색상과 불변 모멘트를 이용한 내용기반 이미지 검색

김목련*, 박영호*
*숙명여자대학교 멀티미디어과학과
e-mail:mrkim@sookmyung.ac.kr

Contents-based Image Retrieval using Color and Invariant Moments

Mok-Ryun Kim*, Young-Ho Park*
*Dept of Multimedia Science, Sook-Myung Women's University

요 약

최근 인터넷과 멀티미디어 기술이 발달함에 따라 이미지 데이터의 양이 급속히 증가하고 있다. 증가하는 이미지를 효과적으로 관리하고 검색하기 위해 내용기반 이미지 검색에 대한 연구가 활발히 진행되고 있다. 대부분의 내용 기반 이미지 검색 시스템은 색상, 모양, 질감 특징을 이용한 유사도-기반 검색에 초점을 맞추고 있다. 따라서 본 논문에서는 이미지에 나타나는 주요 색상과 색상의 공간적 특성을 포함하는 픽셀샘플링, 그리고 이미지의 외형적 변경에 강인함을 갖는 불변 모멘트 값을 이용한 내용기반 이미지 검색 시스템을 제안한다. 첫 번째 유사성 검사 단계에서는 이미지의 영역별로 가중치를 부여하여 추출한 대표색상을 사용하여, 유사하지 않은 이미지를 제거하여 검색대상의 수를 줄이며, 두 번째 유사성 검사 단계에서는 첫 번째 단계에서 선별된 후보 이미지에 색상의 공간적 정보를 포함한 픽셀샘플링을 이용하여 색상의 공간적인 위치까지 유사한 이미지만을 선별한다. 두 번째 유사성 검사단계에서 이미지가 외형적으로 변형된 유사이미지의 검출이 어려운 점을 보완하는 방법으로 이미지의 불변 모멘트를 이용하여 검색의 정확성을 높인다. 제안한 이미지 검색 방법은 10000개의 다양한 이미지로 구성된 데이터베이스에서 검색을 효율성을 실험하였다.

1. 서론

최근 멀티미디어 디바이스 및 기술의 발달로 이미지나 비디오와 같은 멀티미디어 데이터를 쉽게 제작할 수 있게 되었다. 그로 인해 이미지 데이터의 양이 매우 급증하게 되어 멀티미디어 데이터의 효율적인 저장, 브라우징(browsing), 인덱싱(indexing), 검색(retrieval)을 위한 많은 연구들이 진행되고 있다[1][2][3][4]. 초창기 내용기반 검색 방법은 주석에 의한 이미지의 시멘틱 정보를 이용하는 방법이었다. 그러나 이 방법은 이미지 주석 작성 시, 주관적인 특성에 기인한 관점 불일치로 사용자 질의에 대하여 신뢰성 있는 검색 정보를 제공하기 어렵다는 단점이 있으며, 데이터베이스 구축에 있어서 많은 시간과 비용이 소요되는 단점이 있었다. 위와 같은 문제를 해결하기 위하여 현재는 색상, 모양, 질감등과 같은 이미지가 가지고 있는 내용에 기반한 속성을 이용한 내용기반 검색기술의 개발이 진행되어 왔으며, 검색을 위한 다양한 접근 방법들이 연구되었다. 특히 색상은 영상을 구성하는 가장 지배적인 요소로써 추출과 활용이 쉽고 빠르다는 장점을 가지고 있어서 다른 두 가지에 비해 훨씬 더 많이 사용한다. 색상을 이용한 이미지검색 방법의 하나로 색상 히스토그램을 이용한 방법이 있다[5][6][7]. 이 방법은 이미지가 갖는 색상구성을 히스토그램을 이용하여 쉽고 간편하게 이미지검색을 실현한다. 그러나 공간적인 상호관계를 적절히 표현하지 못한다는 단점이 있다. 따라서 본 논문에서는 이 문제를 해결하기 위하여 이미지의 대표색상의 유사성뿐 만 아니라 픽셀 샘플링방법을 이용하여 색상의 공간적 위치의 일치성을 검사하는 이미지를 검색방법을 제안한다. 그리고 시각적으로는 동일한 이미지지만 이미지의 회전, 이동, 크기의 변형 시, 샘플링 된 픽셀의 값이 달라지는 픽셀 샘플링방법의 단점을 보완하기 위하여 이미지의 불변모멘트를 이용하여 이미지의 회전, 이동, 크기가 변형된 동일한 이미지도 검색할 수 있도록 유사성 검사를 보완하는 새로운 내용기반 이미지 검색 방법을 제안하고, 다음에서 본 논문의 공헌을 제시한다.

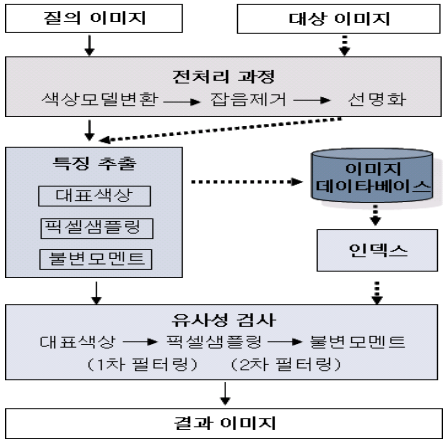
- 이미지의 대표색상 뿐만 아니라 픽셀 샘플링방법 제안하여 색상의 공간적 위치의 일치성을 검사하여 이미지 검색의 정확도를 높인다.
- 이미지의 회전, 이동, 크기의 축소, 확대와 같은 외형적 변경에 강인한 불변모멘트를 이용하여 이미지 검색의 회수율을 높인다.

본 논문의 구성은 2장에서는 제안하는 내용기반 이미지 검색을 위해 필요한 각 과정을 설명한다. 3장에서는 제안하는 방법의 실험을 통해 그 결과를 분석하고, 마지막으로 4장에서는 결론 및 향후 연구 방향에 대해 살펴본다.

2. 제안하는 내용기반 이미지 검색 방법

그림 1은 본 논문에서 제안하는 이미지 검색 시스템의 구성도 이다. 먼저, 대상 이미지는 전처리 과정을 거친 후, 각 이미지들 마다 대표색상, 샘플링 된 픽셀의 값, 불변 모멘트 값, 총 3가지의 특징을 추출한다. 추출된 값들을 이미지의 메타정보와 함께 데이터베이스에 저장하고, 인덱스를 구축한다.

질의 이미지가 입력되면, 질의이미지도 전처리과정을 거친 후, 3가지 특징 값을 추출한다. 그런 후, 데이터베이스에 저장된 대상 이미지들의 대표 색상 값과의 유사성 비교를 통해 1차적으로 필터링 된 대상 이미지 집합을 구하고, 필터링 된 대상이미지 집합의 샘플링 픽셀의 HSI 값과의 비교를 통해 질의 이미지와의 유사도를 계산한다. 마지막으로 1차 필터링 된 대상 이미지 집합 중에 최종적인 결과 이미지로 선택되지 못한 이미지와 질의 이미지의 불변모멘트를 비교하여 불변모멘트 값이 매우 유사한 이미지를 결과 이미지에 추가하여 최종적인 검색 결과 이미지를 출력한다.



(그림 1) 이미지 검색 시스템구성도

다음 소분류에서 각 단계별로 자세한 과정을 설명한다.

2.1 전처리과정

이미지의 전처리 과정을 총 3가지로 구성된다. 첫 번째 과정은 이미지 색상모델의 변환이다. 본 논문에서는 색상을 표현하기 위해서 인간의 인지도와 유사한 HSI 색상 모델을 사용하였다. RGB에서 HSI 모델로의 변환 식은 아래 (1)-(3)과 같다.

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

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

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

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

두 번째 과정은 잡음제거 과정이다. 이미지에 잡음을 제거함으로써, 이미지 특징 추출 시, 오류를 줄인다. 본 논문에서는 비선형 공간적 필터링 기법인 미디언 필터를 이용하여 잡음을 제거한다. 미디언 필터는 입력이미지의 (x, y)좌표 주변의 픽셀들의 값들을 오름 또는 내림차순으로 정렬하여 그 중앙에 있는 픽셀값을 이용하는 필터이다.

세 번째 과정은 이미지 선명화 과정이다. 이미지 선명화 과정은 이미지내의 경계선 부분을 강조하여 영상에 포함된 객체의 모양을 더욱 정확하게 나타내준다. 본 논문에서는 이미지 선명화를 위하여 언샤프 마스크 필터를 이용한다. 언샤프 마스크 필터는 아래 수식 (4),(5)로 계산된다.

$$g(x,y) = f(x,y) - \bar{f}(x,y) \quad (4)$$

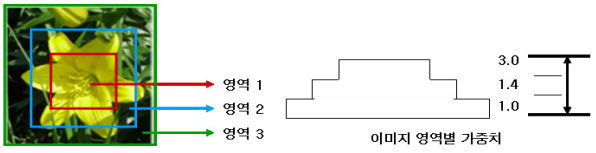
$$h(x,y) = f(x,y) + g(x,y) \quad (5)$$

위 식에서 $f(x,y)$는 입력 영상을 의미하고, $\bar{f}(x,y)$는 입력영상을 부드럽게 변환한 영상이다. 결과영상 $g(x,y)$는 입력 영상에서 경계선 부분에 해당하는 픽셀에서만 큰 값을 가지는 영상이 된다. 이를 이용하여 최종적으로 경계선 부분만 강조한 영상을 만들기 위해서 수식 (5)과 같이 $g(x,y)$를 다시 원래 영상과 더해주는 연산을 수행하면, 최종적으로 선명한 영상 $h(x,y)$를 만들 수 있다.

앞선 3가지의 전처리 과정 후, 이미지에서 몇 가지 주요한 색상만을 표현하기 위하여 양자화(Quantization)과정을 실시한다. 일반적으로 사람이 채도(Saturation)나 명도(Intensity)보다는 색조(Hue) 값에 훨씬 민감하게 반응함으로 채도와 명도는 각각 단 4개의 범주로 나누고 색조는 총 16개의 범주로 나누어 양자화한다.

2.2 이미지 특징 추출

질의 이미지에서 대상 이미지와의 유사성 검사를 위한 지표로 3가지의 특징을 추출한다. 첫 번째는 이미지에 대표적으로 나타나는 대표색상의 추출이다. 일반적으로 이미지는 몇 가지의 주요 색상만으로 대표되는 성질이 있음으로[8]. 이들 색상만을 갖고 유사한 이미지인지 여부를 구분 짓는 인덱스키로 활용이 가능하다. 너무 많은 색상을 고려하면 불필요한 계산이 늘어나게 되고 또한 의미 없는 색상마저도 포함되기 때문에 노이즈(Noise) 같은 역할을 해 오히려 성능을 저하시킨다. 따라서 본 논문에서는 단 4가지로 색상의 수를 제한하고 이를 주요 색상으로 추출한다. 대표색상을 추출할 때, 일반적으로 사람이 이미지의 중앙부에 잇는 색상은 외곽부에 있는 색상에 비해 중요하게 지각함을 반영하여, 그림 2과 같이 이미지의 중앙부와 외곽부에 가중치를 차별화하여 보다 인지도에 가까운 이미지검색을 실시한다.



(그림 2) 이미지 영역별 가중치

두 번째로 추출하는 이미지의 특징은 앞서 추출한 이

미지의 대표색이 색상의 공간적인 위치를 반영하지 못한다는 단점을 보완하기 위한 특징이다. 색상의 공간적 위치를 이미지 검색에 반영하기 위해, 이미지 샘플링방법을 통해 위치정보를 포함한 색상 정보를 2차로 추출한다. 이때, 이미지의 크기로 부터 독립된 색상정보 추출이 필요한데, 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[8].

세 번째로 추출하는 이미지의 특징은 불변 모멘트 값이다. 불변 모멘트[9]는 이미지의 특성을 수학적 계산을 통해 도출된, 이미지의 회전, 이동, 크기변화에 변하지 않는 7개의 값이다. 불변 모멘트는 기하학적 모멘트에서 중심모멘트를 구하고, 중심모멘트를 정규화 시킨 중심 모멘트를 생성하며, 이를 조합하면 불변 모멘트를 구하는 식을 도출할 수 있다. 기하학적 모멘트는 2차원 다항식 $g(x,y)=x^p y^q$를 기저함수로 사용하여 영상의 특성을 수학적으로 표현하는 방법으로 2차원 영상을 함수 $f(x,y)$ 형태로 나타낼 경우, 수식 (6)과 같이 정의 된다.

$$m_{pq}=\sum_{y=0}^{N-1}\sum_{x=0}^{M-1}x^p y^q f(x,y) \tag{6}$$

위 식에서 p와 q는 0보다 같거나 큰 정수이며, $(p+q)$를 모멘트의 차수라고 부른다. M과N은 각각 영상의 가로와 세로 픽셀 크기를 나타낸다. 일반적인 기하학적 모멘트의 경우, 영상의 이동 변환 시 그 값이 크게 변한다는 단점이 있다. 그러므로 영상의 무게 중심을 고려하여 모멘트를 계산하는 중심모멘트를 구한다. 2차원 이미지의 중심모멘트는 수식 (7)로 정의한다.

$$\mu_{pq}=\sum_{y=0}^{N-1}\sum_{x=0}^{M-1}(x-\bar{x})^p(y-\bar{y})^q f(x,y) \tag{7}$$

위의 식에서 $\bar{x}$와 $\bar{y}$는 영상의 무게 중심의 좌표이다. 무게중심의 좌표 $(\bar{x},\bar{y})$은 낮은 차수의 기하학적 모멘트를 이용하여 다음과 같인 수식 (8)로 구할 수 있다.

$$\bar{x}=\frac{m_{10}}{m_{00}},\ \bar{y}=\frac{m_{01}}{m_{00}} \tag{8}$$

차수가 3차 이하인 중심모멘트는 기하학적 모멘트만을 이용하여 수식 (9)과 같이 계산 될 수 있다.

$$\begin{aligned}
&\mu_{00}=m_{00}\\
&\mu_{10}=0\\
&\mu_{01}=0\\
&\mu_{11}=m_{11}-\bar{x}m_{01}=m_{11}-\bar{y}m_{10}\\
&\mu_{20}=m_{20}-\bar{x}m_{10}\\
&\mu_{02}=m_{02}-\bar{y}m_{01}\\
&\mu_{30}=m_{30}-3\bar{x}m_{20}+2\overline{x^2}m_{10}\\
&\mu_{03}=m_{03}-3\bar{y}m_{02}+2\overline{y^2}m_{01}\\
&\mu_{21}=m_{21}-2\bar{x}m_{11}-\bar{y}m_{20}+2\overline{x^2}m_{01}\\
&\mu_{12}=m_{12}-2\bar{y}m_{11}-\bar{x}m_{02}+2\overline{y^2}m_{10}
\end{aligned} \tag{9}$$

중심모멘트를 조합하여 불변모멘트를 계산할 수 있는데, 이는 크기를 정규화한 중심모멘트를 비선형으로 조합하여 만들어진다. 정규화 된 중심모멘트는 수식 (10)과 같이 생성할 수 있고, 정규화 된 중심모멘트를 이용하여 수식 (11)과 같이 7개의 불변 모멘트가 정의된다.

$$\eta_{pq}\frac{\mu_{pq}}{\mu_{00}^{\gamma}},\qquad \gamma=\frac{p+q}{2}+1\ (p+q=2,3,....) \tag{10}$$

$$\begin{aligned}
\Phi_1&=\eta_{20}+\eta_{02}\\
\Phi_2&=(\eta_{20}-\eta_{02})^2+4\eta_{11}^2\\
\Phi_3&=(\eta_{30}-3\eta_{12})^2+(3\eta_{21}+\eta_{03})^2\\
\Phi_4&=(\eta_{30}+\eta_{12})^2+(\eta_{21}+\eta_{03})^2\\
\Phi_5&=(\eta_{30}-3\eta_{12})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]\\
&\quad+(3\eta_{21}-\eta_{03})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]\\
\Phi_6&=(\eta_{20}-\eta_{02})\left[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]\\
&\quad+4\eta_{11}(3\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03})\\
\Phi_7&=(3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{12})\left[(\eta_{30}+\eta_{12})^2-3(\eta_{21}+\eta_{03})^2\right]\\
&\quad+(3\eta_{21}-\eta_{30})(\eta_{21}+\eta_{03})\left[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2\right]
\end{aligned} \tag{11}$$

이렇게 구해진 이미지의 불변 모멘트 값은 이미지의 크기 변형, 회전, 위치의 이동에 불변한 값을 가지므로, 형태적으로 다른 유사한 이미지도 검색한다.

2.3 유사성 검사

추출된 이미지의 3가지 특징을 이용하여 데이터베이스에 저장된 대상 이미지들의 동일한 특징 값과의 비교를 통해 유사성 검사를 수행한다. 데이터베이스에 저장된 많은 이미지들의 특징값을 가져오는 것은 매우 시간이 오래 걸리는 작업이다. 따라서 본 논문에서는 이미지의 각 특징값들에 인덱스를 구성하여 빠르게 데이터를 가져올 수 있도록 한다.

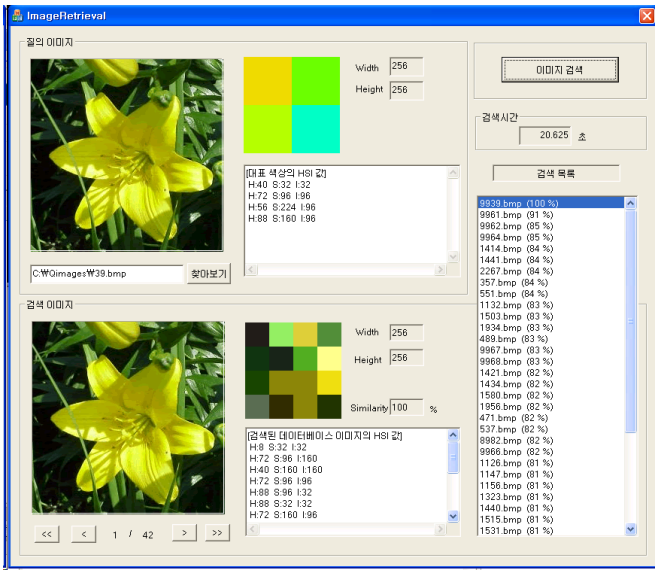
유사성 검사의 첫 번째 단계는 대표색상 값 유사성 검사이다. 이 단계에서는 질의 이미지의 4가지 대표색상과 동일한 대표색상 갖는 대상이미지만을 선별하여 2차 유사성 검사 집합으로 삼는다.

유사성 검사의 두 번째 단계는 1차로 필터링 된 대상이미지 집합과 질의 이미지에서 매치되는 셀에서 샘플링된 픽셀의 HSI값의 차이를 합산하여, 차이의 최대값을 10으로 했을 때, 8이상의 차이 값을 갖는 대상 이미지를 필터링하고, 8이하의 차이 값을 대상이미지의 정보만을 링크드 리스트로 유사도 순으로 정렬하여 검색 이미지 정보를 유지한다.

두 번째 유사성 단계에서 이미지의 외형적 변형에 강인한 특성을 갖지 못함을 보완하기 위해, 픽셀의 HSI 차이 값이 8이상인 이미지만을 대상으로 질의이미지와의 불변 모멘트값을 비교한다. 불변 모멘트 값 중 5번째와 7번째 모멘트 값은 그 값의 변화가 미세하여서 유사성 비교에 큰 영향을 미치지 않으므로, 계산양의 증대를 피하기 위해 유사성 검사의 계산에서 제외한다. 만일 이 값이 질의이미지와 매우 유사한 경우, 해당 이미지를 결과이미지 정보를 구성한 링크드 리스트에 포함한다.

3. 실험 및 분석

제안한 방법의 유용성을 검증하기 위해, SIMPLIcity[10] 이미지 검색시스템에서 제공하는 10000장의 서로 크기가 다른 이미지를 실험에 사용하였다. 제안한 방법의 성능측정의 기준으로 질의이미지와 유사한 이미지가 얼마나 검색결과로 도출 되었는지의 정도를 나타내는 회수율(Recall)과 검색된 이미지에서 실제 질의 이미지와 유사한 이미지가 포함된 정도를 나타내는 정확도(Precision)를 사용한다. 총 10개의 질의 이미지를 임의로 선택하여 유사이미지를 검색한 결과, 평균 회수율은 0.68, 정확도는 0.34를 나타내었다. 그림 (3)은 본 논문에서 제안하는 검색방법을 실험하기 위해 구현한 사용자 인터페이스의 모습이다. 상단에 질의 이미지와 질의 이미지에서 추출된 대표색상, 추출된 특징의 텍스트 정보를 나타낸다. 하단부에서는 질의 이미지와 유사한 이미지로 검색된 대상 이미지와 대상이미지에서 추출된 샘플링 픽셀의 색상 정보를 출력한다. 오른쪽부분에는 검색에 걸린 시간과, 검색된 대상이미지들의 목록이 유사도 순으로 정렬하여 제공된다.



(그림 3) 제안하는 이미지 검색시스템의 인터페이스

4. 결론 및 향후연구

본 논문에서는 이미지의 색상정보와 불변 모멘트를 이용한 내용기반 이미지 검색방법을 제안하였다. 제안한 방법은 이미지로부터 대표색상, 픽셀 샘플링, 불변 모멘트값을 추출하여, 유사성 비교의 특징으로 이용했다. 먼저, 질의 이미지에 나타난 대표색상을 이미지 영역의 가중치를 차별화하는 방법으로 추출하여, 대상이미지의 대표색상과의 비교를 통해 동일한 대표색상을 갖는 대상 이미지만을 추출하여 1차로 필터링 된 대상 이미지 집합을 생성하였다. 그리고 이미지의 대표색상이 색상의 공간적 위치 정보를 표현하지 못한다는 점을 보완하여, 이미지를 특정크기의 셀로 나누고, 각 셀별로 픽셀을 샘플링하여 이를 두 번째 유사성 검사에 이용하였다. 마지막으로 불변모멘트값을 이용하여 이미지의 외형적 변형으로 인해 검색되지 못한 유사 이미지를 결과 이미지로 포함될 수 있도록 하였다. 실험을 통하여 결과 평균 회수율은 0.68, 정확도는 0.34임을 확인하였다.

향후 연구로는 검색의 회수율과 정확도를 높이기 위해 이미지 특징추출 및 유사성 검사단계를 발전시키고, 또한 데이터베이스의 이미지 정보가 증가함에 따라 검색시간이 매우 증가함으로 이를 해결하기 위해 기존 인덱스 방법을 개선시킨 새로운 이미지 데이터베이스 인덱스의 개발을 향후 연구 과제로 한다.

참고문헌

[1] A. Pentlang et al., "Photobook: Tools for content-based manipulation of image databases," SPIE Proc. Storage and Retrieval for image and video databases, vol, 1908, pp. 34-47, Feb. 1994.

[2] M. Flicker et al., "Query by image and video content: The QBIC system," IEEE Computer, vol. 28, pp. 23-32, Sept. 1995.

[3] J. Bach et al., "The virage image search engine: An open framework for image management," SPIE Proc. Storage and Retrieval for image and video database, vol. 2670, pp. 1-4, Feb. 1996

[4] J. R. Smith and S. F. Chang, "Exploring image functionalities in WWW applications-development of image/video search and editing engines," IEEE Proc. ICIP, pp. 1-4. Oct. 1997.

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

[6] 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, February1996.

[7] 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.

[8] 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.

[9] M. K. Hu, "Visual Pattern Recognition by Moment Invariants", IRE Trans. Info. Theory, vol. IT-8, pp.179-187, 1962.

[10]James Z. Wang, Jia Li, Gio Wiederhold, ''SIMPLIcity: Semantics-sensitive Integrated Matching for Picture LIbraries,'' IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 23, no.9, pp. 947-963, 2001.