

논문-00-5-1-07

사용자 피드백 기반의 적응적 가중치를 이용한 정지영상 검색

이진수*, 김현준*, 윤경로*, 이희연*

Image Retrieval using Adaptable Weighting Scheme on Relevance Feedback

Jin-Soo Lee*, Heon Jun Kim*, Kyoungro Yoon*, and Hee-Youn Lee*

요 약

사용자 피드백은 일반적으로 사용자가 의도하는 정지영상 검색 조건을 기술하는 데만 주로 사용되어 왔다. 그러나, 본 논문에서는 사용자 피드백을 정지영상의 일반적인 특징을 기술하는데 사용함으로써 사용자에 의존적이지 않은 정지영상 검색에 적용하였다. 그리고 본 논문에서는 사용자 피드백을 사용하여 각 정지영상마다 고유한 특징을 반영하도록 특징 정보와 관련된 가중치를 전문가에 비중을 두어 학습시킴으로써, 일반적인 검색 성능을 향상시킬 수 있다. 이러한 시스템을 구축하기 위해 본 논문에서는 칼라 기술자와 텍스처 기술자를 기반으로 한 전역 특징 정보와 지역 특징 정보, 그리고 각 기술자들간의 가중치와 기술자 내의 요소 가중치로 구성된 정지영상 기술 구조를 제안하고, 또한 잘못된 학습을 방지하기 위해 신뢰도에 기반한 가중치 학습 방법을 소개한다.

Abstract

Generally, relevance feedback reflecting user's intention has been used utilized to refine the query conditions in image retrieval. However, in this paper, the usage of the relevance feedback is extended to the image database categorization so as to be accommodated to the user independent image retrieval. In our approach, to guarantee a desirable user-satisfactory performance descriptors and the elements of the descriptors corresponding unique features associated with of each image are weighted using the relevance feedback where experts can more lead rather than beginners do. In this paper, we propose a proper image description scheme consisting of global information, local information, descriptor weights and element weights based on color and texture descriptors. In addition, we also introduce an appropriate learning method based on the reliability scheme preventing wrong learning from abusive feedback.

I. 서 론

일반적으로 정지영상 검색 시스템에서는 칼라, 텍스처, 모양정보 등 서로 다른 종류의 특징 정보(feature)를 사용한다. 하지만 정지영상마다 그 정지영상을 구분 지을 수 있는 특징이 다르기 때문에, 해당 정지영상과 유사한 정지영상을 찾는데 보다 효율적인 특징 정보는 각 정지영상마다

다 다르다. 이러한 이유로 각 정지영상마다 기술자(descriptor)의 가중치를 달리하여 정지영상을 검색함으로써 보다 높은 검색 성능을 얻기 위한 연구가 보고되고 있다^{[1][2][3]}. 가중치를 결정하는 방법에는 사용자가 적합한 정지영상(relevant image) 또는 부적합한 정지영상(irrelevant image)에 대한 정보를 시스템에 피드백함으로써 시스템이 적합한 가중치를 계산하도록 하는 방법이 최근 들어 다수 소개되고 있다^[2]. 이와 같은 방법에서는 일반적으로 갱신된 값을 유지하지 않기 때문에 각 질의 때마다 같은 참조정지영상이라도 피드백을 통해 새로 가중치 값을 갱신하게

* LG전자기술원
LG Electronics Institute of Technology

된다. 하지만 각 정지영상마다 이전에 갱신된 가중치 값을 정지영상의 또 다른 기술자로서 유지함으로써, 이전에 갱신된 가중치값을 이후에 재사용할 수 있을 뿐 아니라, 새로운 피드백을 사용하여 현재의 가중치 값을 계속 학습시킬 수 있다. 이와 같이 정지영상 기술자로서의 가중치 값은 서로 다른 사용자, 또는 서로 다른 검색 엔진에 의해 학습되기 때문에 정지영상 고유의 특징을 잘 반영하여 보다 높은 검색 성능을 가능하게 한다.

본 논문에서는 이와 같이 사용자 피드백을 통해 정지영상 특징과 관련된 가중치 값을 효과적으로 학습할 수 있는 방법 및 적합한 정지영상 기술 구조를 제안한다. 사용자 및 검색 엔진과 무관하게 사용자 피드백을 통해 가중치를 지속적으로 학습시키기 위해서 고려해야 될 사항을 제시하고 이를 해결하기 위한 방법을 함께 소개한다.

본 논문의 구조는 다음과 같다. 제 2장에서는 사용자 피드백을 사용한 가중치 학습 시 고려사항 들을 기술하고 제 3장에서는 2장에서 기술한 고려사항 들을 해결하여 구성한 새로운 정지영상 기술 구조 (Description Scheme) 및 방법을 소개하며, 제 4장에서는 테스트에 사용된 정지영상 검색기와 사용된 기술자를 소개한다. 제 5장에서는 우리가 제안한 방법의 우수성을 실험을 통해 증명하며 마지막으로 제 6장에서는 결론을 맺는다.

II. 사용자 피드백을 사용한 정지영상 검색에서의 고려사항

사용자 피드백을 반영하여 각 정지영상마다 검색 성능을 높이기 위한 가중치를 학습시키기 위해서는 다음과 같이 세가지 고려해야 할 사항이 있다.

1. 권한 등급

한 사용자가 각 질의 때마다 피드백을 통해 가중치를 갱신한 후 그 값을 검색에 사용하는 기존 방법과는 달리, 여러 사용자의 피드백을 반영하여 가중치를 학습시킬 경우 잘못된 피드백에 의해 잘못된 가중치 학습이 발생할 수 있다. 따라서, 피드백을 주는 사용자의 권한등급을 달리함으로써, 시스템이 권한을 준 전문가의 경우에는 그 피드백에 의해 가중치가 보다 많이 영향을 받도록 하고 반대로 일반 사용자의 경우에는 그 피드백에 의해 가중치가 덜 영향을 받도록 하는 것이 필요하다. 또한 초기에는 어느

정도 안정된 가중치 값을 갖을 때까지 전문가의 피드백에 의해 학습시킴으로써 잘못된 가중치 학습을 방지할 수 있다. 전문가는 일반적으로 데이터베이스 서버에서 지정할 수 있으며, 전문가에 의해 학습된 초기 가중치 값을 시작으로 일반 사용자의 피드백을 이용한 추가적인 학습이 가능하다.

2. 신뢰도

특정 정지영상의 가중치 값이 보다 많은 피드백에 의해 학습되었다면, 그 가중치는 보다 신뢰적이라고 할 수 있다. 따라서 이 경우에는 다음 새로운 피드백에 의해 영향 받는 정도가 상대적으로 작게 함으로써 새로운 피드백에 의해 쉽게 현재의 가중치 값이 변형되지 않도록 해야 한다. 그러기 위해 현재의 가중치 값이 어느 정도 신뢰적인지를 표현할 수 있는 신뢰도 정보가 가중치와 함께 정지영상 기술자로서 기술되어야 한다. 이러한 신뢰도는 현재의 가중치 값이 얼마나 신뢰적인지를 나타내는 값으로서 학습에 참여한 피드백 수와 피드백을 준 전문가의 등급에 비례하여 수치로 표현될 수 있다. 신뢰도 값은 앞서 설명하였듯이 새로운 피드백에 의해 현재의 가중치를 학습시킬 때 사용될 뿐만 아니라, 같은 정지영상의 두 복사본이 서로 다른 경로로 학습되어 다른 가중치 값을 가질 경우, 이들 가중치 값을 조합하거나 선택할 때 사용된다.

3. 다양한 질의의 만족성

참조 정지영상을 사용하여 유사한 정지영상을 검색할 경우 유사성의 척도가 되는 특징 요소들은 각 정지영상마다 다양하게 나타날 수 있다. 따라서 다양한 참조정지영상을 사용하는 모든 질의를 만족 시키려면, 우선 각 질의에서 주요하게 사용되는 특징을 묘사하는 기술자들이 모두 포함되어 있어야 하며, 또한 각 질의마다 어떠한 기술자가 얼마만큼 주요한 요소인지를 기술하기 위한 가중치 정보가 포함되어야 한다.

이와 같은 조건을 만족시키기 위해 본 논문에서는, 전역적인 정지영상 특징을 나타내는 전역 특징 기술자와 정지영상의 각 부분별로 특징을 기술하는 지역 특징 기술자와 함께, 기술자간 가중치, 요소 가중치, 위치 가중치의 3가지 유형의 가중치들로 구성된 정지영상 기술 구조를 제안한다.

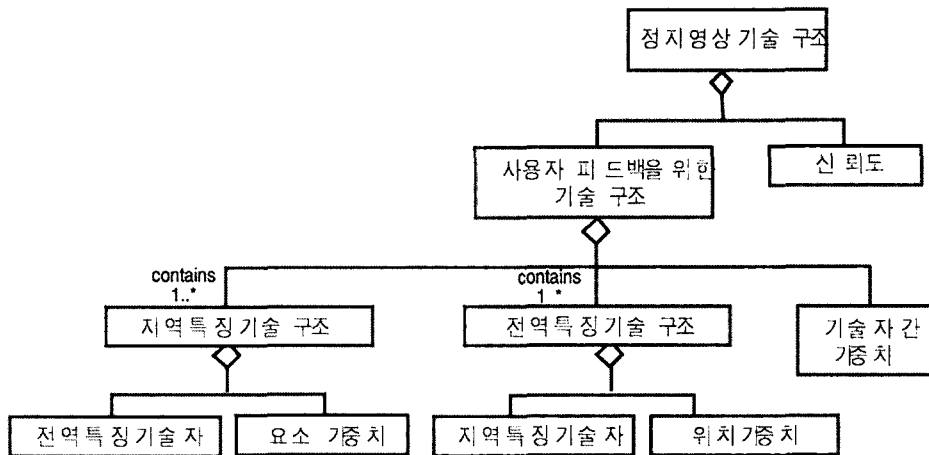


그림 1. 정지영상 기술 구조
Fig. 1. Description Scheme for Still Image

III. 적응적 가중치를 적용한 정지영상 검색

우리는 앞서 2장에서 기술한 권한등급, 신뢰도, 그리고 다양한 질의에 대한 만족성을 해결하기 위해 다음과 같은 정지영상 기술 구조를 제안한다.

3.1. 정지영상 기술 구조

우리가 본 논문에서 제안하는 정지영상 기술 구조를 UML(Unified Modeling Language)[5]로 표현하면 그림 1과 같다. 먼저 정지영상 기술 구조는 사용자 피드백을 위한 기술 구조와 신뢰도로 구성된다. 신뢰도는 앞서 기술하였듯이 해당 정지영상의 가중치 값이 얼마나 신뢰적인지를 나타내는 특징 정보이다. 사용자 피드백을 위한 기술 구조는 전역 특징 기술 구조와 지역 특징 기술 구조 그리고 기술자간 가중치로 구성되는데, 기술자간 가중치는 본 구조에 포함되어 있는 다양한 기술자들의 중요도를 기술하는데 사용된다. 전역 특징 기술 구조는 칼라히스토그램과 같이 정지영상의 전체적인 특징을 기술하는 전역 특징 기술자와 각 기술자 내의 요소별 중요도를 기술하는 요소 가중치로 구성된다. 지역 특징 기술 구조는 지역 대표 칼라와 같이 정지영상의 각 부분별 특징을 기술하는 지역 특징 기술자와 정지영상의 어느 부분이 중요한지를 기술하는 위치 가중치로 구성된다. 신뢰도 기술자는 그림 2와 같이 각 권한 등급의 레벨별로 부여된 학습 영향 정도 (e_k)와 각 권한 등급별로 학습에 참여한 피드백 수로 구성된다. 이러한 신뢰도 정보를 이용하여 현재 가중치 값의

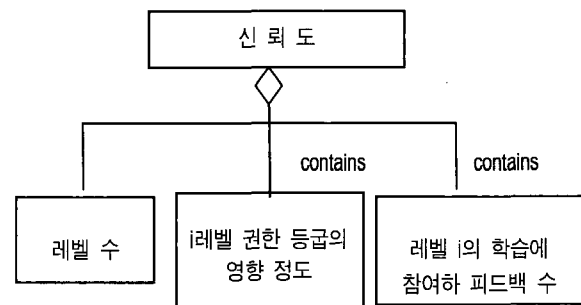


그림 2. 신뢰도 기술자
Fig. 2. Description Scheme for Reliability

신뢰도(Reliability)값을 구하는 방법은 식 1과 같다. 식 1에서 신뢰도 값은 권한 등급과 피드백 수에 비례하므로, 보다 높은 권한 등급을 가진 사용자의 피드백이 주어질수록 가중치에 대한 신뢰도 값은 높아진다.

$$Reliability = \sum_{k=1}^n e_k N_k$$

n : the number of authority levels
 e_k : effect value of authority level k
 ($e_1 > e_2 > \dots > e_{n-1} > e_n$)
 N_k : the number of use by authority level k

(1)

3.2. 사용자 피드백을 통한 가중치 학습

사용자가 피드백으로서 유사 정지영상, 또는 비유사 정지영상에 대한 정보를 주면 시스템은 다음 식 2를 사용하여 상대적인 가중치(w_c)를 구하게 된다. 기술자간 가중치의 경우 각각의 기술자를 사용하여 참조 정지영상과 피드

백된 정지영상의 유사도를 계산하여, 만일 피드백된 정지영상이 유사 정지영상이면 유사도가 클수록 높은 가중치를 할당하고 반대로 비유사 정지영상이면 유사도가 작을수록 높은 가중치를 할당한다. 요소 가중치와 위치 가중치 또한 비슷한 방법으로 구할 수 있다.

$$W_c = \alpha \text{ Sim}(RI, FI), \text{ if}(FI \text{ is the relevant image})$$

$$W_c = \alpha \text{ Dist}(RI, FI), \text{ else if}(FI \text{ is the irrelevant image})$$

$$W_c = 0, \text{ otherwise}$$

W_c is one of type weight(W_k), element weight(W_e), or position weight(W_p)

α is one of α_k, α_e or α_p which is the normalization coefficient for W_k, W_e, W_p respectively. (2)

$\text{Sim}(RI, FI)$: similarity between RI, FI using Feature k element e , or the information at position p

$\text{Dist}(RI, FI)$: distance between RI, FI using feature k element e , or the information at position p

RI : reference image

FI : feedback image

식 (2)에서 구해진 상대적인 가중치(W_c)를 사용하여 기존 가중치를 갱신하는 방법은 식 (3)을 사용한다. 새로 갱신된 가중치(W_{new})는 기존의 가중치와 피드백에 의해 구해진 상대적 가중치의 조합으로 구성되는데, 식 (3)에서 알 수 있듯이 기존의 가중치는 신뢰도(R_i)에 비례하여 더 큰 비중을 갖게 되며, 상대적 가중치는 피드백을 준 사용자의 권한 등급에 따라 부여된 학습 영향 정도(e_c)에 비례하여 더 큰 비중을 갖는다. 따라서 신뢰도가 클수록 기존 가중치는 덜 영향을 받으며, 사용자의 권한 등급이 클수록 피드백에 대한 갱신 정도는 커지게 된다.

$$W_{\text{new}} = \frac{R_i}{R_i + e_c} W_i + \frac{e_c}{R_i + e_c} W_c$$

$$0 < e_c < 1$$

W_i : weights currently stored in an image

W_c : weights calculated from feedback

W_{new} : updated weights of an image

R_i : current Reliability of image

e_c : effect value of a user giving relevance feedback

(3)

IV. 정지영상 검색 시스템(ArtShot) 및 기술자

우리는 정지영상 검색 성능 테스트를 위하여 자체 개발

한 전자 앨범인 Artshot을 사용하였다. Artshot은 자체 개발한 검색 엔진을 포함하고 있으며 사용자 피드백을 통한 가중치 학습 및 관리 기능을 제공한다.

1. 정지영상 특징 정보

정지영상 검색을 위해 본 시스템에서는 다음과 같이 4개의 기술자를 사용한다.

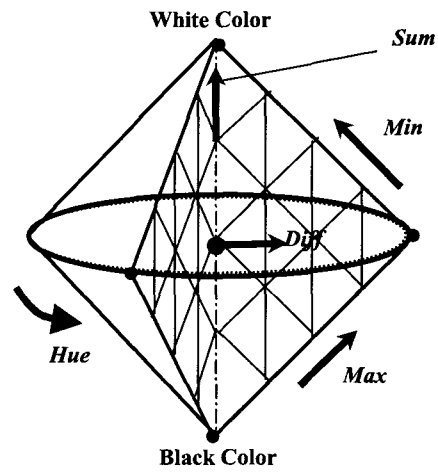
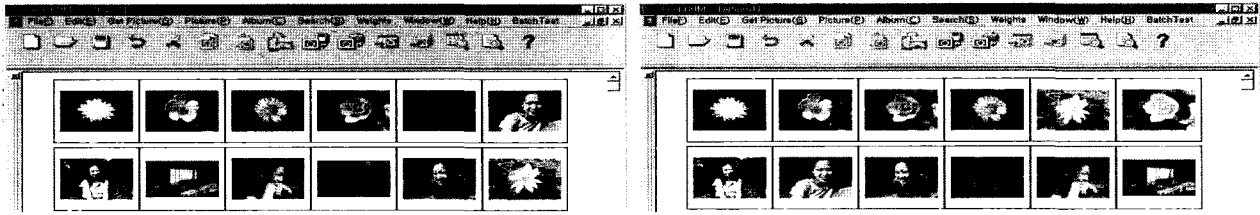


그림 3. HMMD 색공간 구조
Fig. 3. HMMD Color Space

- ▶ 칼라 히스토그램 : 정지영상의 각 픽셀을 HMMD 색공간(color space)^[4]에서 표현한 후 이를 52레벨로 양자화하여 칼라 히스토그램으로 표현하였다. HMMD 색공간은 그림 3에서 보는 바와 같이 두개의 원추가 서로 상하 방향을 달리하여 겹쳐진 모양이며 이 색공간으로부터 인간의 색인지에 쉽게 연관된 Hue(색상), Max(쉐이드), Min(틴트), Diff(순도) 그리고 Sum(밝기)이라 칭하는 5개의 파라미터를 쉽게 얻을 수 있다. 우리는 이들 중 Hue, Max, Min 등 3개의 파라미터를 기준으로 색공간을 분할하였다.
- ▶ 텍스처 히스토그램 : 수평, 수직, 두 대각선 방향의 4방향 성분들을 나타내는 4개의 33 마스크를 이용하여 각 방향 성분의 크기를 추출 후, 크기에 따라 일정 임계치를 기준으로 두 개의 경우로 분할하여 모두 16 가지의 텍스처 성분 중 하나로 추출하였다. 마스크는 정지영상 전체를 대상으로 이동하면서 텍스처 성분을 추출하였으며, 가로와 세로, 각 1 픽셀씩 오버랩하여 이동하였다. 이와 같이 추출된 텍스



(a) 가중치를 사용하기 전의 검색 결과 (b) 학습된 가중치를 사용했을 때의 검색 결과
 (b) The results of image retrieval without weights (b) The result of image retrieval with weights

그림 4. 학습된 가중치를 사용한 검색의 예
 Fig. 4. The examples of the retrieval with weights

쳐 성분들의 분포를 히스토그램으로 표현하여 16레벨의 빈으로 구성된 텍스처 히스토그램을 구성하였다.

- ▶ 지역 대표 칼라 : 정지영상을 8×8개의 부분 영역으로 분할한 후, 각 부분 영역에서 나타나는 대표칼라를 앞서 기술한 52레벨로 양자화된HMMD 색공간에서 표현하였다. 대표 칼라는 각 부분 영역에서 지역 칼라 히스토그램을 구한 가장 큰 분포 값을 갖는 칼라로 정해지, 일정 값에 미치지 못하거나 뚜렷한 대표 칼라가 존재하지 않을 경우는 미지정(Don't care)로 지정하여 그러한 영역은 매칭 시에 제외하였다.
- ▶ 지역 대표 텍스처 : 정지영상을 88개의 부분 영역으로 분할한 후, 각 부분 영역에서 나타나는 대표 방향 성분 특성을 앞서 기술한 16레벨의 텍스처 성분으로 기술하였다. 대표 텍스처의 선택은 지역 대표 칼라의 선택 방법과 동일하다.

V. 실험

1. 실험 환경

우리는 본 정지영상 검색 방법의 테스트를 위해 1,250개의 정지영상으로 구성된 테스트 세트와 1,250개의 정지영상으로 구성된 검증 세트를 준비했다. 테스트 세트는 각 참조 정지영상의 가중치를 학습하는데 사용되었고 검증 세트는 검색 테스트만을 위해 사용되었다.

2. 실험 결과

그림 4는 학습된 가중치를 사용하여 검색 성능이 증가한 예를 보여준다. 그림 4 (a)는 학습되기 전 초기 가중치를 사용하였을 때의 결과를 나타내는데, 그림에서 알 수

있듯이 참조 정지영상의 배경 영역이 꽃 영역보다도 넓으므로 배경이 상이한 정지영상은 같은 꽃 정지영상이라도 유사하지 않은 정지영상으로 계산된다. 하지만 그림 4 (b)와 같이 학습된 가중치를 사용하였을 경우 배경이 상이한 꽃 정지영상도 매우 유사한 정지영상으로 검색되었음을 알 수 있다. 그림 5는 이와 같은 결과를 가능하게 한 가중치 값을 나타낸다. 그림에서 바둑판 모양으로 이루어진 두 개의 영역은 지역 대표 칼라와 지역 텍스처 칼라의 위치 가중치를 나타내는데 그 색이 노란색에 가까울수록 높은 가중치 값을 의미한다. 이와 같이 사용자 피드백에 의해 자동으로 꽃의 영역이 배경 영역보다 중요한 영역임을 학습할 수 있다.

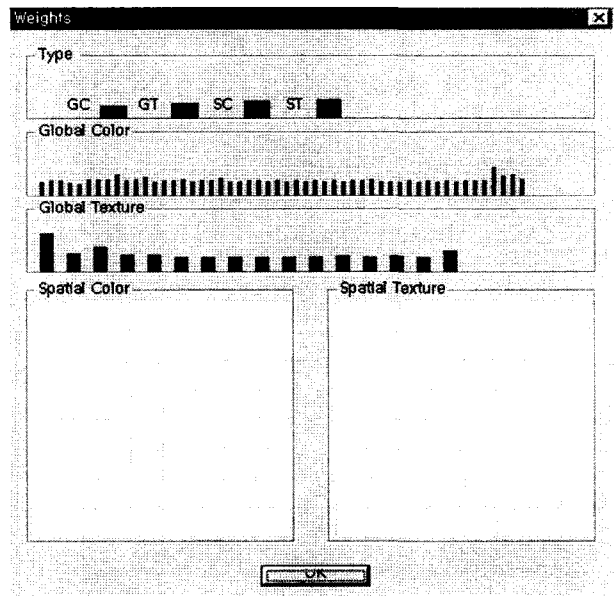


그림 5. 학습된 3가지 유형의 가중치
 Fig. 5. 3-types of weights which have been learned by relevance feedback

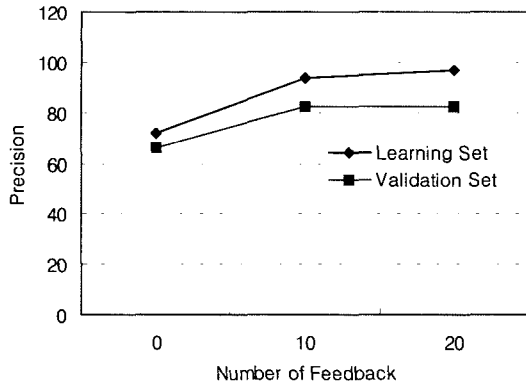


그림 6. 학습된 가중치를 사용한 방법의 성능 비교
Fig. 6.

그림 6은 전체적인 성능 결과를 도표로 보여준다. 그림에서 약 10번의 피드백 이후 학습된 가중치를 사용하였을 때 약 22% 정도의 검색 성능이 증가되었음을 알 수 있다.

VI. 결론

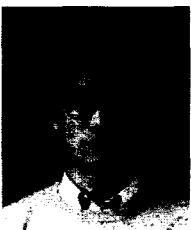
본 논문에서는 사용자 피드백을 이용하여 각 정지영상 특성을 기술하는 가중치를 학습함으로써 보다 높은 검색 성능을 가능하게 하는 방법을 제시하였다. 또한 이와 같이 가중치를 사용한 정지영상 검색에 적합하도록 전역 특징 정보, 지역 특징 정보, 그리고 3가지 유형의 가중치로 구성된 정지영상 기술 구조를 제시하였다. 본 논문에서는 사용자의 잘못된 학습을 방지하기 위하여 권한 등급과 가중치의 신뢰도 정보를 이용한 효과적인 학습 방법을 사용하

였으며 자체 개발한 전자 앨범에 적용하여 그 성능의 우수성을 검증하였다.

참고 문헌

- [1] Myron Flickener et al., "Query by Image and Video Content," *The QBIC System, IEEE Computer*, vol. 28, no. 9, pp. 23-32, 1995.
- [2] Yong Rui, Thomas S. Huang, Michael Ortega, Sharad Mehrotra, and Relevance Feedback: "A Power Tool for Interactive Content-Based Image Retrieval," *IEEE Trans. On Circuit and Video Technology*, Sep. 1998.
- [3] Ana B. Benitez, Mandis Beigi, and Shih-Fu Chang, "Using Relevance Feedback in Content Based Image Metasearch," *IEEE Internet Computing*, pp. 59-69, Jul.-Aug. 1998.
- [4] Jin-Soo Lee, Sung-Bae Jun, and Hyeon-Jun Kim., "Color Quantization for Unconstrained Images, Proceedings of International Conference On Information System," *Analysis and Synthesis*, vol. 2, pp. 537-542, Jul. 1998.
- [5] Harmon Paul, and Sawyer Brian, "UML for Visual Basic 6.0 Developer" *Using Visual Modeler and Rational Rose 98*, Nov. 1998

저자 소개



이진수

1995년 : 동국대학교 전자계산학과 학사
1997년 : 포항공과대학교 전자계산학과 석사
현재 : LG 전자기술원 연구원
주관심분야 : 패턴인식, 컴퓨터 비전



김 현 준

1988년 : 연세대학교 금속공학과 학사
1991년 : Stevens Institute of Technology 석사
1996년 : Stevens Institute of Technology 대학원 박사
현재 : LG 전자기술원 책임연구원
주관심분야 : 컴퓨터 비전, 멀티미디어 콘텐츠 내용기반 검색



윤 경 로

1987년 : 연세대학교 전자공학 학사
1989년 : 미시간 대학 전기공학 석사
1999년 : 시라큐스대학 전산학 박사
현재 : LG 전자기술원 선임연구원
주관심분야 : 멀티미디어 정보검색, 멀티미디어 데이터베이스



이 희 연

1979년 : 서울대 전기공학과 학사
1981년 : 한국과학원 산업전자과 석사
1993년 : Univ. of Southern California 전기공학 박사
현재 : LG 전자기술원 연구위원
주관심분야 : 인공지능, 신호처리