

# 내용 기반 멀티미디어 정보 검색 시스템의 설계 (Design of a Content-based Multimedia Information Retrieval System)

박 민 식\*      유 기 형\*\*  
(Min-Sheik Park) (Gi-Hyoung Yoo)

## 요 약

최근, 여러 가지 멀티미디어 데이터베이스를 통한 영상 정보의 인터넷 검색에 대한 많은 관심을 가지고 있으며, 영상 정보 검색에 대한 연구가 활발히 이루어지고 있다. 본 논문에서는 웨이블릿 변환과 상관관계(correlation)를 이용한 내용기반 검색 시스템을 제안한다. 정합 방법은 완전 정합이 아닌 유사 정합을 사용하는 효율적인 특징벡터 추출 알고리즘을 제안한다. 모의 실험한 결과 기존의 알고리즘보다 데이터베이스에 저장된 영상간의 유사성을 추정하는데 빠르고 정확하게 영상 정합을 수행함을 알 수 있었다. 개선된 결과를 통해 제안한 특징 벡터 사용이 원 영상에 비해 256 : 1 만큼의 압축 효과를 가지며, 상관관계가 영상의 정합 방법에 효율적임을 알 수 있었다.

## ABSTRACT

Recently, issues on the internet searching of image information through various multimedia databases have drawn an tremendous attention and several researches on image information retrieval methods are on progress. By incorporating wavelet transform and correlation matrixes, we propose a novel and highly efficient feature vector extraction algorithm that has an capability of a robust similarity matching. The simulation results have yielded a faster and highly accurate candidate image retrieval performance in comparison to those of the conventional algorithms. Such an improved performance can be obtained because the used feature vectors were compressed to 256 : 1 while the correlation matrixes are incorporated to provide a finer information for the better matching.

## 1. 서론

최근 정보화가 급속히 발전하고 그에 따른 다양한 멀티미디어 데이터베이스를 통해 영상정보를 찾는 일이 빈번해지고 있다. 따라서, 영상정보를 저장하고 검색하기 위한 연구가 활발하게 진행되고 있다. 지금까지 영상검색에 대한 연구는 두 가지 방법으로 진행되어 왔다.

첫 번째 방법은 먼저 대상이 되는 각각의 영상에 사람이 문자 주제어(keyword)를 일일이 부과하고 이를 이용하여 영상을 검색하는 텍스트 기반 검색기법이다. 이 방법은 제한된 범위 내에서는 정확도가 높다는 장점을 가지고 있지만 방대한 영상정보에 대해 사람이 일일이 주제어를 부여하기가 쉽지 않고, 같은 영상에 대해 다른 주제어가 주어졌다면 검색

\* 정희원 : 전주공업대학 컴퓨터정보과 조교수

\*\* 정희원 : 백제예술대학 겸임교수

※ 본 연구는 전주공업대학 2000년도 학술연구조성비에 의하여 연구되었음.

논문접수 : 2001. 7. 27.

심사완료 : 2001. 8. 6.

이 불가능하다는 단점이 있다. 두 번째 방법은 영상 자체 내에서 대표할 수 있는 특징을 추출하여 이를 기반으로 색인 및 검색하는 내용기반 검색기법이다. 이 방법은 영상 내에서 고유한 특징을 추출하기 때문에 시간 및 인력 소모를 줄일 수 있다는 장점을 가지고 있기 때문에, 최근 효율적인 특징벡터를 찾는 데 많은 연구가 진행되고 있다.

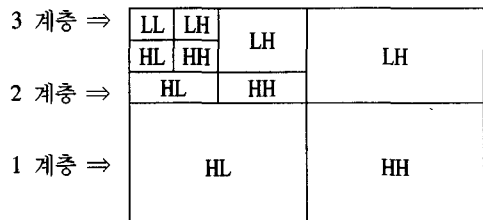
영상을 이용한 내용 기반 정보검색 시스템은 최근 디지털 도서관에 필요한 멀티미디어 정보검색의 필요성에 따라 Columbia 대학의 VisualSeek 시스템 [1], MIT Media Lab의 Photo Book[2] 등 미국 및 캐나다[3] 등지에서 광범위하게 연구되고 있으며 그 응용성 역시 다양해지고 있는 추세이다. 한편, 영상 표현의 일관성을 위해 MPEG-7에서는 비디오 표현에 대한 표준화 작업에 착수하고 있으며, 향후 2~3년 안에 표준화 작업이 완성될 것이다. 최근 국내에서도 대학 및 연구소에서 대상체 이미지의 형태, 색, 질감 등의 특징을 기반으로 부분적으로 연구하고 있으나 아직 초기 단계로 볼 수 있다.

내용 기반 검색 시스템 개발은 크게 두 가지 영역으로 나누게 된다. 첫째, 이미지 프로세싱(Image Processing) 관점에서 영상(Image) 데이터의 특성과 변환(Transform) 방법을 연구하고, 이미지 특징 벡터 (Feature Vector)를 추출하여 검색하는데 사용하는 것이다. 둘째, 데이터베이스 관점에서 영상으로부터 추출된 특징벡터를 사용하여 효과적으로 영상 데이터를 인덱싱(Indexing)하는 것이다. 위에서 언급한 국내의 연구 분야는 후자에 해당이 된다. 본 연구에서는 영상 데이터의 특성과 변환(Transform) 방법을 연구하고, 영상 데이터의 특징 벡터를 추출하여 영상을 검색하는데 초점을 맞춘다. 최근 신호처리와 영상처리에 널리 사용되고 있는 웨이블릿 변환을 이용하여 영상을 변환하고 특징벡터를 추출한 후 특징 벡터들을 사용하여 예제 영상과 후보 영상 간의 정합을 시도한다. 이는 특징벡터들이 원 영상에 비해 크기가 작기 때문에 검색의 효율성을 높일 수 있다는 장점이 있다.

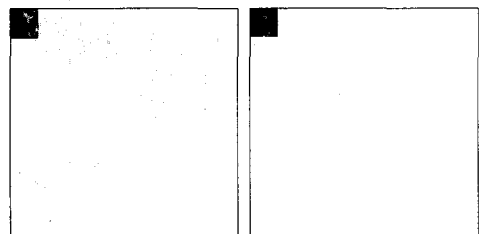
## 2. 웨이블릿의 특징과 전처리 과정

### 2.1 웨이블릿 변환의 특징

웨이블릿 변환은 원 영상을 저주파 성분과 고주파 성분으로 분리하여 분해하게 된다. 우리가 사용하는 영상은 보통 2차원의 영상인데, 웨이블릿 변환으로 분할된 상태를 나타내면 [그림 1]과 같다. 이렇게 분할된 영상에는 저주파 성분(LL대역) 쪽으로 대부분의 에너지가 집중되어 있다. [그림 1]은 Adelson의 5 tap QMF filter[4,5,6]를 이용하여 Lenna 영상과 Cronkite 영상을 각각 3계층으로 분할한 것이다. 그림을 통해서 대부분의 에너지들이 저주파 영역으로 집중되어 있고 고주파 영역은 영상의 윤곽선을 나타내고 있음을 알 수 있다. 여기에서 말하는 계층은 하나의 영상에 웨이블릿 변환을 수행하는 횟수와 관계가 있다.



[그림 1] 3계층으로 분할한 웨이블릿 변환  
[Fig. 1] 3 layer decomposed wavelet transform



(a) Lenna (b) Cronkite

[그림 2] 웨이블릿 분할된 영상  
[Fig. 2] Decomposed image by wavelet

## 2.2 전처리 과정(Preprocessing)

전처리 과정에서는 후보 영상들에 대해서 웨이브렛 변환을 수행하고 정합하는데 필요한 특징벡터를 추출한다. 먼저 영상의 크기가 다양하기 때문에 크기를 표준화하는 작업이 필요하다. 본 논문에서는 256×256으로 영상의 크기를 표준화할 것이다.

특징벡터를 구하는 과정은 다음과 같다.

- 1) 웨이브렛 변환을 통해서 4계층까지 저주파 대역을 분해한다.
- 2) 각각의 후보 영상들에 대해 16×16 크기의 4계층 LL대역의 에너지 값을 특징벡터로 추출한다.
- 3) 4계층에서 LL대역을 제외한 나머지 대역으로 윤곽선 영상을 만들고 이를 특징벡터로 추출한다.
- 4) 4계층 LL대역 자체를 특징벡터로 추출하여 Euclidean distance를 계산하는데 사용한다.

웨이브렛 특성상 LL대역은 원 영상의 특징을 대부분 가지고 있으면서도 크기는 256 : 1만큼 압축된 효과를 주기 때문에 검색효율을 높일 수 있는 훌륭한 특징벡터가 된다. 좀더 검색효율을 높이기 위해서는 간단하면서도 영상의 특성을 반영하는 값을 사용하여 유사 정합되는 영상의 수를 줄일 필요가 있다. 본 논문에서는 그러한 값으로 16×16크기의 4계층 LL대역의 에너지 값을 사용한다. 에너지 값을 구하는 식은 다음과 같다[5].

$$E = \sum_x \sum_y |f(x, y)|^2 \tag{1}$$

에너지 값에 더하여 윤곽선 정보를 특징벡터로 사용한다. 윤곽선을 추출하는 방법은 다양하다. 하지만, 본 논문에서는 윤곽선 추출을 위해 별도의 작업이 필요치 않다. [그림 2]를 통해서 알 수 있는 바와 같이 저주파 대역을 제외한 나머지 대역이 가로와 세로, 대각선 방향의 윤곽선을 잘 나타내고 있다. 좀더 뚜렷한 윤곽선 정보를 얻기 위해 세 대역(LH, HL, HH)을 조합한다. 즉, 세 개의 대역이 같은 해상도를 가지고 있고 같은 위치의 영상정보를 포함하고 있으므로, 각 pixel에 대해 최대값을 선택한 후 0보

다 크면 1로 셋팅한다. 이렇게 구해진 영상은 가로, 세로, 대각선 방향의 윤곽선 정보를 모두 포함하고 있다. 본 논문에서는 4계층의 대역들을 사용하여 16×16 크기의 윤곽선 영상을 만들었다. 만들어진 윤곽선 정보는 상관도를 사용하여 영상간의 유사도를 비교한다. 상관도는 신호처리나 영상처리에 많이 사용되며 두 영상간에 얼마만큼의 유사도가 있는지를 조사하여 비슷한 영상을 찾을 수 있는 방법이다. 이를 식으로 나타내면 다음과 같다.

$$C = \sum_x \sum_y f(x, y) w(x, y) \tag{2}$$

$f(x, y)$ 는 예제 영상이며,  $w(x, y)$ 는 후보 영상을 나타낸다. 영상에서의 상관도는 자기 자신과의 상관도(Autocorrelation)가 가장 높다.

예제 영상과 후보 영상의 각 pixel 간의 차분을 계산하는 Euclidean distance는 식 (3)에서와 같이 정의된다[6,7,8,9,10].

$$D = \sum_x \sum_y |f(x, y) - w(x, y)| \tag{3}$$

식 (3)을 통해서 알 수 있듯이 비교하는 영상이 비슷할수록 값은 작아질 것이고, 만약 비교하는 두 영상이 같다면 값은 0이 될 것이다. 따라서, Euclidean distance값이 가장 작은 것부터 M개 만큼을 최종적인 검색 결과로 디스플레이한다.

## 3. 예제 영상과 후보 영상 간의 정합 과정

영상에서 특징벡터를 추출하는 내용기반 검색에서는 키워드를 이용한 검색처럼 완전정합 되는 경우는 드물다. 따라서 본 논문에서는 정합된 결과에 대해 M개의 영상들을 보여주는 유사정합을 이용한다. 본 논문에서 제안한 검색 과정은 <알고리즘 1>과 같다.

### <알고리즘 1>

1. 예제 영상이 들어오면, 웨이브렛 변환을 4계층까지 수행한 후, 후보 영상에서 처럼 4계층 LL대역의 에너지 값, LL대역 그 자체, 그리

고 4계층에서 고주파 대역을 이용한 윤곽선 영상을 특징벡터로 추출한다.

- 예제 영상과 후보 영상의 에너지 값을 비교하여 식(4)의 조건을 만족하는 후보 영상들을 선택한다.

$$E - E(K_1/100) \leq E^* \leq E + E(K_1/100) \quad (4)$$

$E$  : 예제 영상의 에너지 값.

$E^*$  : 후보 영상의 에너지 값

- 식(4)에 의해서 선택된 후보 영상과 예제 영상의 윤곽선 정보간의 상관도 값을 계산하여 식(5)를 만족하는 후보 영상들을 선택한다.

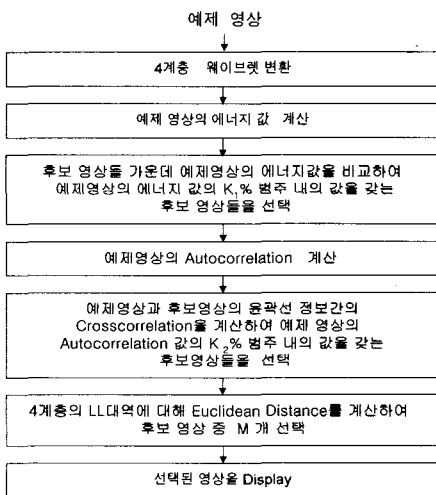
$$C - C(K_2/100) \leq C^* \leq C \quad (5)$$

$C$  : 후보 영상 가운데 가장 큰 상관도 값.

$C^*$  : 후보 영상들의 상관도 값.

- 선택된 후보 영상들과 예제 영상의 4계층 LL 대역에 대해 Euclidean distance를 계산하여 M개를 최종적인 검색결과로 디스플레이한다.

[그림 3]은 위에서 설명한 예제 영상과 후보 영상 간의 검색 과정을 블록도로 나타낸 것이다.



[그림 3] 제안된 알고리즘  
[Fig. 3] The proposed Algorithm

## 4. 실험 결과

내용 기반 검색 시스템의 여러 성능을 측정하기 위한 뚜렷한 표준이 아직 없기 때문에 대부분의 내용 기반 이미지 검색 시스템들은 성능을 평가하지 않거나, 평가를 하더라도 자기들 나름대로 기준을 만들어 자체적으로 평가하고 있는 것이 전부이다.

본 논문에서는 1300개 이상의 256×256 영상을 이용하여 실험하였다. 실험은 두 가지로 나눌 수 있는데, 첫 번째 실험은 일반적인 영상에 대한 검색이며 두 번째 실험은 크기가 변하며, 이동하는 물체에 대한 검색을 보기 위하여 연속되는 영상들을 사용하였다.

### 4.1 일반 영상에 대한 검색 결과

다음은 예제 영상으로 Lenna 영상을 사용한 결과이다.

#### 4.1.1 에너지 값에 의한 검색 결과

- lenna 영상의 에너지 값 = 4348219
- 검색 결과 = 164개

<표 1> 에너지 값으로 검색된 결과

<Table 1> Retrieval result by energy value

Image	에너지 값
ls0202	3914445
ls0107	3915592
...	
getvising	4241818
lenna	4348219
sailboat	4415167
...	
car5_385	4777968
car2_030	4781940

#### 4.1.2 상관도에 의한 검색 결과

- 검색 결과 = 13개

<표 2> 상관도를 이용하여 검색한 결과  
 <Table 2> Retrieval result by Correlation

Image	Correlation	Image	Correlation
lenna	248	mars2	239
bridge	248	light	236
ape	246	kiel256r	231
peppers	244	kiel256	229
goldhill	244	chiesa	229
clown	242	unob	226
sailboat	241		

4.1.3 Euclidean Distance에 의한 검색 결과

<표 3> 최종 검색 결과

<Table 3> The final retrieval result

Image	에너지 값	Correlation	Distance
lenna	4348219	248	0
kiel256	3795840	229	10927
kiel256r	3447745	231	11292
light	3528707	236	12372
ape	4513524	246	12373
sailboat	4415167	241	12512
peppers	3959392	244	13053
unob	3316016	226	13127

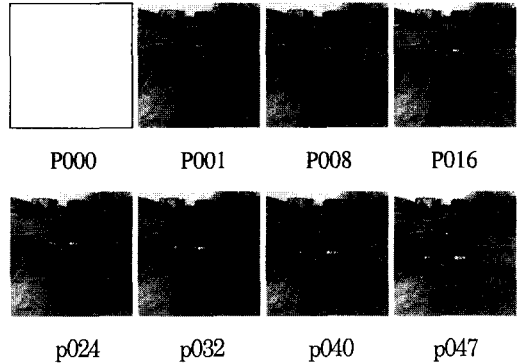
4.2 연속 영상에 대한 검색 결과

<표 4>와 [그림 4]는 예제 영상으로 p000을 사용한 최종 결과이다. 68개의 영상이 최종적으로 검색되었고, 그 중에서 8개의 영상을 선택했다.

<표 4> 연속 영상에 대한 검색 결과

<Table 4> Retrieval result for moving images

No.	Image	에너지 값	Correlation	Distance
1	p000	3643989	227	0
4	p001	3604507	216	289
9	p008	3623096	226	641
17	p016	3609933	225	1004
25	p024	3591683	225	1294
33	p032	3564635	226	1576
41	p040	3526776	222	2133
48	p047	3474362	224	2849



[그림 4] 최종 검색 결과

[Fig. 4] The final retrieval result

5. 결론 및 고찰

본 논문에서는 간단하면서도 영상의 특성을 잘 반영하는 효율적인 특징벡터 추출 기법을 제안하였다. 영상 데이터는 웨이브렛 변환을 사용함으로써 공간 영역의 정보가 잘 보존되었고, 웨이브렛 변환 영역에서 계산된 저대역의 에너지 값과 상관도를 사용하여 유사한 영상을 검색하였다.

제안한 알고리즘으로 모의 실험한 결과 기존의 알고리즘보다 빠르고 정확한 결과를 얻을 수 있음을 확인하였다. 사용된 특징벡터들이 원 영상에 비해 256 : 1만큼의 압축효과를 얻을 수 있었기 때문에, 속도면에서 검색 효율을 높일 수 있었다. 그리고, 상관도를 이용하여 크기가 변하면서 이동하는 물체에 대해서도 별도의 객체 추출없이 검색이 가능하였다. 앞으로 압축된 영상에 대한 검색기법의 개발과 예제 영상 뿐만 아니라 사용자의 그림에 의한 영상 검색에 대한 연구가 계속되어야 할 것이다.

※ 참고문헌

박 민 식

[1] John R. Smith and Shih-Fu Chang, "Visual-SEEK : a fully automated content-based image query system," [http:// www. ctr.columbia.edu/ VisualSeek](http://www.ctr.columbia.edu/VisualSeek).

[2] R.W. Picard, "A Society of Models for Video and Image Libraries," <http://www.media.mit.edu/~picard/>.

[3] Zw-Nian Li and Bing Yan, "Recognition Kernel for Content-based Search," IEEE International Conf. on SMC, vol. 1, pp. 472 - 477, Beijing, China, Oct. 1996.

[4] E.H. Adelson, "Orthogonal Pyramid transform for image coding," In Proc. SPIE, Vol 845, pp.50-58, Cambridge, MA, October 1987.

[5] Rafael C. Gonzalez and Richard E. Woods, Digital Image Processing, Addison-Wesley Publishing Company, 1993

[6] James Ze Wang, Gio Wiederhold, Oscar Firschein, and Sha Xin Wei, "Applying Wavelets in Image Database Retrieval", [http:// www-leland.sanford.edu/~zwang](http://www.leland.sanford.edu/~zwang)

[7] I. Daubechies, "Ten lecture on wavelets," SIAM, Philadelphia, PA, 1992.

[8] M. Vetterli, Wavelet and Subband Coding, Prentice Hall Signal Processing Series, pp. 72-79, 1995.

[9] Committee Draft for MPEG-4, Swiss, 1997.

[10] G.K. Wallace, "The JPEG still picture compression standard," Comm. of ACM, 34(4): 30-44, 1991.

1983년 2월  
전북대학교 전기공학과(학사)

1985년 2월 전북대학교  
전기공학과(석사)

1986년 9월~1990년 2월  
금성통신(주) 연구소

1994년 2월 전북대학교  
전자공학과(박사)

1994년 3월~현재  
전주공업대학 컴퓨터정보과  
조교수

관심분야 : 영상신호처리,  
정보검색, Fuzzy system 등

유 기 형

1989년 2월 전북대학교  
컴퓨터공학과(학사)

1991년 2월 전북대학교  
전자공학과(석사)

1991년 11월~1996년 6월  
전북대학교  
반도체물성연구소

1999년 8월 전북대학교  
컴퓨터공학과 박사수료

1999년 3월~현재  
백제예술대학 겸임교수

관심분야 : 영상처리,  
패턴인식