

논문-03-08-1-11

MPEG-7 디스크립터들의 조합을 이용한 영상 검색

강희범*, 원치선**

Image Retrieval Using a Composite of MPEG-7 Visual Descriptors

Hee Beom Kang* and Chee Sun Won**

요약

본 논문에서는 MPEG-7 영상 디스크립터 중 에지 히스토그램 디스크립터(EHD), 컬러 레이아웃 디스크립터(CLD), 그리고 호모지니어스 텍스처 디스크립터(HTD)로 구성된 각각의 영상 데이터베이스를 조합하여 relevance 피드백을 적용한 영상 검색 방법을 제안한다. 본 논문에서 사용한 에지 히스토그램 디스크립터는 영상의 국부적인 방향성 에지 분포를 표현한 것으로서 영상에 대하여 그 내용물의 형태를 잘 표현하는 디스크립터이다. 컬러 레이아웃 디스크립터는 구조적인 단순함과 빠른 동작 속도에 의해 영상 검색에 넓게 사용되어지며 컬러의 공간적 분포로 표현된다. 호모지니어스 텍스처 디스크립터는 영상의 질감에 대하여 정밀한 통계상의 분할로 서술된다. 앞에서 언급한 디스크립터들은 각각의 특징을 반영한 영상 검색에 적용되어진다. relevance 피드백은 영상 검색에 있어 사용자가 요구하는 정보를 반영할 수 있어 영상의 검색 효율을 높일 수 있다. 제안한 방법은 사용자가 relevance 피드백으로 결정한 영상의 특징 정보가 각각의 디스크립터들에 새로운 가중치를 부여한다. 따라서, 사용자의 선택적 요구가 반영된 특징 정보 갱신을 통해 검색 효율을 높인다. 자연 영상에 대한 실험 결과로 제안한 방법이 검색 성능을 향상시켜주는 것을 확인할 수 있다.

Abstract

In this paper, to improve the retrieval performance, an efficient combination of the MPEG-7 visual descriptors, such as the edge histogram descriptor (EHD), the color layout descriptor (CLD), and the homogeneous texture descriptor (HTD), is proposed in the framework of the relevance feedback approach. The EHD represents spatial distribution of edges in local image regions and it is considered as an important feature to represent the content of the image. The CLD specifies spatial distribution of colors and is widely used in image retrieval due to its simplicity and fast operation speed. The HTD describes precise statistical distribution of the image texture. Both the feature vector for the query image and the weighting factors among the combined descriptors are adaptively determined during the relevance feedback. Experimental results show that the proposed method improves the retrieval performance significantly for natural images.

I. 서론

디지털 방송 도입에 따른 채널 수의 증가와 인터넷의 보

급 및 디지털 카메라, 캠코더, 스캐너 등의 영상 입력 장비의 일반화에 따라 영상 정보 데이터의 증가는 기존의 텍스트 정보 증가보다 훨씬 빠르고 그것의 유통 빈도 역시 크게 증가하고 있다. 그러나, 방대한 양의 정보라 할지라도 체계적인 분류 및 검색 등의 효과적인 도움이 없다면 결국 사용자는 대부분의 시간을 원하는 정보를 찾는 데 소비하게 될 것이다. 따라서, 텍스트 정보의 검색과 같이 영상에서도

* 삼성전자 디지털미디어연구소
Digital Media R&D Center, SAMSUNG ELECTRONICS CO., LTD.

** 동국대학교 전자공학과
Dept. of electronic engineering, Dongguk Univ.

그것이 가지고 있는 특징 정보를 효율적으로 추출하고 그것들을 체계적으로 정리 및 저장함으로써 영상 데이터베이스를 구축할 수 있다. 내용 기반 멀티미디어 정보 검색을 효율적으로 지원하기 위한 기술(description)을 표준화한 것이 MPEG-7이다. MPEG-7은 특정 응용 분야를 염두에 두고 있는 것이 아니라 MPEG-7에서 표준화한 요소들이 가능한 많은 응용 분야에 쓰일 수 있도록 하는데 목표를 두고 있다. 따라서, MPEG-7은 멀티미디어 데이터에 대한 인터페이스, 표현 등에 대한 국제 표준을 정한 것으로서 영상 및 소리 데이터에 대하여 디스크립터(Descriptor), 디스크립션 스킴(Description Scheme), DDL(Description Definition Language)을 표준화했다. 이 중 디스크립터는 영상 및 소리가 가지는 가장 기본적인 특징을 표현한다. 영상 디스크립터는 컬러, 텍스처, 에지, 형태, 움직임 정보 등으로 나누어져있고 본 논문에서는 그 중 에지 히스토그램 디스크립터^{[1][2][5]}, 컬러 레이아웃 디스크립터^{[1][3][4]}, 그리고 호모지니어스 텍스처 디스크립터^{[1][5]}를 조합한 영상 검색을 구현하였다. 본 논문에서 제안하는 방법은 MPEG-7 디스크립터들의 기본 검색 조합 결과에 사용자의 선택적 요구가 반영된 특징 정보 갱신을 통한 랠러번스 피드백^{[6][7][8][9]}을 적용하여 각 디스크립터들의 가중치가 자동적으로 갱신되고 사용자가 선호하는 디스크립터가 부각될 수 있도록 하였다. 본 논문의 구성은 다음과 같다. II장에서는 실험에 사용한 MPEG-7 디스크립터에 대하여 기술하고, III장에서는 MPEG-7 디스크립터들을 조합하는 제안한 방법에 대하여 설명하며, IV장에서 실험 결과 및 성능을 보이고 마지막으로 V장에서 제안한 방법의 결론을 맺는다.

II. 실험에 사용한 MPEG-7 디스크립터

에지는 상대적으로 다른 명암도를 가진 두 영역간의 경계이다. 에지 특징 정보는 영상의 콘텐츠를 표현하기 위한 중요한 요소로 고려되어진다. 이것은 특별히 컬러의 단점을 보완하기 위해서 유용하다. 반면 컬러 특징 정보는 구조적인 단순함과 빠른 동작 속도에 의해 영상 검색에 넓게 사용되어진다. 그러므로, 위에서 언급한 두 특징 정보를 잘 표현하는 에지 히스토그램 디스크립터와 컬러 레이아웃 디스크립터는 서로 보완적인 관계로 사용됨으로써 향상된 영상 검색 결과를 만들 수 있다. 또한, 에지 히스토그램 디스크립터와 컬러 레이아웃 디스크립터는 각

특징을 공간적 분포로 표현하는 공통점을 가지고 있다. 텍스처 특징 정보는 영상 콘텐츠의 부드러움, 거침, 규칙적임과 같은 특성을 계측하여 이용된다. 호모지니어스 텍스처 디스크립터는 영상 질감에 대한 정밀한 통계상의 분할을 서술한다. 호모지니어스 텍스처 디스크립터는 텍스처 특징이 많은 패턴 영상들에 대하여 에지 히스토그램 디스크립터와 서로 보완적인 관계로 사용되어 역시 향상된 영상 검색 결과를 만들 수 있다. 본 논문에 사용된 MPEG-7 디스크립터는 MPEG-7 XM S/W에서 제공하는 알고리즘을 사용하였다. MPEG-7 디스크립터의 성능 평가를 위하여 영상에서 특징 정보를 추출 후 부호화 하는데 걸리는 시간을 측정하고 그 결과를 표 1에 나타내었다. 실험에 사용된 컴퓨터의 성능은 펜티엄 3-450MHz 이며 영상의 크기는 256×384로 CT4 DataSet^[14]의 영상을 사용하였다. 표 1과 같이 MPEG-7 디스크립터들 중 실험에 사용한 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 그리고 호모지니어스 텍스처 디스크립터는 영상 1개 당 특징 정보 추출 시간이 각 0.152초, 0.231초, 1.4초로 확인되었다.

표 1. MPEG-7 디스크립터의 특징 추출 시간 비교

Table. 1. Comparison of the feature extraction time for MPEG-7 descriptors

MPEG-7 Descriptor	DB image	Query image	Extraction time for a image(sec)
Edge Histogram	Natural Image	Natural Image	0.152
Color Layout	Natural Image	Natural Image	0.231
Homogeneous Texture	Texture Image	Texture Image	1.4

1. 에지 히스토그램 디스크립터 (EHD)

에지 히스토그램 디스크립터는 다섯 가지 에지 형태를 국부(local) 영상 영역에서 공간적 분포로 표현한다. 각 국부 영역은 부영상(sub-image)이라 하며 네 개의 방향성 에지와 한 개의 비방향성 에지로 표현된다^[1]. 보다 명확하게 말하면 16개의 겹치지 않는 각각의 부영상에 대해 5개의 빈(bins)으로 구성된 국부 에지 히스토그램이 생성된다. 각각의 부영상에 대하여 5가지 형태의 에지가 존재하므로 총 $16 \times 5 = 80$ 개의 히스토그램 빈(histogram bins)을 갖는다. 여기서, 부영상은 이미지 블록(image block)으로 나뉘지고 에지 형태 정보는 이미지 블록으로부터 추출되어진다^{[1][2][5]}.

2. 컬러 레이아웃 디스크립터 (CLD)

컬러 레이아웃 디스크립터는 컬러의 공간적 분포로 표현된다. 영상 또는 비디오 매칭에 사용되는 것 외에도 스케치 대 영상의 매칭과 같이 컬러에 대하여 레이아웃 기반 검색 (layout based retrieval)을 목표로 하고 있다. 이것은 전체 영상 또는 일정부분의 영상에 사용되어질 수 있다. 추출은 다음과 같이 진행된다. 우선 영상을 8x8개의 블록으로 나눈다. 각 블록 내의 픽셀에 대하여 하나의 우세한 컬러(dominant color)가 선택된다. 선택된 우세한 컬러는 8x8 행렬로 DCT 변환을 사용하여 계수의 열로 변환된다. 이것은 할당된 비트의 수에 적합하게 양자화된다. 계수는 기본값으로 Y에 대하여 6개, Cb와 Cr에 대하여 각각 3개를 사용한다^{[1][3][4]}.

3. 호모지니어스 텍스처 디스크립터 (HTD)

호모지니어스 텍스처 디스크립터^{[1][5]}는 영상 질감에 대한 분할된 주파수 대역에서의 통계적 특성을 서술한다. 이것은 질감의 특징을 추출하기 위해서 영상에 대해 라돈(Radon) 변환을 하고, 이어서 그 투영된 데이터에 대해 1차원 푸리에(Fourier) 변환을 한다. 그리고 각도 방향과 방사 방향에 대해 색인과 함께 30개의 채널이 만들어지며, 채널 내의 첫 번째와 두 번째 모멘트에 의하여 채널의 주파수 특징이 기술된다. i 번째 주파수 채널의 에너지 e_i 는 식 (1) 및 식 (2)와 같이 정의된다.

$$e_i = \log[1 + p_i] \quad (1)$$

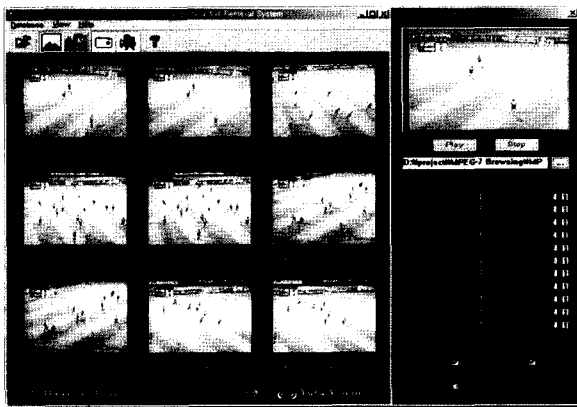


그림 1. 슬라이더에 의한 가중치 조절
Fig. 1. Controlling the weighting factors via the sliders

$$p_i = \int_{w=0}^1 \int_{\theta=0}^{360} [G_{P_{i,r}}(w, \theta) \cdot |w| \cdot F(w, \theta)]^2 dw d\theta \quad (2)$$

여기서, $G_{P_{i,r}}(w, \theta)$ 는 가보(Gabor) 함수이고 $F(w, \theta)$ 는 2D 극 주파수 영역에서 영상 표현이며, $|w|$ 는 직교와 극좌표계 공간들 사이의 자코비안(Jacobian) 항이다. i 번째 채널의 에너지 편차는 식 (3)과 식 (4)와 같이 정의된다^{[1][5]}.

$$d_i = \log[1 + q_i] \quad (3)$$

$$q_i = \sqrt{\int_{w=0}^1 \int_{\theta=0}^{360} \{ [G_{P_{i,r}}(w, \theta) \cdot |w| \cdot F(w, \theta)]^{2-p_i} \}^2 dw d\theta} \quad (4)$$

질감의 밝기 정보와 전체 이미지의 표준편차는 채널의 첫 번째와 두 번째 모멘트와 더불어 질감 특징 정보들로 간주된다. 즉, 채널의 이미지 평균 f_{DC} , 표준 편차 f_{SD} , 에너지 e_i 와 에너지 편차 d_i 는 다음의 순서에 따라 식 (5)와 같이 구성한다.

$$HTD = [f_{DC}, f_{SD}, e_1, e_2, \dots, e_{30}, d_1, d_2, \dots, d_{30}] \quad (5)$$

III. MPEG-7 디스크립터 조합

1. MPEG-7 디스크립터 조합 방법에 대한 고찰

디스크립터들을 조합하기 위하여 기본적으로 각 디스크립터로부터 얻어진 거리값은 정규화되어야 한다. 그 이유

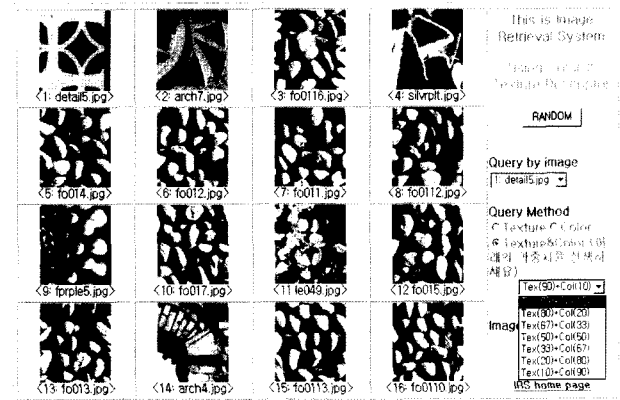


그림 2. 리스트 박스에 의한 가중치 조절
Fig. 2. Controlling the weighting factors vis the list box

는 디스크립터들의 특징 영역과 거리값 행렬이 서로 다르므로 인하여 결과적으로 거리값의 범위 차이가 매우 크기 때문이다. 기존의 조합 방법으로는 그림 1, 그림 2와 같이 사용자가 각 디스크립터의 가중치를 선택하는 방법이 있다. 그림 1은 사용자가 취향에 따라 각 디스크립터의 가중치를 조절할 수 있게 슬라이더를 구성한 GUI^[11]이며 그림 2는 가중치가 정해져 있는 리스트 박스를 구성한 GUI^[12]이다.

앞에서 언급한 슬라이더와 리스트 박스에 의한 가중치 조절 방법은 사용되는 검색 시스템을 잘 알지 못하는 일반 사용자들에게는 적합하지 않다. 그 이유는 사용자가 검색 시스템에 사용된 MPEG-7 디스크립터들의 특징에 대해서 상당 부분 많은 지식을 가지고 있어야 하기 때문이다. 또한, 마찬가지로 이유로 일반적인 사용자들은 어떤 특징을 가지고 있는 MPEG-7 디스크립터에 얼마만큼의 가중치를 주어야 하는지 알지 못하는 문제도 있다. 이런 문제를 감소시키기 위하여 조합된 검색 결과에 사용자의 선택적 요구가 반영된 특징 정보 갱신을 통한 랠러번스 피드백^{[6][7][8][9]}을 적용하여 얻은 검색 결과로부터 가중치를 자동적으로 조절하는 방법을 제안한다.

2. 제안된 MPEG-7 디스크립터들의 조합 방법

MPEG-7 디스크립터들을 조합하는 방법에 의한 검색은 기본 검색 조합 과정, 랠러번스 피드백 과정, 차후 검색 조합 과정으로 나누어진다. 기본 검색 조합 과정은 최초 한번 실행되고 랠러번스 피드백 과정과 차후 검색 조합 과정은 검색 결과에 대한 사용자의 만족 여하에 따라 반복된다.

2.1 MPEG-7 디스크립터들의 기본 검색 조합

디스크립터들의 기본 검색 조합 방법은 각각의 디스크립터를 독립적으로 사용하여 검색된 결과를 이용한다. 에지 히스토그램 디스크립터^{[1][2][5]}, 컬러 레이아웃 디스크립터^{[1][3][4]}, 호모지니어스 텍스처 디스크립터^{[1][5]}에 대하여 각각의 검색 결과로부터 가장 거리가 먼 D_{\max_EHD} , D_{\max_CLD} , D_{\max_HTD} 를 구한다.

$$D_{\max_EHD} = \max \{ dist_{EHD}(Q, i), \quad i = 1, \dots, N \} \quad (6)$$

$$D_{\max_CLD} = \max \{ dist_{CLD}(Q, i), \quad i = 1, \dots, N \} \quad (7)$$

$$D_{\max_HTD} = \max \{ dist_{HTD}(Q, i), \quad i = 1, \dots, N \} \quad (8)$$

여기서, $dist_{EHD}(Q, i)$, $dist_{CLD}(Q, i)$, $dist_{HTD}(Q, i)$ 는 질의 영상 Q와 i 번째 영상과의 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터에 의해 측정된 거리 값을 각각 나타낸다. N은 데이터베이스 내 영상의 개수를 나타낸다.

각 영상의 $dist_{EHD}(Q, i)$, $dist_{CLD}(Q, i)$, $dist_{HTD}(Q, i)$ 값들을 구해진 D_{\max_EHD} , D_{\max_CLD} , D_{\max_HTD} 값으로 식 (9), (10), (11)과 같이 각각 나누어 정규화 된 거리 값 $D_{EHD}(Q, i)$, $D_{CLD}(Q, i)$, $D_{HTD}(Q, i)$ 를 구한다.

$$D_{EHD}(Q, i) = \frac{dist_{EHD}(Q, i)}{D_{\max_EHD}}, \quad i = 1, \dots, N \quad (9)$$

$$D_{CLD}(Q, i) = \frac{dist_{CLD}(Q, i)}{D_{\max_CLD}}, \quad i = 1, \dots, N \quad (10)$$

$$D_{HTD}(Q, i) = \frac{dist_{HTD}(Q, i)}{D_{\max_HTD}}, \quad i = 1, \dots, N \quad (11)$$

최종적으로 각 디스크립터 별로 정규화 된 거리 값 $D_{EHD}(Q, i)$, $D_{CLD}(Q, i)$, $D_{HTD}(Q, i)$ 의 평균을 통하여 각 영상에 대해 새로운 거리 값 $D_{Total}(Q, i)$ 를 식 (12)에 서와 같이 구할 수 있다.

$$D_{Total}(Q, i) = \frac{D_{EHD}(Q, i) + D_{CLD}(Q, i) + D_{HTD}(Q, i)}{3}, \quad i = 1, \dots, N \quad (12)$$

각 영상 별로 구해진 $D_{Total}(Q, i)$ 가 낮은 순으로 기본 검색 조합의 결과 영상을 보여준다. 디스크립터들의 조합 방법을 이용한 영상 검색의 결과가 단지 한 개의 영상 디스크립터를 이용한 검색의 결과보다 더 좋은 결과를 낳는다. 이것은 실험 결과에서 제시한 그림 4의 검색 횟수 1인 경우를 통해 확인된다.

2.2 랠러번스 피드백

랄러번스 피드백은 사용자와 질의 응답 루프 과정을 통하여 검색 효율을 향상시키기 위한 메커니즘이다. 이는 초기 질의 영상에 의해 검색된 기본 검색 결과 영상에 대하여 사용자가 관련 영상을 선택한 후 재 검색한다. 즉, 사용자의 의향에 근거해서 랠러번스 피드백 시스템은 초기 질의 특징을 사용자 입력 요구에 가깝게 반복적으로 갱신한다^{[6][7][8][9]}.

일반적인 렐러번스 피드백 과정은 다음의 단계들로 이루어져 있다^[6].

1. 주어진 k번째 피드백과 관련된 질의 특징 벡터 Q(k)를 계산한다.
2. Q(k)는 데이터베이스 내에 저장된 영상들의 특징 벡터와 비교된다.
3. 결과에서 질의와 유사하게 나온 영상은 적절한 거리 계산 방법으로 순위가 정해진다.
4. 사용자가 일정순위 이내의 영상에 대하여 관련 있음 혹은 관련 없음의 표시를 한다.
5. 벡터 Q(k)는 사용자가 제공한 정보에 의해서 수정된다. 관련된 특징 벡터의 현재 가중치는 증가되고, 반면에 관련 없는 특징 벡터의 가중치는 감소한다.
6. 새로운 질의가 수정된 벡터 Q(k+1)에 의해서 제시된다. 즉, k번째 피드백에서 k+1번째 피드백으로 진행되면서 특징 벡터는 식 (13)과 같이 갱신된다.

$$Q^{(k+1)} = Q^{(k)} + \frac{\delta}{N_{rel}} \sum_{i=1}^{N_{rel}} X^{(k)}(i) - \frac{\epsilon}{M_{notrel}} \sum_{j=1}^{M_{notrel}} Y^{(k)}(j) \quad (13)$$

여기서, $Q^{(k)}$ 는 k번째 질의와 관련된 특징 벡터, X(k)와 N_{rel} 은 각각 관련 있음으로 표시된 영상의 특징 벡터와 개수이며 Y(k)와 M_{notrel} 는 관련 없음으로 표시된 영상의 특징 벡터와 개수이고 δ 와 ϵ 은 가중치 상수이다.

본 실험에 사용한 렐러번스 피드백 방식은 식 (14)와 같이 사용자가 관련 없음이라고 표시한 영상에 대하여 특징 정보의 통계적 특성을 구하지 않는 방식을 적용하였다. 즉, 이것은 식 (13)에서 네거티브 피드백을 주는 Y(k) 특징 정보를 배제한 것이다.

$$Q^{(k+1)} = Q^{(k)} + \frac{\delta}{N_{rel}} \sum_{i=1}^{N_{rel}} X^{(k)}(i) \quad (14)$$

그 이유는 사용자가 선택하는 제거 영상들이 가지는 특징 정보가 랜덤한 분포를 가져 각 번의 분산값들이 모두 아주 크게 나타나기 때문이다. 따라서, 본 실험에 사용한 렐러번스 피드백 방식은 사용자가 제거 영상으로 표시한 경우 단순히 그 영상들을 순위에서 제거시켰다^[9]. 즉, 데이터베이스 내의 하위 순위로 지정된 것으로 상위 순위에 올라오지 못하게 처리한 것이다. 가중치 δ 값은 1로 설정하여 가중치를 적용하지 않았다^[10]. 렐러번스 피드백을 이용하여 MPEG-7 영상 디스크립터들의 기본 검색 조합 결과에 사용자의 선택적 요구가 반영된 특징 정보 갱신을 통한 렐러

번스 피드백을 적용하여 각 디스크립터들의 가중치를 계산하고 차후 검색 조합에 의한 영상 간의 거리 측정에 각 디스크립터의 가중치를 적용하였다.

2.3 MPEG-7 디스크립터들의 차후 검색 조합

MPEG-7 디스크립터들의 기본 검색 조합 방법과 같이 단순히 초기 질의 영상에 대한 디스크립터들의 정규화 된 거리 값을 통합하므로써 어느 정도 검색 성능을 향상시킬 수 있지만 사용자의 의함에 따라서 조합되는 디스크립터들의 가중치가 달라질 수 있다. 이와 같은 상황을 적응적으로 적용하므로써 검색의 성능을 더욱 향상시킬 수 있다. 따라서, 기본 검색 조합 결과 영상들에 대하여 사용자가 관련 영상들을 선택한다. 선택된 관련 영상 별로 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터에 의해 검색된 영상의 개수 N_{EHD} , N_{CLD} , N_{HTD} 를 구한다. 각 영상의 개수는 렐러번스 피드백 후에 디스크립터들의 거리 값을 합할 때 적용된다. 즉, 각각의 디스크립터 별로 렐러번스 피드백 후에 검색된 영상들의 정규화 된 거리 값을 합하여 새로운 거리 값을 만들 때 각 디스크립터의 가중치로 식 (15)와 같이 적용된다.

$$D_{Total_{ref}}(Q, i) = \frac{N_{EHD} \times D_{EHD_{ref}}(Q, i) + N_{CLD} \times D_{CLD_{ref}}(Q, i) + N_{HTD} \times D_{HTD_{ref}}(Q, i)}{N_{EHD} + N_{CLD} + N_{HTD}} \quad (15)$$

$i = 1, \dots, N$

여기서, $D_{EHD_{ref}}(Q, i)$, $D_{CLD_{ref}}(Q, i)$, $D_{HTD_{ref}}(Q, i)$ 는 렐러번스 피드백 후에 각 디스크립터 별로 정규화 된 거리 값을 각각 나타낸다. 각각의 디스크립터에 의해 검색된 영상의 개수 N_{EHD} , N_{CLD} , N_{HTD} 로 $D_{EHD_{ref}}(Q, i)$, $D_{CLD_{ref}}(Q, i)$, $D_{HTD_{ref}}(Q, i)$ 를 곱하여 줌으로써 더 영향력이 있는 디스크립터의 거리 값에 더 큰 가중치를 적용하여 새로운 거리 값 $D_{Total_{ref}}(Q, i)$ 를 만든다. 따라서, 사용자가 선택한 관련 영상들에 의하여 각 디스크립터의 가중치가 반복적으로 결정되어 새로운 검색 결과를 보여주게 된다. 이후의 과정은 7검색 결과에 대하여 사용자의 만족 여하에 따라 렐러번스 피드백과 MPEG-7 디스크립터들의 차후 검색 조합 방법이 반복된다. 이때 반복적으로 사용자의 의함에 따라 디스크립터들의 가중치가 결정되므로써 사용자 의도가 자동적으로 검색에 반영되고 사용자의 의도에 맞는 디스크립터가 더욱 강조되는 결과가 된다.

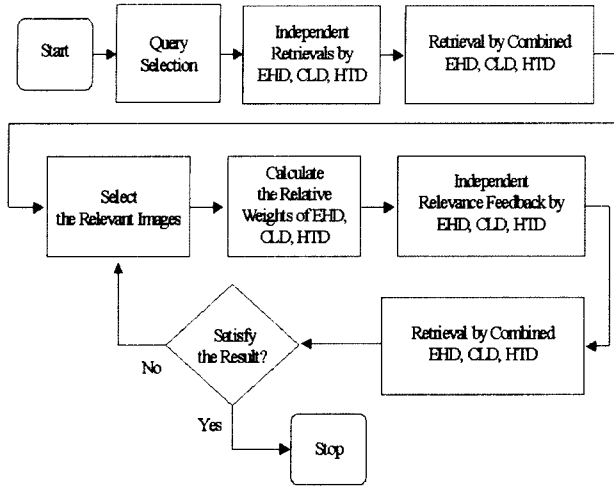


그림 3. MPEG-7 디스크립터 조합 방법의 블록도
Fig. 3. Block diagram of the combination method for the MPEG-7 descriptors

그림 3은 제안하는 MPEG-7 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터들의 조합 방법에 대한 전체 블록도이다. 제안하는 방법은 사용자가 선택한 관련 영상들에 의하여 자동적으로 디스크립터들의 가중치가 결정된다. 그러므로, 제안하는 방법은 질의 영상에 따라 MPEG-7 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터 각각의 장점을 강화하고 단점을 보완하여 향상된 검색 결과를 보여준다.

식 (15)에서 사용된 가중치 적용 방법의 성능을 확인하기 위해 식 (16), (17)과 같은 두 가지의 가중치 적용 방법과 비교하였다. 즉, 각 디스크립터의 검색 결과 검색된 영상의 개수 N_{EHD} , N_{CLD} , N_{HTD} 에 의한 가중치 적용 방법과 NMRR(Normalized Modified Retrieval Rank)^[13]에 의한 $NMRR_{EHD}(Q)$, $NMRR_{CLD}(Q)$, $NMRR_{HTD}(Q)$, 그리고 각 디스크립터에 의해 검색된 영상 개수의 제곱 N_{EHD}^2 , N_{CLD}^2 , N_{HTD}^2 을 이용하여 가중치 적용 방법의 성능을 비교하였다. NMRR과 검색된 영상 개수의 제곱을 각 디스크립터의 가중치로 사용한 경우 식 (16), (17)과 같이 각각 적용된다.

$$D_{Total_{ref}}(Q, i) = \frac{\alpha \times D_{EHD_{ref}}(Q, i) + \beta \times D_{CLD_{ref}}(Q, i) + \gamma \times D_{HTD_{ref}}(Q, i)}{\alpha + \beta + \gamma} \quad (16)$$

$$\alpha = (1 - NMRR_{EHD}(Q)), \quad \beta = (1 - NMRR_{CLD}(Q)),$$

$$\gamma = (1 - NMRR_{HTD}(Q)) \quad i = 1, \dots, N$$

질의영상에 대하여 검색된 결과영상 중 관련영상 모두 상위순위로 검색되었을 경우 NMRR은 0의 값을 가지게 되고 관련영상 모두 검색되지 않았을 경우 1의 값을 가지게 된다. 따라서, 식 (16)에서 가중치로 1-NMRR을 취하였다.

$$D_{Total_{ref}}(Q, i) = \frac{N_{EHD}^2 \times D_{EHD_{ref}}(Q, i) + N_{CLD}^2 \times D_{CLD_{ref}}(Q, i) + N_{HTD}^2 \times D_{HTD_{ref}}(Q, i)}{N_{EHD}^2 + N_{CLD}^2 + N_{HTD}^2} \quad i = 1, \dots, N \quad (17)$$

식 (17)과 같이 각 디스크립터에 의해 검색된 영상 개수의 제곱 N_{EHD}^2 , N_{CLD}^2 , N_{HTD}^2 을 이용하여 적용되는 가중치의 차이를 크게 하여 보았다.

검색 성능 비교를 위한 NRE(Normalized Retrieval Efficiency)는 식 (18), (19), (20), (21)을 통하여 구하였다.

$$NRE_{EHD}(k) = \frac{1}{N_q} \sum_{i=1}^{N_q} \frac{N_{EHD_{q_i}}(k)}{N_{\max_{q_i}}} \quad (18)$$

$$NRE_{CLD}(k) = \frac{1}{N_q} \sum_{i=1}^{N_q} \frac{N_{CLD_{q_i}}(k)}{N_{\max_{q_i}}} \quad (19)$$

$$NRE_{HTD}(k) = \frac{1}{N_q} \sum_{i=1}^{N_q} \frac{N_{HTD_{q_i}}(k)}{N_{\max_{q_i}}} \quad (20)$$

$$NRE_{PM}(k) = \frac{1}{N_q} \sum_{i=1}^{N_q} \frac{N_{PM_{q_i}}(k)}{N_{\max_{q_i}}} \quad (21)$$

여기서, $N_{EHD_{q_i}}(k)$, $N_{CLD_{q_i}}(k)$, $N_{HTD_{q_i}}(k)$, $N_{PM_{q_i}}(k)$ 는 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터, 그리고 제안하는 방법 별로 질의 영상 Q_i 에 대하여 검색 횟수 k 에 따라 검색된 관련 영상의 수를 나타낸 것이다. $N_{\max_{q_i}}$ 는 질의 영상에 대하여 검색된 관련 영상의 최대 수이다. 따라서, $NRE_{EHD}(k)$, $NRE_{CLD}(k)$, $NRE_{HTD}(k)$, $NRE_{PM}(k)$ 는 N_q 개의 질의 영상에 대하여 정규화 된 검색 성능(Normalized Retrieval Efficiency) 값을 나타낸다. 정규화 된 검색 성능 값이 1에 근접할수록 더 좋은 검색 성능을 의미한다.

표 2는 가중치 적용 방법에 따른 성능 차이를 보이고 있다. 표 2에서 보여지는 바와 같이 검색 횟수 1의 경우는 기본 검색 조합 결과로서 랠러번스 피드백에 의한 가중치가 적용되지 않은 상태이다. 따라서, 동일한 NRE 값을 갖는

표 2. 가중치 적용 방법에 따른 검색 성능 비교

Table 2. Comparison of the retrieval performance with various methods for weighting factors

적용된 가중치 별 NRE	검색 횟수				
	1	2	3	4	5
N_{EHD} , N_{CLD} , N_{HTD} 에 의한 NRE	0.68	0.84	0.89	0.95	0.98
N_{EHD}^2 , N_{CLD}^2 , N_{HTD}^2 에 의한 NRE	0.68	0.80	0.84	0.90	0.93
$NMRP_{EHD}$, $NMRP_{CLD}$, $NMRP_{HTD}$ 에 의한 NRE	0.68	0.81	0.88	0.94	0.97

다. 검색 횟수 2 ~ 검색 횟수 5의 경우는 차후 검색 조합 결과로서 적용된 가중치 별로 NRE 값의 차이를 가지며 검색 성능이 향상된다. 표 2에 제시한 바와 같이 각각의 디스크립터에 의해 검색된 영상의 개수 N_{EHD} , N_{CLD} , N_{HTD} 를 가중치로 적용한 제안하는 방법이 더 나은 검색 성능을 보여주고 있다. 표 2의 실험은 동일한 조건에서 가중치 적용 방법의 차이를 두었다. 총 11639장의 자연 영상을 가지고 질의 영상 100장에 대하여 실험한 결과이다.

IV. 실험 결과

MPEG-7 디스크립터들을 이용한 관련 영상 검색을 위하여 제안한 방법에 MPEG-7의 CE(Core Experiment)에 사용한 11639개의 자연 영상^[14]을 사용하였다. 그림 4는 100개의 질의 영상에 대하여 검색 횟수에 따라 검색된 관련 영상의 수를 식 (17), (18), (19), (20)을 통하여 구하여진 정규화 된 수치로 나타내고 있다. 여기서, NRE(Normalized Retrieval Efficiency)는 정규화 된 검색 성능을 의미한다.

실험 결과 영상들은 몇 가지 질의 영상에 대하여 각 디스크립터 및 제안한 방법의 21위까지 검색된 결과 영상들이다. 그림 5~그림 8, 그림 10~그림 13, 그림 15~그림 18은

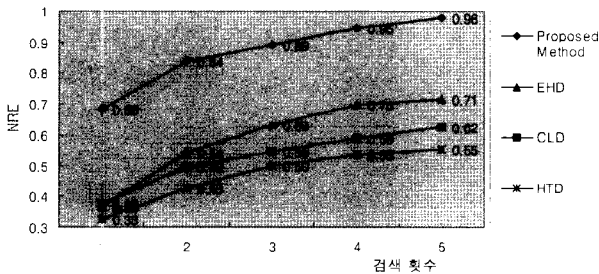


그림 4. 정규화 된 검색 성능 비교
Fig. 4. Comparison with normalized retrieval efficiency

질의 영상 1

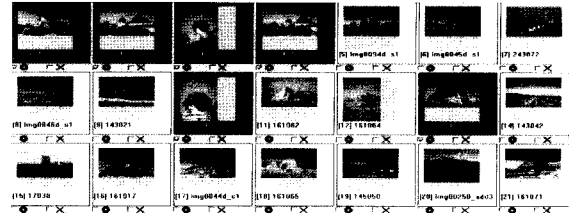


그림 5. 기본 검색 조합 결과 (k=1)
Fig. 5. Combined result 1 for initial retrieval (k=1)

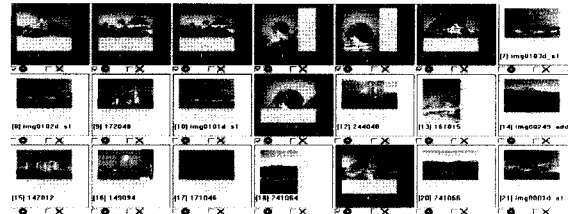


그림 6. 차후 검색 조합 결과 (k=2)
Fig. 6. Combined retrieval result 1 with relevance feedback (k=2)

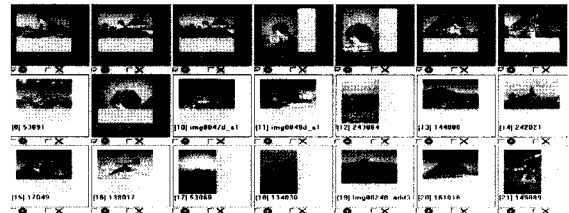


그림 7. 차후 검색 조합 결과 (k=3)
Fig. 7. Combined retrieval result 1 with relevance feedback (k=3)

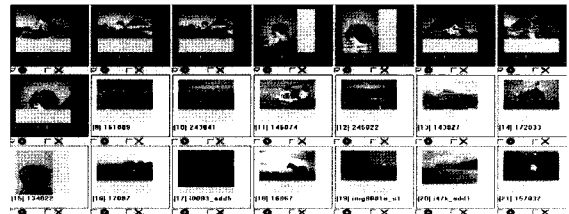


그림 8. 차후 검색 조합 결과 (k=4)
Fig. 8. Combined retrieval result 1 with relevance feedback (k=4)

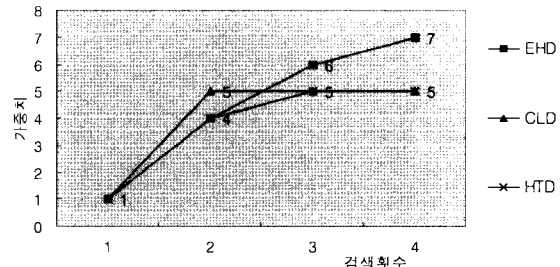


그림 9. 질의 영상1에 대한 디스크립터들의 가중치 변화 추이
Fig. 9. Changes of descriptor weighting factors for query image 1

질의 영상 2

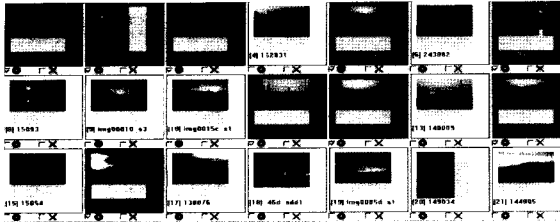


그림 10. 기본 검색 조합 결과2 (k=1)
Fig. 10. Combined result 2 for initial retrieval (k=1)

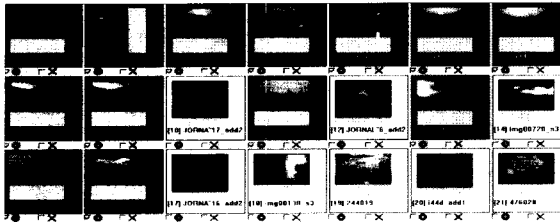


그림 11. 차후 검색 조합 결과2 (k=2)
Fig. 11. Combined retrieval result 2 with relevance feedback (k=2)

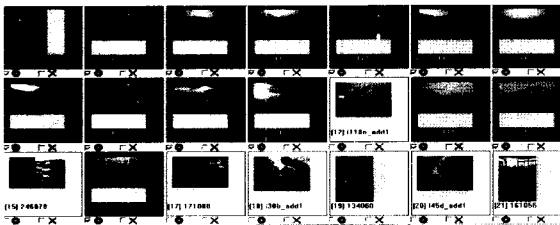


그림 12. 차후 검색 조합 결과2 (k=3)
Fig. 12. Combined retrieval result 2 with relevance feedback (k=3)

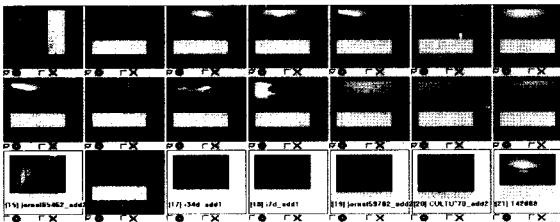


그림 13. 차후 검색 조합 결과2 (k=4)
Fig. 13. Combined retrieval result 2 with relevance feedback (k=4)

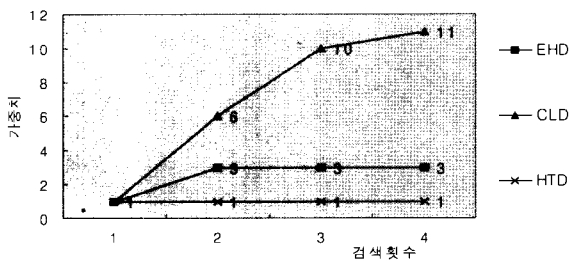


그림 14. 질의 영상2에 대한 디스크립터들의 가중치 변화 추이
Fig. 14. Changes of descriptor weighting factors for query image 2

질의 영상 3

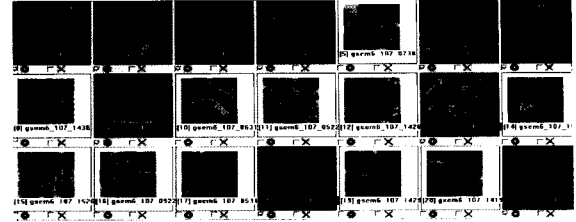


그림 15. 기본 검색 조합 결과3 (k=1)
Fig. 15. Combined result 3 for initial retrieval (k=1)

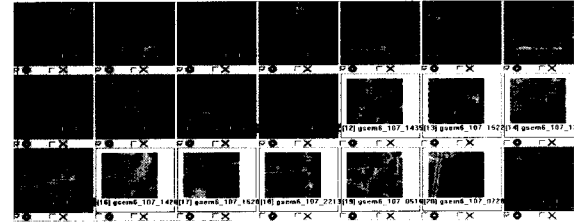


그림 16. 차후 검색 조합 결과3 (k=2)
Fig. 16. Combined retrieval result 3 with relevance feedback (k=2)

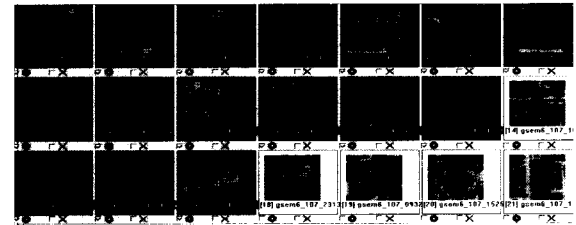


그림 17. 차후 검색 조합 결과3 (k=3)
Fig. 17. Combined retrieval result 3 with relevance feedback (k=3)

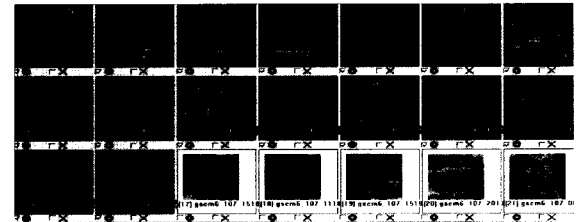


그림 18. 차후 검색 조합 결과3 (k=4)
Fig. 18. Combined retrieval result 3 with relevance feedback (k=4)

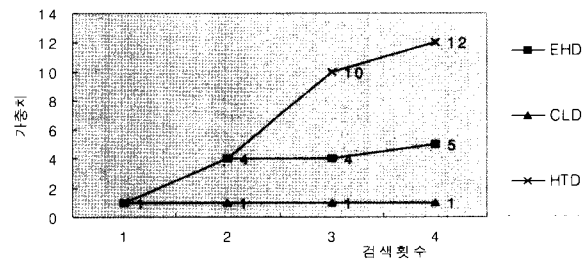


그림 19. 질의 영상3에 대한 디스크립터들의 가중치 변화 추이
Fig. 19. Changes of descriptor weighting factors for query image 3

질의 영상 1, 질의 영상 2, 질의 영상 3에 대한 기본 검색 조합 결과와 차후 검색 조합 결과들을 7 나타낸 것이며 그림 9, 그림 14, 그림 19는 각 질의 영상들에 대하여 적용된 디스크립터들의 가중치 변화 추이를 검색 횟수에 따라 나타낸 것으로 그림에서 보이는 바와 같이 검색의 피드백 횟수가 증가함에 따라 특정 디스크립터의 가중치가 강조되는 것을 확인할 수 있다. k는 검색 횟수를 나타낸다. 결과 그림에서 관련 영상들은 테두리가 어두운 색으로 표현되고 있다.

V. 결 론

본 논문에서는 검색 성능을 향상시키기 위하여 MPEG-7 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터를 조합하여 검색 성능을 향상시키는 방법을 제안하였다. 실험에 사용한 각각의 MPEG-7 디스크립터 만을 사용한 검색도 좋은 성능을 보이지만 조합에 의한 거리 값의 재 구성 및 렐러번스 피드백을 이용한 디스크립터의 가중치 적용 등의 후처리를 통하여 보다 효율적인 영상 검색 결과를 얻을 수 있었다. 특히, 제안하는 방법은 사용자가 선택한 관련 영상들에 의하여 자동적으로 디스크립터들의 가중치가 결정되어 질의 영상에 따라 MPEG-7 에지 히스토그램 디스크립터, 컬러 레이아웃 디스크립터, 호모지니어스 텍스처 디스크립터 각각의 장점을 강화하고 단점을 보완하여 향상된 검색 결과를 보여주었다.

참 고 문 헌

- [1] "FCD 15938-3 Multimedia Content Description Interface - Part 3 isual," *ISO/IEC JTCl/SC29/WG11/W4062*, Singapore, Mar, 2001.
- [2] D.K. Park, Y.S. Jeon, C.S. Won and S.J. Park, "Efficient use of local edge histogram descriptor," *ISO/IEC JTCl/SC29/WG11/W4062*, Int. Workshop on Standards, Interoperability and Practices, ACM, pp. 52~54, Marina del Rey, CA, Nov, 4, 2000.
- [3] Greg Pass, Ramin Zabih and Justin Miller, Comparing images using color coherence vectors, In Proc. ACM Conf. On Multimedia, 1996.
- [4] J. Huang, S.Kumar, W.-J. Zhu and R. Zabih, Image indexing using color correlogram. In Proc. Of IEEE Conf. On Computer Vision and Pattern Recognition, 1997.
- [5] Peng Wu, Yanglim Choi, Yong Man Ro and Chee Sun Won "MPEG-7 Texture Descriptors," *International Journal of Image and Graphics*, Vol. 1, No. 3, pp 547-563, 2001.
- [6] Eugenio Di Sciascio and Marina Mongiello, "Query by Sketch and Relevance Feedback for Content-Based Image Retrieval over the Web," *Journal of Visual Language and Computing*, No.10, 1999.
- [7] Y. Rui, T. S. Huang and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in *Proc. IEEE int. Conf. on Image Proc.*, 1997.
- [8] Zhong Su, Hongjiang Zhang and Shaoping Ma, "Using Bayesian Classifier in Relevant Feedback of Image Retrieval," *IEEE*, 2000.
- [9] Su Jung Yoon, Dong Kwon Park, Soo Jun Park and Chee Sun Won, "Image Retrieval Using a Novel Relevance Feedback for Edge Histogram Descriptor of MPEG-7," *ICCE 2001*, L.A, Jun, 2001.
- [10] 윤수정, 박동권, 원치선, "영상 검색을 위한 적합성 피드백의 개선", *전자공학회 논문지 제39권 CI권 제4호*, pp 28-37, 2002.
- [11] <http://vision.hanyang.ac.kr/researchframeset.htm>
- [12] http://ivory.icu.ac.kr/retreival_with_file.html
- [13] "Color and Texture Core Experiments," *ISO/IEC/JTCl/SC29/WG11, MPEG document N3090*, Maui, Dec, 1999.
- [14] "Core Experiment Results for Spatial Intensity Descriptor(CT4)," *ISO/IEC/JTCl/SC29/WG11, MPEG document M5374*, Maui, Dec, 1999.

 저 자 소 개

**강 희 범**

- 2001년 2월 : 동국대학교 전자공학과 졸업(학사)
- 2001년 9월~2002년 11월 : 한국전자통신연구원 휴먼정보검색팀 위촉연구원
- 2003년 2월 : 동국대학교 전자공학과 졸업(석사)
- 2003년 1월~현재 : 삼성전자 디지털미디어연구소 연구원
- 주관심분야 : 영상처리, MPEG-7, TV-Anytime, 디지털 방송

**원 처 선**

- 1982년 2월 : 고려대학교 전자공학과 (공학사)
- 1986년 2월 : Univ. of Massachusetts-Amherst (공학석사)
- 1990년 2월 : Univ. of Massachusetts-Amherst (공학박사)
- 1989년 11월~1992년 8월: 금성사 선임연구원
- 1992년 9월~현재: 동국대학교 전자공학과 교수
- 2001년 7월~2002년 8월: Stanford University 객원교수
- 주관심분야 : 영상분할, 영상검색, 워터마킹