

## 특징 영역 추출을 통한 내용 기반 영상 검색

이근섭, 정승도, 조정원, \*최병욱  
한양대학교 전자통신전파공학과, \*전자전기컴퓨터공학부  
전화 : 02-2290-0363 / 핸드폰 : 019-204-2941

### Content-based Image Retrieval by Extraction of Specific Region

Geun-Seop Lee, Seung-Do Jeong, Jung-Won Cho, Byung-Uk Choi  
Dept. of Electrical and Computer Engineering, Hanyang University  
E-mail : sousage@mlab.hanyang.ac.kr

#### Abstract

In general, the informations of the inner image that user interested in are limited to a special domain. In this paper, as using Wavelet Transform for dividing image into high frequency and low frequency, We can separate foreground including many data. After calculating object boundary of separated part, We extract special features using Color Coherence Vector.

According to results of this experiment, the method of comparing data extracting foreground features is more effective than comparing data extracting features of entire image when we extract the image user interested in.

#### I. 서론

최근 멀티미디어 데이터의 양이 폭발적으로 증가함에 따라 데이터의 효과적인 분류 및 검색 작업의 필요성이 요구되고 있다. 특히 영상에 대해서 사용자가 원하는 정보를 객관적인 기준을 통해 자동으로 제공할 수 있는 내용 기반 영상 검색 방법에 관한 연구가 활발히 진행 중이다. 내용 기반 영상 검색은 질의 영상(query image)에 대해서 영상에 포함된 내용을 기반으로 시각적으로 유사한 영상들을 검색하고 결과를 제시

하는 방법이다. 유사도는 영상에서 얻은 특징 정보를 비교하여 계산하게 되는데, 이러한 특징들은 주로 색상이나 형태, 질감 등이 이용되고 있다. 특히 색상 정보의 경우는 특징 정보 추출 과정이 비교적 쉽고 영상의 특성을 잘 나타내고 있기 때문에 영상을 구별하는 특징 정보로써 많이 사용되고 있다. 대표적인 색상 기반의 영상 검색 기법으로 히스토그램 비교법[1], 모멘트를 비교하는 방법[2], Color Correlogram[3] 등이 있다. 그러나 보통 영상 내에서 사용자가 원하는 정보는 영상의 전체적인 특성보다는 영상 내의 특정한 부분에 국한되어 있다. 기존의 내용 기반 영상 검색 방법은 영상의 전체 영역에서 특징 정보를 추출하였기 때문에 사용자가 원하는 정보를 추출하는데 있어 효과적이지 못한 단점이 있었다. 예를 들면 사용자가 원하는 정보가 영상 내의 어느 특정한 사물이지만 배경의 색상들로 인해 잘못 표현되는 경우가 많다. Color Correlogram은 색상 정보와 함께 공간적인 정보도 포함하지만 특정 사물에 대한 정보를 나타내지는 못한다. 또한 특징 추출을 위한 계산 시간이 오래 걸리고 특징 벡터의 차원수가 크다는 단점이 있다[4]. 본 논문에서는 이러한 점을 개선하기 위해 영상을 이산 웨이블릿 변환을 통해 높은 해상도를 갖는 부분과 낮은 해상도를 갖는 부분으로 분리해 낸 후 높은 해상도를 갖는 부분에서 특징 정보로써 Color Coherence Vector (CCV)를 추출하였다.

## II. 영상 내 특징 영역 선정

영상 검색시 사용자가 원하는 정보는 보통 영상 내에서 특정한 영역에 국한되어 있다. 특히 영상 내에서도 배경 부분보다는 해상도가 높은 전경 부분에 집중되어 있다. 그렇기 때문에 영상에서 특징 정보를 추출할 때 전체 영역에서 특징 정보를 추출하게 되면 불필요한 정보들이 포함되게 된다. 따라서 사용자가 원하는 정보를 포함하는 영상을 검색하는데 있어 효과적이지 못한 단점이 있다. 본 논문에서는 이산 웨이블릿 변환(Discrete Wavelet Transform)을 이용하여 영상 내에서 해상도가 높은 전경 부분과 해상도가 낮은 배경 부분을 분리하였다. 웨이블릿 변환[5]은 2차원 영상 데이터를 주파수와 시간의 국부성을 갖도록 분해함으로써 데이터를 효과적으로 처리할 수 있는 구조를 제공하여 준다. 웨이블릿 변환된 영상 데이터는 각 부밴드별로 서로 다른 특성을 나타낸다. 주파수가 가장 낮은 밴드는 원 영상의 데이터와 같은 통계적 특성을 나타내고 있으며 나머지 부밴드들은 영상의 경계 영역에 에너지가 집중되어 있는 특성을 나타낸다. 또한 영상 데이터에 수직, 수평 두 방향으로 필터를 적용시켜서 부밴드를 만들어 내기 때문에, 분해된 각 밴드들은 수직, 수평, 대각의 방향 특성을 갖게 된다. 사용한 웨이블릿 변환은 계산 속도가 빠른 Harr 웨이블릿 변환이다. 식 (1)은 Harr 웨이블릿 변환의 스케일링 함수와 웨이블릿 함수를 각각 나타낸다.

$$\begin{aligned} \Phi_{m+1,n} &= \frac{1}{\sqrt{2}} [\Phi_{m,2n} + \Phi_{m,2n+1}] \\ \Psi_{m+1,n} &= \frac{1}{\sqrt{2}} [\Phi_{m,2n} - \Phi_{m,2n+1}] \end{aligned} \quad (1)$$

영상을 웨이블릿 변환하면 저대역 부밴드와 고대역 부밴드로 나뉘므로 영상 내의 높은 해상도를 갖는 계수들과 낮은 해상도를 갖는 계수들이 분리된다. 이때 높은 해상도를 갖는 계수들 가운데 에너지 값이 큰 계수를 추출하여 이를 다시 원 영상의 위치로 변환하면 추출된 계수의 위치가 대략적으로 영상 내에서 높은 해상도를 갖는 전경 부분 주위에 분포하게 된다. 한편 웨이블릿 계수를 추출하는 과정에서 각 웨이블릿 계수간의 대비차를 크게 하기 위해 Non-Maximum Suppression[6]을 이용하였다. Non-Maximum Suppression은 최대값을 가질 픽셀은 그 이웃의 픽셀 값보다 높은 그래디언트(gradient)값을 가져야 한다는 것이 기본 원리이다. 이 과정을 거치면 지역적으로 최대값이 아닌 픽셀들은 제거된다. 그림 1은 원 영상에

서 추출된 계수를 포함하는 사각형을 그린 결과이다. 본 논문에서는 계수를 포함한 사각형을 특징 영역으로 간주하고 사각형의 좌표를 구하여 사각형 내에서만 특징 정보를 추출하였다. 그리고 웨이블릿 계수 추출시 오검출을 막기 위해 추출된 계수들의 위치값들을 무감독 학습 신경망의 일종인 Kohonen Network[7]를 이용하여 계수 무리의 중심점을 찾았다. Kohonen Network는 각각의 입력 벡터에 대해서 가중치 벡터 중 가장 가까운 것을 승자 뉴런으로 삼아 그 뉴런의 가중치 벡터만이 갱신되어지는 규칙이다. 승자 뉴런을 결정하고 난 후에는 식 (2)와 같이 뉴런의 가중치를 조정한다.

$$W_{new} = W_{old} + \alpha (X - W_{old}) \quad (2)$$

식 (2)에서 X는 입력 벡터,  $\alpha$ 는 증가율을 나타내며  $W_{old}$  와  $W_{new}$ 는 각각 승자 뉴런의 가중치와 갱신된 가중치를 나타낸다. 이렇게 찾은 중심점에서 일정한 비율로 멀리 떨어져 있는 계수는 잘못 판단한 계수로 간주하여 무시하였다.

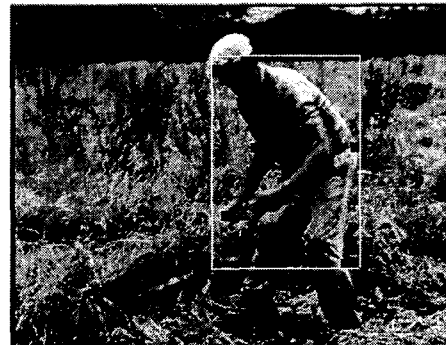


그림 1. 전경 부분을 분리해낸 영상

## III. 특징 정보 추출

영상에서 색상 정보를 이용한 검색 방법에서 주로 사용되는 히스토그램 비교법은 특징 정보를 추출하는 방법이 비교적 간단하고 빠르다는 이점이 있지만 영상의 공간적인 특성을 반영하지 못한다는 단점이 있다. 본 논문에서는 이러한 단점을 보완하기 위하여 CCV[8]를 특징 정보로 삼았다. CCV는 각각의 픽셀들을 같은 값을 가진 픽셀들의 지역에 포함되는지 포함되지 않는지를 분류하는 방법이다. CCV는 두 가지 히스토그램을 생성한다. 첫 번째 히스토그램은 영상의 동진 정도를 표현한 것이고 두 번째 히스토그램은 영상이 흩어진 정도를 표현한 것이다. CCV를 계산하기

## 특징 영역 추출을 통한 내용 기반 영상 검색

위해서는 먼저 영상에서 급격하게 변화하는 색을 제거하는 작업이 필요하다. 모든 픽셀에 대해 주위의 8개의 픽셀의 평균값으로 대체함으로써 급격하게 변화하는 픽셀을 제거하였다. 그 후에 영상을 제한된 수의 색상으로 양자화한다. 양자화한 영상에 대해서 연결 성분 라벨링(*connected component labeling*) 과정을 수행한다. 연결 성분 라벨링은 같은 색상값을 가지고 있고 서로 연결되어 있는 픽셀의 집합에는 같은 라벨을 할당하고 다른 집합에서는 서로 다른 라벨을 할당하는 방법이다. 연결 성분 라벨링은 상당한 시간이 걸리는 단점이 있지만 본 논문에서는 영상의 전체 영역에 적용하는 것이 아니라 이산 웨이블릿 변환을 통해 선정된 특징 영역 내에서만 적용하기 때문에 계산 시간을 단축할 수 있었다. 라벨링을 거친 픽셀들에 대해서 각각의 픽셀들이 결합력이 있는지(*coherence*) 없는지(*incoherence*)를 판단한다. 만약 어떤 픽셀이 속해 있는 연결 성분의 크기가 일정한 값( $\tau$ ) 이상으로 크면 그 픽셀은 결합력이 있는 것으로 판단한다. 그림 2는 영상 내에서 CCV를 추출하는 예시를 보인 것이다.

2	1	2	2	1	1	<i>B</i>	<i>C</i>	<i>B</i>	<i>B</i>	<i>A</i>	<i>A</i>
2	2	1	2	1	1	<i>B</i>	<i>B</i>	<i>C</i>	<i>B</i>	<i>A</i>	<i>A</i>
2	1	3	2	1	1	<i>B</i>	<i>C</i>	<i>D</i>	<i>B</i>	<i>A</i>	<i>A</i>
2	2	2	1	1	2	<i>B</i>	<i>B</i>	<i>B</i>	<i>A</i>	<i>A</i>	<i>E</i>
2	2	1	1	2	2	<i>B</i>	<i>B</i>	<i>A</i>	<i>A</i>	<i>E</i>	<i>E</i>
2	2	1	1	2	2	<i>B</i>	<i>B</i>	<i>A</i>	<i>A</i>	<i>E</i>	<i>E</i>

(a) 양자화된 영상                      (b) 라벨링 결과

그림 2. CCV 예제

그림 2의 (b)를 히스토그램으로 나타내면 표 1과 같다.

표 1. 라벨의 히스토그램

Label	A	B	C	D	E
Color	1	2	1	3	1
Size	12	15	3	1	5

표 1에서 A, B, E가  $\tau$  값보다 크고 C, D는  $\tau$  값보다 작다고 가정하면 A, B, E는 결합력이 있는 정보이고 C, D는 결합력이 없는 정보라 판단할 수 있다. 그러므로 이 영상에서의 CCV는 표 2와 같다.

표 2. 예제 영상의 CCV 정보

color	1	2	3
$\alpha$	17	15	0
$\beta$	3	0	1

이렇게 계산된 CCV 특징 벡터들은 내적을 통해 영상간의 유사도를 판별하게 된다.

## IV. 실험 결과

본 실험에서 사용한 테스트 영상들은 배경색이 비슷하지만 전경의 색이 전혀 다른 그룹들과, 배경색이 다르지만 전경의 색이 비슷한 그룹들을 섞어 실험하였다. CCV 특징 벡터의 추출을 위해 영상을 64색으로 양자화 하였으며, 픽셀의 뭉친 정도를 나타내는 히스토그램 64차원과 흩어진 정도를 나타내는 히스토그램 64차원을 합하여 총 128차원의 특징 벡터가 추출되었다.



그림 4. 질의 영상



그림 5. 전체 영역에서 검색한 결과 영상



그림 6. 특징 영역을 선정하여 검색한 결과 영상

그림 5는 그림 4를 질의 영상으로 입력하였을 때, 전체 영상에서 CCV 특징 벡터를 추출하여 검색한 결과이다. 질의 영상에서 주된 정보는 검은 옷을 입은

남자에 대한 정보이지만. 그림 5에서는 이러한 정보를 적절하게 반영할 수 없었다. 이러한 결과는 배경의 색이 유사한 것이 원인이 되어 발생한다. 한편 그림 6의 3번째와 4번째 그림의 경우 질의 영상과 배경색이 전혀 다르지만 특징 정보로서 추출된 전경 부분의 색이 유사하기 때문에 결과 영상으로 추출되었다.

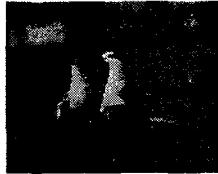


그림 7. 질의 영상



그림 8. 전체 영역에서 검색한 결과 영상

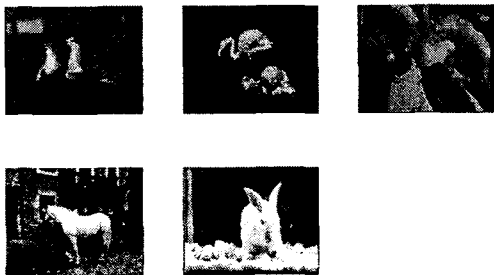


그림 9. 특징 영역을 선정하여 검색한 결과 영상

또 하나의 예제를 그림 7에서 보았다. 그림 7을 질의 영상으로 주어졌을 때 그림 8에 비해 그림 9는 다양한 배경에서도 전경 부분이 유사하기 때문에 사용자의 요구를 충족시키는 결과를 도출하였다.

또한 전체 영상에서 특징 정보를 추출하는 것 보다 특징 영역을 선정하여 특징 정보를 추출하는 방식이 평균 40% 정도의 속도 향상의 이점을 얻을 수 있었다.

## V. 결론

사용자가 원하는 정보는 영상 내에서도 해상도가 높은 전경 부분에 집중되어 있다. 기존의 방법으로는 영상에서의 전체적인 특성을 강조하였기 때문에 영상 내의 중요한 부분에 대한 특징 추출에 있어서 어려움이 많았다. 본 논문에서는 웨이블릿 변환을 이용하여 전

경 부분을 선정함으로써 사용자가 원하는 영역에 대한 특징 정보를 추출할 수 있었다. 선정된 영역에서 Color Coherence Vector를 사용하여 영역 내의 색상 정보뿐 아니라 공간적인 정보까지도 고려한 특징 정보를 추출하였다. 실험 결과, 전체 영상에서 특징 정보를 추출하여 영상을 검색하는 것 보다 특징 영역을 선정하여 특징 영역 내에서 특징 정보를 추출하여 비교하는 것이 영상을 검색하는데 있어 사용자의 요구를 보다 잘 반영하는 것이 입증되었다.

향후 연구 과제로는 영상 내의 특징 영역에서 특징 정보를 추출할 뿐만 아니라 배경에서도 간단한 특징 정보를 추출한다면 검색 성능을 높일 수 있을 거라 기대된다. 또한 특징 벡터가 고차원화 됨에 따라 검색 속도가 저하될 수 있는데 클러스터링 방법을 통해 검색 속도를 높이는 연구가 진행되어야 할 것이다.

## 참고문헌

- [1] M. J. Swain and D. H. Ballard, "Color Indexing", *Int. Journal of Computer Vision*, Vol, 7, No.1, pp.11-32, 1991.
- [2] M. Stricker and M. Orengo, "Similarity of Color Images", *Storage and Retrieval for Image and Video Databases (SPIE)*, pp. 381-392, 1995
- [3] J. Huang, et al. "Image indexing using color correlogram," *IEEE Int. Conf. on Computer Vision and Pattern Recognition*, pp. 762-76, Puerto Rico, June 1997.
- [4] Wei-Ying Ma and HongJiang Zhang, "Benchmarking of Image Features for Content-based Retrieval" *Signals, Systems & Computers*, 1998. *Conference Record of the Thirty-Second Asilomar Conference on*, Volume: 1, 1998 *Page(s):* 253 -257 vol.1
- [5] C. S. Burrus, R. A. Gopinath, and H. Guo, *Introduction to Wavelets and Wavelet Transforms*, Prentice Hall Inc, 1998.
- [6] J. R. Parker, *Algorithm for Image Processing and Computer Vision*, Wiley Computer Publishing, 1997.
- [7] Kishan Mehrotra, Chilukuri K. Mohan, Sanjay Ranka, *Elements of Artificial Neural Networks*, The MIT press, 1997.
- [8] G. Pass, R. Zabih, and J. Miller, "Comparing image using color coherence vectors," *4-th ACM Conf. on Multimedia*, Boston, November 1996.