

BDIP와 BVCL의 질감특징을 이용한 영상검색

천영덕*, 서상용**, 김남철*

*경북대학교 전자전기공학부, **한국통신 멀티미디어연구소

Image Retrieval Using Texture Features BDIP and BVCL

Young Deok Chun*, Sang Yong Seo**, and Nam Chul Kim*

*Department of Electronic Engineering, Kyungpook National University

**Multimedia Technology Laboratory, Korea Telecom

E-mail: ydchun@vcl.knu.ac.kr

Abstract

In this paper, we first propose new texture features, BVLC (block variation of local correlation coefficients) moments, for content-based image retrieval (CBIR) and then present an image retrieval method based on the fusion of BDIP and BVLC moments. BDIP uses the local probabilities in image blocks to extract valley and edges well. BVLC uses the variations of local correlation coefficients in images blocks to measure texture smoothness well. In order not to be affected with the movement, rotation, and size of an object, the first and second moments of BDIP and BVLC are used for CBIR. Corel DB and Vistex DB are used to evaluate the performance of the proposed retrieval method. Experimental results show that the presented retrieval method yields average 12% better performance than the method using only BDIP or BVLC moments and average 13% better performance than the method using wavelet moments.

I. 서론

컴퓨터 및 네트워크 통신의 발달로 문자이외에 음성, 영상 등의 멀티미디어 데이터의 양은 급속하게 증가하고 있다. 이에 따라 방대한 양의 멀티미디어 데이터의 효율적인 관리와 검색을 위하여 내용기반 검색법(content-based retrieval)이 활발하게 연구되고 있다. 내용기반 검색방법은 칼라(color), 모양(shape), 질감(texture) 등과 같은 영상의 내용과 관련된 시각적 특징을 자동으로 추출하여 데이터 베이스(DB)를 구축하고 질의영상의 특징과 유사도를 비교하여 영상을 검색한다[1]. 내용기반을 위한 특징 중에서 질감특징은 객체의 표면이 가지는 특성을 표현하는 것으로 영상의 내용 정보를 기술하는 가장 기본이 되는 특징 중의 하나라 할 수 있다. 영상검색을 위하여 사용된 질감특징으로는 GLCM(gray level co-occurrence matrix)에 근거한 통계적 특징들이나 마코프 랜덤 필드(Markov random

field) 모델링 계수 등이 있으며, 인간의 시각 특성을 반영하는 가보(gabor)변환이 있다[1]. 그리고 영상에 웨이브렛(wavelet)변환을 사용하여 다해상도로 분해된 부대역 계수들의 절대치에 대해 모멘트를 구한 웨이브렛 모멘트를 이용한 방법[2]이 있다. 특히, 웨이브렛 모멘트를 이용한 방법은 영상의 질감정보를 주파수 대역별로 추출함으로써 비교적 우수한 검색결과를 내는 것으로 알려져 있다. 또한, 최근에 엔트로피 특성을 나타내는 질감특징으로 PIM(picture information measure) 특징을 이용한 검색기법[3]과 BDIP(block difference inverse probabilities) 특징을 이용한 검색기법[4]이 수행된 바 있다. 특히, BDIP 특징을 이용한 검색 기법은 영상내의 국부적 밝기변화도와 밸리, 에지 등을 잘 검출하므로 비교적 좋은 성능을 나타낸다.

본 논문에서는 저자 등이 질감특징으로 제안한 바 있는 BDIP 모멘트와 새로이 질감특징으로 제안하는 BVLC(block variation of local correlation coefficient) 모멘트를 이용하여, 이들을 융합하여 검색하는 기법을 제시한다. BDIP는 엔트로피 연산자의 일종으로 블록내의 국부확률을 이용하여 밸리 및 에지를 잘 추출하는 특성을 나타내고 BVLC는 이웃 블록들과의 국부상관계수들의 변화량으로 영상내의 질감성분을 추출하는 성능을 나타낸다. 또한, 대상물의 이동, 회전, 크기변화에 영향을 받지 않기 위해 BDIP와 BVLC의 1, 2차 모멘트를 내용기반검색으로 사용하였다. 시험영상으로는 Corel DB[5]와 Vistex 질감영상[6]을 사용하였다. 실험에서는 제안한 방법이 일반영상과 질감영상에 대해 기존의 방법보다 성능이 향상됨을 확인한다.

II. 내용기반 영상검색

1. 내용기반 영상검색의 개요

내용기반 검색방법에서는 칼라(color), 모양(shape), 질감(texture) 등과 같은 영상의 내용과 관련된 시각적 특징을 자동으로 추출하여 DB를 구축하고 질의영상의 특징과 유사도를 비교하여 영상을 검색한다. 그림 1은 내용기반 영상검색 시스템을 나타낸다.

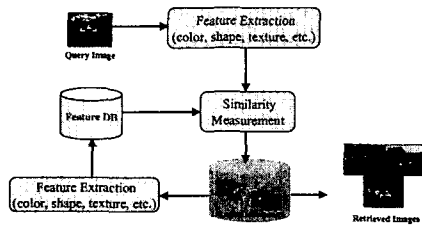


그림 1. 내용기반 영상검색 시스템

2. 영상의 특징

효율적인 영상 검색을 위해서는 무엇보다도 각 영상을 대표하는 특징 추출이 매우 중요하다. 영상 검색에 사용되는 몇 가지 특징들에 대해 알아보기로 한다

(1) 칼라 히스토그램

공간영역에서 히스토그램은 영상내의 밝기값(gray level)의 빈도를 나타낸다. 밝기값이 $L-1$ 인 영상의 히스토그램은 다음과 같이 L 차원의 특징벡터로 볼 수 있다.

$$h(i) = \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} Bin[I(x,y)], \quad i=0,1,\dots,L-1 \quad (1)$$

$$Bin[I(x,y)] = \begin{cases} 1, & I(x,y)=i \\ 0, & I(x,y) \neq i \end{cases}$$

여기서, $I(x,y)$ 은 (x,y) 위치의 화소값이며, $Bin[]$ 은 해당 화소의 이진화를 나타낸다. 이 특징은 영상의 회전이나 이동에는 상당히 안정된 특성을 나타내지만 대비변화에 민감하고 공간 정보를 가지지 못한 단점이 있다.

(2) 웨이브렛 모멘트

웨이브렛 변환을 이용한 내용기반 영상검색은 인간 시각 특성에 적합하게 공간-주파수 영역인 다해상도로 영상을 분해하여 대역별 웨이브렛 계수들의 통계적인 특성을 이용한 방법이다. 그림 2는 입력영상에 대한 웨이브렛 분해 과정을 보인 것이다. $W_{LL}^0(x,y)$ 를 입력영상이라 하면 이를 반복적으로 2차원 웨이브렛 변환하여 다해상도 피라미드를 구성한다. 각 n 번째 웨이브렛 변환에서는 입력 영상 $W_{LL}^{n-1}(x,y)$ 에 수평 방향으로 저역 통과 필터 및 고역 통과 필터를 통과시키고, 변환된 계수를 간축(decimation)한다. 다음으로 수평방향으로 처리된 계수들에 수직방향으로 저역 및 고역 통과 필터를 통과시키고, 변환된 계수를 간축하면, 한 단계 더 근사화된 영상 $W_{LL}^n(x,y)$ 와 수직, 수평 및 대각선 방향의 영상 ($W_{HH}^n(x,y)$, $W_{LH}^n(x,y)$, $W_{HL}^n(x,y)$)을 얻게 된다.

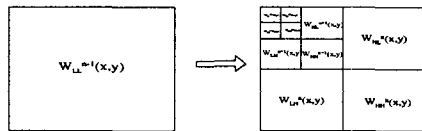


그림 2. 웨이브렛 분해과정

웨이브렛 모멘트는 웨이브렛 변환영역에서 절대치의

1, 2차 중심모멘트(central moment)로서 다음과 같이 표현된다[2].

$$\mu_m^n = \frac{1}{N_{mn}} \sum_x \sum_y |W_m^n(x,y)| \quad (2)$$

$$\sigma_m^n = \sqrt{\frac{1}{N_{mn}} \sum_x \sum_y |W_m^n(x,y) - \mu_m^n|^2} \quad (3)$$

여기서, n, m 은 웨이브렛 변환 단계와 부밴드를 나타낸다. N_{mn} 은 mn 부밴드의 화소수를 나타내며, μ_m^n, σ_m^n 은 변환계수의 절대치 평균과 표준편차를 나타낸다.

(3) PIM(picture information measure)

PIM은 영상의 엔트로피를 측정하는 방법으로 영상의 공간 정보를 반영하였다. PIM의 정의는 다음과 같다.

$$PIM = \sum_{i=0}^{L-1} h_B(i) - \max_{i \in \{0,1,L-1\}} h_B(i) \quad (4)$$

여기서, $h_B(i)$ 는 각 해당 블록의 히스토그램을 나타내며, $\max h_B(i)$ 는 빈도수가 가장 높은 히스토그램 값을 나타낸다. 따라서, PIM은 해당 블록 내의 히스토그램 중 가장 큰 도수의 히스토그램 값과 그 해당 블록 전체 화소수와의 차이로 계산된다.

(4) BDIP(block difference of inverse probabilities)

BDIP는 블록내의 화소수에 블록의 가장 큰 화소값과 블록의 화소들의 합과의 비를 뺀 값으로 정의되며 다음과 같이 표현된다.

$$BDIP = N_B - \frac{\sum_{(i,j) \in B} I(i,j)}{\max_{(i,j) \in B} I(i,j)} \quad (5)$$

여기서, $I(i,j)$ 는 (i,j) 위치의 화소 밝기를 나타내고, N_B 는 블록내의 화소수를 나타낸다. BDIP는 블록내의 밝기 변화가 많으면 커지고 변화가 없으면 작아진다. 특히, 국부적 밝기를 고려하므로 밸리나 에지를 잘 추출하고, 또한 어두운 영역의 밝기 변화를 잘 추출한다.

(5) BVLC(block variation of local correlation coefficients)

영상내의 질감성분을 잘 나타낼 수 있는 질감특징을 영상검색에 적용하기 위하여 새로이 BVLC 모멘트 특징을 제안한다. 이 질감특징은 이웃화소들과의 국부상관계수들 변화량으로 특징을 추출하는 VLCC[7]를 블록 단위로 수정한 것이다. 국부상관계수는 다음과 같이 표현된다.

$$\rho(k,l) = \frac{\frac{1}{M^2} \sum_{(i,j) \in B} f_{i,j} f_{i+k,j+l} - \mu_{i,j} \mu_{i+k,j+l}}{\sigma_{i,j} \sigma_{i+k,j+l}} \quad (6)$$

여기서, B 는 크기가 $M \times M$ 인 블록을 나타내며 $\mu_{i,j}, \sigma_{i,j}$ 는 블록의 밝기 평균과 국부분산의 제곱근을 나타낸다. $\rho(k,l)$ 는 블록 내의 화소 (i,j) 와 수평 방향으로 k , 수직방향으로 l 만큼 떨어진 화소간의 곱의 평균을 나타낸다. BVLC는 블록 내의 화소 (i,j) 를 중심으로 $180^\circ, -45^\circ, 135^\circ, 45^\circ$ 방향으로 구한 네 방향의 국부상

관계수에서 최대값에서 최소값을 뺀 값을 나타내며 다음과 같이 정의한다.

$$BVLC = \max_{(k, \theta) \in O_1} [\rho(k, \theta)] - \min_{(k, \theta) \in O_2} [\rho(k, \theta)] \quad (7)$$

여기서, $O_1 = \{(0, 1), (1, 0), (1, 1), (1, -1)\}$ 이다. BVLC값이 큰 화소들이 많이 모여 있는 영역은 질감영역으로, 그렇지 않은 영역은 평탄영역으로 볼 수 있다.

3. 특징의 차원

영상에서 추출한 특징은 일반적으로 벡터의 형태로 차원을 갖는다. 일반적으로 같은 특징이라도 그 차원이 높을수록 영상의 내용을 잘 표현한다고 할 수 있다. 그러나 특징의 차원이 높을 경우 그만큼 특징 DB에 색인할 데이터량이 많아진다. 이는 검색시 속도면에서 부담이 된다. 따라서 내용기반 검색에서는 검색성능, 색인 저장공간, 그리고 검색속도 등을 적절히 고려하여 특징의 차원을 결정하는 것도 매우 중요하다.

4. 유사도 측정

영상에 대한 특징과 특징의 차원 등이 결정되면 질의 영상과 DB영상들의 특징들 사이의 거리를 효율적으로 계산하는 유사도 척도가 필요하다. 이때 흔히 사용하는 거리 식은 다음과 같다.

$$D_{L_M}(q, d) = \| f_q - f_d \|^M \quad (8)$$

여기서, f_q 는 질의영상의 특징벡터, f_d 는 DB영상의 특징벡터, M 은 메트릭(metric)의 차수를 나타낸다. 그러나 이러한 유사도는 입력영상의 특징벡터의 각 성분별 분포 편차가 크게 다를 경우에는 안정된 결과를 내지 못하게 된다. 이를 보완하기 위하여 전체 DB에서 특징벡터의 각 성분에 대한 표준편차로 각 성분들을 정규화하는데, 이를 마하라노비스(Mahalanobis distance)거리[8]라 하며 다음과 같이 표현된다.

$$D_{L_M}(q, d) = \left\| \frac{f_q - f_d}{\sigma} \right\|^M \quad (9)$$

여기서, σ 는 전체 DB에서 특징벡터의 각 성분에 대한 표준편차를 나타낸다, 여기서, 벡터끼리의 나누기는 각 벡터의 성분별로 나누기 연산이 수행된 결과 벡터를 뜻한다. 본 논문에서는 유사도 척도로써 마하라노비스 거리를 사용하였다.

5. 성능평가

영상 검색방법의 성능을 평가하기 위해 흔히 사용되는 척도로는 정확도(precision) 대 재현률(recall)이 있다. 정확도는 검색된 영상의 총 개수와 검색된 영상 중에 질의영상과 유사한 영상 개수의 비이며, 재현률은 검색된 유사영상의 개수와 DB내에서 질의영상과 유사한 영상의 총 개수의 비로서 다음과 같이 표현된다.

$$R_p = \frac{N_s}{N_r} \quad (10)$$

$$R_r = \frac{N_s}{N_i} \quad (11)$$

여기서, N_r 과 N_i 는 각각 검색결과로 출력되는 영상

의 수와 DB내에서 질의영상과 유사한 영상의 수를 나타내고, N_s 는 검색된 영상들 중에서 질의영상과 유사한 영상의 수를 나타낸다.

III. BDIP와 BVLC의 융합을 이용한 영상검색

그림 3은 제안한 BDIP와 BVLC의 융합을 이용한 영상검색 시스템을 나타낸 것이다.

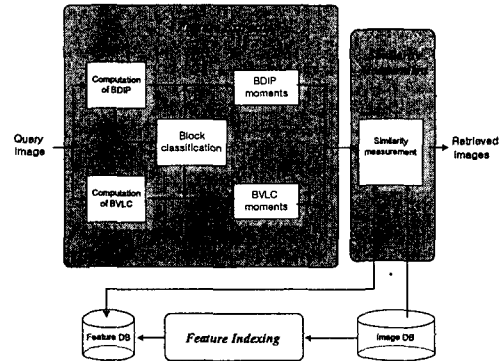


그림 3. BDIP와 BVLC의 융합을 이용한 영상검색 시스템

먼저 입력 영상을 블록단위로 BDIP를 구한 다음 BDIP 값에 따른 영역으로 나눈다. 영역을 나누는 과정은 다음과 같다. BDIP값으로 이루어진 영상을 BDIP의 평균값을 기준으로 BDIP가 높은 영역과 낮은 영역으로 나눈다. 이렇게 나눈 영역을 각 영역에 따른 BDIP 평균값을 구해 다시 높은 영역과 낮은 영역으로 나눈다. 이런 방법으로 영역을 8개의 영역으로 나눈다. 그리고 각 영역에 해당하는 BDIP와 BVLC의 1, 2차 모멘트를 구한 후, 이들을 융합하여 특징 DB내의 대상 영상특징들과의 유사도를 계산한다. 그 결과에 따라 영상 DB내에서 질의 영상과 유사도가 높은 일정 개수의 영상을 검색결과로 출력한다. 유사도 계산은 다음과 같다.

$$Dis(f_q, f_d) = \left(\left\| \frac{D_q - D_d}{\sigma_D} \right\|^M + \left\| \frac{V_q - V_d}{\sigma_V} \right\|^M \right)^{\frac{1}{M}} \quad (12)$$

여기서, D 와 V 는 각각 BDIP와 BVLC 모멘트 특징의 특징벡터이고 σ 는 전체 DB에서 각 특징벡터의 성분이 가지는 표준편차를 나타낸다. q 와 d 는 각각 질의영상과 DB영상을 나타낸다.

IV. 실험결과 및 검토

실험에 사용한 영상 DB는 Corel Draw Photo DB에 분류되어 있는 영상 (192x128) 중에 11부류 각 90장을 선택하여 총 990장으로 구성하였으며, 또한 MIT 대학의 Vision and Modeling Group에서 제공하는 Vistex 질감 영상의 16가지 부류에 속한 영상들 중에 75개의 영상을 선택하였다. 그리고 선택된 원 영상(512x512)을 128x128

크기의 영상 16개로 나누어 총 1200장의 질감영상을 만들어 실험DB로 사용하였다. 그림 4의 (a), (b)는 실험에서 사용한 Corel DB와 Vistex DB에 포함된 영상을 예로 들어 보인 것이다.

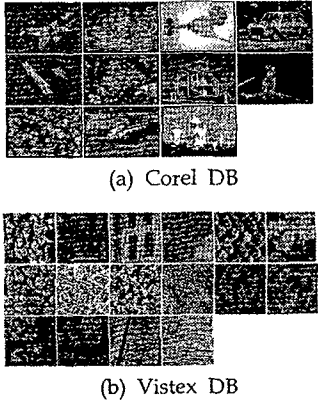


그림 4. 실험영상들

그림 5는 BDIP와 BVLC 모멘트 특징의 융합의 효과를 보이기 위해 차원에 따른 검색결과를 나타내었다. 여기서, 블록의 크기는 2x2로 하였고 매트릭의 차수 M은 가장 높은 검색성능을 나타내는 2를 사용하였다. 같은 차원에서 각 방법들의 성능은 제안한 질감특징들의 융합방법은 이 질감특징들을 단독으로 사용했을 때보다 Corel DB에서는 평균 8%, Vistex DB에서는 평균 15% 정도 검색율이 상승하여 융합의 효과가 매우 큼을 알 수 있다.

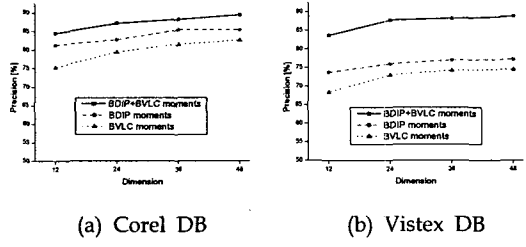


그림 5. 제안한 방법의 차원에 따른 검색결과

그림 6은 기존의 방법과 제안한 방법을 정확률 (precision) 대 재현률(recall)로 성능평가를 나타내었다.

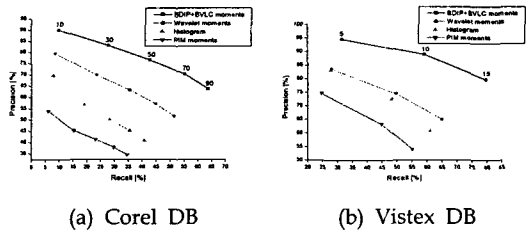


그림 6. 여러 특징들의 검색결과

각 특징들의 차원은 제안한 방법과 웨이브렛 모멘트는 72, 칼라 히스토그램은 96, PIM 모멘트는 48이다. 제안

한 방법은 웨이브렛 모멘트 방법보다 Corel DB에서는 약 10%, Vistex DB에서는 15%정도로 매우 우수한 성능을 나타내었다. 또한 칼라 히스토그램과 PIM 모멘트와는 두 DB 모두에서 각각 20%와 30%정도의 차이를 나타내었다. 더욱이 제안한 방법은 검색된 개수가 많을수록 기존의 방법과의 성능차이가 많이 났으며, 정확률 (precision) 대 재현률(recall)의 이상적인 그래프에 가까움을 볼 수 있다.

V. 결론

본 논문에서는 저자 등이 질감특징으로 제안한 바 있는 BDIP 모멘트와 새로이 질감특징으로 제안하는 BVLC 모멘트를 이용하여, 이들을 융합하여 검색하는 기법을 제시하였다. BDIP는 엔트로피 연산자의 일종으로 블록내의 국부확률을 이용하여 벨리 및 에지를 잘 추출하는 특성을 나타내었고 BVLC는 이웃 블록들과의 국부상관계수들의 변화량으로 영상내의 질감성분을 추출하는 성능을 나타내었다. 실험결과, 제안한 질감특징들의 융합을 이용한 방법은 이 질감특징들을 단독으로 사용했을 때보다 Corel DB에서는 평균 8%, Vistex DB에서는 평균 15% 정도 검색율이 상승하여 융합의 효과가 매우 큼을 확인하였다. 또한 기존의 방법인 웨이브렛 모멘트를 이용한 방법보다는 Corel DB에서는 평균 10%, Vistex DB에서는 평균 15% 정도 높은 결과를 나타내었고 칼라 히스토그램과 PIM 모멘트와는 두 DB 모두에서 각각 20%와 30%정도의 차이를 나타내었다.

참고문헌

- [1] Y. Rui and T. S. Huang, "Image retrieval: current techniques, promising directions, and open issues," *J. Visual Communication and Image Representation*, pp. 39-62, Oct. 1999.
- [2] M. K. Mandal, S.Panchanathan, and T. Aboulnasr, "Illumination Invariant Image Index using Moment and Wavelet," *J. Electronic Imaging*, Apr., 1998.
- [3] S. Y. Jeong, K. H. Kim, B. T. Chun, and J. Y. Lee, "Entropy and color correlation feature for image indexing," in *Proc. ICIP vol. 2*, pp. 895-899, 1999.
- [4] 서상용, 천영덕, 김남철, "엔트로피 특징을 이용한 영상검색," 한국통신학회 논문지 게재 예정.
- [5] Corel Draw Photo DB, <http://dlp.cs.berkeley.edu/photos/corel/>.
- [6] Vistex Texture DB, <http://www.whitemedia.mit.edu/vismod/imagery/VisionTexture/vistex.htm>.
- [7] S. Y. Seo, C. W. Lim, Y. D. Chun, N. C. Kim, "Extraction of texture regions using region-based correlation," *SPIE Proc. VCIP2001*, vol. 4315, pp. 694-701, 2001.
- [8] W. Y. Ma and B. S. Manjunath, "A comparison of wavelet transform feature for texture image annotation," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 1997.