

영상의 텍스처 정보를 이용한
내용 기반 영상 검색에 관한 연구

박경식*, 박강서*, 홍민석*, 정태윤**, 박상희*
*연세대학교 전기·컴퓨터 공학과, **삼성전자 중앙연구소

A Study on Content-based Image Retrieval Technique
using Texture Information

Kyung-Shik Park*, Kang-Seo Park*, Min-Suk Hong*, Tae-Yun Chung**, Sang-Hui Park*

*Dept. of Electrical and Computer Eng. Yonsei Univ., **Corporate R&D Center Samsung Electronics Co. Ltd.

Abstract - 본 논문에서는 영상의 텍스처 정보를 이용하여 일반 영상에 대한 내용기반 영상 검색을 수행할 수 있는 알고리즘을 제안한다. Gabor 웨이브렛 변환을 이용하여 Gabor 필터 뱅크 내의 각 필터에 의해 필터링된 대역의 평균과 표준편차를 영상의 특징 벡터(Gabor Texture Feature)로 추출하여 영상들간의 유사성을 계산하는데 사용한다. 논문의 목적이 영상에 가해진 외적 변형, 즉 잡음 첨가, 블러링, 샤프닝 등과 같은 변형에 강인하게 동작할 수 있는 텍스처 특징 기반 영상 검색 기법을 제안하는 것이므로, 기존의 Gabor 필터만을 사용하여 텍스처 특징을 추출하여 검색의 기준으로 삼을 경우에 발생할 수 있는 주파수 성분의 변화에 대한 민감성을 Daubechies의 웨이브렛 필터를 사용하여 낮은 해상도에서 영상을 해석함으로써, 외적 변형에 대하여도 강인하게 동작할 수 있는 알고리즘을 제시하였다. 기존의 텍스처를 이용한 검색이 주로 텍스처 영역(textured region)에 대한 해석만을 하였지만, 본 논문에서는 이를 일반 영상에 적용하였으며, 일반 영상에 대해서도 효율적인 검색을 수행할 수 있음을 보였다.

1. 서 론

최근 멀티미디어 기술 발달과, 영상의 입력 장치 및 대용량 저장 장치의 발달로 인해 많은 영상들이 획득되어 저장되고 있다. 그 결과 다양한 목적의 영상 데이터베이스들이 생성되어 사용되어지고 있는데, 많은 수의 영상들로 구성되어 있기 때문에 사용자가 원하는 영상을 검색하는데 상당한 어려움이 따른다. 따라서, 데이터베이스로부터 영상을 검색하는 효율적이고 자동적인 검색 알고리즘이 요구된다. 초기의 검색 시스템에서는 입력된 키워드를 사용하는 텍스트를 기반으로 하는 영상 검색이 주류를 이루었으나, 이러한 방법은 몇 가지 문제점을 내포하고 있다. 데이터베이스 내의 영상에 대해서 검색 질의어로 사용하는 키워드를 설정할 때, 개인의 주관성이 포함되기 때문에 고유한 키워드를 설정할 수 없는 경우가 있다. 또한, 검색시 사용자가 주어진 영상에 대한 고유의 키워드를 기억하고 있어야 하는 단점도 지니고 있다. 따라서, 보다 일반적인 경우에 있어서 내용을 기반으로 하는 영상 검색 기법이 요구되어지고, 이와 관련된 여러 방법들이 연구되어지고 있다[1][2][5].

특히, 이러한 내용 기반 영상 검색 기법에서 가장 중요한 것은 각 영상이나 객체를 기술하는데 사용되어지는 특징의 집합을 어떻게 선택하느냐 하는 것이다. 현재까지 고려되어지고 있는 영상의 특징으로는 색상(color), 형태(shape), 그리고 텍스처(texture) 등을 들 수 있다. 본 논문에서는 전처리 과정인 특징 추출 과정에서 영상의 텍스처 정보를 잘 표현할 수 있는 특징 벡터를 추출하여 일반적인 영상에 적용하여 그 검색 효율을 높일 수 있는 방법을 제안하려고 한다.

먼저 영상의 텍스처 정보를 추출하기 위하여 Gabor 웨이브렛 변환을 사용한다. Gabor 필터 뱅크를 설계하고, 각 필터에 의해 필터링된 대역의 평균과 표준편차를

영상의 특징 벡터로 추출하여 영상들간의 유사성을 계산하는데 사용한다[2]. 또한, 검색 성능의 향상과 효율성을 위하여 Daubechies의 8-taps 웨이브렛 필터[4]를 이용하여 영상의 해상도를 낮추어 저대역 신호만을 잘라내어, Gabor 필터를 적용함으로써, 특징 벡터 추출에 필요한 연산량을 줄이고, 잡음(noise) 첨가, 블러링, 샤프닝 등의 변형과 JPEG 압축 등에도 좋은 검색율을 보이는 내용 기반 영상 검색 기법을 제안한다.

2. 본 론

2.1 Gabor 텍스처 특징

텍스처(texture)는 영상 해석이라는 측면에서 매우 중요한 영상의 한 특성이다. 즉, 영상의 영역들에 대한 표현, 인식 등의 작업을 수행하기 위해서는 텍스처로서의 영상 영역들의 특성을 고려하여야 한다. 본 논문에서는 텍스처를 다음과 같이 정의한다. 텍스처란 특정 패턴이 유사적 규칙성을 나타내는 지역적인 영역이다. Gabor 필터는 텍스처 인식에 적당한 필터로 알려져 있다. 1946년 Gabor가 제안하였고, 1985년 Daugman이 2차원으로 확장한 Gabor 필터가 텍스처 인식을 위해 많이 사용되어 왔다. 이 필터는 사람의 시각 피질이 방향과 주파수에 선택적인 수용계(Receptive fields)를 갖고 있다는 생물학적인 근거에 잘 부합한다[3]. 그러나 Gabor 필터는 대역폭이 좁기 때문에, 대역폭이 큰 자연의 텍스처를 처리하기 위해서는 대역이 서로 다른 여러 개의 Gabor 필터를 사용하여야 한다. 즉 서로 다른 대역의 Gabor 필터로 구성된 큰 필터 뱅크가 필요하다. 여러 개의 필터를 사용하는 이런 방식을 다해상(multiresolution) 방식이라 한다. 그림 1은 사람의 시각 피질의 수용계와 2차원 Gabor 함수와의 유사성을 나타내는 것이다[3].

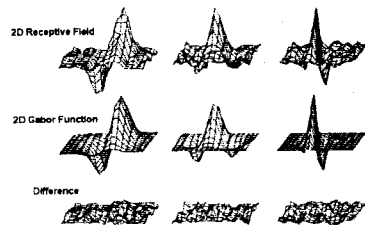


그림 1 시각 피질의 수용계와 2차원 Gabor 함수

2.2 2차원 Gabor 함수와 웨이브렛

Gabor 함수는 가우시안 함수와 복소 sinusoidal 함수의 곱으로 나타내어진다. 2차원 Gabor 함수의 일반적인 형태는 다음과 같다[2]. $G(u, v)$ 는 $g(x, y)$ 를 푸리

에 변환한 것이다. 여기서 σ_x 와 σ_y 는 각각 x 방향과 y 방향의 가우시안 함수의 범위를 결정한다.

$$g(x, y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2}\right) + j2\pi Wx\right] \quad (1)$$

$$G(u, v) = \exp\left\{-\frac{1}{2}\left(\frac{(u-W)^2}{\sigma_u^2} + \frac{v^2}{\sigma_v^2}\right)\right\} \quad (2)$$

여기서 $\sigma_u = \frac{1}{2\pi\sigma_x}$, $\sigma_v = \frac{1}{2\pi\sigma_y}$ 이다. W 는 평면파의 주파수 성분으로, 주파수 영역에서의 필터의 중심 주파수이다.

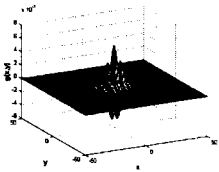


그림 2 Gabor 함수의 공간 영역

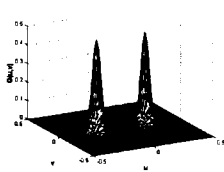


그림 3 Gabor 함수의 주파수 영역

Gabor 함수는 기저 함수(basis set)를 형성한다. 이러한 기저 함수를 사용하여 전개시킨 신호들로 지역화된 주파수를 표현할 수 있다. 자신과 유사한 함수들로 이루어진 이러한 집합을 Gabor 웨이블릿이라 하며, $g(x, y)$ 는 이러한 집합의 모함수가 된다. Gabor 웨이블릿은 이러한 모함수 $g(x, y)$ 의 회전과 팽창에 의해서 만들어질 수 있는데, 다음과 같은 식을 사용하여 Gabor 필터 뱅크를 생성하게 된다[2].

$$g_{mn}(x, y) = a^{-m} g(x', y') \quad (3)$$

$$x' = x \cos n\theta + y \sin n\theta \quad (4)$$

$$y' = -x \sin n\theta + y \cos n\theta \quad (5)$$

$$a = \left(\frac{U_h}{U_l}\right)^{\frac{1}{S-1}} \quad (6)$$

m 과 n 은 각각 크기(scale)와 방향(orientation)을 나타낸다. 단, $m=0, 1, \dots, S-1$ 이고, $n=0, 1, \dots, K-1$ 이다. U_h 와 U_l 은 필터뱅크에 포함되는 대역의 최대와 최소 중심주파수를 나타낸다. 본 논문에서는 그림 4.5과 같이 4개의 크기($S=4$)와 6개의 방향($K=6$)을 가지고 있고, $U_l=0.05$, $U_h=0.4$ 인 총 24개의 Gabor 필터로 이루어진 Gabor 필터 뱅크를 사용한다.

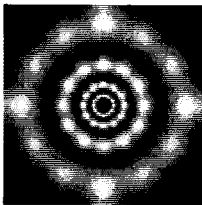


그림 4 Gabor filter의 주파수 영역 응답

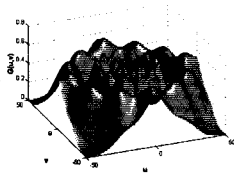


그림 5 Gabor filter 주파수 응답 (3차원)

2.3 웨이블릿 변환

웨이블릿은 신호처리에서 사용해왔던 서브밴드나 필터 뱅크와 같은 개념을 수학적으로 정리한 것이다. 웨이블릿(wavelets)이란 어느 특정 시간에 집중되어 있는 에너지를 가지고 있는 작은 파형을 의미한다. 이를 이용한 웨이블릿 변환은 특정 시간이나 공간에서의 신호의 주파수 특성을 알려준다. 웨이블릿 변환을 정의하기 위해서

는 스케일링 함수 ϕ 와 모웨이브릿 함수 ψ 이 필요하다. ϕ 와 ψ 은 각각 식(7)과 식(8)로 나타낼 수 있다.

$$\phi(t) = \sum_n h(n)\sqrt{2}\phi(2t-n) \quad (7)$$

$$\psi(t) = \sum_n g(n)\sqrt{2}\phi(2t-n) \quad (8)$$

원래의 신호를 두 개의 필터를 이용해서 고주파와 저주파 성분을 갖는 신호로 분리하면, 원신호의 해상도의 절반의 해상도를 지닌 두 개의 신호를 만들어 낼 수 있다. 이것을 반복하게 되면, 원하는 대역의 주파수 성분을 알아낼 수 있으며, 원래의 해상도보다 낮지만 다해상도를 지닌 신호를 만들 수 있다. 영상신호에 대한 웨이블릿의 경우 저대역 분할 신호는 원영상 신호가 가지고있는 에너지의 대부분을 소유하고 있으며 복원 영상에 미치는 영향이 가장 크다. 한편, 고대역 분할 신호는 에너지 집중도가 떨어지지만 인간의 시각에 민감한 영상의 가장자리 정보를 보유하고 있게 된다.

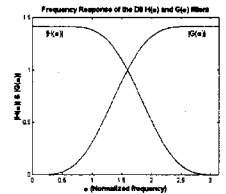
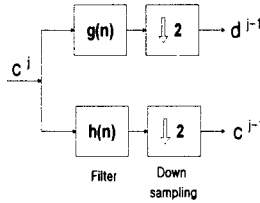


그림 6 웨이블릿 분해 그림 7 D-8staps 필터

웨이블릿 종류는 어떤 기본 함수를 사용하는가에 따라 결정된다. 본 논문에서는 Daubechies 웨이블릿 변환을 사용하였다[4]. 그림 7은 D-8staps 필터의 저주파 필터 계수($h(n)$)와 고주파 필터 계수($g(n)$)의 주파수 응답을 나타낸 것이다.

2.4 특징 벡터 표현

영상의 텍스처 특징을 추출하기 위하여, 입력 영상에 대하여 Gabor 필터 뱅크 내의 24개의 필터 $g_{mn}(x, y)$, 각각에 대하여 필터링을 하여, $h_{mn}(x, y)$ 를 얻는다. 필터링된 영상에서, 그 영상의 화소의 값의 평균(μ_{mn})과 표준편차(σ_{mn})를 Gabor 특성으로 추출해 낸다. N_x 와 N_y 는 입력 영상의 공간 영역에서의 넓이와 높이이다.

$$\mu_{mn} = \frac{1}{N_x N_y} \sum_x \sum_y |h_{mn}(x, y)| \quad (9)$$

$$\sigma_{mn} = \left[\frac{1}{N_x N_y} \sum_x \sum_y (|h_{mn}(x, y)| - \mu_{mn})^2 \right]^{1/2} \quad (10)$$

추출된 특징 벡터들을 이용하여 질의 영상과 데이터베이스 내의 저장된 영상들 간의 유사성을 계산하기 위하여 다음과 같은 거리(distance)의 식을 이용한다. 두 개의 영상 i 와 j 를 가정하고, 각각의 특징벡터를 $\vec{f}^{(i)}$ 와 $\vec{f}^{(j)}$ 라 가정하면, 이 두 영상간의 거리는 다음과 같다.

$$d(i, j) = \sum_m \sum_n d_{mn}(i, j) \quad (11)$$

$$d_{mn}(i, j) = \left| \frac{\mu_{mn}^{(i)} - \mu_{mn}^{(j)}}{\alpha(\mu_{mn})} \right| + \left| \frac{\sigma_{mn}^{(i)} - \sigma_{mn}^{(j)}}{\alpha(\sigma_{mn})} \right| \quad (12)$$

$\alpha(\mu_{mn})$ 와 $\alpha(\sigma_{mn})$ 은 특징 벡터의 성분 각각에 대한 영상 데이터베이스 전체의 표준편차이다. 이 값들은 각각의 특징 벡터의 성분을 표준화하기 위하여 사용된다.

2.5 실험 및 결과

실험을 위하여 FAMSF[6](Fine Arts Museums of San Francisco) 영상 데이터베이스의 200×200 크기의 컬러영상 150개를 사용하였다. 고려해야 할 계수들

의 수를 최소로 줄이기 위해 세 개의 채널(RGB)을 갖는 컬러영상을 하나의 채널을 갖는 그레이 영상으로 변환하였다.

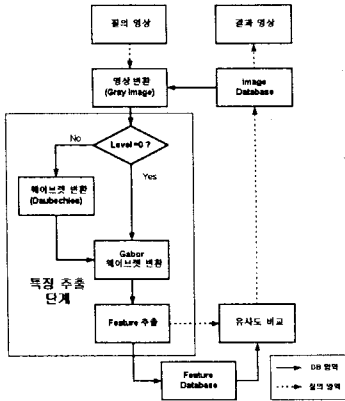


그림 8 제안하는 영상 검색 시스템의 구조

표 1 JPEG 압축에 대한 실험 결과

	Gabor필터	Level = 1	Level = 2
JPEG 70	100 % ²⁾	100 %	100 %
JPEG 50	97.33 %	100 %	100 %
JPEG 30	99.33 %	100 %	100 %
JPEG 10	98.67 %	100 %	100 %

표 2 잡음 첨가에 대한 실험 결과

	Gabor필터	Level = 1	Level = 2
잡음30dB	99.33 %	100 %	100 %
잡음20dB	99.33 %	99.33 %	99.33 %
잡음10dB	38.00 %	99.33 %	99.33 %

표 3 블러링, 샤프닝에 대한 실험 결과

	Gabor필터	Level = 1	Level = 2
블러링3*3	24.00 %	93.33 %	99.33 %
블러링5*5	6.67 %	40.00 %	98.67 %
샤프닝3*3	1.33 %	30.00 %	99.33 %

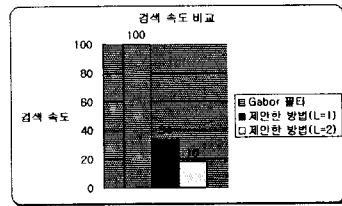


그림 10 검색 속도 비교

위의 데이터베이스를 기본으로 잡음 첨가, 블러링, 샤프닝 등의 외부 조작과 JPEG 압축에 의한 변형과 같은 내용 변형에 대한 실험을 하였다. 웨이브렛 변환을 사용한 경우에는, 원 영상을 Daubechies의 8-taps 웨이브렛 필터를 사용하여 웨이브렛 변환을 한 후, LL 대역만을 잘라내어 새로운 영상을 만들어 이 영상에 대하여 Gabor 필터링을 하여 특징벡터를 추출하여 비교하였다. 실험 결과는 먼저 Gabor 필터만을 사용하여 텍스처 특징을 추출한 경우와, Daubechies의 8-taps 필터를 사용하여 웨이브렛 변환을 한 후에 LL 대역만을 Gabor 필터를 사용하여 특징을 추출한 경우를 비교하였다. 이 경우, 웨이브렛 변환의 레벨(level)을 1 또는 2로 하여 그 결과를 비교하였다. 그림 9는 이러한 실험에 질의 영상으로 사용되어진 영상의 예를 나타낸 것이다.

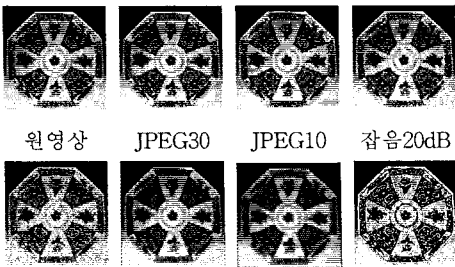


그림 9 실험 질의 영상의 예

JPEG 압축의 경우나 잡음첨가의 경우 Gabor 필터만을 사용한 방법과 제안한 방법 모두에 대하여 좋은 검색율을 보였다. 블러링과 샤프닝한 영상의 경우에는 영상의 주파수 성분이 크게 변화하므로, 텍스처의 특징을 나타내는 성분들이 변화하여, Gabor 필터만을 사용할 경우에는 검색 성능이 좋지 못하다. 그러나, 웨이브렛 변환을 하여 저대역 신호만을 잘라내어, 이 신호에 Gabor 필터를 사용하여 특징 벡터를 추출함으로써, 검색율이 개선되었고, 그림 10에서 보는 것처럼 영상의 검색속도 면에서도 향상되었음을 알 수 있다.

3. 결 론

본 논문에서는 영상의 텍스처 정보를 이용하여 일반 영상에 대한 내용 기반 영상 검색을 수행할 수 있는 알고리즘을 제안하였다.

(참 고 문 헌)

- [1]W. Niblack, R. Barber, W. Equits, et al., "The Qbic Project : Querying Images," proceeding of SPIE 1908, pp.173-187, 1993.
- [2]B.S. Manjunath and W.Y. Ma, "Texture Features for Browsing and Retrieval of Image Data," IEEE Trans. PAMI, Vol. 18, No. 8, pp.837-842, 1996.
- [3]J.G.Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," J. Optical Society of America, A/Vol. 2, No. 7, July 1985.
- [4]Ingrid Daubechies, Ten Lectures on Wavelets, SIAM, Philadelphia, PA 1992.
- [5]김익재, "색상의 공간 분포 정보 및 형태 특징을 이용한 내용 기반 영상 검색 기법에 관한 연구," 연세대학교 대학원, 석사학위 논문, 1998.2.
- [6]FAMSF, <http://www.thinker.org/imagebase/>