

## 색상특징과 웨이블릿 기반의 특징을 이용한 영상 검색

박 종현\*, 박 순영\*, 조 완현\*\*

\*목포대학교 전자공학과, \*\*전남대학교 통계학과  
전화 : (0636) 450-2878 / 팩스 : (0636) 454-0841

### Image Retrieval Using the Color Feature and the Wavelet-Based Feature

Jong Hyun Park\*, Soon Young Park\*, Wan Hyun Cho\*\*

\*Dept. of Electronics Engineering Mokpo National University

\*\*Dept. of Statistics Chonnam National University

E-mail : jhpark@chungkye.mokpo.ac.kr

#### Abstract

In this paper we propose an efficient content-based image retrieval method using the color and wavelet based features. The color features are extracted from color histograms of the global image and the wavelet based features are extracted from the invariant moments of the high-pass band image through the spatial-frequency analysis of the wavelet transform. The proposed algorithm, called color and wavelet features based query(CWBQ), is composed of two-step query operations for efficient image retrieval: the coarse level filtering operation and the fine level matching operation. In the first filtering operation, the color histogram feature is used to filter out the dissimilar images quickly from a large image database. The second matching operation applies the wavelet based feature to the retained set of images to retrieve all relevant images successfully.

The experimental results show that the proposed algorithm yields more improved retrieval accuracy with computationally efficiency than the previous methods.

#### I. 서론

정보검색에서 영상 및 비디오의 효율적인 검색을 위해 멀티미디어가 가지고 있는 속성을 효율적으로 추출하고 이를 이용한 데이터베이스의 구축, 색인, 그리고 검색이 수행되어야 한다. 현재 영상 데이터베이스들에서 가장 널리 보급된 검색 기술은 주석(text-annotation)에 의한 영상의 시멘틱 정보를 묘사하는 방법이다. 그러나 이러한 방법은 많은 문제점들을 내포하고 있는데, 예를 들어 영상에 대한 관점불일치로 사용자 질의에 대하여 신뢰성 있는 검색 정보를 제공하기 어렵다는데 있으며, 멀티미디어 속성을 구현하기가 어렵다는데 있다. 또한 데이터베이스 구축에 있어서 많은 시간과 비용이 소비되는 단점이 있다.

반면에 내용기반 영상검색은 색상(color), 모양(shape), 질감(texture) 등과 같은 영상의 속성을 이용하여 이러한 문제점을 해결할 수가 있다.

현재 이러한 내용기반 멀티미디어 검색 시스템이 많이 개발되어 수행되고 있는데 잘 알려진 시스템은 IBM의 Almaden Research Center에서 개발된 QBIC 시스템[1], Columbia 대학의 Visual SEEK와 Web SEEK[2], 그리고 MIT Media Lab.에서 개발된 PhotoBook[3] 시스템 등이 있다.

이와 같이 내용기반 영상검색 시스템은 광범위한 응용성을 가지고 있어 많은 연구가 수행되고 있는데 중요한 것은 자동적인 색인과 효율적인 검색을 수행하기 위해서는 영상의 속성을 잘 표현할 수 있는 특징을 추출하여야 한다. 실제로 내용기반 영상검색에서 객체의 모양 및 질감에 의한 특징 추출은 어려움이 많이 있으므로 최근에는 통계적인 특성을 이용하여 효율성을 높일 수 있는 방법들을 연구하고 있다. 이러한 시도 중에 하나로 웨이블릿 변환(wavelet transform)이 영상검색 분야에서 많은 관심을 보이고 있다.

본 논문에서는 내용기반 영상검색에 있어서 효율적이고 정확한 검색을 위하여 전체 영상에 대한 색상 특징과 웨이블릿 변환을 적용하여 high-pass band의 부 영역으로부터 질감 특징을 추출하여 검색에 적용하고자 한다.

본 논문의 구성은 II장에서 제안된 내용기반 영상 검색 방법에 대하여 설명하고 또한 전체 영상으로부터 색상 특징 정보의 추출과 웨이블릿 변환을 적용한 질감 특징 정보의 추출에 대하여 설명하였으며, III장에서는 추출된 특징을 이용한 두 단계 유사성 정합을 기술한다. 그리고 IV장에서는 제안된 방법과 기존의 방법에 대한 실험 결과를 분석하고 마지막으로 V장에서 제안된 방법에 대한 결론을 맺는다.

#### II. 제안된 내용기반 영상검색 방법

본 논문에서 제안된 영상검색 방법은 색상과 웨이블릿 특징을 기반으로한 질의(Color and Wavelet Features Based Query: CWBQ) 시스템으로 그림 1에 블록 다이어그램이 나타나 있다.

검색과정은 크게 두 단계에 의해 수행되는데 1단계에서는 색상 특징을 이용한 계략적인 필터링 작업을 수행하여서 질의 영상과의 상관도가 적은 영상들을 제거하고 2단계에서는 웨이블릿 특징 정보를 이용하여 정교한 유사성 정합을 하여 내용기반 영상검색을 수행하게 된다.

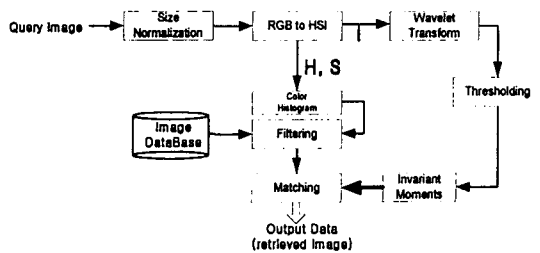


그림 1. 제안된 질의 시스템의 블록 다이어그램

### 2.1 색상 특징 추출

본 논문에서는 색상 특징을 추출하기 위하여 R.G.B(red, green, blue) 세 개의 색상으로 구성되어 있는 입력 영상의 각 화소를 인간의 시각적인 면을 가장 잘 나타낼 수 있는 H.S.I(hue, saturation, intensity) 모델로 변환한 후에 색상 특징을 추출하여 검색에 이용하였다.

색상 특징은 H.S.I 색상 공간에 대하여 Intensity에 관계없이 Hue는 20° 단위로 양자화하여 총 18개, Saturation은 2개의 그룹(bin)으로 분할하여서 총 36개 그룹으로 양자화 하였으며, 특히 Hue 성분은 인간의 시각체계에 더 민감하게 반응을 하기 때문에 세밀하게 분할하였다.

### 2.2 웨이블릿 특징 추출

#### 2.2.1 웨이블릿 변환을 이용한 영상의 분해와 분석

웨이블릿 변환을 통한 영상분석을 위하여 변환된 H.S.I 색상 모델로부터 intensity 성분  $I(x, y)$ 에 대하여 웨이블릿 변환을 적용하였으며, 2-D 웨이블릿 변환은 다음과 같다[4].

$$W_f(a, b_x, b_y) = \int_y \int_x I(x, y) \psi_{a, b_x, b_y}(x, y) dx dy \quad (1)$$

여기서  $\psi_{a, b_x, b_y}(x, y)$ 는 웨이블릿의 기저함수로서 orthogonal하며 전위(translation)와 팽창(dilation)에 의하여 얻어진다.

$$\psi_{a, b_x, b_y}(x, y) = \frac{1}{|a|} \psi\left(\frac{x-b_x}{a}, \frac{y-b_y}{a}\right) \quad (2)$$

여기서 a는 스케일링을 나타내는 양의 실수를 b는 시간 천이를 나타내는 실수이다.

다음 그림 2는 웨이블릿 변환에 의한 분할 형태를 보여주는 것으로  $N \times N$ 의 영상을 2-level 변환한 것을 나타낸다. 다해상도 웨이블릿 변환은 low-pass band 필터와 high-pass band 필터를 모두 적용하여 하위

영상을 얻을 수 있으며, 이를 반복적으로 적용하여 다해상도 영상을 얻을 수 있다.

1-level 분해에서는 1개의 low-pass band 영역 ( $O_2$ )과 3개의 high-pass band 영역 ( $WT_{2H}, WT_{2V}, WT_{2D}$ )으로 분해가 되며 2-level에서는 1-level에서의 low-pass band 영역이 더 분해가 되어  $O_4$ 의 low-low-pass band 영역과 3개의 low-high-pass band 영역( $WT_{4H}, WT_{4V}, WT_{4D}$ )으로 분해된다.

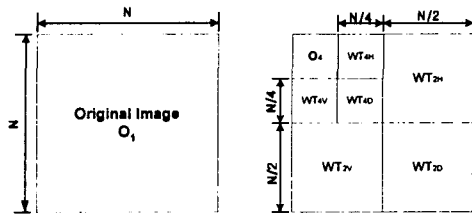


그림 2. 2-level 웨이블릿 분해

웨이블릿 변환에 의해 얻어지는  $O_4$ 는 원 영상을 변환 레벨만큼 압축한 것으로 모든 웨이블릿의 부 영역 중에서 가장 중요한 정보를 포함하고 있다. 그리고 high-pass band된 부 영역은 수직, 수평, 대각선 방향의 영상에 대한 에지 정보를 나타내고 있다.

#### 2.2.2 Invariant moments

웨이블릿 분해에 있어서 high-pass band의 부 영역에 나타나는 에지 부분의 계수는 다른 부분보다 높게 나타나므로 적당한 threshold 값을 부여하여 binary 형태로 각 부영역의 웨이블릿 계수들을 변환한 후에 특징벡터 얻기 위하여 invariant moments[5]를 추출한다. 예를 들어서 6개의 부영역중에서  $WT_{4H}$  영역에 있는 웨이블릿 계수들을 이진화 시킨 결과를  $B_{4H}(x, y)$ 라 가정하면 함수  $B_{4H}(x, y)$ 에 대한 차수  $(p+q)$ 의 moment는 다음 수식과 같이 정의된다.

$$m_{pq} = \sum_x \sum_y x^p y^q B_{4H}(x, y) \quad (3)$$

1차 central moment  $\mu_{00}$ 는  $m_{00}$ 에 의하여 표시할 수 있으며, 차수  $(p+q)$ 의 central moments는 다음 (4)식으로부터 정의된다.

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q B_{4H}(x, y) \quad (4)$$

여기서  $\bar{x} = m_{10}/m_{00}$ ,  $\bar{y} = m_{01}/m_{00}$  이다.

Normalized central moments는  $\eta_{pq}$ 로 나타내며

$$\eta_{pq} = \mu_{pq} / (\mu_{00})^r \text{로 정의된다.}$$

여기서  $r = \frac{1}{2}(p+q) + 1$  for  $p+q=2, 3, \dots$  이 되며 2차, 3차 moments로부터 부영역  $WT_{4H}$ 에 해당되는 7개의 invariant moments[10] 계수를 추출할 수가 있다.

### III. 두 단계 유사성에 의한 정합

유사성 정합에 있어서 커다란 데이터베이스로부터 영상의 특징들을 모두 비교한다면 많은 시간이 소모되므로 두 단계 유사성 정합을 위하여 1단계에서는 개략적인 유사성 정합을 그리고 2단계에서는 정교한 유사성 정합 방법을 사용하였다. 즉, 색상 특징을 이용한 개략적 단계의 유사성 정합은 커다란 데이터베이스로부터 빠르게 유사하지 않은 영상들을 제거하여 후보 영상을 압축하고, 정교한 단계의 유사성 정합에서는 선별된 후보 영상들로부터 웨이블릿 특징을 적용하여 질의 영상과 관련된 모든 영상들을 세밀하게 검색하게 된다.

#### 3.1 개략적 단계의 유사성 정합

개략적 단계의 유사성 정합은 초기 단계의 정합으로 적당하지 않은 영상을 제거하기 위한 처리 과정이다. 유사성 정합은 추출된 질의 영상의 히스토그램 특징  $\{Q\}$ 와 데이터베이스 영상의 히스토그램 특징  $\{T\}$ 의 유클리디언 거리에 의하여 측정되며 다음 식(5)과 같이 나타낼 수 있다.

$$coarse\_dist(\{Q\}, \{T\}) = \sqrt{\sum_{i=1}^N (q_i - t_i)^2} \quad (5)$$

여기서  $q_i$ 와  $t_i$ 는  $i$ 번째의 특징 벡터를 나타내며  $N$ 은 색상 히스토그램에 해당하는 그룹의 개수이다.

#### 3.2 정교한 단계의 유사성 정합

정교한 단계의 유사성 정합은 영상을 더 세밀하게 검색하기 위한 단계로 1단계에서 선별된 후보 영상들을 재차 검색하여 질의 영상을 대표하는 영상들을 검색하여 준다. 유사성 정합은 추출된 벡터 영상의 웨이블릿 변환에 의해 추출된 특징벡터  $\{Q\}$ 와 데이터베이스 영상의 특징벡터  $\{T\}$ 의 유클리디언 거리를 사용하며 식(6)과 같이 나타낼 수 있다.

$$fine\_dist(\{Q\}, \{T\}) = \sum_a \sum_{a'} \sqrt{\sum_{i=1}^N (q_{ai}^L - t_{a'i}^L)^2} \quad (6)$$

여기서  $L$ 은 웨이블릿 변환 레벨을  $a$ 는 분할된 영역을 나타내며  $q_{ai}^L$ 와  $t_{a'i}^L$ 는  $i$ 번째의 특징 벡터를  $N$ 은 정규화된 모우먼트 계수의 개수이다.

### IV. 실험 및 결과

본 논문에서 제안된 내용기반 영상검색 시스템의 효율성을 실험하기 위해 영상을 잘 표현할 수 있는 색상과 웨이블릿 특징을 추출하여서 질의 영상과 데이터베이스 영상 사이의 유사성에 의한 정합을 하였다.

실험에 사용한 데이터베이스 영상은 Corel Photo CD 영상 240장을 혼합하여 데이터베이스로 사용하였으며 그림 3에 데이터베이스 영상 일부를 보여주고 있다. 데이터베이스 영상의 크기는  $120 \times 80$ 으로 총 8개의 class로 구성되어 있으며, 각 class는 30개의 유사한 형태를 가지고 있는 영상들로 구성되었다. 질의 방법

은 class 내의 영상들로부터 임의적으로 한 개씩 선택하여 질의 영상으로 삼고 나머지 영상들로 데이터 베이스를 구성하는 실험을 반복하였다.

본 실험에서는 웨이블릿 특징을 추출하기 위하여 4개의 탭을 갖는 Daubechis' 웨이블릿 필터를 사용하였으며 2-level로 분해하여 공간-주파수 분석을 하였다.

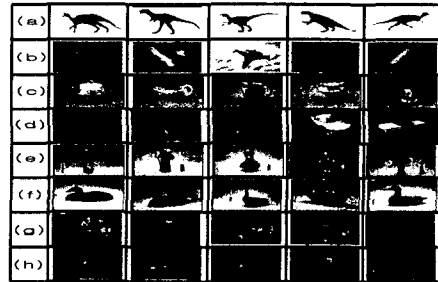


그림 3. 실험에 사용된 데이터베이스의 일부

제안된 CWBQ 시스템의 동작 특성을 분석하기 위하여 James Ze Wang의 WBIS[6]와 Changliang Wang이 제안한 spatial-frequency(S/F) analysis방법[4]을 사용하여서 검색의 효율성을 비교하였다.

그림 4의 (a)는 James Ze Wang이 제안한 WBIS 방법을 사용하여 airplane 영상을 질의로 하였을 때 유사성이 높은 순서로 상위 15개의 검색된 결과를 보여주고 있는데 모양과 색상 측면에서 유사하지 않은 duck\_doll 영상이 다수 검색되었음을 볼 수 있다. 그림 4의 (b)는 Changliang Wang이 제안한 S/F 방법으로 질의 영상과 전혀 관계가 없는 영상이 첫 번째로 검색되었으며 car 영상이 많이 검색되어 검색의 효율성이 떨어짐을 볼 수 있다. 그림 4의 (c)는 제안된 방법을 적용한 결과로 1단계에서는 색상 특징 벡터를 이용하여 추출된 90개의 후보 영상을 선별한 후에 2단계에서는 웨이블릿의 특징을 적용하여 유사성이 높은 상위 15개의 영상을 보여주고 있는데 같은 종류의 airplane 영상이 우선 순위가 높게 검색됨을 볼 수 있다.



(a)



(b)



(c)

그림 4. Airplane 영상의 질의에 대한 검색 영상  
(a) WBIS 방법, (b) Changliang S/F 방법,  
(d) 제안된 CWBQ 방법

표 1은 검색을 수행한 후 class-confusion matrix 형태[7]로 변환시킨 결과를 보여주고 있다. 왼쪽 column에 있는 영상들은 실험에 사용된 질의 영상들로서 데이터베이스에는 포함되지 않으며 이들을 이용하여 각각의 검색을 수행한 후에 유사성이 높은 30개의 영상을 데이터베이스로부터 추출하여 각 class에 소속된 영상들의 개수를 조사하는 방법으로 질의 영상을 바꾸어 가면서 5번 수행한 후에 평균값을 백분율로 환산한 결과가 오른쪽 column에 나타나 있다. 예를 들어 CWBQ의 경우 dinosaur는 88%가 같은 class의 영상들로부터 정확히 분류되었으며 그 외는 1%의 airplane, 10%의 car 영상이 클래스로부터 잘못 분류되었음을 나타내고 있다.

결과적으로 제안된 CWBQ 시스템에서 영상내 색상 특징 및 웨이블릿 특징을 검색에 이용함으로써 색상 히스토그램에서의 단점들을 극복할 수 있으며 검색에 있어서 모양이 유사한 종류의 영상이 우선 순위가 높게 나타나 효율적인 영상검색이 수행됨을 관찰할 수가 있다.

표 1. class-confusion matrix

(a) CWBQ(단위:%)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	88	1		10				1
b)Airplane		30		1	5	41		3
c)China	6	4	62	26				1
d)Car	27	5	31	20	2	7		7
e)Doll	1	1			60	33		5
f)Duck_doll		8			15	75		1
g)Crystal	1	5	8	1	28	7	48	
h)Bird	25	1	1	7	23	9	32	

(b) WBIS(단위:%)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	95	1		3	3			
b)Airplane	21	23	1	5	18	27		1
c)China		5	32	7	31	14	6	3
d)Car		17	26	35	14	5		1
e)Doll			24	1	35	37		
f)Duck_doll		3	16	4	40	38		1
g)Crystal		6	9	5		1	73	6
h)Bird		5	6		1	5	10	7

(c) Changliang의 S/F(단위:%)

class	classification							
	a	b	c	d	e	f	g	h
a)Dinosaur	61	7	10	4	5	8		2
b)Airplane		23	3	32	6			3
c)China		14	19	15	6	11	8	25
d)Car		23	4	28	6	4	2	31
e)Doll		9	21	8	23	13	5	19
f)Duck_doll	1	13	12	13	19	17	3	19
g)Crystal		6	13	11	8	3	32	24
h)Bird		15	2	33	7	1	5	35

## V. 결론

본 논문에서 우리는 영상을 구성하고 있는 속성중 영상의 H.S.I 색상 모델로부터 색상 특징의 추출과 intensity 성분으로부터 공간-주파수 분석을 통한 웨이블릿 특징을 추출하여 영상검색에 적용함으로써 검색의 효율성을 증대시키기 위한 방법을 제안하였다.

유사성 정합은 색상 히스토그램을 이용한 개략적인 정합과 웨이블릿 특징을 이용한 정교한 정합을 단계적으로 수행하는 두 단계 특징 기반 유사성 정합을 수행함으로써 빠른 검색을 수행할 수 있었으며 후보 영상으로부터 질의 영상과 관련된 영상을 더 세밀하게 분석함으로써 효율적인 내용기반 영상 검색을 수행할 수 있었다.

제안된 CWBQ 방법을 기존의 방법들과 서로 비교 분석하였으며 실험결과 전반적으로 제안된 방법이 우수한 결과를 보여 주었으며 또한 웨이블릿 특징 정보를 검색에 적용함으로써 효율적으로 검색 능력을 향상시킬 수가 있었다.

## 참고문헌

- [1] W.Niblack, R. Barber, W. Equitz, M. Flickner, E. Glasman, D. Petkovic, and P.Yanker, "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape, Storage and Retrieval for Image and Video Databases," Vol. 1908, pp. 173-187, SPIE, Feb. 1993.
- [2] John R. Smith and Shi-Fu Chang, "VisualSEEK: a fully automated content-based image query system," ACM Multimedia 96, Nov. 20, 1996.
- [3] A. Pentland, R.W. Picard, and S. Sclaroff, "Photobook: Tools for content-based storage and manipulation of image databases," Storage and Retrieval for Image and Video Databases II, pp. 34-47, San Jose, CA, SPIE, Feb. 7-8 1994.
- [4] Changliang Wang, Kap Luk Chan, Stan Z Li, "Spatial-Frequency Analysis for Color Image Indexing and Retrieval," ICARCV 98', vol. 2, pp. 1461-1465, Apr. 1998.
- [5] Babu M. Mehtre, Mohan S. Kankanhalli and Wing Foon Lee, "Content-based Image Retrieval Using a Composite Color-Shape Approach," Information Processing and Management, Vol. 1. pp. 109-120, 1998.
- [6] James Ze Wang and Sha Xin Wei, "Wavelet-Based Image Indexing Techniques with Partial Sketch Retrieval Capability," Proceeding of the 4th Forum on Research and Technology Advances in Digital Libraries, pp.13-24, Washington D.C. May 1997.
- [7] Serge Belongie, Chad Carson, Hayit Greenspan, and Jitendra Malik, "Color- and Texture-based Image Retrieval," IEEE Computer Society, pp. 675-682, Jan. 1998.