론

본 논문에서는 공간주파수 특성과 다중 해상도 특성을 모두 갖는 웨이블릿 변환된 영상의 특징을 모두 이용하여 비주얼 특징을 추출하고 이를 CBIR에 이용하는 새로운 방법을 제시하였다. LL 대역에서는 퍼지 동질성을 이용하여 N개의 FHV를 추출하였고, LH, HL, HH 대역은 에너지 값을 이용하여 이를 CBIR의 특징 벡터로 사용하였다. 질의 시에는 데이터베이스내의 L개의 FHV 중 가장 큰 10개의 FHV를 사용함으로써 영상간의 유사도 검색 시 필요한 연산량을 줄였다. 본 논문에서 제시한 방법을 텍스처 영상 검색에 이용한 결과 좋은 정확성을 보였다. 그러나 좀 더 다양한 응용과 검색의 효율을 높이기 위해 특징 추출에 색상 정보를 포함하는 연구로 앞으로 해야 할 것이다.



(b)

참고 문헌

- [1] Vittorio Castelli, Image Databases, Wiley Publishers, pp. 11-14, 2002.
- [2] M. Swain and D. Ballard, "Color indexing," International Journal of Computer Vision 7(1), pp. 11-32, 1991.
- [3] K. C. Liang and C. -C. J. Kuo, "Retrieval and Progressive Transmission of Wavelet Compressed Images," in Proc. ISCAS, Hong Kong, pp. 1464-1467, 1997.
- [4] C.E. Jacobs, A. Finkelstein, and D.H. Salesin. "Fast

multiresolution image querying," in Proc. SIGGRAPH 95, Vol. 29, pp. 277-286, 1995.

- [5] J. Z. Wang, et al., "Wavelet-based image indexing techniques with partial sketch retrieval capability," in Proc. Forum on Res. and Tech. Adv. in Dig. Lib., pp. 13-24, May 1997.
- [6] M. H. Gross, R. Koch, and L. Lippert, "Multiscale Image texture analysis in wavelet spaces," in Proceeding IEEE Int. Conference on Image Processing, Vol. 3, pp. 412-416, Nov. 1994.
- [7] A. Kundu, "Texture classification using QMF bank-based subband decomposition," CVGIP, Vol. 54, No. 5, pp. 369-384, Sept. 1992.
- [8] H.D. Cheng, C.H. Chen, "Fuzzy homogeneity approach to multilevel thresholding," IEEE Trans. Image Processing, vol.7, pp.1084-1086, July 1998.

송 영 준(Young-Jun Song)

정회원



1996년 2월 : 충북대학교 정보통신공학과(공학석사)

1996년 ~ 1998년 : LG전자 멀티미디어 사업본부 주임연구원

1998년 ~ 2000년 : LG반도체

메모리사업본부 주임연구원

2000년 ~ 2003년 : 한국전자통신연구원 네트워크연구소 홈네트워크팀 선임연구원

2000년 ~ 현재 : 충북대학교 정보통신공학과(박사과정)

<관심분야> : 영상인식, 영상처리, 얼굴 인식

박 원 배(Won-Bae Park)

정회원



1988년 2월 : 충북대학교 정보통신공학과(공학석사)

1998년 2월 : 충북대학교 정보통신공학과(공학박사)

2003년 ~ 현재 : LG전자 선임연구원

<관심분야> : CBIR, Wavelet

류 은 주(Eun-Ju Ryu)

정회원



2003년 2월 : 충북대학교 정보통신공학과(공학사)

2003년 ~ 현재 : 충북대학교 정보통신공학과(석사과정)

<관심분야> : CBIR