

다채널 Gabor 필터와 Log-Polar 변환을 사용한 내용기반 영상 검색

박 현, 문영식

한양대학교 전자계산학과

전화 : (0345) 407-8991 / 팩스 : (0345) 419-1162

Multichannel Gabor Filter and Log-Polar Transform for Content-Based Image Retrieval

Hyun Park, Young Shik Moon

Department of Computer Science and Engineering, Hanyang University

E-mail : hpark@cse.hanyang.ac.kr, ysmoon@cse.hanyang.ac.kr

Abstract

In this paper, we propose new features for describing texture images by using multi-channel Gabor filter and log-polar transform based on human visual system (HVS). Gabor features are extracted by the mean and standard deviation of energy in Gabor response, followed by Fourier series extension. Log-polar features are extracted by log-polar transform and projection.

The proposed texture descriptor performs reasonably well with less number of features than other texture descriptors, which has been verified by experiments using some texture images of MPEG-7 data set.

I. 서론

본 논문에서는 인간의 시각 특성을 기반으로 한 다채널 Gabor 필터와 Log-Polar 변환을 이용하여 영상에서의 텍스처 특징을 추출하는 방법을 제시한다[1-8].

다채널 Gabor 필터를 이용한 주파수 영역에서의 분해(decomposition)는 인간의 시각구조(Human Visual System, HVS)와 상당히 유사한 다중해상도(multiresolution) 특성을 가지고 있다[2-3][8]. Gabor 필터는 영상을 공간적으로 국부화(localization)시키는 특성을 지

니고 있을 뿐만 아니라 주파수 영역에서도 특정 주파수와 방향에 대해서 선택적으로 반응하는 대역 통과 필터의 특성을 가지며 입력 영상의 크기 변화와 회전 변화에 둔감한 특성을 가지고 있기 때문에 영상처리에서 널리 연구되고 있는 방법이다[2][3][6]. 본 논문에서는 먼저 텍스처 영상에 FFT 변환을 수행한 후 Gabor 분해에 의한 각 채널의 Gabor 응답(response)를 IFFT 변환하여 에너지의 평균과 표준편차를 구하고 이를 푸리에 시리즈(Fourier series)로 확장 변환하여 텍스처 특징들을 추출한다. 다음으로 FFT 변환된 영상의 에너지 스펙트럼을 HVS를 기반으로 한 log-polar 사상으로 변환하고 샘플링 된 에너지 스펙트럼을 log-polar 좌표계의 각 축에 대해서 투영(projection)을 한다. 투영된 각각의 정보를 1D FFT 변환하여 텍스처 특징을 추출한다[5].

제안된 방법은 기존의 MPEG-7 에서 제안된 방법들보다 특징값의 수를 줄이면서 이동 및 회전변환에 보다 둔감한 성질을 갖고 있으며 검색 면에서도 성능이 뒤떨어지지 않는다[11][12].

II. 텍스처 특징 벡터의 추출

Gabor 필터에 의해 얻어진 특징값을 Gabor 특징값이라 하고 Log-Polar 변환에 의해서 얻어진 특징값을 Log-Polar 특징값이라 할 때 두 특징값을 구하는 전체

구조는 그림 1, 2 와 같다.

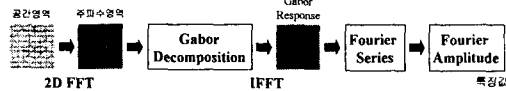


그림 1 공간 영역에서의 텍스처 구조도를 표현

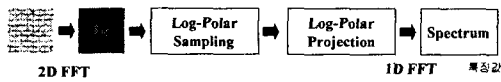


그림 2 주파수 영역에서의 텍스처 특징을 표현

먼저 푸리에 영역(Fourier domain)에서 Gabor 분해를 수행하기 위한 Gabor 필터 설계 방법은 그림 3 및 식 (1)과 같다.

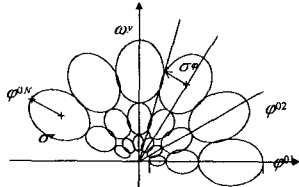


그림 3 Fourier 영역에서의 Gabor 분해

$$G_j(\omega_r, \omega_\phi) = \exp\left\{-\frac{(\omega_r - \omega_r^0)^2}{2\sigma_r^2}\right\} \cdot \exp\left\{-\frac{(\omega_\phi - \omega_\phi^0)^2}{2\sigma_\phi^2}\right\} \quad \text{식(1)}$$

N Orientation Band : $\sigma_r = \frac{r}{2N}$	M Frequency Band : $\sigma_\phi = \frac{(\omega_\phi^0 - \omega_\phi^1)}{2(2^M - 1)}$
Filter Angular Bandwidth : $\frac{\pi}{N}$	Filter Radial Bandwidth : $\sigma_r = \sigma_r \cdot 2^{M-1}$
Filter Orientation Centre : $\omega_\phi = \omega_\phi^0 + \frac{2\pi}{N} \cdot (n-1)$	Filter Frequency Centre : $\omega_r = \omega_r^0 + \frac{2\pi}{M} \cdot (m-1)$

Gabor 분해에서 주파수와 방향 채널(channel)의 수를 몇 개로 할 것이냐를 선택하는 것은 매우 중요한 요소이다. 본 논문에서는 중심 주파수를 4개, 방향을 6방향으로 한 총 24개의 채널을 사용하여 텍스처의 특징값들을 계산한다.

2.1 Gabor 특징값

공간 영역(spatial domain)에서의 Gabor 함수 표현이 $h(x, y; f, \theta)$ 이고 영상이 $i(x, y)$ 일 때 Gabor 응답의 평균값을 계산하면 식(2), (3) 과 같다.

$$q(x, y; f, \theta) = h(x, y; f, \theta) \otimes i(x, y) \quad \text{식(2)}$$

$$Q(f, \theta) = \frac{1}{\Omega} \iint_{\Omega} q(x, y; f, \theta) dx dy \quad \text{식(3)}$$

Gabor 변환은 푸리에 변환과 유사하며 180°의 주기를 갖기 때문에 푸리에 시리즈로 확장할 수 있다. 푸리에 급수는 회전에 불변이 성질을 가지고 있고 Gabor 응답의 에너지 값은 이동에 불변이 성질을 가진다.

따라서 본 논문에서는 Gabor 응답의 에너지 평균과 에너지 평균의 표준편차를 구하고 푸리에 시리즈로 확장 변환하여 얻은 푸리에 급수(DC, AC, 1차 고조파,

2차 고조파)를 텍스처 특징으로 사용한다.

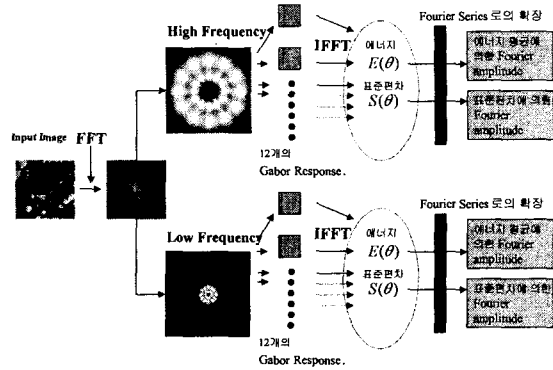


그림 4 Gabor 특징값 추출

그림 4와 같이 24개의 채널에 의해서 필터링 된 영상의 에너지 평균값과 표준편차 값을 고주파(high frequency)부분과 저주파(low frequency) 두 부분으로 나누어서 총 16개의 특징값을 계산한다.

$$F = [\mu C_{00} \quad \sigma C_{00} \quad \mu C_{01} \quad \sigma C_{01} \quad \dots \quad \mu C_{13} \quad \sigma C_{13}] \quad \text{식(4)}$$

2.2 Log-Polar 특징값

Gabor 특징값이 공간 영역에서 텍스처의 구조도(structuredness)를 표현한다면 Log-Polar 특징값은 주파수 영역에서 텍스처의 특징을 표현한다.

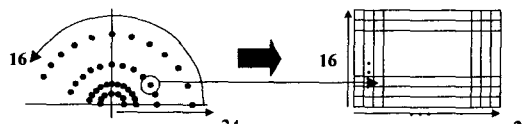


그림 5 Log-Polar 샘플링 구조

$$(\omega_r, \omega_\phi) = (\log(r), \phi) \quad : r = \sqrt{x^2 + y^2} \quad \text{식(5)}$$

$$r^n = \max \left(\begin{matrix} x \\ y \end{matrix} \right) = \begin{pmatrix} x^n \\ y^n \end{pmatrix} + r^n \begin{pmatrix} \cos \phi \\ \sin \phi \end{pmatrix} \quad : n = 0, 1, \dots, 5c$$

r^n : maximum radius (or 수평축 최대 크기)

Gabor 분해와 마찬가지로 HVS를 기반으로 한 Log-Polar 샘플링을 적용하여 Log-Polar 사상으로 변환을 한다. 본 논문에서는 16개의 방향에 대해서 각각 24개로 에너지 스펙트럼을 샘플링한다.

Log-Polar 변환의 속성을 보면 직각 좌표계에서 크기 변화는 수평이동으로 되고 직각 좌표계에서 회전변화는 수직 순환이동이 된다. 위와 같은 속성을 이용하여 회전에 크기 변화에 불변인 특징값을 추출하기 위해서 각 축에 대해 투영(projection)을 한다.

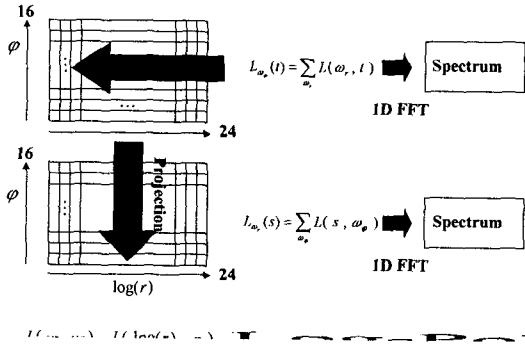


그림 6 Log-Polar 투영

그림 6처럼 먼저 ϕ 축으로 투영을 하면 16개의 스펙트럼을 얻을 수 있고 주파수 영역의 대칭 특성에 의해서 8개의 텍스처의 주파수 특징값을 추출할 수 있다.

$$P_{\phi_0}, P_{\phi_1}, \dots, P_{\phi_7} \quad \text{식(6)}$$

다음으로 $\log(r)$ 축으로 투영을 하면 24개의 스펙트럼을 얻을 수 있고 주파수 영역의 대칭 특성에 의해서 12개의 텍스처 주파수 특징값을 추출할 수 있다.

$$P_{r_0}, P_{r_1}, \dots, P_{r_{11}} \quad \text{식(7)}$$

Log-Polar 특징값은 식(8)과 같이 정리할 수 있다.

$$F = [P_{\phi_0}, P_{\phi_1}, \dots, P_{\phi_7}; P_{r_0}, P_{r_1}, \dots, P_{r_{11}}] \quad \text{식(8)}$$

III. 유사도 측정

텍스처 특징 벡터들에 의한 거리 함수(distance function)는 Gabor 필터를 이용하여 추출한 텍스처 특징값들에 대해선 Mahalanobis 거리 함수를 사용하고 Log-Polar 변환에 의해서 추출된 텍스처 특징값들은 유클리디안(Euclidean) 거리 함수를 사용한다.

두 개의 텍스처 영상이 I_q (query image) 와 I_d (database image) 일 때 Mahalanobis 거리 함수는 식(9) 와 같다.

$$Dissimilarity(i, j) = \sum_m \sum_n d_{mn}(i, j) \quad \text{식(9)}$$

$$d_{mn}(i, j) = \left| \frac{F_{mn}^{(i)} - F_{mn}^{(j)}}{\alpha(F_{mn})} \right| + \left| \frac{M_{mn}^{(i)} - M_{mn}^{(j)}}{\alpha(M_{mn})} \right|$$

$\alpha(\bullet)$ 는 전체 데이터베이스에 있는 각각의 텍스처 특징 값들의 표준편차이고 각각의 특징 값들을 정규화(normalize)하기 위해서 사용된다.

유클리디안 거리 함수는 식(10) 과 같다.

$$Dissimilarity(i, j) = \sqrt{\sum_{m=0}^7 (p_m^i - p_m^j)^2} + \sqrt{\sum_{n=0}^{11} (p_n^i - p_n^j)^2} \quad \text{식(10)}$$

IV. 실험 및 결과

본 논문에서의 검색율(Retrieval Rate : RR)은 텍스처를 16개로 분할 했을 때 식(11)과 같다.

$$RR_i = \frac{\text{나머지 15개의 텍스처 중 검색된 텍스처의 수}}{15} \quad \text{식(11)}$$

전체 데이터베이스에 대한 검색율은 식(12)와 같다.

$$\text{Total RR} = \frac{1}{\text{Texture Set 개수}} \sum_{i=1}^{\text{Texture Set 개수}} RR_i \quad \text{식(12)}$$

실험은 MPEG-7 실험 데이터로 제안되어 있는 128X128 Bordatz 영상 (T1), ICU 영상(T2), Rotate Image(T3)들 중 일부분을 사용하여 테스트하였다.

사용한 특징값들을 정리하면 식(13) 과 같다.

GF1 : Gabor Response 의 평균값을 이용한 Gabor Feature

GF2 : Gabor Response 의 에너지 평균값을 이용한 Gabor Feature 식(13)

LP : Log-Polar 변환을 이용한 Frequency Feature

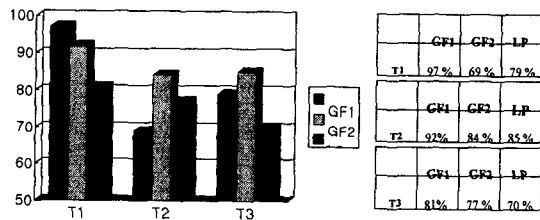


그림 7 특징값들의 검색율

표 1 기존 방법들과의 검색율 비교

Dataset	Hyundai	HHI	UCSB/Samsung	ETRI/ICU
T1	60%	68%	74%	76%
T2	72%	80%	87%	89%
T3	88%	84%	92%	86%

Dataset	GF1	GF2	LP
T1	97%	92%	81%
T2	69%	84%	77%
T3	79%	86%	70%

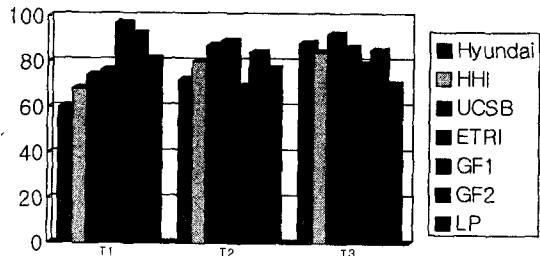


그림 8 기존 방법들과의 검색율 비교

제안된 텍스처 기술자의 성능은 표 1, 그림 7, 및 그림 8에서 제시된 바와 같이, 기존의 방법들과의 검색율면에서 뒤떨어지지 않는 것을 볼 수 있다.

* 이 논문은 1999년도 두뇌한국21사업 핵심분야에 의하여 지원되었음.

V. 결론 및 추후 연구

본 논문은 텍스처를 표현하는 기존방법들 보다 텍스처 특징값의 수를 줄이면서도 검색율을 유지할 수 있도록 하고 또한 이동 및 회전변환에 보다 둔감한 성질을 갖게 하는 것으로 그 효율성을 입증하였다.

Gabor 특징값에 의한 검색은 (T1), (T2), (T3) 영상에서 전반적으로 우수한 검색 성능을 보였고, Log-Polar 특징값에 의한 검색은 Gabor 특징값보다 검색율이 떨어지지만 (T1), (T2), (T3)에 균일한 검색율을 보이고 있다. 추후 연구로는 Gabor 필터를 이용하여 텍스처 특징을 추출할 때 몇 개의 채널을 사용해야 가장 우수한 성능을 나타내는지, 또 Log-Polar 샘플링의 개수는 어느 정도 해야하는지 실험해야 한다. 또한 본 논문에서 제안된 두 가지 서로 다른 영역에서의 특징값들을 혼합하여 사용할 수 있도록 최적화 된 가중치(weight)를 결정하는 방법도 연구되어야 할 것이다.

참고문헌

- [1] J. Bigün and J. M. Hans du Buf "N-folded Symmetries by Complex Moments in Gabor Space and Their Application to Unsupervised Texture Segmentation", IEEE Trans. Pattern Anal. Machine Intell., vol. 16. NO. 1, January 1994
- [2] A. C. Bovik, M. Clark, and W. S. Geisler, "Multichannel texture analysis using localized spatial filters", IEEE Trans. Pattern Anal. Machine Intell., vol. 12. pp 55-73, 1990
- [3] A. K. Jain, S. Prabhakar, L. Hong, "A Multi-channel Approach to Fingerprint Classification", IEEE Trans. Pattern Anal. Machine Intell., vol. 21, NO. 4, April 1999
- [4] B. S. Manjunath and W. Y. Ma, "Texture Features for Browing and Retrieval of Image Data", IEEE Trans. Pattern Anal. Machine Intell., vol. 18, NO. 8, August 1996
- [5] C. H. Teh, R. T. Chin, "On Image Analysis by the Methods of Moments", IEEE Trans. Pattern Anal. Machine Intell., vol 10, NO. 4, July 1988
- [5] R. Milanese and M. Cherbuliez, "A Rotation, Translation, and Scale-Invariant Approach to Content-Based Image Retrieval", Journal of Visual Communication and Image Representation 10. pp 186-196, 1999
- [6] T. Weldon, W. E. Higgins, "Designing multiple Gabor filters for multitexture image segmentation", Optical Engineering, vol. 38, NO. 9, September 1999
- [7] T. Tanden, J. H. Husoy, "Filtering for Texture Classification: A Comparative Study", IEEE Trans. Pattern Anal. Machine Intell., vol. 21, NO. 4, April 1999
- [8] B. Duc, S. Fischer and J. Bigün, "Face Authentication with Gabor Information On Deformable Graphs", IEEE Transactions On Image Processing, vol. 8, NO. 4, April 1999
- [9] T. S. Lee, "Image Representation Using 2D Gabor Wavelets", IEEE Trans. Pattern Anal. Machine Intell., vol. 18, NO. 10, October 1996
- [10] T. N. Tan, "Rotation Invariant Texture Feature and Their Use In Automatic Script Identification", IEEE Trans. Pattern Anal. Machine Intell., vol. 20, NO. 7, July 1998
- [11] 유기원, 김문철, 김진웅, 노용만, "HVS특성에 기반한 텍스처 검색 알고리즘의 연구", 제12회 신호처리 합동학술대회, 1999
- [12] 김문철외 3명, "MPEG-7 Tutorial", 제 12 회 영상처리 및 이해에 관한 워크샵 발표 논문집, 2000