
주 색상에 의한 객체 영역을 이용한 내용기반 영상 검색

Content-based Image Retrieval Using Object Region With Main Color

김동우*, 장언동*, 객내정*, 송영준**

충북대학교 영상통신연구실*, 충북대학교 전기전자컴퓨터공학부**

Dong-Woo Kim(dubssi@paran.com)*, Un-Dong Chang(udchang@naver.com)*,
Nae-Joung Kwak(knj0125@lycos.co.kr)*, Young-Jun Song(songyjorg@dreamwiz.com)**

요약

본 논문은 기존의 컬러 히스토그램 방법들의 단점을 극복하고자 객체 영역을 이용한 내용기반 영상 검색 방법을 제안한다. 기존 컬러 히스토그램 검색 방법들은 양자화 오류 등의 이유로 정확성이 떨어지고, 공간정보가 부족한 단점이 있다. 이를 해결하기 위해 제안 방법은 색상 정보를 HSV 공간으로 변환하여 순수 색상 정보인 hue 성분만을 양자화하여 히스토그램을 구해 명암, 이동, 회전 등에 강인한 검색 특징으로 사용한다. 한편 공간정보가 부족한 문제점을 해결하기 위해 색상 특징과 영역간의 상관관계를 고려하여 객체영역을 선정한다. 선정된 객체 영역에서는 에지와 DC를 이용하여 검색한다. 자연 컬러 영상 1,000개를 가지고 실험한 결과 기존 방법들보다 precision과 recall이 우수하였다.

■ 중심어 : | 내용기반검색 | 특징추출 | 컬러 히스토그램 |

Abstract

This study has proposed a method of content-based image retrieval using object region in order to overcome disadvantages of existing color histogram methods. The existing color histogram methods have a weak point of reducing accuracy, because these have both a quantization error and an absence of spatial information. In order to overcome this problem, we convert a color information to a HSV space, quantize hue factor being pure color information, and calculate histogram. And then we use hue for retrieval feature that is robust in brightness, movement, and rotation. To solve the problem of the absence of spatial information, we select object region in terms of color feature and region correlation. And we use both the edge and the DC in the selected region for retrieving. As a result of experiment with 1,000 natural color images, the proposed method shows better precision and recall than the existing methods.

■ keyword : | CBIR | Feature Extraction | Color Histogram |

1. 서론

현재 우리는 인터넷과 컴퓨터 기술의 발전으로 인하여

다양한 멀티미디어 정보를 획득 저장할 수 있게 되었다. 특히 영상정보는 여러 분야에서 사용이 급증하고 있다. 이러한 영상정보는 사용이 많아진 반면 관리하는 방법이

는 아직도 많은 한계가 있다. 따라서 영상을 효율적으로 관리하기 위한 영상 검색 방법이 필요하게 되었다.

영상 검색 방법은 초기에는 입력한 키워드를 이용하여 원하는 영상을 검색하는 문자기반 검색 방법(text-based retrieval)[1][2]들이 제안되었다. 그러나 이 방법은 검색의 객관성이 떨어지고, 사람이 키워드를 직접 부여하여야 한다는 단점이 있다. 따라서 객관적이며 자동화된 영상검색이 가능한 내용기반 영상검색(CBIR: Content Based Image Retrieval)[3][4] 방법들이 제안되었다.

내용기반 영상검색은 일반적으로 영상의 색상, 질감, 모양의 세 가지 특징을 주로 사용한다. 색상 정보는 영상의 정보 중 가장 큰 부분을 차지하며 널리 사용되는 특징이다. 질감 정보는 색상이나 명암도의 존재만으로 산출되지 않는 유사 성질을 가지는 시각적 패턴이다. 모양 정보는 표현하는 방법의 어려움으로 인해 색상이나 질감보다는 이용이 제한되어 한정된 시스템에서 사용된다.

본 논문은 순수 색상 정보인 hue 성분만을 양자화하여 히스토그램을 구해 명암, 이동, 회전 등에 강인한 특징으로 사용한다[5]. 그리고 16개의 영역으로 나눈 각 영역에서 중앙영역에서 제1 객체 색상을 선택하고, 그 밖의 영역에서 제2 객체 색상을 선택한다. 제1 객체 색상의 컬러 히스토그램의 빈도가 높은 6개의 영역을 우선 선택하고, 그중 빈도가 낮은 영역은 제2 객체 색상으로 선택한 영역으로 교체하여 6개의 객체 영역을 선정한다. 선정된 객체 영역에서는 검색정보를 추가하기 위해 모양 특징인 에지와 질감 특징인 DC(Direct Current)를 이용하여 검색한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 내용기반 검색 방법에 대하여 기술하고, 3장에서는 제안한 방법에 대하여 기술한다. 4장에서는 실험결과를 살펴보고, 5장에서 결론과 향후 과제에 대하여 논한다.

II. 기존 방법

색상은 내용기반 영상검색에서 가장 광범위하게 사용되는 특징이다. 색상 특징은 복잡한 배경에서 다른 특징보다 상대적으로 강건하고 영상의 크기와 이동 및 회전

에 민감하지 않다. 일반적으로 내용기반 영상 검색에서 주로 사용하는 것은 컬러 히스토그램 방법이다.

컬러 히스토그램은 영상 내의 각 화소에 대해 동일한 계 양자화된 색상을 계수(count)함으로써 얻을 수 있다. 이렇게 얻어진 컬러 히스토그램은 해당 영상의 색상 특징으로 사용된다.

Swain[6] 등이 제안한 히스토그램 인터섹션은 계산이 간단하다는 장점이 있다. 그러나 이 방법은 조명의 변화와 영상내의 물체 크기에 민감하다는 단점이 있다. 또한 영상의 공간정보를 이용하지 않는다는 최대 단점이 있다.

Pass[7] 등은 컬러 히스토그램 방법이 공간정보가 부족한 것을 해결하기 위해 CCV(Color Coherence Vector)를 제안하였다. 그러나 CCV는 공간정보를 얻기 위해 주변 화소의 관계를 고려하기 때문에 계산량이 많은 단점이 있다.

Song[8] 등은 컬러 양자화를 수행함으로써 발생하는 양자화 에러를 줄이기 위해 R, G, B 성분을 특징벡터로 활용하는 방법을 제안했다. 이 방법의 경우도 검색도가 떨어지는 단점이 있다.

류은주[9] 등은 영상에 있어 객체가 주로 중앙에 위치한다는 것을 바탕으로 중앙 영역 색상 특징을 이용한 내용기반 검색 방법을 제안하였다. 이 방법은 객체가 중앙에 있지 않는 경우에는 오류가 발생할 수 있는 단점이 있다.

한편 오늘날 내용기반 검색은 색상만을 이용하기 보다는 질감 특징, 모양 특징을 같이 사용한다. 질감은 색상이나 명암도의 존재만으로 산출되지 않는 유사 성질을 가지는 시각적 패턴이며 모든 물체 표면의 고유한 성질로 색상으로 표현할 수 없는 정보를 가진다. 또한 모양특정도 제약사항이 많아 널리 적용하기는 힘들지만 적절한 가중치를 적용하게 되면 검색률을 효과적으로 높일 수 있다.

최근에는 색상, 질감, 모양 특징이 각각의 장단점을 보완하여 줄 수 있기 때문에 2-3가지의 특징을 혼용하여 검색률을 높이는 방법들[10][11]이 많이 제안되고 있다.

III. 제안 방법

기존의 컬러 히스토그램 방법들의 단점인 공간정보가 부족한 문제점을 해결하기 위해 객체 영역에 질감 특징과 모양 특징을 추가하면 검색률을 높일 수 있다. 제안한 방법의 전체적인 구성은 [그림 1]과 같다.

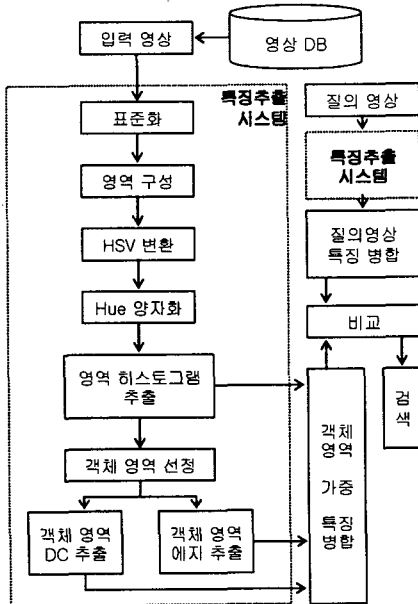


그림 1. 제안 방법 구성도

1. 전처리

제안 방법은 입력영상을 전처리 단계로 256×256 크기의 영상으로 표준화한다. 표준화된 영상을 영역으로 구성하기 위해 N개의 영역으로 분할한다. 실험에서는 64×64 크기로 블록화 하여 N을 16으로 하였다. 영역 구성 후 순수 색상 정보인 hue를 얻기 위해 영상을 변환하여 HSV 색상 공간을 만든다.

2. 1차 특징 추출

전처리 단계를 거쳐 나온 HSV 공간의 색상 영상에서 색상 정보를 담당하는 hue 성분만을 가지고 양자화를 한다. hue는 0-359까지의 값을 가지게 되며 계산량을 줄이기 위해 M개 레벨로 양자화한다. 양자화한 레벨명 H_i는 영역 내의 대표값이라 할 수 있는 평균값으로 설정된다.

영상에서 추출한 색상 특징은 각 영역 R_k별로 구하여 RCFV(Region Color Feature Vector)라 하고 식(1)과 같이 표현된다.

$$RCFV = \{R_k, H_i, P_i\} \quad (1)$$

여기서 k = 1, 2, ..., N이고, i = 1, 2, ..., M이다. 이때 N은 영상의 영역을 나눈 블록의 총 수를 나타내고 M은 양자화 색상 총 수를 나타낸다. 그리고 P_i는 해당 hue 값을 갖는 색상의 비율 값이다.

3. 객체 영역 선정

기존 검색 방법은 컬러 히스토그램을 사용하여 검색할 경우 배경의 색상 비중이 가장 클 경우가 많다. 이 경우 중요한 객체보다 배경에 의해 검색되는 문제점이 있다. 이러한 문제점을 해결하기 위해 객체가 있는 영역에서는 정보를 추가하여 검색하면 검색률을 높일 수 있다. [그림 2]와 같이 영역이 만들어진 경우 객체가 주로 중앙에 있으므로[9], 중앙 영역(6, 7, 10, 11번)에서 가장 높은 컬러 히스토그램을 가진 색상을 영상의 제1 객체 색상으로 정하고, 외각 영역(1, 2, 3, 4, 5, 8, 9, 12, 13, 14, 15, 16번)에서는 가장 높은 컬러히스토그램을 가진 색상은 배경이므로 이를 제외하고 그 다음으로 많은 색상을 제2 객체 색상으로 정한다. 16개의 영역 중에서 제1 객체 색상의 빈도수가 높은 6개의 영역을 1차 객체 영역으로 선택한다. 선택된 1차 객체 영역 중에서 제1 객체 색상의 빈도가 20% 이하인 영역은 객체 영역에서 제외한다. 제외된 영역은 제2 객체 색상이 80% 이상인 영역으로 교체한다. 제2 객체 색상이 80% 이상인 영역의 수가 제외된 영역의 수보다 작은 경우에는 제외되었던 영역을 다시 객체 영역으로 선정한다. 영역 선택 결과 최종 객체 영역은 6개가 된다. [그림 2]의 경우 제1 객체 색상이 흰색이 되고, 제2 객체 색상은 갈색이 된다. 따라서 최초 영역은 6, 7, 8, 10, 11, 12, 15가 되지만 제1 객체 색상이 20%,미만인 15번은 제외되고, 제2 객체 색상이 80%를 넘는 9번이 추가되어 최종 6, 7, 8, 9, 10, 11번이 객체 영역으로 선정된다.

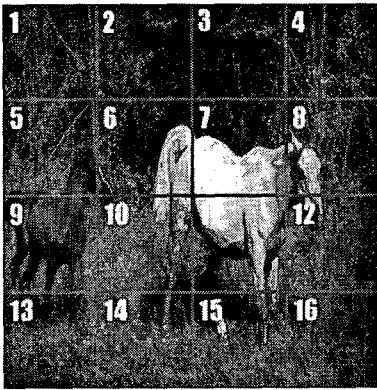


그림 2. 영상의 영역 분할

4. 2차 특징 추출

객체 영역으로 선정된 영역은 색상 특징 외에 질감과 모양 특징을 추가하여 추출한다. 질감 특징은 그레이 영상으로 변환한 후 객체 영역별로 DCT 변환을 하여 DC 정보 D_k 를 얻어 사용한다.

영상에서 추출한 각 객체 영역별 질감 특징은 RTFV(Region Texture Feature vector)라 하고 식 (2)와 같이 벡터 형식으로 표현된다.

$$RTFV = [R_k, D_k], (k = 1, 2, \dots, N) \quad (2)$$

모양 특징은 DC 추출과는 별개로 그레이 영상 변환 후 잡음 제거를 위해 median 필터를 사용하고, Sobel 변환으로 수직방향 에지를 검출하여 사용한다. 검출한 에지 영상에서 에지성분으로 검출된 화소가 4번의 계수 중 2번 이상 나올 시 에지로 인정하고 에지 특징을 증가시킨다. 이렇게 증가시킨 정보 E_k 를 모양 특징으로 사용한다.

영상에서 추출한 각 객체 영역별 모양 특징은 RSFV(Region Shape Feature Vector)라 하고 식 (3)과 같이 벡터 형식으로 표현된다.

$$RSFV = [R_k, E_k], (k = 1, 2, \dots, N) \quad (3)$$

5. 특징 병합과 비교

3가지 특징은 DB에 저장될 때는 병합되어 16개영역

모두의 정보를 저장한다. 그러나 검색할 때는 색상 특징은 모든 영역에서 사용하고, 질감과 모양 특징은 객체 영역의 정보만을 사용한다.

내용기반 검색에서 일반적으로 사용하는 유클리디안 거리 비교는 유사한 영상이 조명의 변화 같은 요인에 의해 검색률이 낮아질 수 있다. 이러한 문제점을 해결하기 위해 RCFV 비교는 퍼지 컬러 거리 함수[9]를 사용하고, 식 (4)와 같다.

$$D^2(I_1, I_2) = \sum_{i=1}^B P_{1i}^2 + \sum_{j=1}^B P_{2j}^2 - \sum_{i=1}^B \sum_{j=1}^B 2\mu(I_1, I_2) P_{1i} P_{2j} \quad (4)$$

여기서 I_1 은 질의영상의 컬러 특징, I_2 는 DB내 영상의 컬러특징, B는 컬러 수를 나타낸다. 한편 RTFV, RSFV의 비교는 색상과 달리 유클리디안 거리를 사용하여도 오류 영향이 적다. 그래서 계산량을 줄이기 위해 각 영역별로 유클리디안 거리 함수를 사용한다.

IV. 실험 결과

1,000장의 자연 영상[12]인 말, 꽃, 버스 아프리카 영상 등을 사용하여 제안한 검색 방법을 실험하였다.

본 논문에서는 제안한 방법의 시각적 검색 성능을 평가하기 위해 Song[8] 등이 제안한 컬러 히스토그램 방법과 류은주[9] 등이 제안한 중앙영역 히스토그램 방법을 가지고 비교하였다.

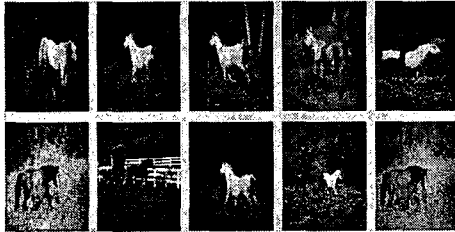
[그림 3]은 [그림 2]에 보인 말 영상을 질의해서 검색된 결과를 유사도가 가장 높은 1번부터 차례로 10번까지 나열한 결과 화면이다.

실험 결과 Song이 제안한 방법은 10번째 나온 결과 영상이 양자화 오차에 따라 검색에 실패하였다. 류은주가 제안한 방법은 모두 말 영상을 검색하였으나 대부분 흰색 말과 갈색 말이 함께 있는 영상을 검색하지 못하였다.

그러나 제안한 방법은 8번을 빼고 모두 흰색과 갈색 말이 함께 있는 영상을 검색하였다. 따라서 제안된 방법이 기존 방법들 보다 원 영상과 시각적으로 더 유사한 영상을 검색함을 알 수 있다.



(a) Song 방법



(b) 류은주 방법



(c) 제안 방법

그림 3. 말 영상 검색 결과

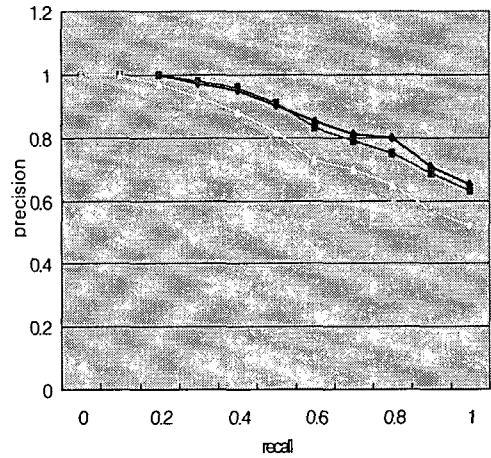
객관적인 평가를 위하여 내용기반 영상 기법에서 주로 사용하는 precision과 recall을 이용하였다. 또한 최근 논문인 Saha[11] 등이 제안한 질감과 퍼지 컬러를 이용한 방법을 추가적으로 비교하였다. 본 논문에서는 말과 꽃 영상의 각 그룹에서 무작위로 10개의 영상을 질의하여, 각 그룹별로 precision과 recall의 평균을 구하였다. 식(5)는 precision P_k 를 구하는 식이고, 식(6)은 recall R_k 를 구하는 식이다.

$$P_k = \frac{A_k}{A_k + B_k} \quad (5)$$

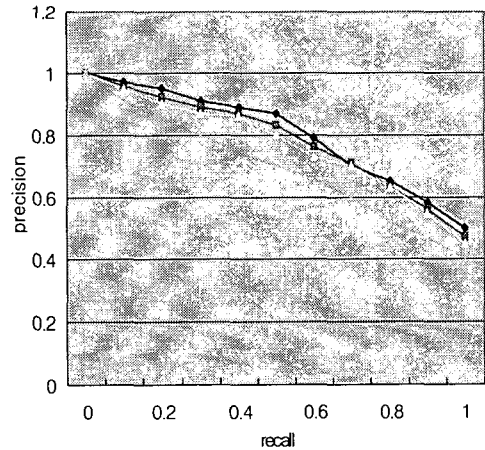
$$R_k = \frac{A_k}{A_k + C_k} \quad (6)$$

여기서 A_k 는 검색된 항목 중에서 질의와 관련된 유사

이미지의 수, B_k 는 검색된 항목 중에서 질의와 관련이 없는 이미지의 수, C_k 는 검색되지 않았으나 질의와 관련된 유사 이미지의 수를 나타낸다.



(a) 말 영상



(b) 꽃 영상

그림 4. 검색 결과

[그림 4]는 각 검색 방법에 대한 비교 결과이다. 말 영상의 경우 객체인 말이 중앙뿐만 아니라 배경 영역에도 존재하기 때문에 중앙 영역 위주로 검색한 류은주 방법

보다 검색률이 전체 평균으로 6% 정도 많이 좋아 졌다. Saha 방법은 질감과 색상 정보만을 사용하여 recall이 커지면 제안방법보다 낮은 검색률을 보이며 전체 평균으로 제안 방법이 2%가량 검색률이 좋았다.

꽃 영상의 경우는 객체인 꽃이 중앙에만 있으므로 류은주 방법보다 검색률이 3% 정도 약간 좋아진 것을 알 수 있다. Saha 방법은 제안방법보다 2%정도 낮은 검색율을 보였다.

한편 제안 방법이 Song의 방법보다는 두 영상 모두 10% 정도 좋은 검색률을 보인다. 결과적으로 제안한 방법의 성능이 기존의 검색 방법들보다 우수함을 알 수 있다.

V. 결론

내용기반 영상 검색 시스템에서 영상의 색상 특징을 추출하여 특징 벡터로 나타내는 대표적인 방법으로 컬러 히스토그램 방법들이 있다. 그러나 이 방법들은 양자화 오류나 공간정보의 부족으로 인하여 비슷한 색상 분포를 가지는 다른 영상을 검출할 수 있다. 이를 개선한 다양한 방법들이 있지만 아직 검색률을 좀더 높여야 한다.

본 논문에서는 기존의 컬러 히스토그램 방법들의 단점을 극복하고자 객체 영역을 이용한 영상 검색 방법을 제안한다. 영상을 일정크기의 영역으로 나누어서 컬러 정보를 HSV로 변환하여 순수 컬러 정보인 hue 성분만을 양자화하여 각 영역에 대하여 히스토그램을 구해 색상 특징벡터(RCFV)를 얻는다. 공간 정보가 부족한 것을 해결하기 위해 객체 영역은 RCFV와 영역 상관관계를 이용하여 선정한다. 객체가 있는 영역에서는 검색율을 높이기 위해 에지를 이용한 모양 특징벡터(RSFV), DC정보를 이용한 질감 특징벡터(RTFV)를 추가로 사용한다. 이렇게 얻어진 영역 특징벡터를 색상은 퍼지 컬러 거리 함수를 이용하고, 모양과 질감은 유클리디안 거리 함수를 이용하여 유사도를 측정한다.

자연 컬러 영상 1000개를 가지고 실험한 결과 기존의 방법들 보다 precision과 recall이 우수하였다. 향후 연구 과제는 객체 영역을 선정하는 부분에 있어서 배경이 복잡한 경우 오류가 발생할 수 있는 부분을 보완하는 것이다.

참고 문헌

- [1] N. S. Chang and K. S. Fu, "Query-by Pictorial Example," IEEE Trans. On Computer, Vol.6, No.6, pp.519-524, 1980.
- [2] S. K. Chang, C. W. Yan, D. C. Dimitroff, and T. Arndt, "An Intelligent Image Database System," IEEE Trans. Software Eng., Vol.14, No.5, pp.681-688, 1988.
- [3] V. N. Gudivada and J. V. Raghavan, "Special Issue on Content-based Image Retrieval Systems," IEEE Computer Magazine, Vol.28, No.9, pp.18-22, 1995.
- [4] J. Eakins and M. Graham, *Content-based Image Retrieval*, JSIC Technology Application Report, 1999.
- [5] J. C. Russ, *The Image Processing Handbook* Third Edition, 1999.
- [6] M. J. Swain and D. H. Ballard, "Color Indexing," International Journal of Computer Vision, Vol.7, No.1, pp.11-32, 1991.
- [7] G. Pass, R. Zabih, and J. Miler, "Comparing Images Using Color Coherence Vectors," ACM Conference on Multimedia, Massachusetts, pp.65-73, 1996.
- [8] Y. J. Song, W. B. Park, D. W. Kim, and J. H. Ahn, "Content-based Image Retrieval Using New Color Histogram," ISPAC 2004, pp.609-611, 2004.
- [9] 류은주, 송영준, 박원배, 안재형, "중앙 영역의 컬러 특징과 최적화된 빈 수를 이용한 내용기반 영상 검색", 한국정보처리학회 논문지B, 제11-B권, 제5호, pp.581-586, 2004.
- [10] 박원배, *내용기반 영상 검색을 위한 시각 특징 추출 기법*, 충북대학교 박사학위논문, 2004.
- [11] S. K. Saha, A. K. Das, and B. Chanda, "CBIR Using Perception Based Texture and Colour Measures," Pattern Recognition ICPR 2004.

Proceedings of the 17th International Conference, Vol.2, pp.985-988, 2004.

- [12] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLiCity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.23, No.9, pp.947-963, 2001.

저자소개

김 동 우(Dong-Woo Kim)

정회원



- 1997년 2월 : 충북대학교정보통신 공학과(공학사)
- 2002년 2월 : 충북대학교정보통신 공학과(공학석사)
- 2002년 3월~현재 : 충북대학교정보통신공학과 박사과정

<관심분야> : 내용기반검색, 멀티미디어 정보처리

장 언 동(Un-Dong Chang)

정회원



- 1996년 2월 : 충북대학교정보통신 공학과(공학사)
- 2002년 2월 : 충북대학교 정보통신공학과(공학석사)
- 2002년 3월~현재 : 충북대학교 정보통신공학과 박사과정

<관심분야> : 영상신호처리, 컴퓨터 비전

곽 내 정(Nae-Joung Kwak)

정회원



- 1993년 2월 : 충북대학교정보통신 공학과(공학사)
- 1995년 2월 : 충북대학교정보통신 공학과(공학석사)
- 2005년 2월 : 충북대학교정보통신 공학과(공학박사)

• 2005년 3월~현재 : 목원대학교 정보통신 공학부 프로그래밍 전문강사

<관심분야> : 해프토닝, 영상분할, 영상정보처리, 컴퓨터 비전, 멀티미디어 콘텐츠

송 영 준(Young-Jun Song)

종신회원



- 1994년 2월 : 충북대학교정보통신 공학과(공학사)
- 1996년 8월 : 충북대학교 정보통신공학과(공학석사)
- 2004년 8월 : 충북대학교 정보통신공학과(공학박사)

• 2005년 9월~현재 : 충북대학교 전기전자컴퓨터공학부 박사 후 연구원

<관심분야> : 영상 인식, 영상 처리, 얼굴 인식