주요 색상의 분포 블록기호를 이용한 영상검색과 유사도 피드백을 통한 이미지 검색

(Image Retrieval using Distribution Block Signature of Main Colors' Set and Performance Boosting via Relevance Feedback)

박 한 수 † 유 헌 우 †† 장 동 식 †††
(Han-Soo Park) (Hun-Woo Yoo) (Dong-Sik Jang)

요 약 이 논문은 색상과 위치정보를 이용한 새로운 내용기반 영상검색 알고리즘을 제안한다. 이를 위해서, 질의가 주어졌을 경우, 데이타베이스의 검색공간을 줄일 목적으로 두 가지 종류의 색인 키(Key)를 제시하는데 하나는 영상의 고유한 색상 구성적 특성을 나타내는 주요 색상세트(*MCS, Main Colors' Set*)이고 다른 하나는 주요 색상마다의 분포 및 위치적 특성을 나타내는 분포 블록기호(*DBS, Distribution Block Signature*)이다. 이 두 가지 필터(Filter)를 연속적으로 적용하면 영상 데이타베이스로부터 잠재성이 높은 유사 후보 영상만을 걸러내게 된다. 이어서 보다 높은 검색성능을 얻기 위해 새롭게 제안한 쿼드 모델 (Quad Modeling)과 유사도 피드백 메커니즘을 이용한다. 이 방법은 색상과 위치정보에 대한 가중치를 역동적으로 조절함으로써 검색성능을 향상시킨다. 실험을 통해서 제안된 알고리즘이 성공적으로 영상검색에 사용될 수 있음을 보인다.

키워드 : 주요 색상세트(MCS), 분포 블록기호(DBS), 쿼드-모델, 내용기반 영상검색, 유사도 피드백

Abstract This paper proposes a new content-based image retrieval algorithm using color-spatial information. For the purpose, the paper suggests two kinds of indexing key to prune away irrelevant images to a given query image; MCS(Main Colors' Set), which is related with color information and DBS(Distribution Block Signature), which is related with spatial information. After successively applying these filters to a database, we could get a small amount of high potential candidates that are somewhat similar to the query image. Then we would make use of new QM(Quad modeling) and relevance feedback mechanism to obtain more accurate retrieval. It would enhance the retrieval effectiveness by dynamically modulating the weights of color-spatial information. Experiments show that the proposed algorithm can apply successfully image retrieval applications.

Key words : MCS(Main Color's Set), DBS(Distribution Block Signature), QM(Quad Modeling), Content Based Image Retrieval, Relevance Feedback

1. 서 론

디지털 영상을 제작할 수 있는 각종 하드웨어 장비가 나날이 발전하고 대중화됨에 따라 영상과 관련된 자료의 양은 기하급수적으로 늘어나고 있다. 이런 증가 추세는 앞으로 더욱 가속화될 것으로 예상되는데 이와 같은 방대한 영상자료를 효율적으로 관리하고 활용할 수 있는 방법에 대한 연구가 필요하다. 본 논문은 영상이 가지고 있는 고유의 내용특징을 이용하여 검색하는 '내용기반 영상검색'(Content Based Image Retrieval)을 다룬다.

일반적으로 내용기반 영상검색을 위해서는 영상의 '색상', '질감', '형태'의 세 가지 특징 요소를 주로 활용한다. 이 중에서도 특히 색상은 영상을 구성하는 가장 지배적인 요소로써 추출과 활용이 쉽고 빠르다는 장점을 가지고 있어서 다른 두 가지에 비해 훨씬 더 많이 사용한다. 색상을 이용한 고전적 영상검색 방법의 하나로 색상 히스토그램을 이용한 방법이 있다[1-3]. 이 방법은 영상이 갖는 색상구성을 히스토그램을 이용하여 쉽고

· 본 연구는 고려대학교 특별연구비에 의하여 수행되었음
† 비 회 원 : 고려대학교 산업시스템 정보공학과
vision@korea.ac.kr
†† 정 회 원 : 연세대학교 인지과학연구소 교수
paulyhw@yonsei.ac.kr
††† 정 회 원 : 고려대학교 산업시스템 정보공학과 교수
jang@korea.ac.kr
논문접수 : 2003년 2월 12일
심사완료 : 2004년 11월 4일

간편하게 영상검색을 실현한다. 그러나 공간적인 상호관계를 적절히 표현하지 못하므로 오늘날의 검색시스템은 색상과 그 색상간의 위치특성도 고려하여 보다 정확한 검색방법을 실현하기 위한 연구가 진행되고있다. 색상-위치 정보를 활용한 기존의 연구들로는 색상 히스토그램(Color Histogram)[5,6], 색상 코헤런스 벡터(Color Coherence Vector)[7], 색상 코릴로그램(Color Correlogram)[8] 그리고 바이너리 색상 세트(Binary Color Set)[9,10]등이 있다. 이러한 방법을 크게 3가지로 구분해서 설명하면,

1. 기호에 의한 방법(Signature-based technique),
2. 분할에 의한 방법(Partition-based algorithm),
3. 클러스터링에 의한 방법(Clustering-based method) 이다.

먼저 '기호에 의한 방법'은 영상을 일정한 크기로 분할한 후 각 셀(Cell) 단위로 특징을 추출해 기호로 만들어주고 이들 기호를 서로 비교함으로써 영상간 유사도를 결정하는 방법이다. '분할에 의한 방법'은 비교하는 두 영상을 일정한 크기로 분할한 후 상응하는 셀끼리 1:1로 쌍을 이루어 색상에 대한 유사도를 측정한 뒤 모든 쌍들의 값을 통합해 영상 전체의 유사도를 결정하는 방식이다. 마지막 '클러스터링에 의한 방법'은 영상을 구성하는 몇 가지의 주된 단일색상 단위로 영상을 분할한 뒤, 각 색상별 위치를 비교해 유사도를 얻는 방법이다. 일반적으로 '클러스터링에 의한 방법'이 비교적 정확하게 색상과 위치 정보를 나타내지만 추출이 어렵고 계산량이 많다는 단점을 갖는다. 이에 반해 '기호에 의한 방법'은 추출과 비교가 매우 간단하다. [11]이 제시한 결과에 따르면 효과성 측면(검색의 정확성)에서는 '기호에 의한 방법'과 '클러스터링에 의한 방법'이 서로 비슷한 수준을 보이고 '분할에 의한 방법'은 이 둘에 비해 다소 떨어지는 것으로 나타났다. 그리고 효율성 측면(검색속도)에서는 '기호에 의한 방법' > '분할에 의한 방법' > '클러스터링에 의한 방법'의 순으로 우수했다. 이 결과로부터 '기호에 의한 방법'이 효율성과 효과성 모두에서 일반적으로 우수하다는 사실을 알 수 있다. 특히 데이타베이스의 크기가 커지고 질의 영상이 아주 복잡한 구성을 갖지 않는 경우에 더욱 유용한 특성을 보여준다.

또한 관련연구로 영상을 빠르고 효율적으로 검색하기 위한 MPEG-7 표준화 작업도 진행중이다[12]. 색상정보를 이용한 표준화 대상으로는 Color Space, Dominant Color, Color Quantization, Color Structure, GoF/GoP Color 등이 있다.

본 논문에는 '기호에 의한 방법'을 채택하여 새로운 영상검색 알고리즘을 제안한다. 2차에 걸친 필터링(Filtering) 작업과 색상-위치 정보만을 이용한 방법으로 검색속도를 향상시키고, QM(Quadrant Modeling)이라고 이름 붙인 초기 가중치(Weight) 설정 기법 및 사용자의 유사도 피드백 과정으로 정확한 검색을 실현하고자 한다.

2. 제안 알고리즘의 개요

본 논문은 웹(Web)이나 대용량 데이타베이스로부터 영상을 효율적으로 검색할 수 있는 시스템의 설계를 그 목적으로 한다. 이를 위해 시스템의 설계 방향을 아래의 표 1과 같이 설정하였다.

표 1 시스템 설계 방향

데이타베이스의 특성	고려 사항	접근 방식
저장/ 검색 요구	텍스트기반 검색 제한	내용기반 영상검색
잦은 자료의 출입	자료의 구조화 곤란	개별 색인화
D/B의 분산화		
검색 공간의 방대	효율성 제한	색상-위치정보 이용 검색공간 필터링
	효과성 제한	유사도 피드백

본 논문에서는 신속한 자동 검색이 가능하도록 내용기반 영상검색의 방법을 택하였다. 또한 데이타베이스는 그 특성상 자료의 출입이 잦고, 웹의 경우에 자료가 여러 곳에 분산되어 있는 특성 때문에 효율적인 검색을 위한 자료의 구조화가 곤란하다. 따라서 개별 색인화라는 방법을 통해 이런 제한을 극복하고 검색 공간을 대폭 줄일 수 있도록 하였다. 그리고 검색 효율성을 향상시키기 위해 색상-위치 특성만을 이용한 간단한 방법을 택하되 동시에 검색의 효과성도 보장하기 위해서 사용자의 능동적인 유사도 피드백 과정을 포함시켰다. 그림 1은 전체 알고리즘의 개요를 보여준다.

일단 시스템 'ANGEL'에 질의 영상이 입력되면 MCS와 DBS라는 두 가지 종류의 색인 키(Key)를 추출해 이를 데이타베이스에 저장된 영상들의 GCS 그리고 DBS라는 키와 각각 비교한다. MCS/GCS는 색상구성

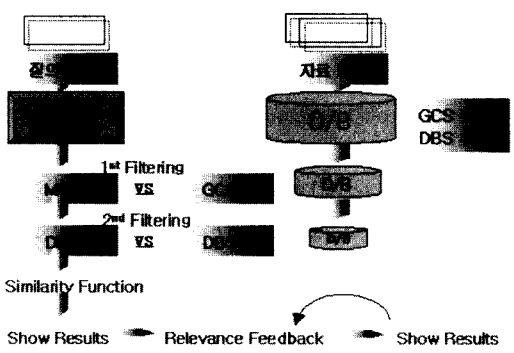


그림 1 알고리즘 개요도

과 관련된 키이고, DBS는 색상분포와 관련된 키이다. 이 두 가지 색인 키의 비교를 통해 방대한 데이타베이스로부터 잠재력이 높은 후보 영상만을 추출한다. 이어서 이 후보 영상만을 대상으로 보다 정밀한 유사도 검색을 실시한 뒤 그 결과를 다시 사용자에게 제시한다. 만약 사용자의 의도가 적절히 반영되지 않았다면 사용자는 다시 능동적으로 유사도 피드백을 해줌으로써 보다 정확한 영상을 얻을 수 있게 된다.

본 연구가 기존의 다른 연구와의 차별성은 아래와 같다.

1. 주요 색상정보를 MCS bit열로 표현하여 기호에 의한 방법의 장점을 살리고 많은 데이타베이스 영상에서 유사한 영상만을 먼저 필터링 할 수 있는 방법을 제시하였다.

2. 주요 색상을 DBS bit열로 표현함으로써 위치정보까지 포함한 보다 정확한 검색을 할 수 있는 방법을 제시하였다.

3. QM(Quadratic Modeling)을 통해서 주요 색상의 민감한 분포, 위치변화, 회전변화 등에 강인한 가중치 설계방법을 제안하였다.

4. 제안된 방법에 적합한 유사도 피드백(relevance feedback)을 제시하여 검색 시 보다 사용자의 의도를 반영하도록 하였다.

3. 영상 전처리

본 논문은 색상을 이용하여 영상을 검색하는 알고리즘을 제안하는 것이다. 색상을 표현하기 위해서 인간의 인지도와 유사한 HSI 색상모델이 사용되었고 몇가지 주요한 색상만을 표현하기 위해 양자화(Quantization)를 실시하였다. RGB에서 HSI모델로의 변환식은 (1)~(3)을 따른다.

$$h = \begin{cases} \theta & g \geq b \\ 2\pi - \theta, & g \leq b \end{cases},$$

$$\theta = \cos^{-1}\left[\frac{\frac{1}{2}[(r-g)+(r-b)]}{\sqrt{(r-g)^2+(r-b)(g-b)}}\right] \quad (1)$$

$$s = 1 - \frac{3}{r+g+b}[\min(r,g,b)] \quad (2)$$

$$i = \frac{1}{3}(r+g+b) \quad (3)$$

일반적으로 사람이 채도(Saturation)나 명도(Intensity)보다는 색조(Hue) 값에 훨씬 민감한 점을 바탕으로 채도와 명도는 각각 단 3개의 범주로 나누고 색조는 매 20도 단위로 나눠 총 18개의 범주로 만들어 주었다. 이렇게 하면 총 162개의 양자화된 색상 빈(Bin)이 만들어지는데 여기에 다시 4개의 그레이 값을 포함시켜서 총 166개의 색상 빈으로 양자화한 값을 사용한다. 이처럼 HSI모델의 공간을 벡터 양자화한 후, 그 중앙값을 해당 빈의 대표 값으로 삼아주게 된다.

크기가 다른 영상들을 서로 비교하기 위해서는 크기로부터 독립된 비교 방법이 필요하다. 이를 위해 영상의 크기를 가로 세로 32등분하여 총 1024(=32×32)개의 셀을 얻는다. 이렇게 영상을 분할하면 셀의 크기는 다르지만 수량은 모두 똑같이 1024개의 셀을 갖는 영상이 된다. 이처럼 영상을 픽셀 단위로 표현하지 않고 셀 단위로 나누어 사용하게 되면 크기에 독립적일 수 있다는 장점뿐만 아니라 영상의 특성을 충분히 표현하면서도 계산량을 크게 줄일 수 있고 또한 각 셀의 좌표 값으로 영상의 색상-위치 정보를 간단히 나타낼 수 있게 된다[11].

4. MCS/GCS 추출

본 논문에서는 고속검색을 위한 목구조(R-Tree, S-Tree등)를 사용하지 않고 개별색인 방법을 이용하고 있다. 대신에 질의 영상에서 추출한 4개의 주요한 색상을 표현하는 MCS와 데이타베이스 영상이 가지고 있는 다양한 색상을 표현하는 GCS간의 AND연산(공통부분 추출)을 이용하여 질의 영상의 주요한 색상을 포함하지 않는 데이타베이스 영상을 미리 제거하는 방식으로 고속검색을 실현한다.

4.1 MCS (Main Colors' Set) 추출

일반적으로 영상은 몇 가지의 주요 색상만으로 대표되는 성질이 있다[11]. 따라서 이들 색상만을 갖고 유사한 영상인지 여부를 구분 짓는 색인 키로 활용이 가능하다. 여기에 다시 공간적 위치 정보를 활용해 보완해 준다면 상당한 효용성을 가질 수 있다. 너무 많은 색상을 고려한다면 불필요한 계산이 늘어나게 되고 또한 의미 없는 색상마저도 포함되기 때문에 노이즈(Noise) 같은 역할을 해 오히려 성능을 저하시킨다. 따라서 본 논문에서는 경험적 방법을 통해 단 4가지로 색상의 수를 제한하고 이를 주요 색상으로 분류한다. 이것이 바로 MCS(Main Colors' Set)이다. 즉 MCS는 총 4가지의 MC(Main Color)로 이루어진다. 4개의 주요 색상을 사용하는 또 다른 이유는 인간이 사물을 인식할 때 보통 2-3개의 주요한 색상만을 가지고 인식하고[13], 영상 중에는 매우 다양한 색상 분포를 가지는 영상도 존재하기 때문에 이런 이유를 고려하여 4개의 색상을 사용하였다.

MCS의 추출 절차는 먼저 32*32개의 셀로 나뉜 질의 영상에 대해서 각 셀별로 히스토그램(H_{CELL})을 확인해 가장 높은 빈도를 갖는 색상으로 해당 셀의 모든 픽셀 값을 바꿔준다. 이 때 색상은 166가지의 양자화된 색상을 의미한다. 이 작업을 마치면 그림 2의 두 번째 그림과 같이 마치 모자이크 처리된 것과 같은 영상을 얻게 된다. 이것은 마치 해상도가 낮은 영상을 보는 느낌을

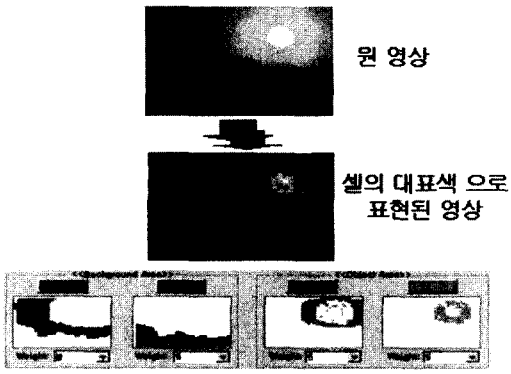


그림 2 MCS의 추출

주지만 영상의 대략적인 색상 및 위치정보를 고스란히 담고 있게 된다. 이제 다시 모자이크 영상 전체에 대한 히스토그램(H_{IMAGE})을 구해 가장 빈도가 높은 색상 4가지만을 내림차순으로 뽑아 내는데 이 네 가지 색상 각각이 바로 MC가 되는 것이다. 그림 2의 맨 아래 영상들은 원영상으로부터 이 네 가지 MC별로 각각의 이진 영상만을 뽑아낸 것이다.

일반적으로 영상에서 배경이 차지하는 면적은 넓은 반면 중심 개체는 상대적으로 적은 면적을 차지한다. 따라서 단순히 히스토그램만을 이용해 MCS를 추출한다면 배경 영역이 차지하는 비중에 따라 중심 개체의 정보가 완전히 사장되는 경우가 생긴다. 이런 문제를 해결하기 위해서는 원영상에 대해 중심 영역과 배경 영역의 두 부분으로 나누어 MCS를 뽑아내는 것이 바람직하다. 그림 3은 영상의 배경 영역과 중심 영역에 대한 구분을 보여주고 있다. 먼저 (a)는 기본 값으로 설정된 중심 영역과 배경 영역을 보여주고 있고 (b)는 사용자가 직접 중심 영역을 재선정한 경우를 보여주고 있다. 중심 개체가 영상의 중앙에 위치하고 있지 않은 경우나 또는 개체의 크기가 너무 크거나 작은 경우에 이와 같은 영역의 재선정이 필요하다. 이렇게 영역을 구분한 뒤 중앙 영역에 대한 히스토그램 H_C와 배경 영역에 대한 히스토그램 H_B를 구해 각 영역에 대한 MC_C와 MC_B를 2가지씩 뽑아내게 된다. 그 순서는 먼저 배경 영역에서 MC_B를 두 가지를 뽑은 다음, 중심 영역에 대한 히스토그램 H_C로부터 이 두 가지 MC_B를 배제한 상태로 나머지 색상만을 대상으로 MC_C 둘을 뽑아낸다. 일반적으로 인간의 영상에 대한 인지정보는 주로 중심영역의 색상에 영향을 많이 미치게 되므로 단순히 영역 구분 없이 영상 내의 4개의 주요 영상을 추출하게 되면 배경영역의 색상들이 주요하게 추출될 가능성이 많다. 따라서 본 연구에서는 중심과 배경간에 독립적으로 겹치지 않는 범위 내에서 주요한 색상들을 추출하는 것이다.

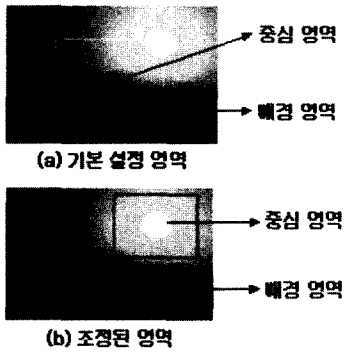


그림 3 배경영역과 중심영역

사람은 색조(Hue), 채도(Saturation), 밝기(Brightness) 중에서 특히 색조 성분에 대단히 민감한 특성이 있다. 따라서 색조 정보만 갖고도 간단히 유사한 영상인지 여부를 판별할 수 있다. 따라서 색조 정보를 표현하는 간단한 기호로 만들어 필터링을 위한 색인 키로 활용할 수 있는데 이것이 바로 MCS 비트열이다. 이 비트열은 앞선 양자화 과정과 마찬가지로 색조를 18개의 빈으로 나누어 각 빈마다 1비트를 할당하되 채도는 무시하고 그레이 값에는 단 1비트만 할당하는 길이 19짜리의 형태를 갖는다. 이 비트열에는 네 가지 MC에 대한 색조나 혹은 그레이 정보가 복합되어 들어간다. 각 비트의 기본값은 0이고 MC들의 색조 값을 확인해 해당 비트 값을 1로 만들어준다. 이 때 두 가지 이상의 MC가 하나의 색조 빈에 들어갈 수도 있다. 4가지 MC중에 r, g, b가 동일한(r=g=b) 그레이가 존재할 경우 g_0를 1로 만든다.

$$MCS_{bit_stream} = \underbrace{h_0, h_1, h_2, h_3, h_4, h_5, h_6, h_7, h_8, h_9, h_{10}, h_{11}, h_{12}, h_{13}, h_{14}, h_{15}, h_{16}, h_{17}}_{Hue}, \overbrace{g_0}^{Gray}$$

4.2 GCS (Global Colors' Set) 추출

GCS(Global Color Set)는 영상에 담겨져 있는 모든 종류의 색상으로 이루어진 집합을 의미한다. MCS가 질의 영상의 1024개의 셀에서 주요한 4개의 색상만을 추출하여 19개의 비트열에 해당 색상의 위치를 갖는 빈을 1로 설정해주는데 반해 GCS는 데이타베이스에 존재하는 각 영상에서 추출하는데, 각 영상의 1024개의 모든 셀에서 색상을 추출하여 해당 색상이 있는 비트열을 1로 설정해 준다. 따라서 색상분포가 다양한 영상일수록 비트열내에 1의 개수가 많아지고 색상분포가 단순한 영상일수록 1의 개수가 적어진다(MCS는 최대 4개 이상의 1을 가질 수 없다).

5. 필터링(Filtering)

5.1 1차 필터링

질의 영상과 유사한 영상을 데이타베이스에서 검색할 때 모든 영상간의 유사도를 비교하면 계산량이 많아진다. 따라서 본 논문에서는 먼저 MCS와 GCS를 가지고 1차 필터링을 함으로써 질의와 유사한 영상만을 먼저 걸러낸다. 필터링은 질의 영상의 MCS 비트열과 비교 영상의 GCS 비트열을 서로 AND 연산을 하여 그 결과값이 MCS 비트열 자체와 같은 영상만을 후보영상으로 먼저 뽑아내는 것이다. 그림 4에서 이 연산 과정을 보여주고 있다.

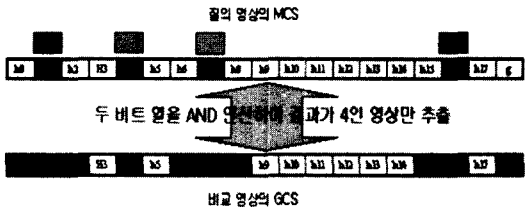


그림 4 1차 필터링 과정도

이와 같이 AND연산을 수행하면 질의 영상의 4개의 MC와 유사한 색상을 가지고 있는 데이타베이스영상만을 먼저 추출함으로써 질의와 유사하지 않은 색상을 가지고 있는 영상은 유사도 계산에서 미리 제외시킬 수 있다.

5.2 DBS(Distribution Block Signature)의 추출

DBS는 색상의 공간적 배치 및 분포를 확인하여 상이한 영상들을 걸러주기 위한 2차 필터링 과정에서 활용되는 기호이다. DBS의 생성과정을 먼저 영상을 4*4개의 블록으로 나누어준다. 이제 임의의 MC에 대해서 각 블록이 동일한 MC를 가진 셀을 단 하나라도 포함하고 있는지 여부를 확인해, 있는 경우에는 해당 블록을 1로, 없다면 0으로 만들어준다. 이렇게 하면 결국 길이 16비트의 DBS 비트열이 만들어진다. MC가 네 종류이기 때문에 길이 16의 비트열이 총 네 개 생긴다. 아래의 그림 5는 이 과정을 보여주고 있다.

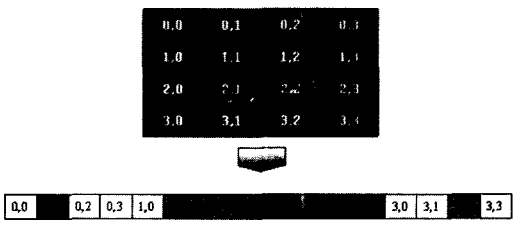


그림 5 DBS 비트열의 생성

5.3 2차 필터링

1차에서 걸러진 후보 영상을 대상으로 각각의 DBS를 비교해 좀 더 잠재성이 높은 영상만을 2차 후보로 걸러내게 된다. 이 비교에서는 XOR 연산을 통해 서로의 불일치도를 결정하는데 그 이유는 각 블록상에 임의의 MC가 있는 곳과 없는 곳이 서로 정확히 일치할수록 유사하기 때문이다. 그림 6은 2차 필터링의 과정을 보여주고 있다. 결과적으로 MC가 네 종류이므로 전체적으로는 [0,64] 범위의 불일치도를 갖게 된다. 본 논문에서는 43을 임계치로 삼아 그 이상의 값을 갖는 영상들은 후보에서 제외시키고 있다.

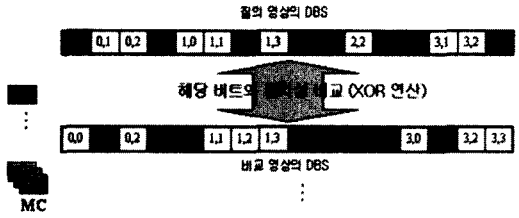


그림 6 2차 필터링 과정도

일반적으로 인간이 영상간의 유사도를 판단할 때 색상 정보 등이 가장 중요하지만 특정 영역에서는 오히려 모양 특징과 같은 다른 정보 등이 보다 중요할 수 있다. 본 연구에서는 색상의 색조(hue)값과 위치정보만을 가지고 필터링을 하기 때문에 상표검색등과 같이 모양이 중요한 영향을 미치는 영역에서는 만족할 만한 결과를 얻지 못할 수도 있다. 또한 HSI 색상 모델의 채도(saturation), 명도(intensity)정보 등을 이용하지 않기 때문에 이러한 정보가 영상간 유사도의 중요한 요소가 될 경우에는 낮은 검색 성능을 나타낼 수도 있다. 그러나 이러한 특정한 영역을 제외한 일반적인 영상검색에는 제안된 방법이 효과적으로 사용될 수 있다.

6. 유사도 비교

6.1 QM을 통한 가중치 설정

그림 7에 QM의 생성 절차를 보여주고 있다. 왼쪽의 (a)계열은 임의의 MC에 대한 이진 영상이고 오른쪽 (b)계열이 가중치 공간을 의미한다. 가중치의 각 원소는 초기 값으로 1을 가진다.

먼저 왼쪽의 이진 영상을 (a2)처럼 4개의 분면으로 나누어준다. 그런 다음 각 분면 안에 해당 MC가 하나라도 존재할 경우 해당 위치의 가중치를 2라는 값으로 새로 갱신한다. 단, 존재유무를 확인하는 데는 앞서 32*32로 나눠준 영상의 각 셀이 그 단위가 된다. 이러한 작업을 그림에서 보듯이 총 다섯 번에 걸쳐 재귀적으로 실시한다. 이러한 작업이 모두 끝나면 각 색상과 위치(셀)별로 차별화된 가중치를 얻게 되는데 이 가중치

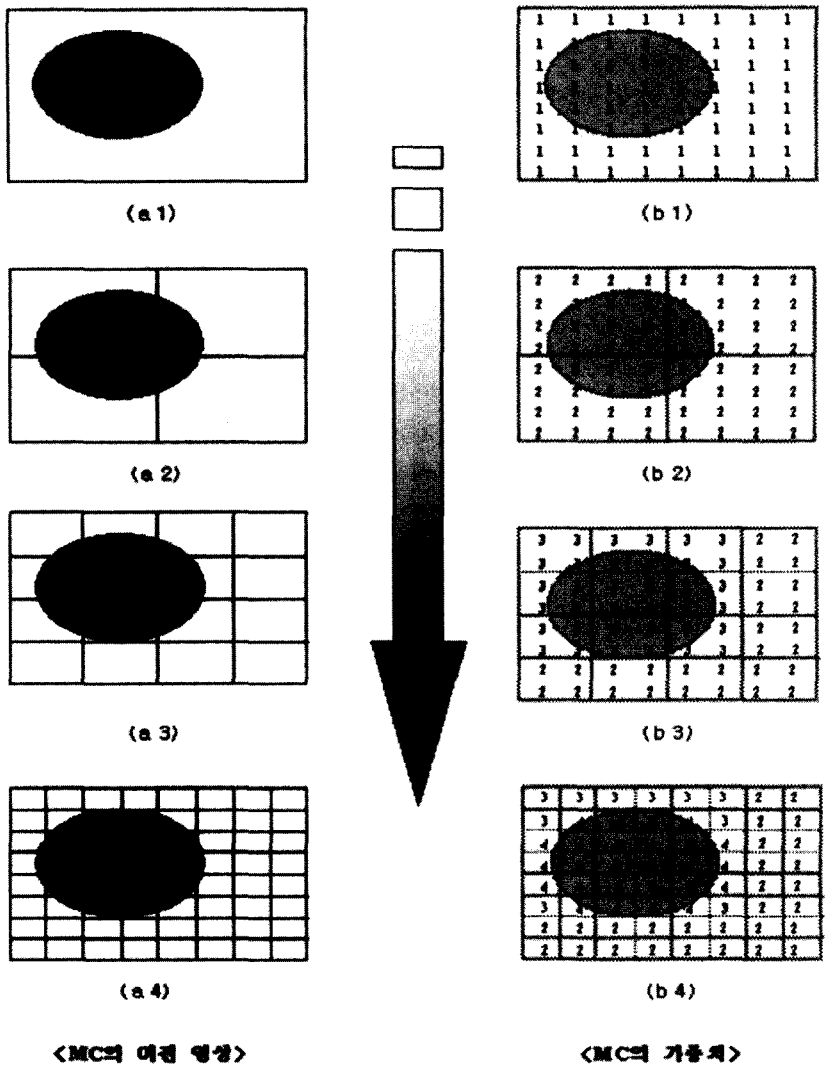


그림 7 QM에 의한 초기 가중치 설정

는 각 색상이 갖는 위치적 특성을 고스란히 표현할 수 있다. 단 그림에서는 설명을 위해 1~4의 순차적인 값으로 가중치를 갱신했지만 실제로는 분할된 분면의 크기에 반비례하는 값으로 갱신한다. 즉 처음에는 1/256, 다음에는 1/64, 1/16, 1/4, 1의 순서로 갱신해주게 된다.

QM은 2차원 영상 평면을 4진 트리 구조로 표현하기 때문에 영상 평면에 대한 입체적 접근이 가능하다. 또한 해당 MC에 가까운 셀이 큰 가중치 값을 갖게 되고 멀어질수록 점진적으로 작은 값을 갖게 되어 색상에 대한 위치 특성을 자연스럽게 표현할 수 있다. 따라서 색상에 대해서 어느 정도 분포 차이, 회전, 위치 변화 등에 대해 강인한 성질을 갖는다.

이와 같이 QM을 통하면 각각의 MC별로 32×32의 셀들에 대해 가중치 행렬을 얻게 된다. 이후에는 이 기본 가중치 행렬에 대해 영상의 중앙영역을 보다 강조하고 외곽영역을 상대적으로 약화시키는 방법으로 가중치를 다시 변경한다.

일반적으로 사람이 영상을 인식할 때 영상의 중앙부에 있는 색상은 외곽부에 있는 색상에 비해 중요하게 지각된다. 따라서 중앙부와 외곽부에 가중치를 차별화하게 되면 보다 인지도에 가까운 영상검색을 실시할 수 있다. 본 논문에서는 일반적인 영상에 대해 그림 8에서처럼 중심 영역: 배경 영역A: 배경 영역B로 구분한 뒤 각각에 대한 가중치를 3 : 1.4 : 1로 보정하였다.

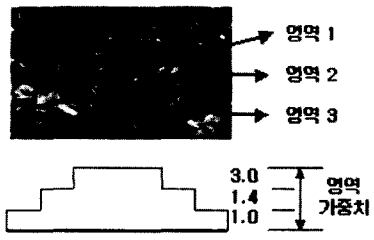


그림 8 영역 가중치 보정

6.2 유사도 비교 함수

영상간 유사도를 비교하는 함수를 S로 정의하고 다음과 같이 계산한다. 먼저 임의의 MC에 대한 질의 영상과 후보 영상간의 유사도는 서로간의 색상중첩정도를 표현하는 아래의 식을 통해 구한다.

$$S_{MC[k]} = \sum_{i=0}^{31}\sum_{j=0}^{31} Wt_{MC[k]}[i][j] \times bImage_{MC[k]}[i][j] \quad (4)$$

여기서 $S_{MC[k]}$는 k에 해당하는 MC의 유사도 비교 함수를 나타내고 $Wt_{MC[k]}[i][j]$는 질의 영상에서 k에 해당하는 MC의 (i,j)번째 가중치 (즉, QM과 영역 보정 후의 가중치)를 의미한다. 물론 보정된 초기 가중치이다. 그리고 $bIMAGE_{MC[k]}[i][j]$은 후보 영상에서 k에 해당하는 MC의 이진 영상 중 (i,j)번째 셀에 대한 비트 값을 의미한다.

마지막으로 후보 영상에 대한 유사도 비교의 결과 4가지 MC를 모두 고려하면 값은 다음과 같다.

$$S = S_{MC[0]} + S_{MC[1]} + S_{MC[2]} + S_{MC[3]}$$
$$= \sum_{k=0}^{3}\sum_{i=0}^{31}\sum_{j=0}^{31} Wt_{MC[k]}[i][j] \times bImage_{MC[k]}[i][j] \quad (5)$$

이 유사도 함수를 적용해서 얻은 결과 값이 클수록 영상간 유사도가 높다고 말할 수 있다.

6.3 유사도 피드백

식 (5)를 통해서도, 검색된 영상이 사용자가 판단하기에 질의와 유사하지 않은 영상이 있을 수 있다. 이는 색상자체가 저차원적인 영상특징이며 인간의 의미론적 유사도 개념을 반영하지 않은 것이기 때문이다. 따라서 본 논문에서도 유사도 피드백을 통하여 보다 고차원적인 검색을 실시하여 검색 성능을 높이고자 한다.

예를 들어 시스템이 검색한 영상 중 3장의 영상이 유사한 영상으로 체크되었을 경우 이들을 바탕으로 각 셀별 중첩도를 구한다. 그러면 3번 중첩되는 셀부터 시작해 2번, 1번 그리고 단 한번도 중첩이 없는 셀로 각각 나누어 진다. 그런 다음 최초 1을 그 초기값으로 갖고 있는 유사도 피드백 행렬의 각 원소를 중첩도에 따라 (6)과 같이 새로 갱신해준다.

$$RF_{matrix}[i][j] = 1 - (n - Conut) * (Step_Size), \ \forall\ i,j = 0,1,2,\ldots,31 \quad (6)$$

여기서, $RF_{matrix}[i][j]$는 피드백을 위한 32*32 크기의 행렬($Matrix$)이고 n은 사용자가 유사하다고 고른 영상의 총 개수, $Count$는 각 셀별 중첩의 횟수를 의미한다. $Step\ Size$는 가중치에 대한 조절의 범위로서 완만한 근사를 위해 0.01로 하였다.

이렇게 유사도 피드백 행렬인 $RF_{matrix}[i][j]$의 모든 원소 값을 얻은 후 (식 (7))와 같이 질의 영상의 MC별 가중치인 $Wt_{MC[k]}$에 곱한다.

$$Wt_{MC[k]}[i][j] = Wt_{MC[k]}[i][j] \times RF_{matrix_MC[k]}[i][j] \quad (7)$$

여기서 $Wt_{MC[k]}[i][j]$는 MC[k]에 대한 (i, j)번째 셀에서의 유사도 결정 가중치를 나타내고 $RF_{Matrix_MC[k]}[i][j]$는 유사도 피드백 행렬의 (i, j)번째 원소가 갖는 가중치 조절 값을 의미한다.

유사도 피드백에 관한 연구로는 [3, 4]에서처럼 벡터 공간모델에서 검색결과 영상의 유사한 것과 유사하지 않은 것들의 특징 벡터들의 가중선형결합으로 새로운 질의를 형성하는 방법과, [14]에서처럼 베이시안 방법(Bayesian Approach)을 통해서 찾고자 하는 목표영상(Target Image)이 정해지면, 사용자의 행동(Action)에 따라 시스템이 찾고자 하는 목표 영상이 무엇인지를 확률모형에 의해 예측해 내는 방법 등이 있지만, 이와 같은 방법은 본 연구에서와 같이 분할된 셀의 색상을 가지고 색상과 위치정보를 표현한 기호에 의한 검색에서는 적용할 수 없으므로 본 연구에서는 위와 같이 분할된 셀에서 동일위치에서의 색상의 중첩도를 가지고 유사도 피드백을 실현하였다.

이제 이와 같이 갱신된 유사도 결정 가중치를 갖고 앞서 설명한 유사도식 (5)에 적용해 다시 후보 영상을 대상으로 검색할 수 있다.

본 논문에서는 또한 1차 검색 후에 사용자가 직접 $S_{MC[0]}$~$S_{MC[3]}$ 각각의 가중치도 조절할 수 있도록 GUI를 설계하였다. (5)의 최종 유사도 함수 S는 $S_{MC[0]}$~$S_{MC[3]}$ 각각의 색상 가중치를 디폴트로 모두 1로 한 경우이다. 그러나 수동으로 질의 영상에서 뽑은 MCS에 대해 사용자가 관심을 갖고 있는 개체에 관한 색상의 가중치는 높여주고 그 반대인 경우는 낮추거나 아예 0으로 하여 유사도 검색에서 배제시킬 수 있다. 이처럼 사용자의 관심 영역에 대한 의도를 유연하게 반영하여 검색의 정확성을 높여준다. (8)은 색상 가중치를 반영한 유사도의 계산식을 보여준다.

$$S = \omega_{MC[0]} * S_{MC[0]} + \omega_{MC[1]} * S_{MC[1]} + \omega_{MC[2]} * S_{MC2} + \omega_{MC[3]} * S_{MC[3]} \quad (8)$$

여기서 $\omega_{MC[0]}$~$\omega_{MC[3]}$는 각각 MC[0]~MC[3]에 대해 사용자가 설정한 색상 가중치이다.

7. 실험 및 결과 분석

7.1 실험 환경과 평가 척도

본 논문에서 제안한 새로운 알고리즘의 영상검색 성능을 평가하기 위해 Window2000 환경에서 Pentium-Ⅲ 1GHz, Ram 256M 사양을 가진 PC를 사용하였고 Visual C++을 이용하여 프로그래밍하였다. 실험 데이타는 크게 14개의 범주(비행기, 치타, 코끼리, 북극곰, 호랑이, 산, 돌고래, 말, 장미, 독수리, 풍경, 일몰등등)로 구성된 1000장의 영상을 사용하였다. 사용자 인터페이스(GUI)는 그림 9와 같다.

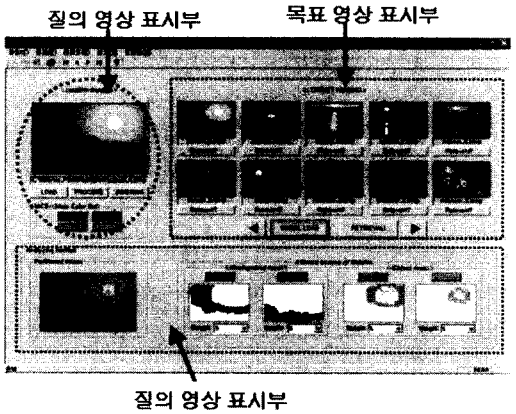


그림 9 GUI(Graphic User Interface) 구성

제안한 알고리즘의 성능을 평가하기 위하여 크게 효율성(검색속도)과 효과성(검색의 정확성)의 두 부분으로 나누어 실험하였다. 효율성에 대한 실험은 실험 데이타 전체를 대상으로 평균적 필터링 효과를 측정해 알아보았다. 필터링은 탐색 공간을 축소해주기 때문에 검색 소요시간의 단축과 비례하는 특성을 갖기 때문이다. 효과성은 정확성(Precision)과 회수율(Recall)의 두 가지 관점에서 평가하였다.

$$\text{정확도(Precision)} = \frac{N_{correct}}{N_{correct} + N_{false}} \quad (9)$$

$$\text{회수율(Recall)} = \frac{N_{correct}}{N_{total}}, \text{ 단 } N_{tatal} = (N_{correct} + N_{missed}) \quad (10)$$

여기서 N_{tatal}은 데이타베이스 안에 있는 질의 영상과 유사한 영상의 총 수량을 나타낸다. 그리고 $N_{correct}$는 검색된 영상 중 질의 영상과 유사한 영상의 수량을 나타내며 N_{false}는 검색된 영상 중에 유사하지 않은 영상의 수량을 나타내는 것이다. 한편 N_{missed}는 유사한 영상임에도 불구하고 검색되지 못한 채 데이타베이스 안에 남아있는 영상의 개수를 의미한다.

마지막으로 사용자의 유사도 피드백이 얼마나 유효한지를 알아 보이기 위해 한 차례 유사도 피드백을 한 결과와 제안된 알고리즘을 [4]와 비교하였다.

7.2 실험 결과 및 분석

7.2.1 효율성(1/2차 필터링)

1차 필터링의 경우, 1000장의 실험 영상을 대상으로 GCS 비트열 구성(비트 값의 합)을 살펴보았는데 각 영상은 3/19~19/19에 이르는 다양한 구성을 가지고 있었다. GCS 비트열 구성의 평균은 12.058/19이고 그 표준편차는 약 3.923였다. 그리고 이들의 MCS 비트열 구성은 약 3.04/19로 나타났다. 이들 평균값을 바탕으로 GCS의 구성이 12/19인 데이타베이스에 대해 MCS가 3/19인 질의영상을 검색할 때 그 필터링 기대치를 계산해 보면 아래와 같다.

GCS의 종류: ${}_{A}C_{B}={}_{19}C_{12}$, MCS의 종류: ${}_{A}C_{C}={}_{19}C_{3}$

$$\text{필터링 기대치} = \left(\frac{{}_{B}C_{C}}{{}_{A}C_{C}}\right) \times \text{record number}$$

$$= \left(\frac{{}_{12}C_{3}}{{}_{19}C_{3}}\right) \times 1000 = (0.227038) \times 1000 \approx 227.04$$

여기서 *record number*는 데이타베이스의 자료량을 나타낸다.

즉, 1차 필터링 후에는 1000장의 영상 중 약 227장의 후보 영상만을 추출할 것으로 기대할 수 있으므로 77.3%의 검색공간 축소 효과를 볼 수 있다.

이제 이 227장의 후보들만이 2차 필터링 과정으로 들어가 각각의 DBS를 비교하게 된다. 이제 영상전체의 불일치도를 확인해 임계치인 43보다 큰 영상들을 배제해주면 대략 33% 정도가 걸러지게 되므로 결국 평균적인 후보 영상의 수는 152장이 된다. 즉, 이로써 1,2차 필터링을 통해 약 84.80%의 탐색공간 축소 효과를 보임을 알 수 있다. 특히 사용자가 능동적으로 중심 영역을 재선정해줄 경우, MCS 비트열이 갖는 구성은 4/19에 더욱 가까워지게 될 것이고 이에 따라 필터링의 효과는 더 커지게 된다.

7.2.2 효과성

1000장의 영상을 대상으로 회수율과 정확성에 대해 실험한 결과를 그림 10에 담았다. 전체적인 평균 정확성은 72.27%이고 회수율은 34.04%이다. 정확성과 회수율이 서로 비슷한 형태를 띠고 있음을 볼 수 있으나 단 11번 범주(장미)에서는 회수율이 급격히 상승한 점이 주목된다. 장미의 경우 검색 결과의 수량어 다른 영상에 비해 상대적으로 적은 반면에 결과의 정확성은 상당히 높았다. 이것은 장미 영상에 포함된 색상이 다른 영상에서는 그리 흔하지 않았기 때문으로 판단된다. 또 검색 결과의 형태를 보면 8번(말), 11번(장미), 12번(일몰) 범주의 정확성이 평균 이상인 것으로 나왔는데 이들은 영

상 전체에서 MC 네 가지가 차지하는 면적이 다른 영상들에 비해 상대적으로 높기 때문에 영상 검색간 변별력이 더 커진 결과로 판단된다. 반면에 1번(비행기)과 4번(돌고래) 범주의 영상은 하늘색과 바다색이 유사하게 판단되어 서로의 유사영상으로 분류되는 현상이 많이 발생했다. 이에 따라 정확성은 평균 이하 였지만 회수율 면에서는 대체로 안정된 면을 보여주었다.

7.2.3 유사도 피드백

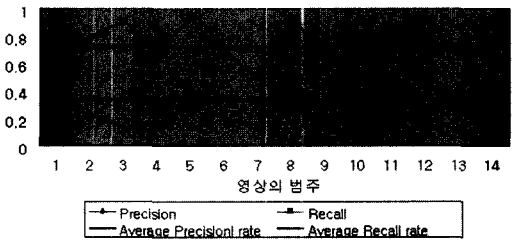
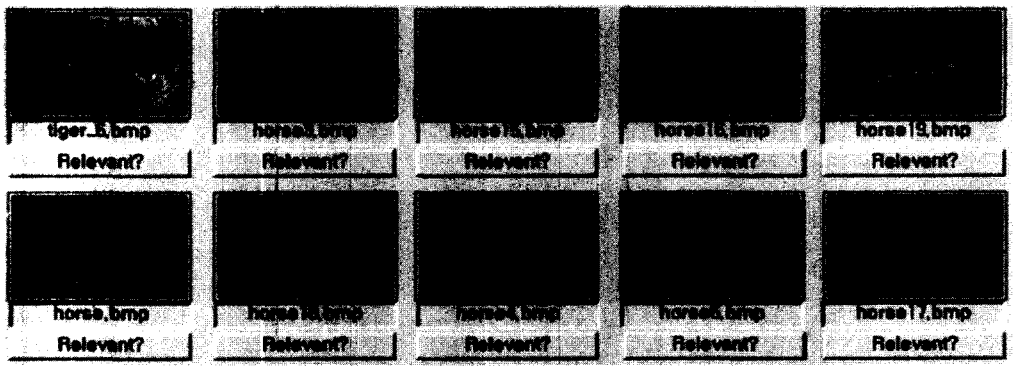
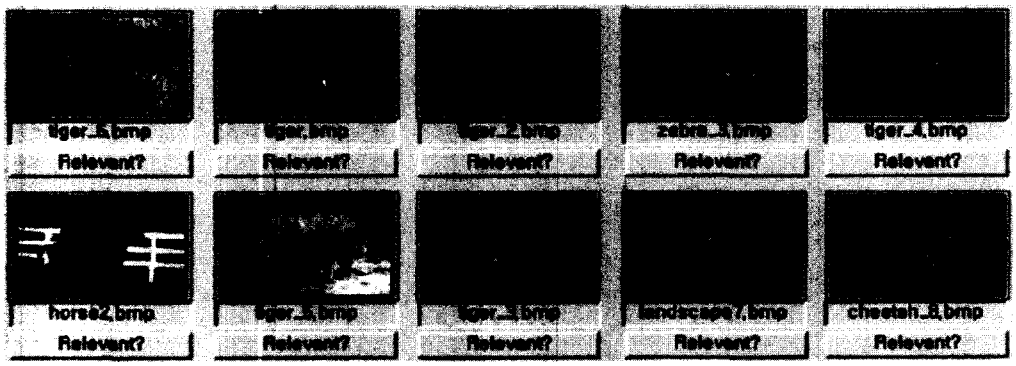


그림 10 정확성과 회수율

유사도 피드백 과정이 갖는 효용성을 확인하기 위해 앞선 실험에서 각 범주별로 최하의 정확성을 보여준 영상들을 대상으로 유사도 피드백 효과를 측정하였다. 실험의 공정성을 기하기 위해 본 시스템에 비전문적인 실험자(Subject)를 선정하여 시스템의 유사도 피드백 과정 및 절차에 대한 소개를 실시한 후 실험에 임하도록 하였다. 또한 유사도 피드백의 횟수는 단 1차례로 한정하여 실험하였다. 실험 영상 중 의미론적 개념이 강한 2번(산), 9번(풍경)의 두 범주는 실험에서는 제외하고 나머지 12개의 범주에 대해서만 실험하였다. 실험 결과, 최초에 정확성 평균이 16.185%이었던 것에 비해 1차례의 유사도 피드백 후에는 25.062%로 향상되었다. 이는 약 8.877% 정도 향상된 것으로 최초의 정확성에 비하면 약 54.85%의 향상효과를 보여주었다. 그림 11은 13번(호랑이) 범주에 대한 유사도 피드백의 예를 보여주고 있다. 여기서 보듯이 최초의 검색에서는 호랑이를 둘러싼 녹색 풀의 배경과 호랑이의 황색이 강하게 반영되어 유사한 계열의 말 영상이 다수 검출되었으나 유사도 피드백



(a) 최초 영상검색 결과



(b) 1차 유사도 피드백 후 영상검색 결과

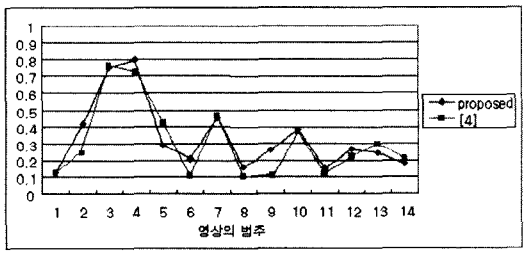
그림 11 Tiger 6의 유사도 피드백

을 통해 색상의 가중치를 조절해 주고 유사한 영상을 체크한 결과 다음 검색에서는 여섯 장의 유사 영상을 검색하고 있다.

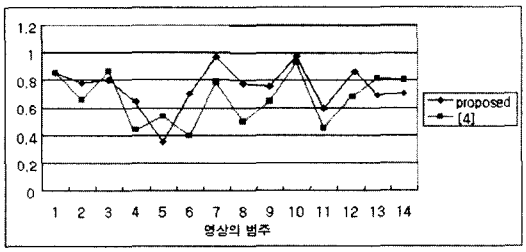
7.2.4 다른 알고리즘과의 비교

제안한 알고리즘의 객관적 성능을 비교, 평가하기 위해 [4]에서 제안한 시스템과 비교하였다. 이 시스템은 색상과 질감, 형태의 세 가지를 모두 고려하여 영상을 검색하는 시스템이다. 먼저 실험 조건은 동일한 데이타베이스를 대상으로 하여 상호간 성능 비교가 가능토록 하였고 또 질의영상은 범주의 크기에 비례하게 무작위로 뽑은 44장만을 대상으로 삼았다.

그림 12는 두 알고리즘에 대한 회수율과 정확성 비교 그래프이다. 그림에서 [4]의 경우 회수율은 65.25%, 정확성은 33.77%을 보여주었고 이에 반해 시스템 Angel은 회수율 면에서는 73.23%, 정확성에서는 37.13%를 나타냈다. 즉 회수율 면에서는 약 7.98% 가량 우수하고, 정확성 면에서는 약 3.36% 가량 앞서는 것으로 나타났다.



(a) 정확성



(b) 회수율

그림 12 제안된 알고리즘과 다른 알고리즘과의 비교

8. 결 론

본 논문에서는 색상-위치 정보만을 이용한 새로운 내용기반 영상검색의 방법을 소개하였다. 그리고 실험을 통해서 제안한 방법이 약 84.80%의 높은 필터링 효과뿐만 아니라 72.27%라는 평균 회수율과 34.04%에 달하는 평균 정확성을 갖는다는 사실을 확인하였다. 보다 정확한 검색을 위하여 제안된 유사도 피드백기법으로 54.85%의 검색성능이 향상되었다. 이와 같이 본 시스템은 색상과 그 위치 정보만을 이용해서도 충분히 유용한 내용기반 영상검색 시스템을 구현할 수 있다는 것을 보여주고 있다. 이처럼 간단한 특징요소만으로도 우수한 성능을 보일 수 있는 것은 본 논문에서 새롭게 제시한 방법 중의 하나인 QM 기법에 있다고 판단된다. 이 기법은 영상을 재귀적으로 분할해가며 32*32개의 셀 각각에 대해 기존의 0 아니면 1의 값을 부여하던 방식을 탈피해 대신 0~1 사이의 가중치를 갖도록 설정하기 때문에 좀 더 융통성이 있고 유연한 영상검색이 가능했다. 이러한 특성 때문에 영상 내 개체의 크기가 다르거나 일부 회전 또는 위치 이동이 있는 영상 그리고 영상의 일부분만 부분 추출한 영상의 검색에도 강인한 면모를 보여주었다.

그러나 본 시스템은 또한 몇 가지 문제점을 안고 있다. 먼저 가장 비중이 큰 것은 검색 효율성 제고를 위한 필터링 과정에서 유사한 영상임에도 불구하고 필터링 절차상의 임계치보다 낮은 값을 가져 검색에서 아예 제외되어 버리는 영상이 생기는 문제이다. 본 시스템은 휴리스틱한 방법을 통해 얻은 임계치를 고정해 놓고 사용하기 때문에 일단 한번 필터링 과정에서 놓쳐버린 영상은 다시 포함시키기가 불가능하다. 따라서 놓치는 영상이 발생하는 것을 막도록 향후에는 질의 영상의 색상 구성 및 분포 특성을 반영한 유동적인 임계치를 산출해 적용할 수 있는 방법을 고안해야 할 것이다. 또한 제안한 색인 키들이 비교적 짧은 비트열로 구성되는 만큼 이들을 구조화 시켜 활용하면 여러 곳에 분산되어 있는 데이타베이스의 검색에서 유용할 것이다.

참 고 문 헌

[1] M. Swain and D. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, pp. 11~32, 1991.

[2] J. R. Smith and S.-F. Chang, "Tools and Techniques for Color Image Retrieval," In *Proc. SPIE Vol. 2670: Storage and Retrieval for Image and Video Databases IV*, pp. 426~437, February 1996.

[3] Y. Rui, T. S. Huang, and S. Mehrotra, "Relevance Feedback Techniques in Interactive Content-Based Image Retrieval," In *Proc. SPIE Vol. 3312: Storage and Retrieval for Image and Video Databases VI*, pp. 25~36, January 1998.

[4] H.-W. Yoo, D.-S. Jang, S.-H. Jung, J.-H. Park, and K.-S. Song, "Visual Information Retrieval System via Content-based Approach," *Pattern Recognition*, vol 35, pp. 749~769, 2002.

[5] L. Cinque, S. Levialdi, K.A. Olsen, A. Pellicano, "Color-Based Image Retrieval Using Spatial-

Chromatic Histograms," In. *Proc. of the Multimedia Systems*, vol. 2, pp. 969~973, June, 1999.
[6] R. Brunelli, O. Mich, "On the Use of Histograms for Image Retrieval," In *Proc. of the Multimedia Systems*, vol. 2, pp. 143~147, June, 1999.
[7] Greg Pass, Ramin Zabih, Justin Miller, "Comparing Images Using Color Coherence Vectors," In *Proc. of the 4th ACM International Conference on Multimedia*, pp. 65~73, November, 1996.
[8] J. Huang, S.R. Kumar, M. Mitra, W.J. Zhu, "Spatial Color Indexing and Applications," In *Proc. of the 6th International Conference on Computer Vision*, pp. 602~607, January, 1998.
[9] J.R.Smith, S.F. Chang, "Integrated Spatial and Feature Image Query," *Multimedia Systems*, vol. 7, pp. 129~140, March, 1999.
[10] L. Cinque, S. Levialdi, K.A. Olsen, A. Pellicano, "Color-Based Image Retrieval Using Spatial- Chromatic Histograms," In *Proc. of the Multimedia Systems*, vol. 2, pp. 969~973, June, 1999.
[11] Kian-Lee Tan, Beng Chin Ooi, Chia Yeow Yee, "An Evaluation of Color-Spatial Retrieval Techniques for Large Databases," *Multimedia Tools and Applications*, vol. 14, pp. 55~78, 2001.
[12] MPEG Requirements Group, "Overview of the MPEG-7 Standard," ISO/IEC/JTC1/SC29/WG11, Geneva, May/June 2000.
[13] I. Biederman, "Human Image Understanding: Recent Research and a Theory," *Computer Vision, Graphics, and Image Processing*, vol. 32, pp. 29~73, 1985.
[14] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathomas, P. N. Yianilos, "The Bayesian Image Retrieval System, PicHunter: Theory, Implementation and Psycophysical Experiments," *IEEE Transactions on Image Processing*, vol. 9, pp. 20~37, 2000.

박 한 수
1970년 7월 30일생. 1994년 육군사관학교 화학과 졸업. 고려대학교 산업시스템정보공학석사과정. 관심분야는 컴퓨터비전, 멀티미디어시스템

유 헌 우
1966년 12월 24일생. 1992년 인하대학교 전기공학과 졸업, 동대학 전기공학 석사(1994). 고려대학교 산업시스템정보공학 박사(2001). LG전자 생산기술센터기술개발 연구소(1994~1997). 코스모 정보통신 수석연구원(2000~2003). 현재 연세대학교 인지과학연구소 연구교수. 관심분야는 컴퓨터비전, 멀티미디어시스템, 제어이론

장 동 식
1956년 12월 16일생. 1979년 고려대학교 산업공학과 졸업. 텍사스 주립대 산업공학석사(1985). 텍사스 A&M 산업공학 박사 (1988). 현재 고려대학교 산업시스템정보공학과 교수. 관심분야는 컴퓨터비전, 멀티미디어시스템, 로봇비전