

# 칼라영상의 감성평가와 이를 이용한 내용기반 영상검색

박 중 수<sup>†</sup> · 엄 경 배<sup>††</sup> · 신 경 해<sup>†††</sup> · 이 준 환<sup>††††</sup> · 박 동 선<sup>††††</sup>

## 요 약

내용기반 영상검색에서 질의는 영상 그 자체이며 질의와 유사한 영상을 찾는 방식으로 검색이 진행된다. 즉 사용자가 검색을 원하는 영상의 색, 형태, 질감 또는 이들의 공간적인 배치 등의 내용을 염두에 두고 있어야 검색이 가능하다. 이러한 검색방법은 사용자가 검색대상의 내용을 상당 부분 파악하고 있어야 검색이 가능하다는 제약을 수반한다. 본 논문에서는 사용자가 영상이 제공하는 감성을 이용하여 칼라영상을 검색하는 방법을 제안한다. 제안된 방법은 검색과 감성평가가 relevance feedback을 통하여 동시에 진행됨으로 기존의 내용기반 영상검색과 통합이 용이하며, 사용자의 주관적인 평가가 이루어질 수 있다는 점에서 과거의 감성기반 칼라영상 검색과 차이가 있다. 본 논문에서 제안한 검색기의 평가를 위해 영상속성으로 MPEG-7의 칼라 기술자(descriptor)를 사용하였으며 “깨끗한”, “밝은”, “재미있는”, “포근한” 등의 감성형용사를 적용한 결과 1500개의 벽지영상을 대상으로 만족할 만한 결과를 얻을 수 있었다.

## Emotion from Color images and Its Application to Content-based Image Retrievals

JoongSoo Park<sup>†</sup> · Kyoung-Bae Eum<sup>††</sup> · Kyung-Hae Shin<sup>†††</sup>  
Joon-Whoan Lee<sup>††††</sup> · Dong-Sun Park<sup>††††</sup>

## ABSTRACT

In content-based image retrieval, the query is an image itself and the retrieval process is the process that seeking the similar images to the given query image. In this way of retrieval, the user has to know the basic physical features of target images that he wants to retrieve. But it has some restriction because to retrieve the target image he has to know the basic physical feature space such as color, texture, shape and spatial relationship. In this paper, we propose an emotion-based retrieval system. It uses the emotion that color images have. It is different from past emotion-based image retrieval in point of view that it uses relevance feedback to estimate the users intend and it is easily combined with past content-based image retrieval system. To test the performance of our proposed system, we use MPEG-7 color descriptor and emotion language such as “warm”, “clean”, “bright” and “delight”. We test about 1500 wallpaper images and get successful result.

**키워드:** 영상검색(Retrieval), 감성(Emotion), MPEG-7 칼라 기술자(MPEG-7 Color Descriptor), relevance 피드백(Relevance Feedback)

### 1. 서 론

인터넷을 통한 멀티미디어의 유통이 매우 증가된 시점에서 광범위하게 분포되어 있는 수많은 영상들의 효율적인 분석과 검색이 필요하게 되었다. 기존의 텍스트 기반 영상검색과 달리 내용기반 영상검색에서는 영상의 색, 형태, 질감 또는 이들의 공간적인 배치 등을 이용하여 질의영상과 비

슷한 영상을 검색한다[1-3]. 최근에는 내용기반 멀티미디어 데이터 검색을 목표로한 MPEG-7의 표준화가 진행되었다. MPEG-7의 칼라, 형태, 질감 등의 표준 기술자(descriptor) 들은 검색을 위한 영상의 기초적인 속성과 표현의 표준화를 제공하고 있다[4-7].

텍스트기반 영상검색과 달리 내용기반 영상검색은 영상의 다양한 물리적 속성들을 검색에 사용함으로써 검색의 정확도를 높이고 있다. 그러나, 내용기반 영상검색은 사용자가 검색대상과 유사한 영상을 가지고 있거나 또는 사용자가 검색대상의 칼라, 형태, 질감 또는 그들의 위치관계 등의 기초속성을 염두에 두고 있어야 가능하다. 이는 사용

\* 본 논문은 부분적으로 한국학술진흥재단 협동과제(2000-042-E00045)의 지원을 받음.  
† 정 회 원 : 우석대학교 전산정보학부 교수  
†† 정 회 원 : 군산대학교 컴퓨터과학과 교수  
††† 준 회 원 : 전북대학교 대학원 영상공학과  
†††† 정 회 원 : 전북대학교 전자정보공학부 교수  
논문접수 : 2002년 9월 26일, 심사완료 : 2003년 3월 19일

자에게 다양한 검색방법을 제공하지 못하는 제약이 있다.

본 논문에서는 사용자가 “깨끗한”, “밝은”, “재미있는”, “포근한” 등의 감성형용사를 사용하여 직접 원하는 영상을 검색하는 감성기반 영상검색 방법을 제안한다. 기존의 감성기반 영상검색 방법은 칼라영상에서 대표 칼라, 질감 등의 물리적 속성들을 추출하고 이를 다중회귀분석, 신경회로망, 적응 퍼지논리 시스템 등의 방법을 이용하여 감성속성 공간으로 변환하고 있다[8-12]. 감성속성 공간에서는 다양한 감성형용사들을 그들의 어의를 분석하여 주된 형용사들로 변환하는 변환기를 통하여 주된 감성형용사의 색인을 구성하고 있다. 구성된 색인들은 접속자, quantifier 등을 사용하여 검색키들의 조합을 사용하여 검색하고 있다. 이 방법은 물리적 속성공간을 다양한 감성형용사를 가진 감성공간으로 변환하는데는 많은 어려움이 있고 형용사의 어의를 분석하여 주된 형용사를 추출하는 데도 어려움이 있다. 이러한 방법들은 검색하려는 대상에 따라 서로 다른 방법을 적용해야 하는 적응적이지 못한 단점이 있다. 또, 감성언어가 하나의 독립적인 속성으로 검색에 사용되므로 기존의 내용기반 영상검색 방법과 통합이 어렵다.

본 논문에서 제안한 방법은 검색대상의 주요 감성형용사와 이를 대표할 수 있는 칼라영상을 미리 저장하여 사용하도록 하며 사용자가 대표영상을 선택하고 평가하면 검색기는 대표영상의 MPEG-7 칼라 기술자를 이용하여 유사영상을 검색하도록 한다. 또, relevance feedback[13-15]을 사용하여 사용자의 주관적인 평가가 검색에 반영되고 이를 바탕으로 필터링이 가능하도록 하여 사용자가 원하는 칼라영상을 검색할 수 있는 방법을 제안한다. 제안된 방법은 감성에 기반한 영상검색 방법이면서 기존의 내용기반 영상검색 시스템과 통합이 용이하므로 다양한 검색대상에 응용할 수 있는 장점이 있다. 또, 사용자의 주관적인 감성이 relevance feedback 과정을 통하여 검색기에 반영되므로 사용자에게 따른 주관적인 감성차이를 극복할 수 있다. Feedback 과정에서 유전자 알고리즘을 적용하여 사용자와 시스템이 상호작용 하려는 시도가 있으나 유전자 알고리즘의 최적화 수렴 속도가 느리므로 본 논문에서 제안된 방법이 검색속도 면에서 보다 우수하다[16, 17].

제안된 검색방법의 실험을 위하여 1500개의 벽지영상을 사용하였다. 검색을 위해서 “깨끗한”, “밝은”, “재미있는”, “포근한” 등의 감성형용사를 질의로 사용하였다. 실험 결과 만족할 만한 검색결과를 제공한다.

본 논문의 2절에서는 MPEG-7 칼라 기술자 및 relevance feedback에 대하여 논의하였고 3절에서는 실험 및 검토에 대하여 논의하였다. 그리고 4절에서는 결론 및 향후 연구계획에 대하여 언급하였다.

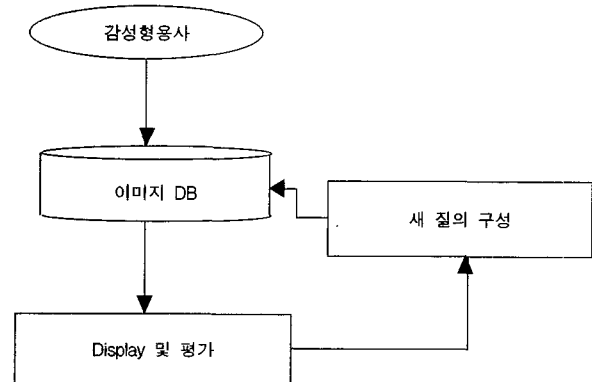
## 2. 시스템 구성

### 2.1 시스템 개요

#### 2.1.1 구성된 시스템의 개요

사용자로부터 감성형용사를 입력받으면 시스템은 입력받은 감성형용사에 대응하는 대표영상 군을 미리 저장된 DB로부터 보여준다. 사용자가 시스템이 보여준 대표영상 군을 평가하면 시스템은 이 평가자료를 근거로 새로운 질의를 구성하고 각 기술자의 가중치를 조정하며 이를 이용하여 새로운 후보영상 군을 검색하여 보여준다. (그림 1)은 시스템의 대략적인 구성을 보여주고 있다.

실험에 사용될 영상들은 자신이 속한 범주에서 여러 사용자들이 자주 사용하는 감성형용사에 따라 분류한다. 분류된 영상들은 다시 각 감성형용사에 대한 해당 정도를 평가하여 그 중에서 대표영상을 추출하고 이를 DB에 저장한다. 감성형용사를 질의로 사용하는 1차 검색에서는 이 DB에 저장된 대표영상 군이 사용자에게 추천된다.

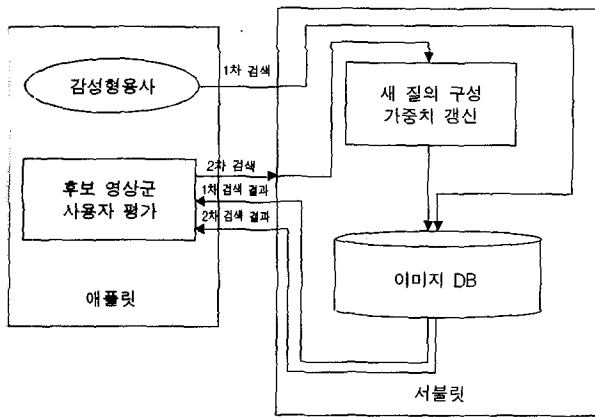


(그림 1) 시스템 구성

추천된 영상들은 사용자로부터 평가를 받게되며 이 평가자료는 시스템에 전달된다. 시스템에 전달되는 평가자료는 평가된 영상의 색인과 그 영상에 대한 사용자의 평가값으로 구성한다. 시스템은 전달된 평가자료를 근거로 새로운 질의를 구성하게 되는데 질의를 구성하는 방법은 Rocchio 방정식[13, 18], mind reader[19], optimal query[20] 등 여러가지가 있을 수 있다. 질의는 MPEG-7에서 제공하는 칼라 기술자를 사용하여 표현한다. 본 논문에서 사용한 질의는 16bin의 scalable 칼라 기술자와 32bin의 칼라 구조 기술자이다. 시스템은 평가자료를 근거로 사용자의 의도를 파악하게 된다. 이는 사용자에게 따라 같은 감성형용사라 할지라도 서로 다른 감성을 가질 수 있으므로 이를 반영하기 위해서이다. 이를 위해서 시스템은 사용자가 평가한 자료를 토대로 어떤 기술자가 사용자로부터 일관되게 좋은 평가를 받는지를 분석하며 또 기술자들의 집합 중에서 어떤 기술자들의 집합

이 사용자로부터 좋은 평가를 받았는지를 판단하여 가중치를 갱신한다. 시스템은 구성된 새 질의와 각 기술자들의 가중치를 고려하여 질의 영상과 DB에 있는 영상들과의 유사도를 측정하여 후보영상을 검색하여 사용자에게 보여주게 된다.

제안한 시스템은 두 부분으로 구성되어 있다. 사용자에게 후보영상 군을 보여주고 평가를 받는 애플릿 부분과 사용자의 평가자료를 입력받아 이를 분석하여 사용자의 의도를 파악하고 후보영상 군을 영상 DB에서 검색하기 위한 서블릿 부분으로 구성한다. (그림 2)는 시스템 내부에서 검색이 이루어지는 과정을 그림으로 나타낸 것이다.



(그림 2) 시스템 내부 동작

2.1.2 기존 감성기반 시스템과의 차이

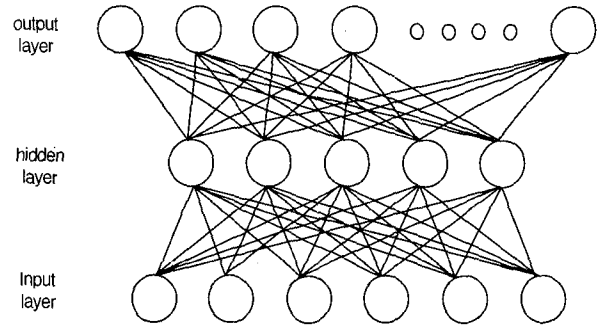
기존의 감성기반 영상검색 방법중 Soen[9]이 시행한 다중 회귀분석 방법은 물리적 속성으로부터 감성속성으로의 변환을 위하여 다음 식 (1)에서

$$\begin{aligned}
 P_i = & c_{i1} + c_{i2}u^* + c_{i3}v^* + c_{i4}L^* + c_{i5}u^{*2} + c_{i6}v^{*2} \\
 & + c_{i7}v^*c_{i8}u^{*3} + c_{i9}u^{*2}v^* + c_{i10}u^*v^{*2} \\
 & + c_{i11}v^{*3} + c_{i12}D_L + c_{i13}D_M + c_{i14}D_H
 \end{aligned}
 \tag{1}$$

와 같은 비선형 방정식 모델을 사용하였다. 여기서,  $c_{ij}$ 는 물리적 속성에 대한 회귀방정식의 계수이고,  $P_i$ 는 13개의 감성 척도 중에서  $i$ 번째 값이다. 14개의 회귀방정식의 계수들은 피실험자들의 평가치와의 에러를 최소화시키는 값이다.

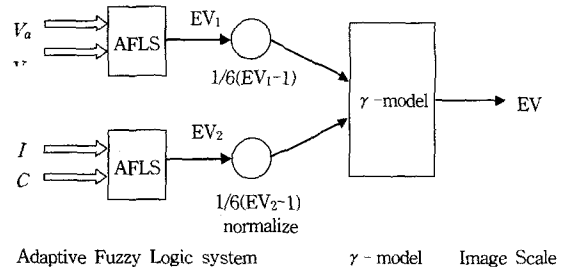
신경망을 이용한 감성평가 시스템[11, 12]은 다중회귀분석과 마찬가지로 6개의 물리적 속성을 입력하고 13개의 감성 척도를 출력하도록 구성되었다(그림 2).

이 방법은 Soen의 방법이 고정된 비선형 방정식을 사용하는데 반하여 입력된 물리적 속성이 은닉층(hidden layer)을 거치면서 다양하게 서로 비선형 조합되어 Soen의 방정식보다 다양한 물리적 특성이 서로 결합하여 더 좋은 결과를 얻을 수 있다.



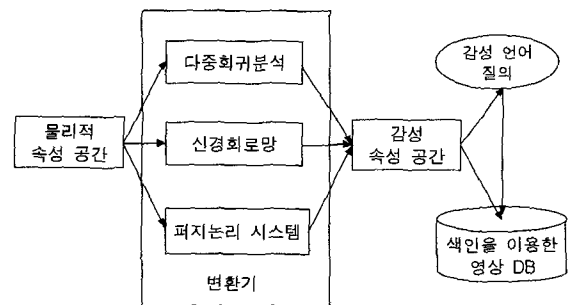
(그림 3) 감성평가를 위한 신경망 구조

한편, 영상의 물리적 속성값에 따라 퍼지 IF-THEN 규칙을 적용한 퍼지논리 시스템을 구성하고 여기에 다시  $\gamma$ -모델을 적용하여 감성속성으로 변환할 수 있다[10]. 퍼지논리 시스템과  $\gamma$ -모델의 혼합 시스템은 다음 (그림 3)과 같다.



(그림 4) 퍼지논리 시스템과  $\gamma$ -모델 혼합 시스템

기존의 감성기반 영상검색 방법은 물리적 속성공간을 감성속성 공간으로의 변환하기 위하여 변환기를 사용한다. 이렇게 변환된 감성속성들은 계층적 근접화를 이용한 트리구조의 색인이나 R-tree을 이용하여 색인을 한 후 이를 DB에 저장한다. 또는, 퍼지논리 시스템에서는 질의어들의 조합을 사용하여 질의를 구성하고 이를 영상 데이터베이스에서 검색한다. 기존의 감성기반 영상검색이 감성 질의어를 이용하여 검색하는 과정을 그림으로 나타내면 다음 (그림 5)와 같다.



(그림 5) 영상의 검색과정

본 논문에서 제안한 방법은 기존에 방법보다 복잡한 변환기의 구성이 필요 없고 영상 DB를 감성속성 공간에서

감성언어의 색인으로 저장할 필요가 없다.

2.2 MPEG-7 칼라 기술자 및 유사성 척도

심리학적인 실험에 의해 인간의 감성에 영향을 주는 보편적인 영상요소로는 칼라와 텍스처라는 사실이 Soen 등에 의해 입증된바 있다[9]. 그들은 영상의 전역적인 칼라속성과 공간주파수 성분 및 밝기 등이 총체적으로 인간의 감성에 영향을 주고, 가장 영향을 많이 미치는 요인은 밝기, 전체적인 칼라, 공간 주파수로 표현되는 칼라 성분임을 실험 데이터와 인자분석(factor analysis)을 통하여 확인하였다. 본 시스템을 구성하는데 있어서 영상의 전체적인 칼라와 밝기 정보는 MPEG-7의 scalable 칼라 기술자(SCD)로 대표될 수 있으며, 칼라의 변화에 해당되는 주파수 성분은 칼라 구조 기술자(CSD)를 통하여 포착될 수 있다는 가정 하에서 이들 두 기술자를 사용하였다.

2.2.1 Scalable 칼라 기술자(SCD)

SCD 칼라 기술자는 HSV 칼라 공간을 256bin으로 균일 양자화(uniform quantization) 한 후 이를 비선형 양자화(non-linear quantization)와 하아변환(Haar transform)을 거쳐 256 bin의 계수를 구한다. 원하는 기술자의 수준에 따라 이를 다시 선형양자화(linear quantization)하여 사용한다[4-6]. 본 논문에서는 256bin의 계수를 16bin의 계수로 양자화하여 사용하였다. 거리척도는 두 영상 A와 B에 대하여

$$\text{dist}(A, B) = \sum_{i=0}^n |h_A(i) - h_B(i)| \quad (2)$$

이다. 여기서,  $n$ 은 총 계수의 수이며,  $h_A(i)$ 는 A영상의  $i$  번째 계수의 값이다.

2.2.2 칼라 구조 기술자(CSD)

CSD는 주어진 영상의 국소적인 구조를 표현하기 위하여

사용한다. HMMD 공간의 Hue, Sum, Diff 요소를 사용하며 이를 비균일(non-uniform) 양자화 한다. 사용자가 원하는 정확도에 따라 bin의 수를 조정할 수 있다[4-6]. 본 논문에서는 32bin을 사용하였다. CSD 계수의 추출 방법은 주어진 영상 위에 structuring element를 sub-sampling factor만큼 씩 왼쪽에서 오른쪽으로 위에서 아래로 이동시키면서 각 bin에 속하는 color가 있는지 관찰하여 있으면 그 bin의 값을 1씩 증가시킨다. Structuring element의 크기와 sub-sampling factor는 보통 8\*8과 1을 사용하지만 영상의 크기에 따라 변한다.  $E$ 를 structuring element의 너비(혹은 높이)라고 하고  $K$ 를 sub-sampling factor라면  $K$ 와  $E$ 는 아래의 식 (3)

$$p = \max(0, \text{round}(0.5 * \log_2 WH - 8)) \quad (3)$$

$$K = 2^p, E = 8 * K$$

에 의해서 정해진다. 여기서  $W$ 와  $H$ 는 원 영상의 너비와 높이이다. 추출하고자 하는 영상들의 크기가 다를 수 있으므로 structuring element의 총 개수로 정규화 한다.

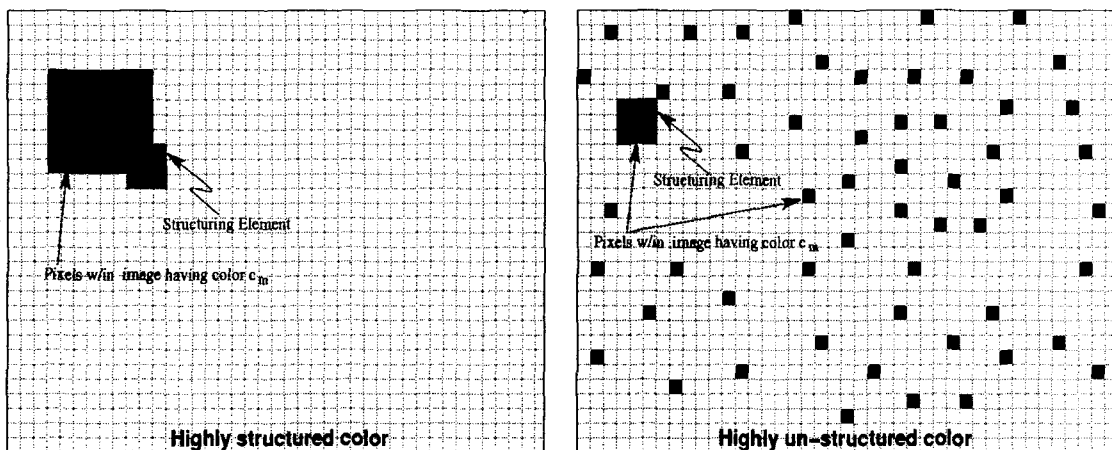
Structuring element는 원 영상의 각 칼라 bin에 대한 국소적 구조를 반영할 수 있도록 하였다. (그림 6)에서 두 영상의 SCD 히스토그램 값이 같으므로 SCD로는 두 영상을 구별할 수 없다. 그러나 structuring element를 이용하여 각 칼라의 히스토그램을 구하면 두 영상의 CSD 값은 아주 다르다.

주어진 두 영상 A와 B에서 두 영상 사이의 거리는 SCD와 같이 L1 노름을 사용하였다.

2.3 Relevance feedback 및 질의 갱신

2.3.1 Relevance feedback

기존의 내용기반 영상검색은 질의 영상과 데이터베이스 영상의 물리적 속성에 근거하여 검색이 이루어진다. 그러나, 같은 물리적 속성을 갖는 영상이라 하더라도 사용자에



(그림 6) CSD 값의 비교

따라 느끼는 감성이 다를 수 있으며 더 큰 문제는 사용자의 감성과 물리적 속성을 연결시키는 적당한 함수를 개발하기 매우 어렵다는 것이다. 본 논문에서 적용한 relevance feedback 알고리즘[13-15]은 사용자의 평가자료를 근거로 사용자의 의도를 파악하기 매우 유리하다. 사용자가 평가한 자료는 실시간으로 시스템에 전달되어 사용자의 감성에 따른 새로운 질의를 구성할 수 있게 하고 사용자 사이의 개인차를 보상할 수 있게 한다. Feedback 과정에서 검색과 평가가 동시에 이루어지므로 검색시간을 단축할 수 있다. 본 논문에서 적용한 relevance feedback과 이를 이용한 검색과정은 다음과 같다.

- [1 단계] 모든 영상을 기술자를 성분으로 갖는 벡터모델로 표현한다.
- [2 단계] 각 기술자의 가중치를 초기화한다.
- [3 단계] 대표영상 군을 평가한다.
- [4 단계] 새 질의를 구성한다.
- [5 단계] 유사성 척도를 이용하여 후보영상 군을 검색하여 사용자에게 보여준다.
- [6 단계] 후보영상 군을 평가한다.
- [7 단계] 각 기술자의 가중치를 갱신한다.
- [8 단계] [4 단계]로 간다.

영상  $I$ 의 SCD와 CSD를 사용한 벡터모델은 각각

$$\begin{aligned} I_1 &= [d_{1,1}, \dots, d_{1,16}], \\ I_2 &= [d_{2,1}, \dots, d_{2,32}] \end{aligned} \quad (4)$$

로 표현하였다. 각 기술자의 가중치에는 두 가지가 있다. 벡터 성분  $d_{ij}$ 의 가중치를  $W_{ij}$ 라하고 초기 가중치  $W_{ij}^0$ 를

$$W_{ij}^0 = \frac{1}{J_i} \quad (5)$$

로 한다. 여기서,  $J_1=16$  이고  $J_2=32$  이다. 다른 가중치는 SCD와 CSD에 대한 가중치로 각각  $W_1$ 과  $W_2$ 로 나타낸다. 초기 가중치는

$$W_i = \frac{1}{2}, \quad i = 1, 2 \quad (6)$$

로 한다. 사용자는 추천된 영상에 대하여 사용자의 의도에 적합한 정도에 따라 “매우 적합”, “적합”, “의견 없음”, “적합하지 않음”, “매우 적합하지 않음”으로 평가하고 평가값은 각각 0, 2, 3, 4, 6이다. 두 영상간의 유사성 척도는

$$\begin{aligned} S_i &= \sum_{j=1}^{J_i} W_{ij} * |d_{ij} - d'_{ij}|, \quad i=1, 2 \\ S &= \sum_{i=1}^2 W_i * S_i \end{aligned} \quad (7)$$

이다. 질의영상이 갱신되면 식 (7)의 유사성 척도  $S$ 를 이용하여 데이터베이스의 영상을 검색한다. [7 단계]에서 가중치를 갱신하기전에 각 기술자의 값의 범위가 다르고 또 SCD와 CSD의 척도가 다르므로 먼저 각 기술자의 값을 정규화할 필요가 있다. 전자에 해당하는 정규화를 내부적 정규화(intra-normalization)라 하고 후자에 해당하는 정규화를 외부적 정규화(inter-normalization)라 한다.

내부적 정규화는 데이터베이스에 있는 모든 영상을 벡터 모델로 표현했을 때 각 열에 있는 모든 성분의 범위를 일정하게 하는 과정이다. 예를 들어, 실험에 사용된 1500개 영상의 SCD 벡터모델은  $1500 \times 16$  행렬

$$\begin{aligned} [d_{1,1}^{(0)}, \dots, d_{1,16}^{(0)}] \\ \dots \\ [d_{1,1}^{(1499)}, \dots, d_{1,16}^{(1499)}] \end{aligned} \quad (8)$$

로 표현되고 이때 각 열의 값이  $[0, 1]$  사이에 있도록 정규화한다. 본 실험에서는  $j$ 열을 정규화 한다고 할 때  $j$ 열의 모든 성분 중에서 최대값을  $\max$  최소값을  $\min$ 이라 하면  $d_{ij}$ 의 정규화 값  $d'_{ij}$ 은

$$d'_{ij} = \frac{d_{ij} - \min}{\max - \min} \quad (9)$$

로 계산하였다. 외부적 정규화는 식 (7)에서 구한  $S_1$ 과  $S_2$  값이  $[0, 1]$  사이에 있도록 하는 과정이다. 다음은 외부적 정규화 과정이다. 각  $i$ 에 대하여

- ① 데이터베이스에 있는 영상 중에서 2개씩을 택하여 거리  $S_i(I_i, J_i)$ 를 계산한다.
- ②  $S_i$ 들의 평균  $\mu_i$ , 표준편차  $\sigma_i$ 를 계산한다.
- ③ 질의 영상  $q = [d'_{11}, \dots, d'_{i, J_i}]$ 가 주어지면,

$$S_i(d, q) = \sum_{j=1}^{J_i} |d_{ij} - d'_{ij}| \quad (10)$$

를 계산한다.

- ④  $3\sigma_i$ 를 이용하여 정규화

$$S_i(d, q) = \frac{S_i(d, q) - \mu_i}{3\sigma_i} \quad (11)$$

한다.

- ⑤ 거리는 0보다 크거나 같아야 하므로

$$S_i(d, q) = \frac{S_i(d, q) + 1}{2} \in [0, 1] \quad (12)$$

로 치환한다.

식 (12)에서 구한  $S_i(d, q)$ 가 질의  $q$ 와 임의의 데이터베이스에 있는 영상  $d$ 사이의 거리  $S_i$ 이다. 외부적 정규화는 feedback 단계마다 가중치  $W_{ij}$ 가 갱신되므로 식 (7)의  $S_i$ 는 매번 정규화해야 한다. 각  $i$ 에 대하여 가중치  $W_i$ 의 갱신을 위해서 다음 ①~④를 시행한다.

- ①  $RT = [RT_1, \dots, TR_M]$ 를 유사성 척도  $S$ 에 의하여 검색된  $M$ 개의 영상이라 하자. 또,  $RT^i = [RT_1^i, \dots, TR_M^i]$ 를 유사성 척도  $S_i$ 에 의하여 검색된  $M$ 개의 영상이라 하자.
- ②  $W_i = 0, i = 1, 2$ 로 초기화한다.
- ③  $l = 1, \dots, M$ 에 대하여

$$W_i = \begin{cases} W_i + Score_l, RT_l^i \in RT \text{ 이면} \\ W_i + 0, \text{if } RT_l^i \notin RT \text{ 이면} \end{cases} \quad (13)$$

$$Score_l = \begin{cases} 0, RT_l^i \text{이 } RT \text{에서 "매우 적합"과 대응하면} \\ 2, RT_l^i \text{이 } RT \text{에서 "적합"과 대응하면} \\ 3, RT_l^i \text{이 } RT \text{에서 "의견 없음"과 대응하면} \\ 4, RT_l^i \text{이 } RT \text{에서 "적합하지 않음"과 대응하면} \\ 6, RT_l^i \text{이 } RT \text{에서 "매우 적합하지 않음"과 대응하면} \end{cases}$$

④  $W_{Ti} = \sum_{i=1}^M W_i$  일때,  $W_i = \frac{W_i}{W_{Ti}}$ 이다. (14)

가중치  $W_{ij}$ 의 갱신을 위해서 각 기술자의 표준편차의 역수를 취한다. 즉, 가중치  $W_{ij}$ 는

$$W_{ij} = \frac{1}{\sigma_{ij}} \quad (15)$$

로 구한다. 여기서,  $\sigma_{ij}$ 는 식 (8)의  $j$ 열의 표준편차이다.

이렇게 구한 가중치  $W_i$ 와  $W_{ij}$ 는 각 기술자들의 중요도를 반영한다고 볼 수 있다. 초기  $\frac{1}{n}$ 이었던 가중치는 feedback이 진행됨에 따라 사용자의 의도에 부응하여 동적으로 변화한다. 또, 이 가중치들은 각 사용자들의 감성에 관한 정보를 가지고 있으므로 이를 누적하여 각 사용자 그룹에 따라 통계치를 구해 놓으면 다음번 이 가중치를 사용하여 초기 검색을 보다 효과적으로 할 수 있을 것이다.

### 2.3.2 질의 갱신

Relevance feedback에서 “매우 적합” 또는 “적합” 영상만 하나이면 그 영상을 질의(query)로 사용한다. 그러나, “매우 적합” 또는 “적합” 영상이 2개 이상이면 질의를 재구성할 필요가 있다. 영상 검색에서 질의는 검색성능에 많은 영향을 주므로 어떻게 질의를 구성하느냐는 문제는 매우 중요하다. 질의를 구성하는 방법은 여러 가지가 있을 수 있

으며 실험 결과 ③과 ④의 클러스터링 방법이 좋은 결과를 보였다. 다음은 본 실험에서 사용한 질의 구성 방법들이다.

- ① “매우 적합”과 “적합” 영상들의 평균치를 택한다.
- ② “매우 적합”과 “적합” 영상들의 가중치 평균을 택한다. 이때, “매우 적합”과 “적합” 영상의 가중치 비를 2 : 1로 하였다.
- ③ “매우 적합”과 “적합” 영상들을 SCD와 CSD에서 즉, 16차원과 32차원에서 클러스터링하여 각 클러스터의 중심을 택한다.
- ④ “매우 적합”과 “적합” 영상들을 SCD와 CSD에서 즉, 16차원과 32차원에서 클러스터링하여 각 클러스터의 중심을 택한다. 이때, 클러스터의 중심을 정할 때 “매우 적합”과 “적합” 영상의 가중치 비를 2 : 1로 하였다.

위의 방법 중에서 ①과 ②는 빠른 질의를 구성하는 장점이 있으나 검색 횟수가 증가할 수록 검색결과가 멍개지는 것을 볼 수 있다. 이는 질의가 계속하여 평균을 취하는 것에 대한 당연한 결과라 생각한다. 반면에 ③과 ④는 사용자의 의도를 충분히 반영할 수 있었다. ③과 ④의 성능차이는 없었으나 DB에 있는 영상의 개수가 현재보다 더 많아진다면 ③ 보다 ④의 방법이 우수할 것으로 보인다.

사용한 클러스터링 알고리즘은 GLA(Generalized Lloyd Algorithm)으로 최대 클러스터의 수까지 split한 후 이를 적당한 threshold로 merge를 한다[21]. 여기서 사용한 threshold 값은 각각의 bin에서 0.2이다. threshold로 0.2를 사용한 이유는 우리가 사용하는 feature값이 [0, 1] 사이로 normalize 되었으므로 두 feature 값이 similar하다는 것은 하나의 feature 값이 다른 feature 값의  $\pm 0.1$  이내에 있는 것으로 본 것이다.

## 3. 실험 및 검토

제안된 검색기의 실험을 위해서 인터넷에서 모은 1500개의 벽지영상을 사용하였다. Visual C++를 사용하여 이 영상들의 MPEG-7 칼라 기술자들을 추출하여 데이터베이스화하였다. 제안된 검색기에는 감성형용사로 “포근한”, “부드러운”, “밝은”, “자연스러운”, “편안한”, “시원한”, “깔끔한”, “화사한”, “품위 있는”을 사용하였다. 이는 시장조사를 통하여 사용자가 벽지를 선택할 때 자주 사용하는 감성형용사에서 대표적인 것을 선택한 것이다.

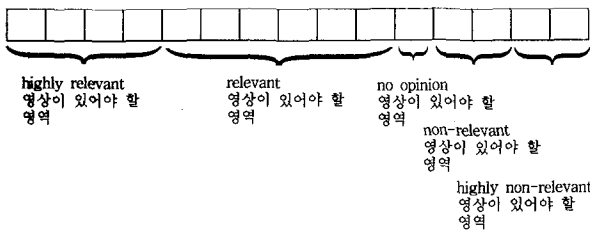
검색기는 애플릿 부분과 서블릿 부분으로 구성하였는데 애플릿 부분은 사용자와의 인터페이스 부분으로 사용자가 Web 상에서 검색기를 사용할 때 볼 수 있는 부분이다. 사용자는 애플릿에서 검색기가 추천한 대표영상 군을 “매우 적합하지 않음”, “적합하지 않음”, “의견 없음”, “적합”, “매우 적합”의 5단계로 평가한다. 애플릿에서 평가한 자료들은

서블릿 부분으로 전달되고 서블릿은 이 전달된 자료들을 분석을 하여 사용자의 의도를 파악하고 새 질의를 구성하고 이를 토대로 검색하고 결과를 애플릿으로 전달한다. (그림 8)은 검색기의 인터페이스이다.

애플릿과 서블릿을 자바로 구현하였다. 제안한 검색기의 성능을 평가하기 위해서 적합률(relevant rate) 척도와 벌점(penalty) 척도를 사용하였다. 적합률 척도는 디스플레이된 영상의 개수에 대한 “매우 적합”과 “적합” 영상의 개수의 비율이다.  $N$ 이 사용자의 평가에서 “매우 적합”과 “적합” 한 영상의 개수를 나타내고  $Sc$ 가 사용자에게 보여준 전체 이미지의 개수라 할 때 적합률  $Relev$  은

$$Relev = \frac{N}{Sc} \times 100\% \quad (16)$$

이다. 벌점 척도는 시스템이 사용자에게 검색결과를 보여줄 때 “매우 적합”, “적합”, “의견 없음”, “적합하지 않음”, “매우 적합하지 않음” 영상의 순서로 보여주므로 이 영상들이 제 위치에 있지 않을 때 벌점을 부과하는 방법이다. (그림 7)에서는 시스템이 15개의 영상후보를 추천한다고 가정하고 “매우 적합” 영상 4개, “적합” 영상 6개, “의견 없음” 영상 1개, “적합하지 않음” 영상 2개, “매우 적합하지 않음” 영상 2개인 경우를 예로 들어 설명한 것이다.



(그림 7) 벌점 척도

구분된 영상들이 제 위치에 있지 않으면 다음 식 (17)과 같이 벌점을 부과한다.

$$Pen = \sum_{hr \in HR} 1 + \sum_{r \in R} 1 + \sum_{no \in NO} 1 + \sum_{hnr \in HNR} 1 + \sum_{nr \in NR} 1 \quad (17)$$

여기서,  $hr$ ,  $r$ ,  $no$ ,  $nr$ ,  $hnr$ 는 각각 “매우 적합”, “적합”, “의견 없음”, “적합하지 않음”, “매우 적합하지 않음” 영상을 나타내며  $HR$ ,  $R$ ,  $NO$ ,  $NR$ ,  $HNR$ 은 각각 “매우 적합”, “적합”, “의견 없음”, “적합하지 않음”, “매우 적합하지 않음” 영상들의 위치를 나타내는 인덱스의 집합이다.

적합률 척도는 값이 클수록 벌점 척도는 값이 작을수록 좋은 검색결과를 나타낸다. 또, 벌점 척도는 검색된 영상의 순서를 반영하므로 보다 정확한 척도를 제공한다. 두 검색  $A$ 와  $B$ 에서  $Relev_A < Relev_B$ 이면 검색  $B$ 의 결과가 검색  $A$ 의 결과보다 우수하다. 또,  $Pen_A < Pen_B$ 이면 검색  $A$ 가 검색  $B$

의 결과 보다 우수하다. 그러나,  $Relev_A < Relev_B$ 이고  $Pen_A < Pen_B$  또는  $Relev_A > Relev_B$ 이고  $Pen_A > Pen_B$ 인 경우는 두 검색의 결과를 비교할 수 없다. 다음 <표 3>과 <표 4>는 사용된 9개의 감성형용사 중에서 7개의 감성형용사에 대한 검색 결과를 20명에게 조사하여 평균치를 취한 것이다.

<표 3> 가중 평균치를 이용한 결과

	1차 검색	2차 검색	3차 검색	4차 검색
	적합률	적합률	적합률	적합률
	벌 점	벌 점	벌 점	벌 점
부드러운	55.0	75.0	80.0	85.0
	13	14	8	7
밝은	55.0	60.0	75.0	80.0
	13	14	10	8
자연스러운	55.0	60.0	80.0	85.0
	13	14	8	7
시원한	70.0	75.0	75.0	80.0
	10	13	12	12
깔끔한	80.0	85.5	95.5	95.5
	9	9	7	4
재미있는	55.0	70.0	75.0	75.0
	13	9	7	7
품위 있는	65.0	65.0	75.0	80.0
	11	10	9	9

<표 4> 클러스터링과 가중 평균치를 이용한 결과

	1차 검색	2차 검색	3차 검색	4차 검색
	적합률	적합률	적합률	적합률
	벌 점	벌 점	벌 점	벌 점
부드러운	55.0	60.0	80.0	90.0
	13	14	8	6
밝은	55.0	60.0	80.0	85.0
	13	14	8	7
자연스러운	55.0	60.0	80.0	85.0
	13	14	8	7
시원한	70.0	75.0	80.0	80.0
	10	13	10	8
깔끔한	85.5	85.5	95.5	95.5
	9	9	7	4
재미있는	55.0	70.0	80.0	85.0
	13	9	7	6
품위 있는	65.0	65.0	75.0	80.0
	11	10	9	7

위의 검색 결과는 두 개의 척도에서 동일하게 가중 평균치를 이용한 방법보다 클러스터링을 이용한 방법이 좋은 결과를 나타내고 있다. 가중 평균치를 이용한 방법은 질의를 구성할 때 평균을 취하므로 feedback을 거듭할수록 검

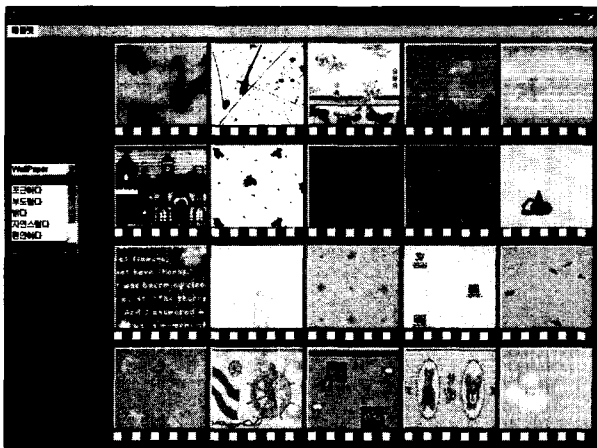
색결과가 공개지게 되므로 좋은 평가를 받지 못했다. 반대로 클러스터링을 이용한 방법은 사용자의 다양한 의도를 feedback 횟수에 관계없이 반영할 수 있는 장점이 있다. 두 검색 결과에서 동일하게 2차와 3차 검색에서 증가하는 비율이 매우 크다. 2차 검색이 relevance feedback을 처음 실시한 결과이며 이는 검색의 수렴속도가 매우 빠르다는 것을 나타낸다. (그림 8)은 클러스터링에 의하여 질의를 갱신할 때 “재미있는”에 대한 각 검색 라운드의 검색 결과이다.

4. 결 론

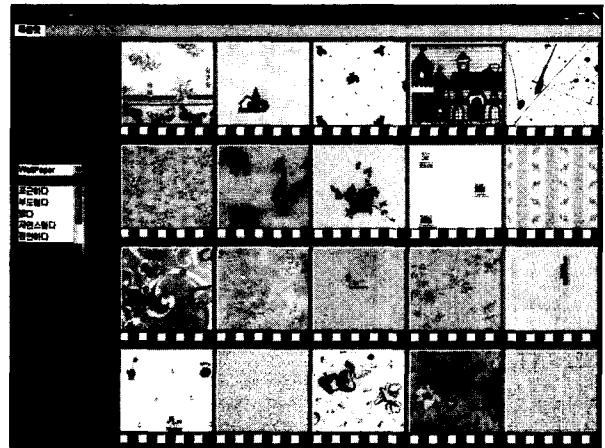
본 논문에서는 영상의 특징을 추출하기 위해서 MPEG-7의 칼라 기술자를 사용하였다. 사용자의 감성을 반영할 수 있도록 감성형용사를 사용한 내용기반 영상검색 방법을 제안하였다. 이는 기존의 감성기반 영상검색 방법이 물리적속성공간을 감성속성 공간으로 변환하여 검색하는 방식과 다르다. 제안된 검색 방법은 검색 대상의 감성 형용사들을 잘

선택한다면 복잡한 변환기의 사용 없이 relevance feedback을 통하여 사용자의 감성속성과 물리적 속성간의 변환을 빠르게 할 수 있게 하며 사용자간의 감성에 대한 개인 차이를 보상하여 검색할 수 있게 한다. 제안된 검색기는 자바로 구현하였으므로 Web 상에서 다양하게 활용될 수 있을 것으로 보인다. 예를 들어, 벽지 검색, 장판 검색, 섬유디자인 검색 등 다양한 분야에서 감성에 기반한 상품 검색기로서 사용이 가능하다고 생각된다. 또, 상품 검색과 함께 사용자들의 평가결과와 relevance feedback 시의 가중치를 축적하고 분석하여 각 사용자 그룹-성별, 나이, 직업, 교육정도, 지역 등에 따른 감성의 취향을 파악하여 상품 개발에 이용할 수 있을 것으로 보인다.

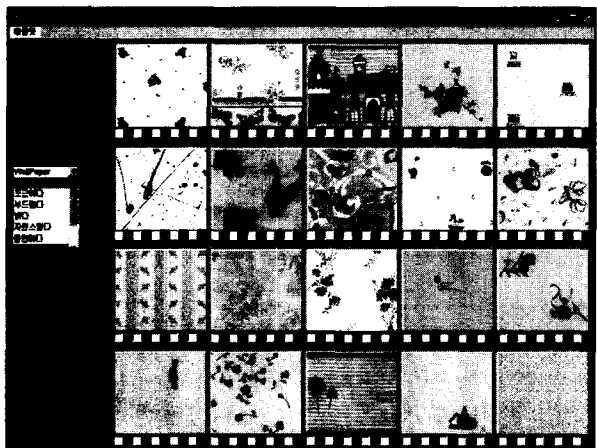
향후 연구계획은 제안된 시스템을 바탕으로 검색기의 사용 속성에 텍스트 속성 포함시켜 더욱 강건한 시스템을 구축하는 것이다. 또, 성별·직업별·연령별 등에 대한 사용자의 감성 DB를 구축하여 1차 검색에서 보다 사용자의 취향에 맞는 후보영상 군을 시스템이 예측하여 보여주도록



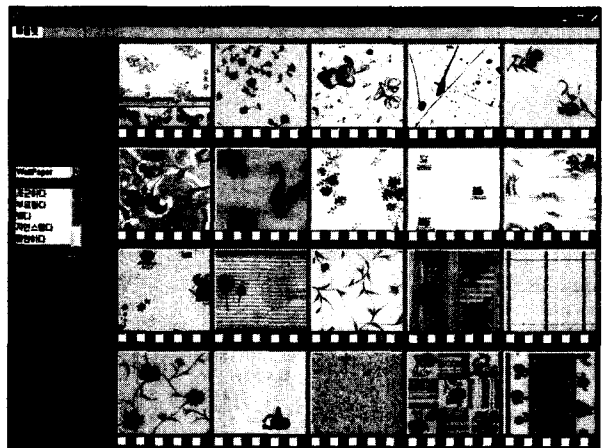
(a)



(b)



(c)



(d)

(그림 8) 클러스터링을 이용한 각 검색 라운드의 검색 영상



할 것이다. 사용한 속성의 bin수를 늘리면 보다 세밀한 검색을 할 수 있을 것으로 생각되나 bin의 수가 증가하면 검색속도가 저하되므로 검색 알고리즘의 향상이 이루어져야 될 것으로 생각된다. 이 또한 향후 연구과제이다.

**참 고 문 헌**

[1] R. K. Shirari, "Automatic indexing and content-based retrieval of captioned images," *IEEE Computer*, Vol.28, No.9, pp. 49-56, 1995.

[2] P. Aigrain, H. Zhang and D. Petkovic, "Content-based representation and retrieval of visual media : A state-of-the-art review," *Multimedia Tools and Applications*, Vol.3, No.4, pp.179-202, 1996.

[3] A. D. Bimbo, "Visual Information Retrieval," Morgan Kaufmann Pub. Inc., 1999.

[4] Text of ISO/IEC 15938-3 Multimedia Content Description Interface-Part 3 : Final Committee Draft, ISO/IEC/JTC1/SC29/WG11, Doc., N4062, Mar., 2001.

[5] MPEG-7 Visual Experimentation Model (XM), Version 10.0, ISO/IEC/JTC1/SC29/WG11, Doc., N4063, Mar., 2001.

[6] B. S. Manjunath, J. R. Ohm, V. V. Vasudevan and A. Yamada, "Color and Texture Descriptors," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.11, No.6, pp.703-715, 2001.

[7] B. S. Manjunath, P. Salembier and T. Sikora, Introduction to MPEG-7, John Wiley & Sons, Inc., 2002.

[8] T. Soen, T. Shimada and M. Akita, "Objective evaluation of color design," *Color Res. Appl.*, Vol.12, pp.187-194, 1987.

[9] N. Kawamoto and T. Soen, "Objective evaluation of color design II," *Color Res. Appl.*, Vol.18, pp.260-266, 1993.

[10] 엄진섭, 이준환, "퍼지이론을 이용한 색채계획 시스템에 관한 연구", 한국 퍼지 및 지능시스템학회논문지, 제7권 제3호, pp.55-64, 1997.

[11] 엄진섭, 이준환, "감각 언어를 이용한 칼라패턴 검색", 한국 퍼지 및 지능시스템학회논문지, Vol.9, No.1, pp.1-19, 1999.

[12] 엄진섭, 칼라패턴의 감성 인식 및 검색에 관한 연구, 전북대 박사학위논문, 2000.

[13] Y. Rui, T. S. Huang and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," *Proc. IEEE Int. Conf. Image Processing*, pp.815-818, 1997.

[14] Y. Rui and T. S. Huang, "Relevance Feedback : A power Tool for Interactive Content-based Image Retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.8, No.5, pp.644-655, 1998.

[15] S. Aksoy, R. M. Haralick, F. A. Cheikh and M. Gabbouj, "A weighted Distance Approach to Relevance Feedback," *IEEE Proc. of 15th Int. Conf. on Pattern Recognition*, Vol.4, pp.812-815, 2000.

[16] 이주영, 조성배, 감성기반 영상검색을 위한 대화형 유전자 알고리즘의 적용, 정보과학회논문지(B), 제26권 제3호, pp.422-430, 1999.

[17] H. Takagi, "Interactive evolutionary computation : Cooperation of computational intelligence and human kansei," *Proc. of 5th Int. Conf. on Soft Computing*, Iizuka, Japan, pp.41-50, 1998.

[18] D. R. Heisterkamp, J. Peng and H. K. Dai, "Feature relevance learning with query shift for content-based image retrieval," *IEEE Proc. of 15th Int. Conf. on Pattern Recognition*, Vol.4, pp.250-253, 2000.

[19] Y. Ishikawa, R. Subramanya and C. Faloutsos, "Mindreader : Query databases through multiple examples," *Proc. of 24th VLDB Conf.*, (New York), 1998.

[20] Y. Rui and T. S. Huang, "Optimal learning in image retrieval," *IEEE Proc. of Conf. on Computer Vision and Pattern Recognition*, Vol.1, pp.236-243, 2001.

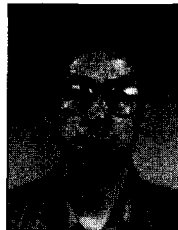
[21] Y. Deng, B. S. Manjunath, C. Kenney, M. S. Moore and H. Shin, "An efficient color representation for image retrieval," *IEEE Trans. Image Processing*, Vol.10, pp.140-147, Jan., 2001.

**박 중 수**



e-mail : jspark@woosuk.ac.kr  
 1980년 공주대학교 사범대학 수학교육학과 (학사)  
 1983년 성균관대학교 대학원 수학과(이학 석사)  
 1987년 미국 오하이오대학 박사과정 수료  
 1993년 전북대학교 대학원 수학과(이학박사)  
 1988년~현재 우석대학교 전산정보학부 교수  
 관심분야 : 영상처리, 정보검색, 응용수학

**엄 경 배**



e-mail : kbeum@kunsan.ac.kr  
 1990년 전북대학교 대학원 전자공학과 (공학박사)  
 1995년 캐나다 토론토대학교 전기 및 컴퓨터 공학과 박사후 연구원  
 1999년 일본 동경대학 생산기술 연구소 객원 연구원  
 1989년~현재 군산대학교 컴퓨터공학과 교수  
 관심분야 : 칼라 영상처리 및 인식, 병렬처리, 퍼지이론 및 신경망

**신 경 해**

e-mail : shinkh828@hotmail.com  
 2000년 중국 연변과학기술대학교 전산학과(학사)  
 2003년 전북대학교 대학원 영상공학과(이학석사)  
 관심분야 : 영상처리, 신경망론, 멀티미디어 검색



### 이 준 환

e-mail : chlee@moak.chonbuk.ac.kr

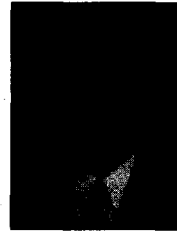
1980년 한양대학교 전자공학과(학사)

1982년 과학기술원 전기전자공학과(공학석사)

1990년 미국 미주리대학 컴퓨터공학과(공학박사)

1982년~현재 전북대학교 전자정보공학부 교수

관심분야 : 영상처리, 컴퓨터 비전, 인공지능



### 박 동 선

e-mail : dspark@moak.chonbuk.ac.kr

1979년 고려대학교 전자공학과(학사)

1984년 미국 미주리대학 컴퓨터공학과(공학석사)

1990년 미국 미주리대학 컴퓨터공학과(공학박사)

1991~현재 전북대학교 전자정보공학부 교수

관심분야 : 인공지능, 멀티미디어 통신, 컴퓨터 네트워크