

논문 2006-43SP-1-5

Borda count 방법을 이용한 다중 MPEG-7 서술자 조합에 관한 연구

(A study on MPEG-7 descriptor combining method using borda count method)

엄 민 영*, 최 윤 식*

(Minyoung Eom and Yoonsik Choe)

요 약

본 논문에서는 정지영상에 대해서 borda count 방법을 적용하여 MPEG-7 서술자기반 검색 결과 리스트의 조합에 대한 방법을 제안한다. 대부분의 검색 엔진의 경우 MPEG-7에서 정의하고 있는 서술자중 하나의 서술자만을 가지고 검색을 수행하고 있다. 이러한 경우 영상의 특성에 따라 검색 성능이 좋을 수도 있지만 그렇지 않을 수도 있다. 따라서 여러 가지 기술자들의 검색 결과를 조합하여 결과 리스트를 작성하는 방법이 요구되고 있다. 이에 따라 본 논문에서는 relevance feedback 알고리즘을 이용한 검색 시스템에 대하여, 여러 가지 서술자에 대한 검색 결과에 대한 결과 리스트를 borda count 방법을 이용하여 각각의 서술자에 의한 검색 결과를 조합하여 하나의 결과 리스트를 작성하는 방법을 제안한다

Abstract

In this paper, search result list synthesis method is proposed using borda count method for still image retrieval based on MPEG-7 descriptors. MPEG-7 standardizes descriptors that extract feature information from media data. In many cases, using a single descriptor lacks of correctness, it is suggested to use multiple descriptors to enhance retrieval efficiency. In this paper, retrieval efficiency enhancement is achieved by combining multiple search results which are from each descriptor. In combining search result, newly calculated borda count method is proposed. Comparing current frequency compensated calculation, rank considered frequency compensation is used to score animage in database. This combining method is considered in Content based image retrieval system with relevance feedback algorithm which uses high level information from system user. In each relevance iteration step, adoptive borda count method is used to calculate score of images.

Keywords : Content base image retrieval, MPEG-7, Relevance feedback, Borda count method, Descriptor

I. 서 론

멀티미디어 데이터는 다양한 정보를 내재하고 있으므로, 데이터 사용자는 기존의 단순한 텍스트 기반의 검색 방법을 통해서 원하는 정보를 쉽게 얻을 수 없

다. 사용자는 검색을 위한 효과적인 검색어를 생각해야 하고, 한 두 개의 검색어로써는 복잡하고 모호한 멀티미디어 데이터를 잘 표현 할 수가 없다. 따라서 텍스트 차원을 넘어선 새로운 검색 방법이 요구되게 되었다. 그 중 하나가 CBIR(Content based image retrieval) 시스템으로서^{[1][2][3]}, 이것의 대표적인 방법인 영상에 의한 질의이다. 이는 영상 자체를 시스템에 입력하는 것으로서, 사용자의 시스템으로의 접근이 용이하고 영상의 특징 정보를 자동으로 추출함으로써 색인 과정이 필요 없는 이점을 지니고 있다.

* 정회원, 연세 대학교, 전기전자 공학과
(Dept. of Electrical and Electronic Engineering, Yonsei University)

※ 본 연구는 한국 과학재단 목적 기초 연구 (R01-2003-000-10412-0) 지원으로 수행되었음.

접수일자 : 2005년6월10일 수정완료일 : 2005년11월14일

내용 기반 검색 시스템은 사용자가 요구하는 다양한 정보를 포함하는 질의 영상(query image), 또는 영상 시퀀스를 전체 데이터베이스와 매칭(matching) 시킴으로써 원하는 데이터를 효율적으로 얻을 수 있는 시스템이고 검색을 위해 여러 형태의 해결 방법이 존재한다. MPEG-7^{[4][5][6]}은 그들 중 하나로서 내용에 관한 정보를 나타내며 MPEG-7내 각각의 서술자(descriptor)들은 서로 다른 알고리즘을 통해 영상의 특징 정보를 추출한다^{[4][6]}. 기술자간 상이한 알고리즘으로 인해 모든 기술자의 검색 정확도는 다르며, 복수의 기술자를 함께 사용함으로써 검색 효과를 높일 수 있다. 그러나 복수의 기술자를 사용해서 검색 정확도를 높이는 데는 한계가 있는데 이는 질의 영상이 사용자가 찾고자 하는 영상부류의 한 예에 불과하기 때문이다. 따라서 보다 효과적인 영상 검색을 수행 하려면 질의 영상의 검색 범위를 갱신해야 할 필요가 있으며 적합성 피드백(relevance feedback)을 통해 선정된 긍정적 예제들은 질의의 위치 이동을 통해 질의 영상의 검색 범위를 갱신한다^{[7][8][9][10]}.

갱신된 검색 범위를 기반으로 검색 결과 리스트가 작성 되는데 일반적인 내용 기반 검색 시스템은 다수의 기술자를 사용하므로 각 서술자마다 도출 되는 결과 리스트들을 모두 고려 해 최종 검색 결과 리스트를 작성해야 한다. 이를 위해 borda count method가 사용된다^[11].

본 연구에서 제안한 시스템에서는 적합성 피드백을 이용하여 긍정적 예제와 부정적 예제를 선별 한 후 이 정보를 borda count method의 빈도수 계산 방법에서 기존의 방법이 가지는 빈도수 보상 값과 순위 정보와의 불일치를 순위 정보를 고려한 보상 값 계산 방법을 제안하였다.

II. MPEG-7 기술자를 사용한 영상검색

1. 내용 기반의 검색 시스템과 MPEG-7 기술자

내용 기반 영상 검색 시스템은 기존의 텍스트 기반의 검색 시스템이 영상 검색에 있어서 비효율적이라는 문제점에서 제기되기 시작했다. 그 결과 영상의 저 수준 정보를 이용한 컴퓨터 척도 검색 방법이 도입되었고, 이를 바탕으로 영상의 자동 검색이 가능하게 되었다. MPEG-7은 이러한 내용 기반 검색을 위한 여러 형태의 해결 방법 중 하나이고 본 연구에서는 총 4가지의 기술자 Edge Histogram, Homogeneous Texture, Dominant Color, Color Layout를 사용 하여 영상을 검색 하였다.

2. 내용기반 검색 시스템을 위한 적합성 피드백

(1) 적합성 피드백의 구조

적합성 피드백이란 기술자를 통한 특징 정보인 저 수준 정보를 인간 지각 내용을 표현한 고 수준 정보와 함께 영상 검색에 활용하기 위한 알고리즘이다. 적합성 피드백은 기본적으로 벡터 공간에서의 검색에 따른 질의의 위치 변동을 결정하는 방법으로서 알고리즘의 구현을 위해 Standard Rocchio 방식이 주로 쓰인다. 아래의 Q 는 원래 영상의 특징 정보를, Q' 는 위치 이동한 특징 정보를, E_P 는 검색 결과 중 긍정적 예제의 집합을, E_N 는 부정적 예제의 집합을 나타낸다. 그리고 n_p 는 E_P 원소의 개수를, n_n 는 E_N 원소의 개수를 나타낸다.

$$Q' = \alpha Q + \beta \frac{1}{n_p} \sum_{x \in E_P} x - \gamma \frac{1}{n_n} \sum_{x \in E_N} x \quad (1)$$

$$\frac{1}{n_p} \sum_{x \in E_P} x : E_P \text{의 평균} \quad (2)$$

$$\frac{1}{n_n} \sum_{x \in E_N} x : E_N \text{의 평균} \quad (3)$$

여기서 α , β 그리고 γ 는 실험적으로 결정되어지는 가중치이다.

3. 영상 검색에서의 borda count method

다수의 결과 리스트를 종합하기 위해 응용된 형태의 borda count method를 이용된다. Borda count method는 다수의 결과 리스트에서 임의의 영상이 위치하는 순위에 반비례하는 점수를 할당하여 최종 점수의 높낮이로 최종 순위를 결정하는 방법이다.

$$B_i = \prod_{k=1}^f g^k(i) + M^{f-1} \quad (4)$$

즉 임의의 영상 B_i 에 대해서 각 리스트 상의 점수를 $g^k(i)$, 사용자가 원하는 검색 순위 범위를 M , 각 결과 리스트 내에 영상 B_i 가 나타나는 빈도수를 f 이라 했을 때, 기존의 방법에서는 점수 $g^k(i)$ 들의 곱에 빈도수 보상 값 M^{f-1} 을 더해주어서 영상 B_i 에 대한 최종 순위를 결정 한다.

III. 제안한 시스템

본 연구에서 제안한 시스템의 구조는 다음과 같다.

먼저 시스템 사용자는 질의 영상과 사용할 기술자들을 결정한 후, 각 서술자에 대해 개별적으로 사용하여 질의 영상과 유사한 영상들을 검색한다. 이때 만들어진 검색 결과 리스트들에 제안한 형태의 borda count method를 적용하여 통합된 리스트를 작성한다. 이 결과가 만족스러우면 시스템은 종료하고, 그렇지 않으면 적합성 피드백을 적용한다.

1. 제안한 리스트 종합 방법

기존의 리스트 종합 방법은 하위 순위 영상이 과도한 빈도 수 보상에 의해 순위가 필요 이상으로 올라간다는 단점을 지닌다. 즉, $g^k(i)$ 가 작은 값들만을 가질 때 고정된 값 M^{f-1} 을 더함으로써 빈도수가 1-2회 적고 순위가 비교적 높은 영상의 경우보다 월등히 높은 순위를 갖게 된다. 이 문제점을 해결하기 위해 본 연구에서는 고정된 값 M^{f-1} 을 더해주는 대신 상위 순위들과 하위 순위들에 차등적으로 빈도수 보상 값을 할당하는 방법을 제안 하였다.

순위의 높낮음을 판단하기 위해 우선 상위 순위들과 하위 순위들을 구분 짓는 대표 값을 설정할 필요가 있다. 본 연구에서는 대표 값의 설정에 있어서 대표 값과 각 순위들에 해당하는 점수들과의 분포 정도를 기준으로 결정 했다. 즉, 점수들과의 에러가 가장 적은 수를 그 순위들, 즉 점수들을 표현할 수 있는 대표 값으로 선정하였다.

대표 값은 x 로, 각 순위 (r_k)들에 해당하는 점수를 $g_k^i = M - r_k + 1$; $g_k^i (1 \leq g_k^i \leq M)$ 로 표현할 때, 자유 변수 G_k^i 는 n 을 영상 i 가 나타난 리스트의 수에 대하여 표본 공간 $G_k^i : \{g_k^i | 1 \leq k \leq n\}$ 을 갖는다. 대표 값 x 와 G_k^i 와의 에러를 함수 $f(x)$ 라고 표현하면, 함수 $f(x)$ 가 최소가 되는 x 를 결정한다.

$$f(x) = E[(x - G^i)^2] \tag{5}$$

$$x = E[G^i] \tag{6}$$

따라서 본 연구에서는 $E[G_k^i]$ 을 점수들의 대표 값으로 사용한다.

빈도수 보상은 순위 정보의 지수 적 증가를 고려하여 동일하게 지수 적으로 증가해야 할 필요가 있고 증가 정도는 빈도수에 의해 결정되어야 한다. 따라서 빈도수 보상 값을 Y^f 의 형태로 제안한다. 대표 값 $E[G_k^i]$ 은 식(7)과 같이 빈도수 보상 값 Y^f 의 밑수 Y 에 비례해야 하므로 빈도수 보상 값은 식(8)과 같이 나타내 진다.

$$Y = \Omega \bar{G}^i \tag{7}$$

$$Y^f = (\Omega \bar{G}^i)^f \tag{8}$$

식(8)의 비례상수 Ω 를 결정하기 위해서 고려해야 할 사항은 다음과 같다.

1. 빈도수 보상 값의 최대값은 순위 정보 값의 최대 값보다 작거나 같아야 한다.
2. 빈도수 보상 값의 최소값은 지수 적 성질에 의해 1보다 크거나 같아야 한다.
3. 리스트 내의 모든 영상이 가질 수 있는 점수는 1이상 M이하의 값을 갖는다.

첫 번째 고려 사항의 이유는 빈도수 보상 값이 순위 정보 값보다 더 의미가 있어야 하기 때문이고, 두 번째 고려사항에서 최소값을 1 이상으로 한정 한 이유는 빈도수가 0일 경우 빈도수 보상 값은 $Y^0 = 1$ 이기 때문이다.

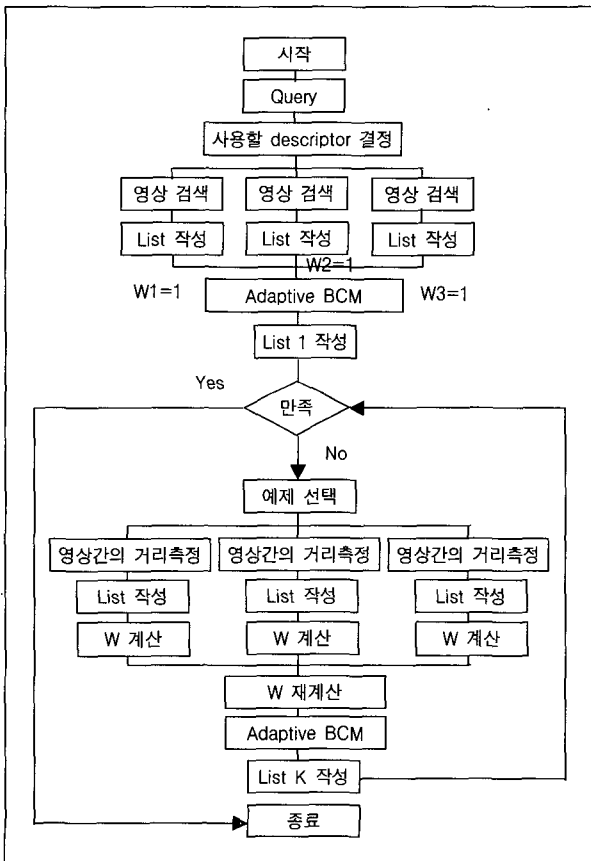


그림 1. 제안한 시스템
Fig. 1. Proposed System.

세 번째 고려 사항에서 의 제한 조건은 값은 모든 순위가 1위 일 때 최대 M점을 모든 순위가 M위일 때 최소 1점을 갖기 때문이다.

순위 정보 값이 가질 수 있는 최대값은 빈도수 f ($1 \leq f \leq M$) 각각에 해당하는 순위가 모두 1위인 경우, 즉 점수가 모두 M점 일 경우이다. 이때 순위 정보 값은 M^f 이므로 $(\Omega \bar{G}^i)^f$ 의 최대값은 M^f 이하로 제한된다. 또한 두 번째 조건에 의해서 $(\Omega \bar{G}^i)^f$ 의 최소값은 1이상으로 제한된다. 따라서 $(\Omega \bar{G}^i)^f$ 의 값은 식(9)와 같은 범위의 값을 갖는다.

$$1 \leq (\Omega \bar{G}^i)^f \leq M^f \quad (9)$$

또한, 세 번째 고려 사항에 의하여 $(\Omega \bar{G}^i)^f$ 의 값은 식(10)과 같이 제한된다.

$$1 \leq \bar{G}^i \leq M \quad (10)$$

식(10)의 \bar{G}^i 와 M의 최소값은 1이므로 f ($1 \leq f \leq M$)에 대하여 $(\bar{G}^i)^f$ 과 M^f 는 단조 증가의 함수이다. 따라서 식(11)과 같은 관계가 성립한다.

$$1 \leq (\bar{G}^i)^f \leq M^f \quad (11)$$

비례상수 Ω 는 $\Omega > 0$ 이므로 1이상의 자연수 f 에 대하여 $\Omega^f > 0$ 이다. 따라서 양수 Ω^f 를 식(11)에 곱해주어도 부등호의 방향은 바뀌지 않는다.

$$\Omega^f \leq (\Omega \bar{G}^i)^f \leq \Omega^f M^f \quad (12)$$

식(12)에서 $(\Omega \bar{G}^i)^f$ 의 최대, 최소값은 M과 1로 결정되어 있다. $(\Omega \bar{G}^i)^f$ 가 최대값 M을 가지므로 비례상수 Ω 는 1이하의 값을 가짐으로써 $(\Omega \bar{G}^i)^f$ 의 최대 범위가 제한된다. 또한 $(\Omega \bar{G}^i)^f$ 의 최소값은 1이상의 값을 가지므로 Ω 를 1이상으로 설정함으로써 $(\Omega \bar{G}^i)^f$ 의 최고값의 범위가 결정된다. 따라서 $\Omega = 1$ 로 결정된다.

또한 최하위 순위에 해당하는 점수가 1이므로 최하위 순위들 간의 빈도수에 의한 구별이 불가능하므로 점수 할당을 1만큼씩 높여서 부여하였다. 따라서 최종 식은 식(13),(14)과 같다.

$$\prod_{k=1}^n g_k^i + (\bar{G}^i)^f \quad (13)$$

$$g_k^i = M - r_k + 2 \quad (14)$$

IV. 실험 및 결과

질의 영상과 데이터베이스 영상의 거리 계산 방법은 MPEG-7의 experimental model인 XM에 의한 계산 방법을 따랐으며 borda count method를 적용하기 위해 영상의 색 정보를 이용하는 color layout, dominant color 과 edge 정보를 이용하는 edge histogram, 질감 정보를 이용하는 homogeneous texture의 총 4개의 기술자를 사용 하였다. 영상 검색을 위한 데이터베이스의 크기는 컬러 세포영상 2000장을 사용 하였으며, 검색 성능의 평가 기준으로서는 총 10가지의 질의 영상에 대한 식 (15)과 같이 NRE (Normalized Retrieval Efficiency)를 이용하였다.

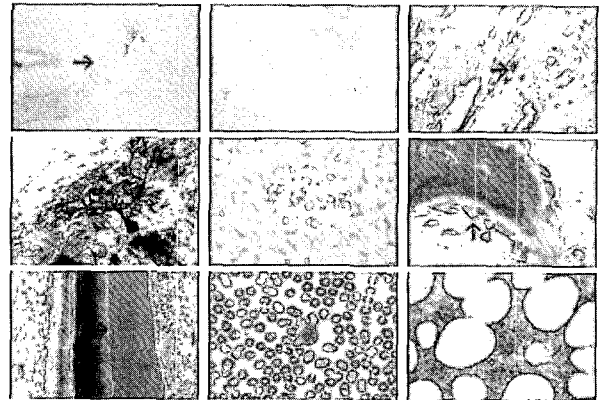


그림 2. 데이터베이스에 있는 영상들의 예
Fig. 2. Example of Images in Database.

표 1. NRE 결과 (M=20)
Table 1. Results of NRE (M=20)

방법 \ 횟수	1	2	3
기존방법 (M^{f-1})	0.9003	0.9642	0.9910
제안한방법 (\bar{G}^f)	0.9003	0.9732	1.000

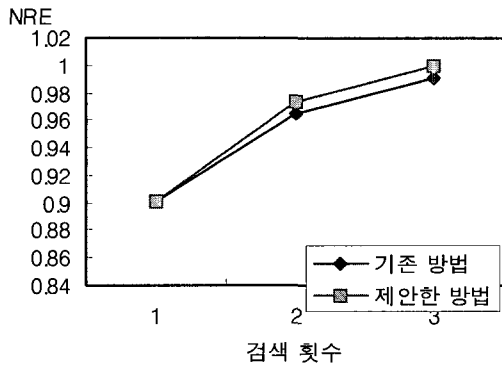


그림 3. 제안한 방법과의 비교 (M=20)
Figure3. Comparison with method(M=20).

$$NRE(k) = \frac{1}{N_q} \sum_{i=1}^{N_q} \frac{N_{Q_i}(k)}{N_{\max(Q_i)}} \quad (15)$$

여기서 $N_{Q_i}(k)$ 는 질의 영상 Q_i 에 대하여 k번째 검색 시 검색된 유사 영상의 수를 나타내고, $N_{\max(Q_i)}$ 는 질의 영상에 대하여 검색된 관련 영상의 최대 수이다.

Fig.2에서는 실험에 사용된 데이터베이스내의 세포 영상을 나타내었다. Table1과 Fig.3은 기존의 방법과 제안한 방법의 최대 검색 순위가 20일 경우의 검색 횟수 증가에 따른 검색 정확도의 차이를 보여준다. Table2와 Fig.4는 최대 검색 순위가 30일 경우의 기존 방법과 제안한 방법의 검색 정확도의 차이를 나타낸다. Table과 Figure에서 가로축에 해당하는 검색 횟수는 적합성 피드백이 사용되면서 검색하는 횟수를 나타내고 사용자가 시스템에 제공 하는 고수준 정보를 검색에 이용함으로써 검색 횟수가 증가할수록 검색되는 질의 대 유사 영상의 수가 증가하게 된다.

Fig.3와 Fig.4내의 두 그래프의 차이에서 확인 할 수 있듯이, 기존의 고정된 밀수에 의한 빈도수 보상 방법에 비해 제안한 가변적 밀수 할당 방법이 NRE측면에서 이득을 가져온다. 이러한 이득의 이유는 여러 검색 결과내의 하위에 위치하는 비 유사영상이 리스트의 종합 과정에서 필요 이상으로 상위에 위치하게 되는 현상

표 2. NRE 결과 (M=30)
Table 2. Results of NRE (M=30)

방법 \ 횟수	1	2	3
기존방법 (M^{f-1})	0.9314	0.9682	0.9761
제안한방법 (\overline{G}^f)	0.9473	1.000	1.000

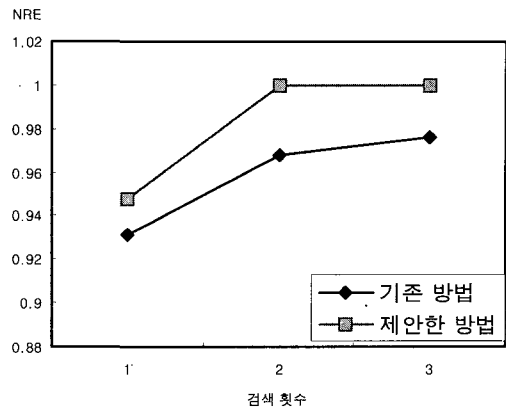


그림 4. 제안한 방법과의 비교 (M=30)
Figure4. Comparison with method(M=30)

을 순위 정보를 포함하는 빈도수 보상 방법에 의해 감소 시켰기 때문이다.

최대 검색 순위 M=20의 경우의 NRE 이득은 초기 검색 시 0, 두 번째 검색 시 0.009, 세 번째 검색 시 0.009이었고, M=30인 경우의 NRE 이득은 초기 검색 시 0.0159, 두 번째 검색 시 0.0318, 세 번째 검색 시 0.0239로서 M=20인 경우 보다 M=30인 경우에 있어서 제안한 방법이 더 높은 검색 정확도를 가져온다. 그 이유는 검색 범위가 커질수록 유사하지 않은 영상이 검색 순위 안에 들어올 확률이 그만큼 커지기 때문이고 제안한 방법에 의해서 이러한 여러 영상이 필요 이상으로 높은 순위를 갖게 되는 현상을 줄일 수 있기 때문이다.

V. 결 론

본 연구에서는 MPEG-7 내용 기반 검색 시스템에서 다수의 기술자를 이용한 검색을 수행하기 위한 borda count method의 빈도수 보상 방법에 대하여 연구하였다. 기존 방법에서 야기되는 과도한 빈도수 보상 문제를 해결하기 위해 새로운 계산 방법에 의거하여 영상 간 순위를 결정하였다.

본 연구에서는 이러한 과도한 빈도수 보상 문제를 해결하기 위해 순위 정보와 빈도수 정보를 함께 고려한 빈도수 보상 계산 방법을 제안하였다. 제안한 방법은 상위 순위들의 빈도수 보상 값은 기존 방법과 같이 비교적 크게 유지하는 동시에 하위 순위들의 빈도수 보상 값을 상대적으로 감소 시켜 하위 순위의 영상이 상위에 위치하는 현상을 감소 시켰다. 이와 동시에 최하위 순위의 점수를 1로 할당함으로써 발생하는 최하위 순위들

의 순위 중복 현상을 순위에 할당된 점수를 일정치 높여줌으로써 해결 하였다.

본 연구에서 제안한 방법에서는 각 기술자들의 가중치 문제를 고려하지 않았다. 기존의 적합성 피드백을 통한 가중치 할당은 각 기술자를 통한 결과 리스트를 사용자가 모두 확인해서 최종 결과 리스트와 비교함으로써 이루어지는데 이러한 방법은 시스템의 자동화 측면에서는 상당히 비효율적이다. 따라서 사용자가 선택한 긍정적 예제와 부정적 예제들을 적절히 활용함으로써 가중치 할당을 자동적으로 해결하는 방법이 수행되어야 한다.

참 고 문 헌

- [1] S. J. Yoon, D. K. Park, S. J. Park and C. S. Won, "Image retrieval using a novel relevance feedback for edge histogram descriptor of MPEG-7," *ICCE International Conference on Consumer Electronics*, June 2001, pp. 354-355.
- [2] S. F. Chang, W. Chen, H. J. Meng, H. Sundaram and D. Zhong, "A fully automated content-based video search engine supporting spatiotemporal queries," *IEEE Transactions on Circuits and Systems for Video Technology*, Sept. 1998, vol. 8, pp. 602-615.
- [3] Y. Rui, T. S. Huang and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," *International Conference on Image Processing*, Oct. 1997, vol. 2, pp. 815-818.
- [4] Text of ISO/IEC 15938-3/FDIS Information technology Multimedia content description interface-Part 3 Visual, 2002
- [5] Information Technology Multimedia Content Description Interface-Part8: Extraction and use of MPEG7 descriptions, July, 2002.
- [6] B. S. Manjunath, P. Salembier and T. Sikora, "Introduction to MPEG-7 Multimedia Content Description Interface," Wiley, 2002.
- [7] D. R. Heisterkamp, J. Peng and H. K. Dai, "Feature relevance learning with query shifting for content-based image retrieval," *International Conference on Pattern Recognition*, Sept. 2000, vol. 4, pp. 250-253.
- [8] R. Brunelli and O. Mich, "Image retrieval by examples," *IEEE Transactions on Multimedia*, Sept. 2000, vol. 2, pp. 164-171.
- [9] J. Assfalg, A. D. Bimbo and P. Pala, "Image retrieval by positive and negative examples," *International Conference on Pattern Recognition*, Sept. 2000, vol. 4, pp. 267-270.
- [10] H. Muller, W. Muller, S. M. Marchand, T. Pun and D. M. Squire, "Strategies for positive and negative relevance feedback in image retrieval," *International Conference on Pattern Recognition*, Sept. 2000, vol. 1, pp. 1043-1046.
- [11] S. Y. Jeong, K. H. Kim, B. T. Chun, J. Y. Lee and Y. L. Bae, "An Effective Method For Combining Multiple Features Of Image Retrieval," *Proceedings of the IEEE Region 10 Conference*, Sept. 1999, vol. 2, pp. 982-985.

저 자 소 개



엄민영 (정회원)
2001년 연세대학교 전파공학과
학사
2004년 연세대학교 전기전자공학과
석사
2004년~현재 연세대학교
전기전자공학과 박사과정

<주관심분야 : 웨이블릿, 비디오, 영상신호처리>



최윤식 (정회원)
1979년 연세대학교 전기공학과
학사
1984년 Case Western Reserve
Univ. 시스템공학과 졸업
1987년 Pennsylvania State Univ.
전기공학과 석사

1990년 Purdue Univ. 전기공학부 박사
1990년~1993년 (주)현대전자 산업전자 연구소
책임 연구원
1993년~현재 연세대학교 전기전자공학부 정교수
<주관심분야 : 비디오, 영상 신호처리, HDTV>