

Block-based Edge Histogram 과 DCT 를 이용한 칼라 영상 검색

이동호*, 유광석, 김희율
한양대학교 전자공학과 영상공학연구소
e-mail: {dhlee*, ksryoo}@vision.hanyang.ac.kr, wykim@email.hanyang.ac.kr

Color Image Retrieval using Block-based Edge Histogram and DCT

Dong-ho Lee*, Kwang-seok Ryoo, Whoi-yul Kim
Image Engineering Lab., Department of Electronic Engineering, Hanyang Univ.

요 약

본 논문에서는 질감 정보를 나타낼 수 있는 Block-based 에지 히스토그램과 색상 정보를 표현할 수 있는 DCT 를 이용한 칼라 영상 검색 방법을 제안한다. 제안된 방법은 최소의 특징량으로 최대의 검색효율을 얻기 위해 YCbCr 칼라 모델상에서 Y 영상으로부터는 전체적인 영상에 대한 히스토그램과 에지 히스토그램을 특징량으로 추출하고 Cb, Cr 영상으로부터는 DCT 계수를 특징량으로 추출하여 칼라 영상을 검색한다. 이는 칼라와 질감을 동시에 고려하면서 특징량의 크기가 적어 웹, 대용량 검색 시스템 및 동영상 검색에 적합하다. 성능 평가는 MPEG-7 의 칼라 특징자들의 성능평가를 위해 사용된 S1 및 S3 그룹 영상을 대상으로 실험하였으며 제안한 복합 특징량은 칼라 영상 검색에서 우수한 성능을 나타냄을 실험으로 확인 하였다.

1. 서론

최근 컴퓨터와 통신 기술의 급속한 발달로 인해 인터넷을 통한 멀티미디어 데이터 전송이 보편화 되면서 정보 통신망 사용자들은 시간과 장소에 상관없이 다양한 정보에 접근할 수 있게 되었고 또한 사용자들이 접근할 수 있는 멀티미디어 정보의 양도 폭발적으로 증가하고 있다. 그러나 정보 이용자들이 접근할 수 있는 정보가 늘어날수록 이용자들이 원하는 정보를 찾기는 더더욱 어려워져 멀티미디어 정보 검색, 저장, 관리 기술에 대한 요구가 증가하고 있다. 영상의 경우 원거리 감지를 위한 항공/위성 사진, 지리정보 시스템, 의료 영상 정보 시스템, 법의학에서의 지문인식, 캐릭터 영상, 상표와 로고 등록에서와 같이 다양한 분야에서 널리 개발 및 연구되고 있다. 따라서 대용량 데이터에서 원하는 정보를 빠른 시간 내에 검색할 수 있는 검색 기법이 필요하게 되었다. 텍스트나 파일명에 의한 기존의 검색 방법에서 영상의 내용 자체를 이용하여 영상 정보를 검색하는 내용기반 검색 방법이 활발히 진행되고 있다 [1].

일반적으로 내용기반 이미지 검색(Content-based

Image Retrieval) 방법은 많은 응용에서 문자기반 영상 검색(Text-based Image Retrieval) 방법의 단점을 보완해주는 방법으로서 영상에서 칼라(Color), 질감(Texture), 모양(Shape), 구성(Composition) 및 주파수 성분과 같은 인지적이고 의미 있는 특징 벡터(Feature Vectors)를 이용하여 사용자가 원하는 영상을 검색하는 방법이다.

1991년 Swain[2]에 의해 발표된 칼라 히스토그램 인터섹션(Histogram Intersection) 기법은 내용기반 영상 검색을 위한 특징 벡터로서 일반적으로 많이 사용된다 [3, 4]. 이는 전체적인 영상의 칼라 분포를 의미하는데, 계산이 쉽고 물체의 회전이나 작은 이동 등과 같은 기하학적인 변형에 강건한 장점이 있다. 그러나 인지적으로 비슷하지 않은 영상도 동일한 칼라 분포를 가질 수 있고 또 특징량이 많아 검색 시 많은 시간이 소요되는 문제가 제기되었고, 이를 해결하기 위한 방법으로 칼라 정보에 공간 정보를 포함시킨 여러 방법들이 제안되었다 [5, 6, 7, 8].

Pass는 영상 내에서 같은 칼라의 응집 여부를 이용한 CCV(Color Coherence Vectors)를 제안하여 기존의 칼라 히스토그램 방법을 개선하였으나, 공간 정보가 극

히 적다는 단점을 가지고 있다 [6]. Jing Huang 은 칼라 간의 상관관계를 이용한 Color Correlogram 이라는 새로운 방법을 제안하였다 [7]. 이 방법은 영상 내의 픽셀로부터 특정 거리만큼 떨어진 픽셀들에 대해 해당 픽셀의 빈도수를 표시한 히스토그램으로 영상을 비교하는 것으로, 칼라 정보와 공간 정보를 효과적으로 결합하였으나, 다른 칼라와의 상관관계를 고려하지 않은 단점이 있다.

대부분 칼라나 질감 특징자들은 단일 특징량으로 검색하고 또 많은 특징량으로 인해 검색 시 많은 시간이 소요되기 때문에 웹, 동영상 그리고 대용량 데이터 베이스 검색 시스템에는 부적합하다.

본 논문에서는 칼라 정보와 질감 정보를 효과적으로 결합하여 매우 적은 특징량으로 뛰어난 검색효과를 보여주며, 빠른 검색을 요구하는 웹, 대용량 검색 시스템 및 동영상 검색에 적합한 방법을 제시한다. 제안된 방법은 JPEG 이나 MPEG 표준에서 칼라 모델인 YCbCr 모델을 적용하고, Y 영상은 전체 영상의 히스토그램과 0°, 45°, 90° 및 135° 에지 히스토그램[12]으로 질감 정보를 추출하고, Cb, Cr 칼라 성분은 DCT 계수를 사용하여 보다 압축된 색상 정보를 얻는 방법을 제안한다 [14].

본 논문의 2 절에서는 칼라 영상 검색을 위한 특징 벡터 추출 방법에 대해 설명하고 3 절에서는 유사도 측정 방법에 대해 설명한다. 4 절에서는 제안된 알고리즘에 대한 실험결과 및 분석을 보여주며 5 절에서는 결론을 맺는다.

2. 칼라 영상에 대한 특징량 추출 방법

칼라를 이용한 이미지 검색에 있어서는 사용할 칼라 모델에 대한 검토가 필요하다. 칼라 모델에는 RGB, HIS, YUV, YIQ 및 HMMD[13] 등 다양한 모델이 사용되고 있다. 본 논문에서는 칼라 영상에서 색상 정보와 질감 정보를 효율적으로 추출 할 수 있고 동영상 검색을 위해 JPEG 이나 MPEG 표준에서 사용하는 YCbCr 칼라 모델을 이용하였다.

본 논문에서 특징량을 추출 하는 알고리즘은 원 영상을 YCbCr 칼라 모델을 적용하여 3 개의 단일 영상으로 분리한다. 분리된 Y 영상은 6 레벨로 양자화 하고 그 전체적인 영상에 대한 히스토그램과 0°, 45°, 90° 및 135° 에지 히스토그램을 구한다. 에지 분포는 객체의 윤곽을 추출 하여 객체에 대한 정보를 추출할 수 있지만 색상 성분과는 전혀 관련이 없다는 특징이 있다. 또 전체적인 영상의 히스토그램은 에지 히스토그램을 보완해줄 수 있는 것으로 그레이 영상(Gray Image)이 전체적으로 밝아지거나 어두워지더라도 에지의 변화는 없기 때문에 전체 영상의 밝기 변화를 알 수 있다. 이것은 그레이 영상을 검색할 때에도 유효한 특징량이 된다.

Cb, Cr 영상은 원 영상을 16×16 블록으로 나누고 그 블록에 대한 주요 칼라 성분을 추출한다. 주요 칼라 성분은 서브 블록 내에 있는 픽셀 값들의 평균값으로 한다 [15]. 구해진 Thumbnail image 는 2D-DCT 변환을

하게 되고 DCT 변환으로 얻어진 계수를 특징량으로 하여 칼라 영상을 검색한다. 그림 1 은 본 논문에서 제안한 알고리즘의 특징량 추출 과정을 도식적으로 나타내었다.

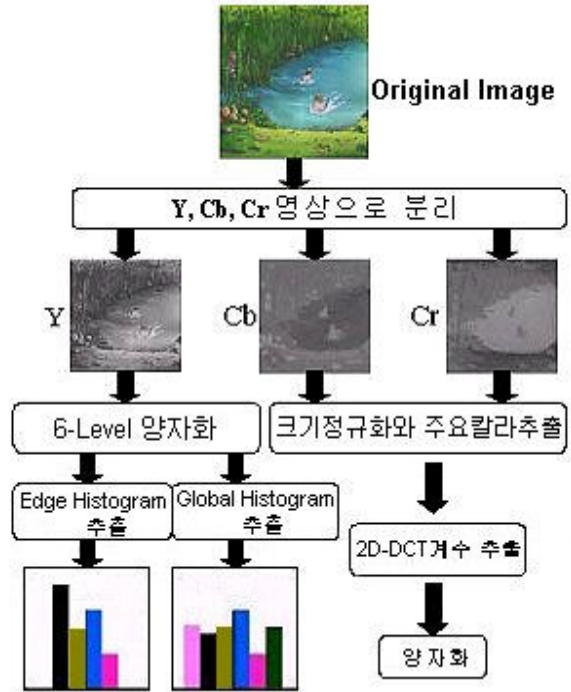


그림 1. 제안된 방법의 특징량 추출 과정

먼저 Y 영상은 선형적인 양자화 방법으로 6 레벨로 양자화 하였다. 양자화된 영상에 대해 6 개의 히스토그램을 구하고 원 영상의 크기에 상관없이 하기위해 식(1) 을 사용하여 히스토그램을 정규화 한다.

$$Normalize_Histogram = \frac{Global_Histogram[i]}{X_Width \times Y_Height} \quad (1)$$

$$i = 0, 1, 2...5$$

여기서 X_Width, Y_Height 는 원 영상의 가로와 세로 크기이다.

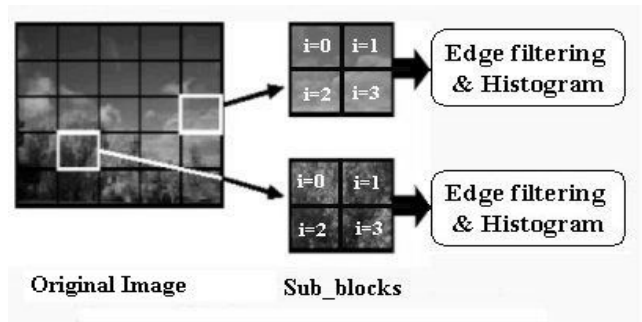


그림 2. 영상분할과 에지 히스토그램 추출 과정

그림 2 과 같이 양자화된 영상을 M×N 블록으로 나누고 그 블록에 대해 다시 2×2 서브 블록으로 나눈다. 각각의 서브 블록에 대해 평균값을 구하고 그 평균값과 0°, 45°, 90° 및 135° 에지 필터를 사용하여 에지를

추출 한 다음 에지 값이 사용자가 정한 임계치 이상인 값에 대해 각각의 에지 히스토그램을 구한다. 각 방향에 따른 각각의 에지를 구하는 식은 다음과 같다.

표 1. 0°, 45°, 90°, 135° 에지 필터의 계수

| | | | |
|----------------------|----|-----------------------|-------------|
| Edge0_filter | | Edge45_filter | |
| 1 | 1 | $\sqrt{2}$ | 0 |
| -1 | -1 | 0 | $-\sqrt{2}$ |
| Edge90_filter | | Edge135_filter | |
| 1 | -1 | 0 | $\sqrt{2}$ |
| 1 | -1 | $-\sqrt{2}$ | 0 |

$$edge\ 0 = \sum_{i=0}^3 (mean(i) \times edge\ 0_filter(i)) \quad (2a)$$

$$edge\ 45 = \sum_{i=0}^3 (mean(i) \times edge\ 45_filter(i)) \quad (2b)$$

$$edge\ 90 = \sum_{i=0}^3 (mean(i) \times edge\ 90_filter(i)) \quad (2c)$$

$$edge\ 135 = \sum_{i=0}^3 (mean(i) \times edge\ 135_filter(i)) \quad (2d)$$

여기서 mean(i)는 그림 2 에서와 같이 각각의 서브 블록 내에 있는 픽셀 값들의 평균값들이고 각 방향의 에지 filter(i)는 좌 상단에서 시계방향으로 i 에 해당하는 필터의 계수이다. 식(3)으로 (2a), (2b), (2c), (2d)식에 의해 구해진 에지 히스토그램을 정규화 한다.

$$Edge_Histogram = \frac{Edge[i]_Histogram}{Block_number} \quad (3)$$

$i = 0, 45, 90, 135$

다음 Cb, Cr 영상에 대해 특징량을 추출하는 과정은 영상 분할, 주요 칼라 선택, DCT 변환, 양자화 등 4 단계의 과정으로 특징량을 추출한다. 처음 영상 분할은 데이터 베이스 영상들의 크기에 상관 없이 16×16 으로 나눈다. 그 블록의 가로 세로 크기는 각각 X_Width/16 이고 Y_Height/16 이 된다. 여기서 X_Width, Y_Height 는 원 영상의 가로 넓이와 세로 높이이다. 만일 원 영상이 16 의 배수가 아닐 경우 마지막 행과 열 블록에 해당하는 블록은 원래 영상에서 남은 부분에 대한 크기가 된다. 나누어진 블록에서 주요 색상 성분 추출 방법은 블록 내의 모든 픽셀 값들의 평균값으로 선택한다. Cb, Cr 영상의 히스토그램 분포를 살펴보면 일정한 값들의 범위 내로 몰려있는 것을 볼 수 있는데 이러한 특징은 DCT 변환을 사용하여 효율적인 색상 정보를 추출하게 한다. 2D-DCT 변환식[9]은 식(4)과 같다.

$$F(u,v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

여기서 $\alpha(w) = \begin{cases} \sqrt{\frac{1}{N}} & \text{for } w=0 \\ \sqrt{\frac{2}{N}} & \text{for } w=1,2,3,\dots,N-1 \end{cases} \quad (4)$

여기서 x, y = 0, 1, 2,N-1 이다.

표 2. 색상성분 DCT 계수의 양자화 표

| | DC | AC1 | AC2 | AC3 |
|----|----|-----|-----|-----|
| Cb | 17 | 18 | 18 | 24 |
| Cr | 17 | 18 | 18 | 24 |

2D-DCT 변환을 한 색상 영상은 지그재그 스캔(Zig-zag Scan)을 통해 4 개의 계수를 추출한다. 지그재그 스캔 방식은 JPEG 압축 알고리즘에서 사용하는 방식이다. 추출된 계수는 1 개의 DC 성분과 3 개의 AC 성분을 나타내고 표 2 의 양자화 테이블을 통해서 양자화된다.

3. 유사도 측정

칼라 영상의 검색을 위해 질의 영상이나 데이터 베이스내의 영상에 대해 특징벡터를 추출하여 그 특징 벡터 값을 이용하여 검색한다.

본 논문에서는 질의 영상과 데이터 베이스 내의 영상과의 거리 D 는 식(5)에 의해 구해지고 유사도 S 를 위해 다음과 같은 식을 사용하여 유사도를 측정한다.

$$D = \omega_1 \sqrt{\sum_{i \in C1} (Q(1)_i - T(1)_i)^2} + \omega_2 \sqrt{\sum_{i \in C2} (Q(2)_i - T(2)_i)^2} + \omega_3 \sqrt{\sum_{i \in C3} (Q(3)_i - T(3)_i)^2} \quad (5)$$

$$S = \frac{1}{1+D} \quad (6)$$

여기서 $Q(r)_i$ 는 질의 영상의 r 번째 칼라 성분 영상의 i 번째 계수를 말하고, $T(r)_i$ 는 데이터 베이스 내에 저장되어 있는 영상의 r 번째 칼라 성분 영상의 i 번째 계수를 말한다. $\omega_1, \omega_2, \omega_3$ 는 칼라 영상 성분의 가중치이다. 가중치를 어떤 쪽에 많이 할당 하느냐에 따라 검색 되는 영상에도 많은 차이가 있다. 유사도 S 는 정규화 과정을 거쳐 0 과 1 사이의 값으로 표현된다.

4. 실험 및 결과

본 논문에서 제안된 방법을 검증하기 위해 MPEG-7 칼라 특징자들의 성능평가에 사용된 테스트 영상 388 개의 50 개 그룹으로 나눈 경우와, 그 영상을 나누어 지지 않은 5,072 에 포함 시켜 전체 5,460 개의 테스트

영상으로 구성하여 이 두 경우에 대해 실험을 하였다. 실험 환경은 팬티엄 III 450Mhz 기종과 윈도우 98 의 운영체제 상에서 진행하였으며, Visual C++ 6.0 컴파일러를 사용하였다.



그림 3. 실험에 사용된 영상

제안된 알고리즘의 실험에서는 Block_Number 를 32×32 로 하고 에지 히스토그램 검출을 위한 임계치는 10 로 정하여 실험하였다. 유사도 측정 시 가중치를 $\omega_1 = 5, \omega_2 = 1, \omega_3 = 1$ 로 하여 실험하였다.

성능 평가를 위해 MPEG-7 표준화 방법으로 제안된 NEC-Proposal 방법과 본 논문에서 제안한 방법과의 검색결과를 서로 비교하되 특징량의 개수는 18 개로 하였다. 성능 분석 지표로는 현재 MPEG-7 칼라와 질감에 대한 성능 분석 방법인 ANMRR(Average of Normalized Modified Retrieval Rank)[11]과 모양에 대한 성능 분석 방법인 BEP(Bull's Eye Performance)[10]로서 성능평가를 하였다.

$$ANMRR = \frac{1}{Q} \sum_{q=1}^Q NMRR(q) \quad (7)$$

여기서 $K = \min(4 * NG(q), 2 * GTM)$

$$AVR(q) = \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)}$$

$$MRR(q) = AVR(q) - 0.5 - \frac{NG(q)}{2}$$

$$NMRR(q) = \frac{MRR(q)}{K + 0.5 - 0.5 * NG(q)}$$

여기서 NG(q)는 질의하는 영상이 속해 있는 그룹내의 총 데이터 수이고, GTM 은 NG(q)중 가장 큰 값이

다. K 가 결정되면 Rank 는 어떤 데이터 셋에 있는 데이터를 질의했을 때 검색순위가 K 값 이하일 경우 검색의 순위이고 K 값보다 클 경우는 검색 순위에 상관없이 항상 K+1 의 값을 가진다. Rank 가 구해지면 식 (10)에 의해 ANMRR 이 구해지고 NMRR(Normalized Modified Retrieval Rank)은 항상 0 에서 1 의 값을 가지며 ANMRR 은 낮은 값일수록 좋은 결과를 나타낸다.

그림 4 와 표 3 은 각각 MPEG-7 칼라 영상 검색자들의 성능평가에 사용된 영상의 NMRR 과 실험 결과를 나타낸 것이다.

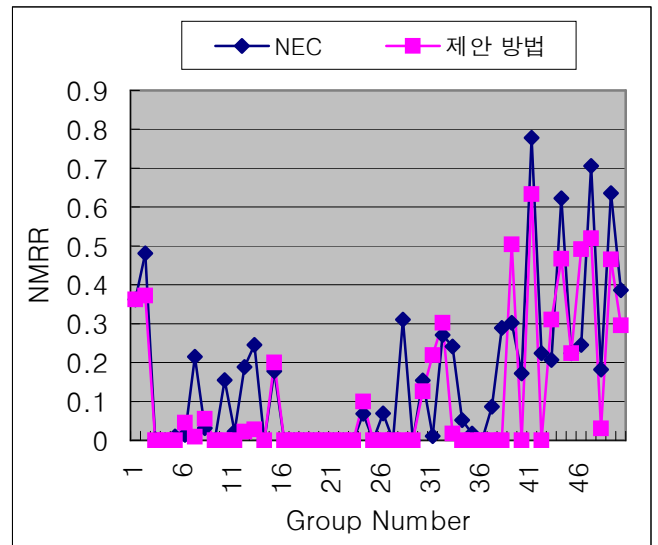


그림 4. NMRR 결과

표 3. 388 개의 영상에 대한 실험결과

| | 제안한 방법 | NEC-12 | NEC-18 |
|---------------|--------|--------|--------|
| BEP(%) | 87.14 | 79.49 | 79.7 |
| ANMRR | 0.1159 | 0.2092 | 0.1868 |

표 4 는 MPEG-7 칼라 영상 검색자를 평가하는데 사용된 영상과 다른 칼라 영상을 합쳐 5,460 개의 영상에 대해 실험한 결과를 나타낸 것이다. 여기서 NEC-12 는 NEC 알고리즘에서의 특징량이 12 개 라는 것을 나타내고 이것은 MPEG-7 Proposal 에 제안된 방법이다 [15]. NEC-18 은 NEC 알고리즘에서의 특징량이 18 개 임을 나타내는 것으로 제안된 알고리즘과 똑같은 특징량으로 했을 때 검색 효율을 비교하기 위해서 사용하였었다. 또 많은 테스트 영상에서 검색 함으로서 제안한 알고리즘의 강인함을 검증 할 수 있다. 표 3 과 표 4 에서 보는 것과 같이 본 논문에서 제안한 방법의 검색 효율이 NEC 방법에 비해 월등히 좋은 것으로 나타났다.

표 4. 5,460 개의 영상에 대한 실험결과

| | 제안한 방법 | NEC-12 | NEC-18 |
|---------------|--------|--------|--------|
| BEP(%) | 75.36 | 70.71 | 71.36 |
| ANMRR | 0.2365 | 0.3183 | 0.3046 |



그림 5. 검색 결과

5. 결론

본 논문에서는 질감 정보를 나타낼 수 있는 Block-based Edge Histogram 과 칼라 성분 영상에 DCT 를 적용한 칼라 영상 검색 방법을 제안하였다.

제안된 방법은 질감 정보와 색상 정보를 효과적으로 결합하여 매우 뛰어난 검색효과를 실험을 통해 입증하였으며 매우 적은 특징량으로 검색하므로 웹, 대용량 검색 시스템, 동영상 검색에 적합한 방법을 제시하였다.

향후 연구 과제로는 공간 정보, 질감 정보, 색상 정보 모두를 고려하면서 더 적은 특징량을 가지는 칼라 영상 검색 특징자를 연구해야 할 것이다.

참고문헌

[1] V. N. Gudivada and V. V. Raghavan, "Content-based image retrieval systems," IEEE Computer, pp. 18-22, Sept. 1995.

[2] M. J. Swain and D. H. Ballard, "Color indexing," International Journal of Computer Vision, 7(1), pp11-32, 1991.

[3] M. Flickner, H. Sawhney, et al, "Query by image and video content: The QBIC system," IEEE Computer, pp. 23-39, Sept. 1995.

[4] V. Ogle and M. Stonebraker, "Chabot: Retrieval from a relational database of images," IEEE Computer, 28(9): 40-48, 1995.

[5] J. Smith and S-F Chang, "Tools and techniques for color image retrieval," SPIE Proceedings, 2670: 1630-1639, 1996.

[6] Greg Pass and Ramin Zabih, "Histogram Refinement for Content-Based Image Retrieval," IEEE Workshop on Applications of Computer Vision, pp. 96-102, December 1996.

[7] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei-Jing

Zhu, and Ramin Zabih, "Image indexing using color correlograms," IEEE Conference on Computer Vision and Pattern Recognition, pp. 762-768, 1997.

[8] Asha Vellaikal and C.-C. Jay Kuo, "Content-Based Image Retrieval Using Multiresolution Histogram Representation," SPIE Digital Image Storage and Archiving Systems, pp. 312-323, Philadelphia, October 1995.

[9] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," 2nd edition, Reading, Massachusetts, Addison-Wesley, 1992.

[10] Kwangseok Ryoo, Whoi-yul Kim, "Content Based Character Image Retrieval with Probability Distribution of Color Feature," The Collection of 11th Workshop's Paper on Image Processing and Understanding, February 1999.

[11] Video Group, Core Experiments on MPEG-7 Color and Texture Descriptors," ISO/IEC JTC1/SC29/WG11 N2929(1999).

[12] Soo-Jun Park, et.al., "Core Experiments on MPEG-7 spatial image intensity distribution descriptors," ISO/IEC JTC1/SC29/WG11 M5454(1999).

[13] Heon Jun Kim, Sung Bae Jun, Jin Soo Lee, Jung Min Song and Hee Youn Lee, LG Corporate Institute of Technology, "Descriptor for quantized color using HMM color model and color histogram," Proposal Doc. P669, MPEG-7 Ad-hoc Group Meeting, Lancaster, Feb. 1999.

[14] Michael Shneier, Mohamed Abdel_Mottaleb, "Exploiting the JPEG Compression Scheme for Image Retrieval," IEEE Trans. Pattern Analysis and Machine intelligence, pp. 849-853, August 1996.

[15] A.Yamada et.al., "Results on Core Experiment of Compact Color Layout D," ISO/IEC JTC1/SC29/WG11 MPEG99/M5356 Oct. 1999.