

# 색깔과 질감을 이용한 영역별 영상 검색

## Regional Image Retrieval by using Color and Texture

곽 정 원 , 조 남 익  
서울대학교 전기공학부

Jung Won Kwak, Nam Ik Cho  
School of Electrical Engineering

Seoul National University

E-mail : [jwonk@ispl.snu.ac.kr](mailto:jwonk@ispl.snu.ac.kr) [nicho@snu.ac.kr](mailto:nicho@snu.ac.kr)

### 요 약 문

많은 정보를 포함하고 있는 영상 자료에서 빠른 검색과 분류를 위해서 색깔이나 질감 등의 특징을 나타내는 기술자가 필요하다. 또한 한 영상 안에서도 각 영역별로 다른 특징을 나타내고 있기 때문에 영역별 검색과 분류를 위한 영역 단위의 특징 추출이 중요하다. 본 논문에서는 색깔 특징으로 영역화된 영상의 각 영역에서 색깔 특징 벡터와 질감 특징 벡터를 추출하고 추출된 특징 벡터를 다른 영역에서 추출된 특징 벡터와의 거리를 이용하여 비슷한 특징을 보이는 영역을 검색한다. 기존의 전체 영상의 색깔이나 질감 어느 하나만을 이용한 검색과 달리 이러한 특징을 공간적 위치와 색깔, 질감을 조합하여 검색함으로써 보다 만족스러운 검색 결과를 얻을 수 있다.

### 1. 서론

멀티미디어 자료가 많이 사용됨에 따라 영상 등에 대한 저장 및 검색은 중요한 연구 사항이 되고 있다. 특히 영상은 단순히 몇 마디의 글로 모든 것을 표현하기에는 부족한 많은 정보를 지니고 있기 때문에 이들 영상 자료의 특징들을 추출하고 단순한 텍스트(text)가 아닌 수치적 기술자(descriptor)로 나타내고 이를 이용하여 검색하는 것이 요즘 MPEG-7등에서 많이 생각하고 있는 연구 분야이다.

본 논문에서는 영상에서 사람의 시각에 가장 민감한 성질인 색깔(color)과 질감(texture)을 고려한 영상 검색 방법을 제안하고자 한다. 좋은 성능을 보여주는 색깔에 대한 검색은 주로 색깔 벡터(color vector)를 추출하고 이를 양

자화(quantization)를 통해 대표 색깔 벡터를 구한 후 이를 이용한 검색 방법을 사용하고 있다. 또한 질감은 Co-occurrence Matrix나 Gaussian Markov Random Field(GMRF) 등을 이용한 방법이 많이 연구되어 왔으나 요즘은 특히 인간의 시각 구조와 가장 비슷한 다채널 접근법(multichannel approach)인 가버 필터(Gabor filter)를 이용한 추출방법이 많이 사용되고 있다. 그러나 대부분의 기존의 색깔과 질감 검색은 각각 독립적으로 이루어지며 또한 영역별로 국한된 정보를 사용하는 것이 아닌 영상 전체에 대한 전반적인 영상의 성질을 추출하게 된다. 특히 질감과 같은 경우에는 대부분의 검색 데이터를 순질감 영상(homogeneous texture image)을 대상으로 국한되었기 때문에 실영상(real-world image) 검색에

서는 실용적이지 못하였다. 그러므로 영상에 여러 가지 물체(object)가 존재하거나 너무 많은 색깔이나 질감의 성질을 포함하고 있는 경우에는 적절히 검색을 할 수 없었을 뿐만 아니라 색깔이나 질감 어느 하나에 대해서만 검색해야만 하였다.

본 논문에서 제안하는 방법은 영상의 부분별 영역에 대한 색깔과 질감의 정보를 동시에 고려함으로써 사람의 눈에 적절한 검색이 되도록 하는 것이다. 즉 영상을 색깔 벡터 양자화(color vector Quantization)를 통한 색깔 영역화로 사람 시각에 알맞게 나눈 후 각 영역에서 대표 색깔 특징을 뽑아내고 또한 임의의 모양인 각각의 영역에서 질감 특징 벡터를 추출한다. 이렇게 추출된 대표 색깔 특징 벡터와 질감 특징 벡터의 가중치 합을 특징 벡터로 정의한다. 결국 이 특징 벡터의 거리함수를 통하여 가장 가까운 영역을 추출하였기 때문에 영상의 부분 영역 정보만을 이용하여 만족스러운 결과를 얻을 수 있다.

본 논문은 다음과 같이 구성되어 있다. 2장에서는 색깔을 이용하여 특징 벡터를 추출하고 검색하는 방법을 소개하고 또한 질감을 이용한 특징 벡터 추출과 검색에 대해서 간단히 설명하겠다. 그리고 색깔과 질감을 동시에 고려하기 위한 임의의 영상의 영역에서 질감을 추출하는 방법과 이 질감 벡터를 색깔 벡터와의 거리를 조합하는 방법을 제안하고 3장에서는 이를 실험을 통해 확인하며 4장에서 결론과 앞으로의 연구 방향에 대해서 밝히고자 한다.

## 2. 색깔과 질감의 추출과 검색

### 2.1. 색깔 특징 추출과 검색

색깔 특징 추출 방법은 Deng *et al.* [1], [2], [3]의 방법을 기반으로 하였다. 우선 주어진 영

상을 색깔 벡터 양자화(color vector quantization) 방법[1]을 거친다. 색깔 벡터 양자화 작업은 영상의 잡음 제거와 스무딩 효과를 가져오는 피어 그룹 필터링(peer group filtering)을 하고 색깔 벡터의 초기 빈(bin)을 그리디(greedy)한 방법으로 일정한 간격으로 잡은 후 RGB 공간에서의 유클리디안 거리(euclidean distance)를 이용하여 입력 데이터의 벡터 분포가 인접한 중앙 벡터에 효율적으로 포함될 수 있도록 하는 알고리즘인 Generalized Lloyd algorithm을 이용하여 색깔 벡터의 코드북(codebook)을 갱신한다. 물론 분산에 따라 분할(split) 알고리즘과 병합(merge) 알고리즘을 사용한다.

색깔 양자화 후 이러한 색깔들의 균일성(J-value)을 정의한다. 양자화를 통해 얻어진 색깔 수로서 이루어진 영상의 각 픽셀들의 균일한 정도(J-value)를 해당 픽셀에 저장함으로써 J-image를 얻게 되고, 이것으로부터 영역을 나누게 된다. 그런 후 반복 작업을 통해 과영역화(over-segmented)된 영역지도(region map)를 얻은 후 마지막으로 색깔에 대한 상관성을 이용하여 적절한 영역은 병합함으로써 최종 영역지도를 얻게 된다[2].

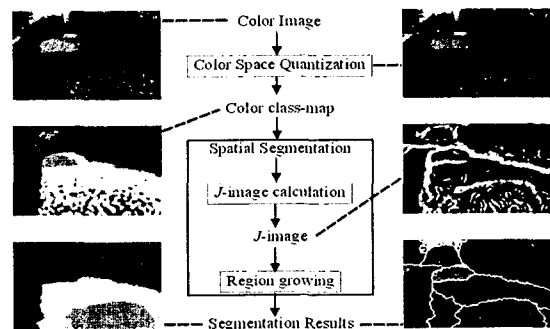


그림 1 color segmentation

이러한 색깔을 통한 영상의 영역에서 얻어진 색깔 벡터는 격자(lattice) 구조로 저장되어 검색된다. 즉 인덱싱은 양자화된 색을 기반으로 수행되며 각각의 색은 독립적으로 인덱싱 된

다. 이 때 각 색의 비율과 영상 내에서의 레이블(label) 또한 양자화된 색과 함께 데이터 베이스 노드에 저장된다[3].

## 2.2 질감 영상의 추출과 검색

질감을 분석하는 접근방법으로서 최근에 다채널 필터링(Multi-channel filtering)을 이용한 방법이 많이 사용되고 있는 이유는 이 방법이 인간의 시각적 특성(Human Visual System)과 많이 흡사하다는 것이 정신물리학적(psychophysical) 실험과 정신생리학적(psychophysiological) 실험에 의해서 증명되었기 때문이다. 포유류의 시각 조직에는 좁은 폭의 주파수(frequency)와 회전(orientation)의 조합들로 조율하여(tune) 인식하는 메카니즘이 있다. 이러한 메카니즘은 각각의 채널로 연관되고 또한 이러한 채널은 특정주파수 필터(band-pass filter)로 해석될 수 있다.

이러한 다채널 필터링 질감 분석 방법에서의 가장 중요한 관건이 되는 것은 우선 채널들의 함수적인 특징과 그 채널의 수를 정하는 것이다. 또한 필터를 거친 영상에서 적절한 질감 특징(texture feature)을 추출하는 작업과 각각의 채널 간의 관계, 또한 각각의 다른 채널에서 나온 질감 특징을 적절하게 조합하는 방법 등이 가장 중요한 고려대상이 된다.

위의 성질을 가장 잘 모델화할 수 있는 것이 가버 함수(Gabor Function)이다. 2차원(2-D) 가버 함수[4]는 2차원(2-D) 가우시안 함수의 변조(modulation)로, 또한 주파수 영역에서 이 2-D 함수의 스펙트럼(spectrum)이 이동(shift)한 것을 보이는 complex sinusoidal function라고 정의할 수 있다. 이 함수의 수학적 표현은 다음과 같다.

$$h(x, y) = g'(x', y') \exp[2\pi j(Ux + Vy)]$$

$$(x', y') = (x \cos \phi + y \sin \phi, -x \sin \phi + y \cos \phi)$$

$$g(x, y) = \frac{1}{2\pi\lambda\delta^2} \exp[-(x^2/\lambda^2 + y^2)/2\delta^2]$$

$\lambda$ : aspect ratio  $\delta^2$ : variance

$$H(u, v) = \exp\{-2\pi^2\delta^2[(u' - U')^2 + (v' - V')^2]\}$$

$$(u', v') = (u \cos \phi + v \sin \phi, -u \sin \phi + v \cos \phi)$$

$$(U', V') = (U \cos \theta + V \sin \theta, -U \sin \theta + V \cos \theta)$$

방사상 대역폭(radial bandwidth) B와 회전 대역폭(orientation bandwidth)  $\Omega$  는 위에서 사용한 다른 독립적인 변수 (free parameter)에 의해 가버 함수의 redundancy를 줄이기 위해서 2차원(2-D) 가버 필터(Gabor filter)의 각 채널 필터의 주파수 응답(frequency response)의 크기의 1/2 값(half-peak magnitude)이 서로 맞닿을 수 있도록 결정된다.

이렇게 얻어진 가버 함수와 영상을 컨벌루션(convolution)을 하여 얻어지는 크기(amplitude) 값의 평균(mean)과 편차(deviation)가 질감의 성질(feature)을 나타낸다고 할 수 있다. 또한 여기에서 얻어지는 위상 포락선(phase envelope) 또한 특징으로 사용할 수 있다.

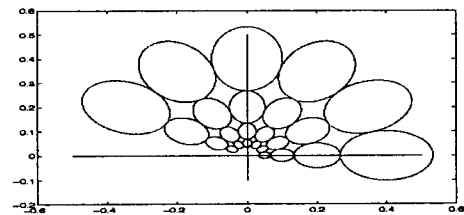


그림 2 spatial frequency response of 2-D Gabor function

질감의 유사도를 측정하는데 각 채널에서 나온 크기(amplitude)의 평균값(mean)과 편차(deviation)의 값을 이용하였는데 이들 간의 성질(feature)의 거리함수가 유사성을 판단하는데 중요한 도구가 된다. 이 실험에서 사용한 거리 함수(distance function)는 T. N. Tan의 Cortical

Channel Modeling(CCM)[5]에서 분류화 작업(classification)에 사용하였던 분류기(classifier)의 거리함수를 이용하였다.

$$d(\vec{F}, \vec{F}_k) = \sum_{n=1}^N \frac{(f_n - f_{kn})^2}{v_{kn}^2}$$

$f_n$ 은 질감 성질 벡터(texture feature vector)  $\vec{F}$ 의  $n$ 번째 요소  
 $f_{kn}$ 과  $v_{kn}$ 은 질감 클래스(texture class)의 대표 성질 벡터(representative feature vector)의  $n$ 번째 요소의 평균값(sample mean)과 편차(sample deviation)

이를 이용하여 두 영상의 질감의 성질간의 유사도는 두 성질 벡터(feature vector)  $\vec{F}_1$  과  $\vec{F}_2$

를 이용하여  $d(\vec{F}_1, \vec{F}_2)$ 가 가까울수록 유사한 영상이라고 정의하였다. 비교적 계산량이 적기 때문에 빠른 정합을 하기에 적합하다.

### 2.3 임의의 모양의 영역에서 질감 추출

2.1에서의 색깔 영역화 또는 [6]등에서와 같은 사용자 선택에 의한 영상의 조각 영역은 그 모양이 불규칙적이다. 색깔의 정보는 그 영역 안의 색깔 픽셀(pixel)의 분포와 그 수를 통하여 얻을 수 있기 때문에 그 영역의 모양이 문제되지 않는다. 그러나 질감의 경우 앞에서 말한 것과 같이 질감(texture)의 정의를 국부적인 공간 주파수(localized spatial frequency)에 집중된 정도에 의해서 구별될 수 있는 발광 패턴(irradiance patterns)이라고 한 것으로 보아 알 수 있듯이 픽셀들 간의 패턴(pattern)이 중요하기 때문에 영역의 모양에 영향을 받을 수 있다. 또한 가버 필터가 영상과 가버 함수와의 convolution을 통하여 얻어지기

때문에 임의의 모양에 대한 convolution은 쉽지가 않다. 그러한 이유로 임의의 모양의 영역에서 동일한 영역의 성질을 가지는 직사각형의 영역을 만들어 내는 작업이 필요로 하게 된다. 또한 이 작업이 이 영역에서 질감(texture)의 정보를 얻어내야 한다는 점에서 그 픽셀의 발광 패턴(irradiance patterns)이 많이 들어갈 수 있는 큰 영역을 취해야만 정확한 질감 정보를 얻을 수 있다는 점을 유의해야 한다.

우선 영상을 YUV 변환을 통하여 Y 영상 즉 흑백 영상(gray image)을 취한다. 영상의 질감이 조도의 변동(fluctuation of intensity)라는 점을 이용하여 흑백에서 색깔의 정보는 버리고 휘도(luminance)의 정보만을 이용하면 된다. 다음은 영상 영역 지도(image region map)를 이용하여 각각 영역에 알맞게 사각형을 취한다. 사각형을 취하는 방법은 각 영역을 포함할 수 있도록 영역보다 크게 취하게 되나 너무 큰 경우 영역에 속하지 않는 부분이 많이 포함되어서 정확성이 떨어지기 때문에 사각형의 선의 30%이상이 영역 안에 포함되도록 직사각형을 취한다. 그리고 이런 방법으로 취해진 직사각형을 따라 각각의 영상 영역 지도(image region map)와 흑백 영상(gray image)을 자른다.

이렇게 얻어진 각각의 영역 지도(region map)와 흑백 영역 영상(gray region image)에 대해서 대하여 격자(grid)를 취한다. 그래서 원하는 영역에 완전하게 포함된 영역을 반사시켜 다른 부분을 이 영역으로 채워 넣는다. 그러나 거울과 같이 그냥 반사시킬 경우 질감의 방향성(direction)이 바뀔 수 있기 때문에 상하좌우의 방향성을 그대로 유지한 채로 상하, 좌우 반사를 거듭하여 그 직사각형에 동일한 성질을 갖는 직사각형의 순 질감 영상을 만들어낸다.

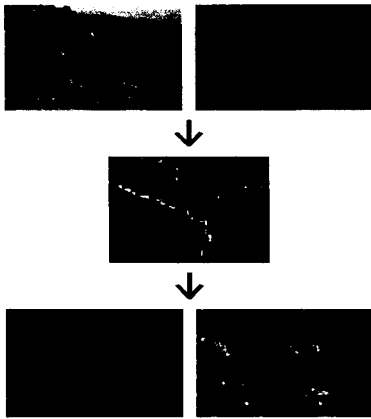


그림 3 regional gray image and image map

### 3. 실험 결과

#### 3.1 실 영상 영역(real-image region)에 대한 질감 검색(texture retrieval) 결과

위의 방법을 이용하여 실 영상(real-image)에서의 영역(region)에 대한 질감 검색(texture retrieval)을 실험해 보았다. 실험 데이터로는 13개의 영상, 총 98개의 영역을 사용했다.

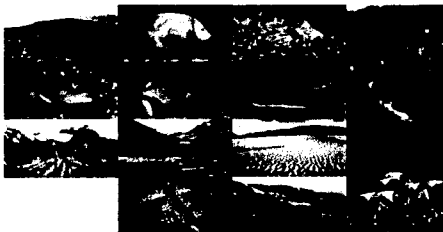


그림 4 test data image

다음 두 그림에서 나타나듯이 숲과 같이 거친 영역에 대한 검색과 하늘과 같이 매우 깨끗하고 매끈한 질감을 가진 영역에 대한 검색에서 좋은 결과를 보여주고 있다. 숲을 쿼리로 한 검색에서는 숲이나 거친 풀 또는 모래와 같이 거친 질감을 가진 영역을 검색하는 것을 알 수 있고 하늘과 같이 미끈한 영상에 대하여 하늘이나 바다 배경, 물고기의 몸표면 등과 같이 비슷한 질감을 가지고 있는 영상 영역을 검색하는 것을 확인할 수 있다. 직사각형 안

에 있는 영상이 쿼리 영역 영상이다.



그림 5 texture retrieval (1)

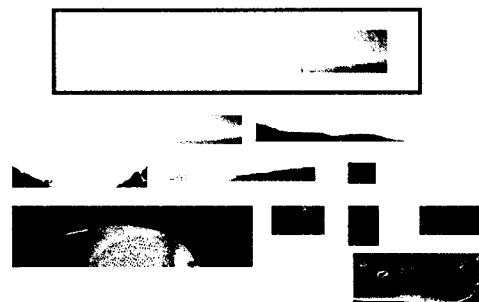


그림 6 texture retrieval (2)

#### 3.2 색깔과 질감을 이용한 영역별 영상 검색

색깔에 대한 유사도의 거리와 질감에 대한 유사도의 거리의 가중치 합을 그 영역의 특징간의 유사도의 거리라고 정의를 할 수 있다.

$$d(I_1, I_2) = d_c(I_1, I_2) + \lambda \cdot d_t(I_1, I_2)$$

$d_c(I_1, I_2)$ : 영상  $I_1, I_2$ 에 대한 색깔 유사도

$d_t(I_1, I_2)$ : 영상  $I_1, I_2$ 에 대한 질감 유사도

$\lambda$ : 가중치

다음은 corel 영상 2500개에 있는 영역에 대한 색깔과 질감을 이용한 영역별 영상 검색의 결과이다. 왼쪽의 검색 결과는 왼쪽 영상을 쿼리로 한 색깔만을 이용한 영역별 검색 결과이고 오른쪽 결과는 같은 영상을 쿼리로 한

색깔과 질감을 동시에 고려한 검색 결과이다.

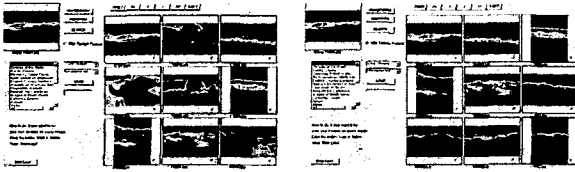


그림 7 regional image retrieval result (1)

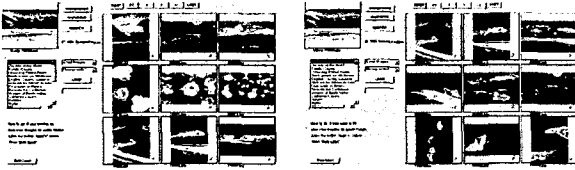


그림 8 regional image retrieval result (2)

위의 두 결과에서 확인할 수 있듯이 단순히 영상의 색깔만을 고려한 검색 결과보다는 질감을 동시에 고려한 검색 결과가 보다 적절하다는 것을 보여준다. 즉 색깔과 질감이라는 더 많은 정보를 조합하여 검색함으로써 사람의 주관적인 판단에 더 가까운 만족스러운 검색 결과를 나타낸다.

#### 4. 결론

위 실험을 통해 영상의 영역별로 추출된 색깔의 정보(color information)와 질감의 정보(texture information)를 동시에 사용함으로써 보다 사람 시각적인 감각에 적합한 검색 결과를 얻을 수 있었다. 또한 영역별 검색으로 한 영상에 있는 많은 물체(object)와 정보를 보다 정밀하게 검색할 수 있었다.

앞으로는 색깔과 질감과 같은 다른 특징(feature)을 적절하게 조합하는 방법이나 색깔과 질감을 동시에 표현하기 용이한 색깔질감(colortexture)에 대한 연구가 수행되어야 한다.

#### 참고 문헌

- [1] Yining Deng, Charles Kenny, Michael S. Moore and B.S. Manjunath, "Peer group Filtering and Perceptual Color Image Quantization", *Proc. of the 1999 IEEE International Symposium on Circuits and Systems(ISCAS 1999)*, vol. 4, pp. 21-24, 1999.
- [2] Yining Deng and B.S. Manjunath, "Color Image Segmentation", *IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1999 (CVPR 1999)*, vol. 2, pp. 446-451, 1999.
- [3] Yining Deng and B.S. Manjunath, "An Efficient Low-Dimensional Color Indexing Scheme for Region-Based Image Retrieval", *Proc. of the 1999 IEEE Intl. Conf. On Accoustics, Speech, and Signal Processing (ICASSP 1999)*, vol. 6, pp. 3017-3020, 1999.
- [4] A.C. Bovik, M. Clark and W.S. Geisler, "Multichannel Texture Analysis using Localized Spatial Filters", *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 1, pp. 55-73, Jan. 1990.
- [5] T.N. Tan, "Texture Feature Extraction via Visual Cortical Channel Modeling", *Proc. of IAPR Inter. Conf. On Pattern Recognition*, pp.607-610, 1992.
- [6] Nicholas R.Howe, Daniel P. Huttenlocher, "Integrating Color, Texture, and Geometry for Image Retrieval", *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp.239-246, 2000.