

# 관련성 피드백을 이용한 효과적인 내용기반 영상검색

손재곤, 김남철

경북대학교 전자전기공학부

## Effective Content-Based Image Retrieval Using Relevance Feedback

Jae Gon Son and Nam Chul Kim

School of Electronic & Electrical Eng., Kyungpook Nat'l University

E-mail: nckim@ee.knu.ac.kr

### Abstract

We propose an efficient algorithm for an interactive content-based image retrieval using relevance feedback. In the proposed algorithm, a new query feature vector first is yielded from the average feature vector of the relevant images that is fed back from the result images of the previous retrieval. Each component weight of a feature vector is computed from an inverse of standard deviation for each component of the relevant images. The updated feature vector of the query and the component weights are used in the iterative retrieval process. In addition, the irrelevant images are excluded from object images in the next iteration to obtain additional performance improvement.

In order to evaluate the retrieval performance of the proposed method, we experiment for three image databases, that is, Corel, Vistex, and Ultra databases. We have chosen wavelet moments, BDIP and BVLC, and MFS as features representing the visual content of an image. The experimental results show that the proposed method yields large precision improvement.

### 1. 서 론

컴퓨터 기술과 인터넷의 발전은 디지털 정보의 양을 폭발적으로 증가시키고 있다. 이러한 정보의 대부분은 멀티미디어 정보로서 디지털 영상, 비디오, 오디오 그리고 문자 등으로 구성된다. 특히 이러한 다양한 매체형태들 가운데 영상은 문자와 함께 가장 많이 사용되는 것중의 하나이다. 따라서 방대한 양의 영상 데이터를 사용하기 위해서는 영상 정보들을 효과적으로 검색

할 수 있는 기술의 개발이 필요하다.

최근에 많이 연구되고 있는 내용기반 영상검색 방법은 질의영상 자체의 내용적인 특징(feature)과 유사한 특징을 가지는 영상들을 DB 내에서 찾는다[1-2]. 영상 자체의 내용적인 특징만을 비교하여 유사영상을 찾기 때문에 영상 저장시 추가적인 수작업이 필요치 않은 장점이 있다. 그러나 내용기반 영상검색에 대한 많은 연구에도 불구하고 검색성능은 만족스럽지 못하다. 이는 대부분의 연구들이 사용자의 고수준의 개념이나 사용자의 유사성 인식의 주관성을 반영하지 못하는 저수준의 시각적인 영상특징을 찾는 것이기 때문이다. 따라서 이러한 단점을 보완하기 위해 사용자를 영상검색과정의 일부로 포함시키는 인터랙티브(interactive) 내용기반 영상검색에 대한 연구가 활발히 이루어지고 있다[3-6]. 이는 사용자에게 의해 피드백(feedback)된 관련성 정보(relevant information)를 이용하여 저수준의 특징들이 사용자의 고수준의 개념과 주관성을 반영하도록 사용자와 검색시스템이 인터랙티브하게 동작하는 것이다.

본 논문에서는 대표적인 관련성 피드백방법으로 검색결과에서 사용자가 선택한 관련성이 있는 영상을 사용하여 성분가중치를 계산하고 이를 유사도 계산에 적용하는 방법과 사용자 피드백을 사용하여 질의영상의 특징벡터를 갱신한 후 반복검색시 입력벡터로 사용하는 방법을 사용하였다. 또한 임의의 질의영상에 대한 초기 검색결과에서 사용자가 질의영상과 관련성이 없다고 판단한 영상들의 경우, 반복검색 결과에서도 질의영상과 유사한 영상으로 나타날 가능성이 높기 때문에 이러한 영상들을 반복검색시에 비교대상에서 제외함으로써 검색성능을 대폭 향상시켰다.

제안된 방법의 성능평가를 위해 Corel DB, Vistex DB 그리고 초음파 영상 DB에 대하여 웨이브릿 모멘트[3], BDIP (block difference of inverse probabilities)와

BVLC (block variation of local correlation coefficient)[7] 그리고 MFS (magnitude frequency spectrum)[8]를 영상특징으로 사용하여 실험하였다. 실험결과 제안한 방법에 의해 검색정확도(precision)가 15-45%정도 대폭 향상되었다. 그리고 첫 번째 반복검색에서 검색정확도가 가장 많이 향상되었으며 검색이 반복되어도 안정된 성능을 유지하였고 높은 검색재현률(recall)에서도 검색정확도의 감소폭이 크지 않았다. 이는 실용적인 측면에서 볼 때 아주 바람직한 결과이다.

## II. 인터랙티브 영상검색

영상의 내용적 특징을 기반으로 사용자의 피드백을 사용하여 유사영상을 검색하는 인터랙티브 내용기반 영상검색의 과정은 일반적으로 영상의 특징을 추출하는 부분, 추출된 특징 정보를 색인하는 부분, 사용자가 질의를 수행하는 사용자 인터페이스(interface)부분과 DB와의 연동부분 그리고 사용자가 결과영상들에서 질의영상과 관련성이 있는 영상을 선택하는 부분과 이를 피드백하여 반복검색에 적용하는 부분으로 나눌 수 있다. 그림 1은 일반적인 인터랙티브 영상검색 과정의 구성도를 보여준다.

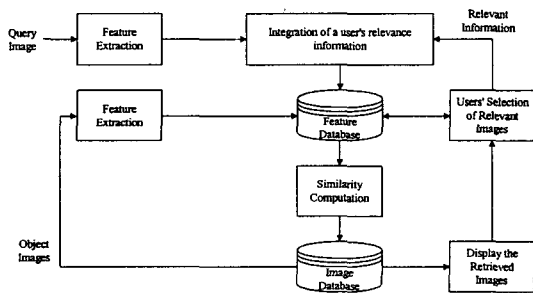


그림 1. 인터랙티브 내용기반 영상검색의 구성도

대상영상들의 특징은 임의의 영상이 질의되기 전에 미리 특징추출과정과 색인과정을 거쳐 특징 DB에 저장된다. 그리고 임의의 영상이 질의되면 초기검색에서 특징을 추출한 후, 이를 특징 DB내의 대상영상 특징들과의 유사도를 계산한다. 그 결과에 따라 영상 DB내에서 질의영상과 유사도가 높은 일정 개수의 영상을 초기 검색 결과로 출력한다. 이러한 초기결과에 대하여 사용자는 질의영상과 관련이 있는 영상을 선택하여 피드백하며 반복검색에서는 이러한 사용자 피드백 정보를 사용한다. 사용자 피드백 정보를 사용하는 대표적인 방법은 다음과 같다.

### 2.1 질의점 이동 방법[4-6]

이 방법은 초기 검색결과에 대한 사용자의 피드백을 이용하여 현재의 질의점이 최적의 질의점에 가까이가도록 추정하는 방법이다. 이 방법은 벡터공간모델을

바탕으로하여 주로 문서정보처리에 적용했으며 Rocchio의 식이라 하는 기본 수식은 다음과 같이 주어진다.

$$f = \alpha f + \beta \sum_{i=0}^{n_1} \frac{R_i}{n_1} - \gamma \sum_{i=0}^{n_2} \frac{S_i}{n_2} \quad (1)$$

여기서  $f$ 는 초기질의 벡터이고  $f'$ 은 갱신된 질의벡터이며,  $n_1$ 은 이전 검색결과에서 질의와 관련성이 있는 것들의 개수이고  $n_2$ 는 관련성이 없는 것들의 개수이다. 또한  $\alpha$ ,  $\beta$ 와  $\gamma$ 는 각각 적절한 상수를 나타낸다.

### 2.2 성분 가중치 갱신 방법[3-4]

성분 가중치 갱신 방법은 초기 검색결과에 대한 사용자의 피드백을 이용하여 유사도를 계산할 때 성분별 가중치를 반영하는 방법이다. 사용자가 관련성이 있는 영상이라고 선택하여 피드백한 영상들에 대한 특징들은 특정 성분, 즉 관련성있는 영상들을 대표하는 성분에서는 변화가 크지 않다. 따라서 유사도를 계산할 때 이러한 성분들에 대해 가중치를 크게 주는 것이 유리하다. 이러한 성분 가중치를 추정하는 대표적인 방법은 각 성분의 표준편차를 이용하는 것으로 다음과 같은 식으로 표현된다.

$$w = \left\{ \frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_i}, \dots, \frac{1}{\sigma_{N_d}} \right\}, \quad (2)$$

여기서,  $w$ 는 성분 가중치 벡터이고  $\sigma_i$ 는 사용자에게 의해 피드백된 관련성있는 영상들의 성분별 표준편차를 그리고  $N_d$ 는 벡터의 차원을 나타낸다. 또한 각 영상의 특징벡터는 다음 식으로 표현된다.

$$f = \{f_1, f_2, \dots, f_i, \dots, f_{N_d}\}, \quad (3)$$

따라서 영상의 특징벡터들과 (2)식으로 표현된 성분가중치 벡터를 사용한 유사도 계산식은 다음과 같이 표현할 수 있다.

$$d(f_q, f_o) = \left( \sum_{i=0}^{N_d} (w_i(f_{qi} - f_{oi}))^p \right)^{1/p}, \quad (4)$$

여기서,  $f_q$ 는 질의영상의 특징벡터,  $f_o$ 는 대상영상의 특징벡터,  $w_i$ 는 성분별 가중치를 나타내며  $p$ 는 매트릭 차수를 나타낸다.

## III. 제안된 인터랙티브 영상검색

제안된 인터랙티브 내용기반 영상검색 알고리즘은 세가지로 구성된다. 첫째로 앞장에서 언급한 방법들 가운데 성분 가중치 갱신방법은 동일하게 적용하고, 둘째로 질의점 이동 방법 수정하여 적용하였으며 마지막으로 초기 검색결과에서 사용자가 비관련 영상이라고 판단한 영상들을 반복 검색시 대상영상에서 제외하는 방법을 사용하였다.

수정된 질의점 이동 방법은 앞장의 Rocchio의 식에서 비관련영상에 의해 구성된 항이 악영향을 미친다고

판단하여 제안한 알고리즘에서는 비관련영상에 의한 항을 제외하였다. 따라서 수정된 질의점 이동 방법은 다음과 같이 표현된다.

$$\bar{f}_{k+1} = \frac{1}{N_r+1} \left\{ \sum_{i=1}^{N_r} f_k^r + f_0 \right\}, \quad (5)$$

여기서,  $f_0$ 는 초기 질의영상의 특징벡터를 나타내며,  $f_k^r$ 는  $k$ 번째 반복검색에서 사용자에게 의해 선택된  $N_r$ 개의 관련영상들의 특징벡터를 나타낸다. 또한

$\bar{f}_{k+1}$ 는  $k+1$ 번째 반복검색에서 사용될 갱신된 질의영상의 특징벡터를 나타낸다. (5)식에 의해 갱신된 질의영상의 특징벡터는 질의의 중심이동 이론[6]에 의해 질의의 최적특징 벡터에 가까워진다. 질의영상과 유사한 대상영상들이 그들의 평균벡터 주위에 몰려있기 때문에 초기 질의영상의 특징벡터 보다는 평균벡터에 의해 갱신된 새로운 특징벡터를 사용하는 것이 성능향상에 도움이 된다.

마지막으로 초기 검색결과에서 사용자에게 의해 비관련 영상으로 판단된 영상들을 반복검색시 검색대상영상에서 제외하는 방법을 사용하였다. 이러한 영상들은 시각적으로는 질의영상과 관련성이 없는 영상이나 유사도 계산결과로는 질의영상과 유사한 영상으로 판단된 것이다. 이러한 영상들은 반복검색에서도 계속해서 질의영상과 유사한 영상을 출력될 가능성이 매우 높다. 따라서 이러한 영상들을 검색 대상영상에서 제외하면 대폭적인 검색성능향상이 가능하다.

관련성 피드백을 이용한 영상검색에서 각 반복검색에서 갱신된 질의특징 벡터와 대상영상들의 특징벡터들 사이의 거리를 효율적으로 계산하는 유사도 척도로 사용한 거리식은 다음과 같다.

$$d(f_k, f) = \left( \sum_{i=1}^{N_s} (w_k(f_{ki} - f_i))^2 \right)^{1/2}, \quad (6)$$

여기서,  $f_k$ 와  $w_k$ 는  $k$ 번째 반복검색에서 사용될 갱신된 질의영상의 특징벡터와 성분별 가중치를 나타내며,  $f$ 는 대상영상의 특징벡터를 나타낸다.

영상 검색방법의 성능을 평가하기 위해 흔히 사용하는 척도에는 정확도(precision)와 재현률(recall)이 있다[5-6]. 정확도는 검색된 영상의 총 개수와 그 가운데에서 질의영상과 유사한 영상 개수의 비이고, 재현률은 검색된 유사영상의 개수와 DB내에서 질의영상과 유사한 영상의 총 개수의 비로서 다음과 같이 표현된다.

$$R_p = \frac{N_s}{N_r}, \quad (7)$$

$$R_r = \frac{N_s}{N_t}, \quad (8)$$

여기서,  $N_r$ 과  $N_s$ 는 각각 검색결과로 출력되는 영상의 수와 영상 DB 내에서 질의영상과 유사한 영상의 수를 나타내고,  $N_t$ 는 검색된 영상들 중에서 질의 영상과 유

사한 영상의 수를 나타낸다. 반복적인 검색에서 성능평가를 위해서는 이러한 정확도와 재현률을 동시에 표현하는 정확도 대 재현률의 그래프를 흔히 사용한다[6].

#### IV. 실험결과 및 고찰

제안된 방법의 성능평가를 위해 사용된 영상 DB는 Corel Draw Photo DB에 분류되어 있는 영상중에 11종류 각 90장씩 총 990장으로 구성된 Corel DB로 이들 영상은 192x128크기의 RGB 칼라영상을 되어 있다. 또한 MIT 대학의 Vision and Modeling Group에서 제공하는 Vistex 질감영상은 다음과 같은 부류 "Bark, Brick, Buildings, Clouds, Fabric, Flowers, Food, Grass, Leaves, Metal, Misc, Sand, Stone, Terrain, Tile, Water"로 나누어져 있다. 이와 같이 16가지 부류에 속한 영상들 중에 100개의 영상을 선택하였다. 그리고 선택된 원 영상(512 x 512)을 128 x 128 크기의 영상 16개로 나누어 총 1600장의 질감영상을 만들어 실험DB로 사용하였다. 마지막으로 770개의 실제 초음파 영상을 사전에 장기별로 분류하여 구성한 DB인 Ultra를 실험에 사용하였다. 각 DB에 대하여 실험에 사용한 영상의 특징들로는 Corel DB와 Vistex DB에 대해서는 웨이브릿 모멘트[1-2] 그리고 BDIP와 BVLC [7]를 사용하였고 Ultra에 대해서는 초음파 영상의 내용적인 특징을 잘 나타내는 MFS를 영상특징으로 사용하여 실험하였다[8]. 또한 실험에 사용된 영상특징들의 차원은 웨이브릿 모멘트가 72, BDIP와 BVLC를 합한 특징이 96 그리고 주파수 크기 스펙트럼이 275를 각각 사용하였다.

그림 2는 Corel DB에 대하여 웨이브릿 모멘트 특징과 BDIP와 BVLC를 합한 특징을 사용하여 실험한 결과이다. 제안한 인터랙티브 검색방법의 사용으로 검색 정확도가 15-40%정도 대폭적으로 향상함을 알 수 있다. 특히 높은 재현률에서 검색정확도의 증가폭이 더 크며, 초기단계의 반복검색에서 검색성능이 가장 많이 향상된다는 것을 알 수 있다. 이러한 것은 실용적인 측면에서 바람직한 현상이라 볼 수 있다. 두 개의 그래프에서 알 수 있듯이 반복검색이 반복됨에 따라 정확도 대 재현률 그래프가 이상적인 경우에 수렴하는 것을 알 수 있다.

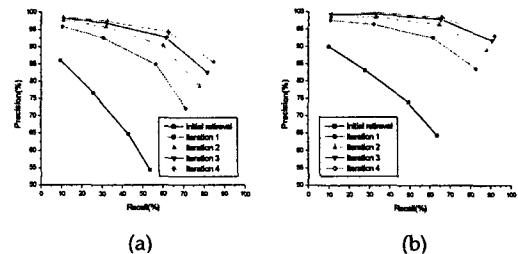


그림 2. Corel DB에 대한 인터랙티브 영상검색 결과: (a) 웨이브릿 모멘트 특징사용, (b) BDIP+BVLC 특징 사용

그림 3는 Vistex DB에 대하여 웨이브릿 모멘트 특징과 BDIP와 BVLC를 합한 특징을 사용하여 실험한 결과이다. 제안한 인터랙티브 검색방법의 사용으로 검색 정확도가 15-30%정도 대폭적으로 향상함을 알 수 있다.

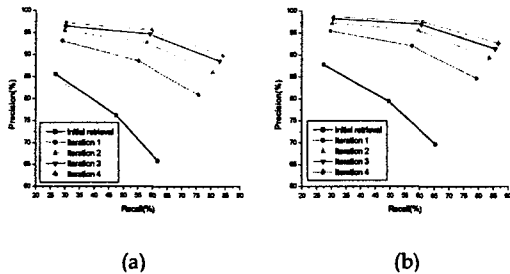


그림 3. Vistex DB에 대한 인터랙티브 영상검색 결과: (a) 웨이브릿 모멘트 특징사용, (b) BDIP+BVLC 특징 사용

그림 4는 Ultra에 대하여 주파수 크기 스펙트럼 특징을 사용하여 실험한 결과이다. 제안한 인터랙티브 검색방법의 사용으로 검색정확도가 25-45%정도 대폭적으로 향상함을 알 수 있다.

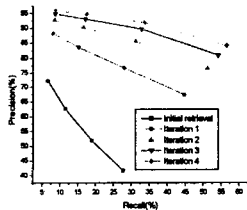


그림 4. Ultra에 대해 주파수 크기 스펙트럼 특징을 사용한 인터랙티브 영상검색 결과

## V. 결 론

본 논문에서는 검색결과에서 사용자가 질의영상과 유사한 것으로 판단한 영상들의 특징벡터와 질의영상의 특징벡터의 평균으로 질의영상의 특징벡터를 갱신하여 대상영상들을 반복검색하며 사용자가 관련성이 없다고 판단한 영상들은 반복검색시 대상영상에서 제외하는 인터랙티브한 내용기반 영상검색 알고리즘을 제안하였다. 제안한 방법에서는 기존에 가장 대표적으로 사용되는 관련성 피드백 방법인 성분 가중치 갱신 방법을 기본적으로 사용하였으며, 질의갱신의 대표적인 식인 Rocchio의 식에서 비관련영상에 의한 항을 제외한 식으로 사용하였다. 마지막으로 초기 검색결과에서 사용자에게 의해 비관련 영상으로 판단된 영상들을 반복검색시 검색대상 영상에서 제외함으로써 대폭적인 검색성능향상을 가능

하게 하였다.

제안한 방법은 Corel DB와 Vistex DB에 대해서는 웨이브릿 모멘트 특징과 BDIP와 BVLC를 합한 특징을 사용하여 실험하였고 초음파 영상 DB에 대해서는 주파수 크기 스펙트럼 특징을 사용하여 실험하였다. 실험결과 제안한 방법에 의해 검색정확도가 15-45%정도 대폭 향상되었다. 그리고 첫 번째 반복검색에서 검색정확도가 가장 많이 향상되었으며 높은 검색재현률에서도 검색정확도의 감소폭이 크지 않아 실용적인 적용가능성을 확인하였다.

## 참 고 문 헌

- [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] A. Yoshitaka and T. Ichikawa, "A survey on content-based retrieval for multimedia databases," *IEEE Trans. Knowledge and Data Engineering*, vol. 11, no. 1, Jan. 1999.
- [3] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feedback: A power tool for interactive content-based image retrieval," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 8, no. 5, pp. 644-655, Sep. 1998.
- [4] D. R. Heisterkamp, J. Peng, and H. K. Dai, "Feature relevance learning with query shifting for content-based image retrieval," *Proceedings of 15th International Conference on Pattern Recognition*, vol. 4, pp. 250-253, 2000.
- [5] Y. Rui and T. Huang, "Optimizing learning in image retrieval," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 236-243, 2000.
- [6] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "Mindreader: Querying database through multiple examples," *Technical Report CMU-CS-98-119*, Carnegie Mellon University, 1998.
- [7] 천영덕, 서상용, 김남철, "BDIP 및 BVLC의 질감특징을 이용한 영상검색," *14회 신호처리합동학술대회 논문집*, 2001년 9월.
- [8] J. G. Son, S. H. Kim, and N. C. Kim, "Content-based ultrasound image retrieval using magnitude frequency spectrum," *SPIE Proc. Medical Imaging*, Feb. 2001.