

# 관련성 귀환을 가진 질감 기반의 영상검색

정성환, 이신주

창원대학교 전자계산학과

## Texture-based Image Retrieval with Relevance Feedback

Sung-Hwan Jung, Sin-Joo Lee

Dept. of Computer Science, Changwon National University

### 요약

본 논문에서는 간단하면서도 효과적인 관련성 귀환을 가진 영상 검색시스템에 대하여 연구하였다. 먼저 영상 데이터베이스 내에 있는 영상들에 대하여 Gabor Wavelet 변환을 이용하여 질감특징을 추출하고, 추출한 특징값을 다양한 형태로 영상검색에 이용하였다. 초기 검색결과에 대하여 관련성 귀환을 영상 검색시스템에 적용하고, 이를 기존의 관련성 귀환을 가진 시스템과 비교하였다. 16종류의 512개의 영상으로 구성된 영상 데이터베이스에 대하여 실험한 결과, 제한된 방법은 INRIA의 방법보다 각 귀환단계에서 약 7~8%의 높은 검색 효율을 보였다.

### 1. 서론

최근 디지털 영상, 비디오, 오디오 그리고 문자정보로 구성된 멀티미디어 정보는 컴퓨터 기술과 웹의 발전으로 수요가 급증하고 있다. 특히 영상은 문자와 함께 가장 많이 사용되는 매체형태 중에 하나이다. 따라서 방대한 영상정보를 검색하기 위한 효과적인 기술 개발이 필요하다.

초기 영상검색의 기본형태는 주로 키워드를 사용한 문자기반 영상검색(Text-based Image Retrieval)이었다. 그러나 정보량의 급속한 증가로 인하여, 문자기반의 검색은 주석 개발의 어려움과 영상내용에 대한 해석의 차이, 그리고 키워드 할당의 일관성 결여 등의 문제점을 가지고 있었다. 이 문제점을 해결하기 위하여 내용기반 영상검색(Content-based Image Retrieval)이 소개되었다[1].

내용기반 영상검색은 영상의 내용인 칼라, 질감, 모양 등의 정보를 검색에 이용한다. 이 방법은 색인과정에서 자동으로 추출된 특징을 사용함으로써 보다 효율적인 관리와 검색이 가능하므로 문자기반 검색의 문제점들을 해결할 수 있다. 이 검색 시스템은 일반적으로 사용자가 선택한 특징과 가중치에 의하여 사용자 정의와 유사한 영상들을 검색하도록 설계되었다. 그러나 이와 같은 접근방법은 검색과정에서 사용자를 포함하고 있지 않기에 사용자의 고 수준의 개념(high level concepts)과 시스템의 저 수준의 특징값 사이에 상당한 차이가 있다는 점과 인간의 유사성 인식의 주관성을 고려하지 못하고 있다[2, 6].

따라서, 최근의 내용기반 영상검색의 연구는 전체적인 영상 정보검색 시스템의 일부로서 사용자를 포함시키는 상호작용 메카니즘(interactive mechanism)인 관련성 귀환(relevance feedback)에 집중하고 있다[2]. 관련성 귀환의 중요한 개념은 저 수준의 특징 값을 근거한 질의를 보다 더 정확하게 표현하기 위하여, 고 수준의 질의들을 재정의 하도록 사용자와 시스템이 서로 상호 작용하는 것이다.

본 논문에서는 관련성 귀환을 가진 영상검색 시스템에 대하여 연구하였다. 먼저 Gabor wavelet 계수를 이용한 질감특징을 사용하여 영상의 내용을 표현하고, 이를 다양한 형태로 영상검색에 사용하였다. 그리고 간단하면서도 효과적인 관련성 귀환을 제안하고 이를 검색 시스템에 적용하여 기존의 관련성 귀환을 가진 검색 시스템과 비교하였다.

### II. 관련성 귀환

대부분의 관련성 귀환에 관한 연구들은 다음과 같은 다양한 접근 방법들을 사용하고 있다.

#### 1. 질의점 이동 방법

이 방법은 사용자의 검색 결과에 대한 귀환(feedback)을 사용하여 현재의 질의점(query point)을 이동함으로써 이상적인 질의점에 대한 추정을 개선하기 위한 접근 방법이다. Rocchio의 벡터 공간 모델(vector space model)을 바탕으로 주로 문서정보처리에 적용했으며, 다음 (1)식과 같은 기본 식을 사용한다.

$$Q' = Q + b (\sum R_i) / N_1 - c (\sum R_i') / N_2 \quad (1)$$

여기서 Q는 초기 질의이고, Q'는 수정된 질의를 나타낸다.  $R_i$  과  $R_i'$ 은 관련된 문서와 관련되지 않은 문서를 각각 나타낸다. 그리고  $N_1$ 과  $N_2$ 는 관련된 문서의 수와 관련되지 않은 문서의 수를 의미하며, b와 c는 적절한 상수이다.

미국 일리노이 대학의 Rui 팀은 영상특징 벡터로부터 의사 문서 벡터(pseudo-document vector)를 생성하여 (1)식을 적용하려고 시도하였다[3]. 그리고 프랑스 INRIA의 Nastar 팀은 인수(parametric) 혹은 비인수(non-parametric) 추정자(estimator)를 사용하여 관련이 있는 개체의 특징벡터의 분포를 추정하는 접근 방법을 사용하였다[4].

#### 2. 가중치 갱신 방법

가중치 갱신 방법은 유사성을 계산할 때 사용하는 가중치를 수정함으로써 이루어진다. 예로서, 만일 j 번째 축을 따라 질의와 관련된 객체들의 분산이 크면, j 번째 축과 관련된 특징 값들의 가중치를 낮게 주어야 한다. 따라서 j 번째 축과 관련된 표준편차의 역수를 가중치 값으로 사용할 수 있다. 즉,  $W_j = 1/\sigma_j$  의  $W_j$ 는 다음 영상 검색의 새로운 유사성 계산을 위한 가중치로 사용될 수 있다.

싱가폴의 Chua 와 미국의 Rui 등은 간단한 경험적인 가중치 값을 사용하여 각 특징의 중요성에 따라 특징의 가중치를 조절하였다[3,5].

3. 기타 방법들

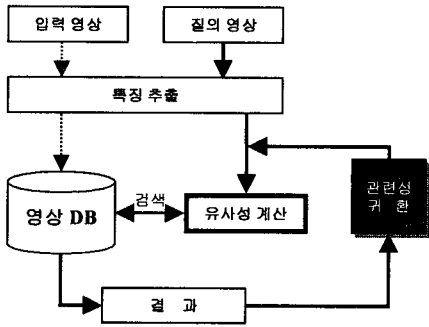
사용자의 귀환은 다양한 수준에 다양한 방법으로 적용이 가능하다. 즉, 저 수준 특징 값의 가중치를 조절하거나 또는 유사성 척도의 인수들을 변경할 수 있다. 또한 특징 값이나 유사성 척도의 인수를 변경하지 않고 고 수준 귀환의 하나로서 검색 시스템에서 나온 검색 결과의 순위를 조정할 수도 있다.

기타 방법의 예로, 일본의 Ishikawa[7], 그리스의 Doulamis[8], 그리고 영국의 Wood[9] 등은 질의 영상과 관련된 영상 혹은 관련되지 않은 영상간의 거리를 측정하는 유사성 척도나 비용함수(cost function)를 조절하는 접근 방법을 사용하고 있다.

본 연구에서는 고 수준 귀환의 한 방법으로서 초기 검색 결과에 대하여 사용자의 평가를 간단하게 유사성 계산에 적용시키는 관련성 귀환을 사용하였다.

III. 영상검색 시스템

제안한 영상 검색 시스템의 전체적인 구성도는 (그림 1)과 같다.



(그림 1) 제안한 영상 검색 시스템

먼저 질의 영상이 들어오면 입력된 데이터베이스의 영상들과 마찬가지로 특징 추출 단계를 거치고, 데이터베이스 내의 영상들과 유사성 계산을 하여 초기 검색영상을 출력한다. 다음, 사용자의 검색출력 영상에 대한 평가를 관련성 귀환을 통하여 1차적으로 시스템에 귀환시킨다. 시스템은 사용자의 관련성 귀환을 고려하여 2차 검색을 수행하고 그 결과를 출력한다. 이러한 사용자의 귀환 과정이 반복되어진다. 다음은 각 세부 과정에 대한 설명이다.

1. 특징 추출

영상 데이터베이스에 입력된 각 영상에 대하여 효과적인 특징을 추출하여 각 영상을 표현하고 이를 검색에 이용한다.

본 연구에서는 128 x 128 영상으로부터 각기 24차원(4 scale and 6 orientation)의 Gabor 특징을 추출하고, 특징 벡터 크기를 줄이기 위하여 KL변환을 이용하여 12차원의 특징벡터로 줄였다.

그리고 추출된 특징벡터(Raw feature)를 그대로 혹은 가우시안 정규화(Gaussian normalization), 평균 정규화(Normalized by mean) 그리고 평균-표준편차 정규화(Normalized by mean\*std dev) 등 다양한 형태로 변형하여 사용하였다.

2. 관련성 귀환 알고리즘

일반적으로 영상 검색시스템이 사용하는 저 수준의 특징값을 근거한 질의는 사용자의 고 수준의 질의 정보를 전부 표현할 수 없으므로, 영상 검색시스템과 사용자간의 관련성 귀환(relevance feedback)을 통하여, 고 수준의 질의를 재 정의하도록 함으로써 최종적으로 원하는 영상들을 검색한다.

본 연구에서는 사용자 귀환의 간편성과 효율성을 위하여 다음 (그림 2)와 같은 간단한 상벌(reward and punishment)알고리즘을 유사성 계산에 적용하여 유사성을 수정하는 방법을 사용하였고, 이를 검색시스템에 적용하여 관련성 귀환을 수행하였다.

```

for (i=0; i< # of image; i++)
  if (image[i] is in the Top Rank List and relevant)
    // Reward weight=Wr
    similarity[i] = Wr*similarity[i];
  else if (image[i] is in the Top Rank List and
non-relevant)
    // Punishment weight=Wp
    similarity[i] = Wp*similarity[i];
  else do not change similarity[i];
Reorder the similarities and display the next image
set
    
```

(그림 2) 상벌 귀환 알고리즘

3. 유사성 계산 및 검색 효율

영상 데이터베이스로부터 질의 영상과 유사한 영상들을 검색하기 위하여 유사성 척도를 사용하여 유사성을 계산한다.

본 연구에서는 질의 영상과 영상 데이터베이스 내에 저장된 영상들간의 유사성을 측정하기 위해, City-block 거리 유사성 척도를 사용하였다. 식 (2)는 City-block 거리 유사성 척도이다.

$$D(Q, I)_{City-block} = \sum_{i=1}^n \sum_{j=1}^m |f_i - f_j| \quad (2)$$

(2)식에서 Q는 질의 영상, I는 데이터베이스 내의 영상을 나타낸다. 그리고 f와 f'은 각각 두 영상의 특징 벡터들이다.

내용기반 영상 검색의 효율성을 분석하기 위하여, Retrieval Precision을 성능 평가 척도로 이용한다. Retrieval Precision은 다음 식(3)과 같이 검색된 전체 영상들 중에서 질의와 관련된 영상들의 비율로 나타낸다.

$$Retrieval\ Precision = \frac{Relevant\ Images}{Returned\ Images} \times 100 \quad (3)$$

여기서 Relevant Images는 질의 영상과 관련이 있는 영상을 뜻하는 것으로, 본 연구에서는 질의 영상과 같은 512 x 512 크기의 원 영상에 속하는 128 x 128 크기의 실험 영상을 뜻한다.

Returned Images는 검색 시스템에 출력된 영상을 뜻하며, 본 연구에서는 질의 영상을 포함하여 검색되어진 최상위 16개의 검색된 영상들을 나타낸다.

IV. 실험 및 토의

본 실험을 위하여 MIT Media lab의 VisTex 영상 데이터베이스에서 16종류의 512 x 512 크기의 원 영상들을 사용하여, 128 x 128 크기의 512 개의 영상들로 구성된 실험 영상 데이터베이스를 구축하고 실험하였다. 질의는 16종류의 영상에서 32개 질의 영상을 선택하여 실험하였다.

<표 1>은 제안된 관련성 귀환 방법을 사용한 각 특징 값 형태에 따른 Retrieval Precision을 나타낸 것이다. 표에서 전체적으로 모든 각 특징 값 형태에서 보면, 관련성 귀환을 한 경우가 관련성 귀환을 하지 않은 경우에 비하여 Retrieval Precision이 약 10~20% 높은 값을 보였다. 관련성 귀환 중에서는 첫 번째 귀환이 가장 효과가 있고, 다음은 두 번째 귀환 등의 순으로 나타났다. 그리고 특징 값들 형태 중에서 Raw feature와 가우시안 정규화(Gaussian normalization)의 경우가 다른 형태의 특징 값들보다 Retrieval Precision이 높은 것으로 나타났다.

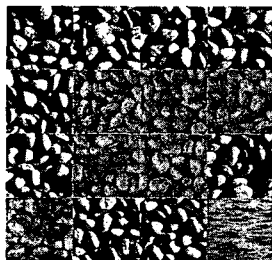
<표 1> 각 특징 값 형태에 따른 Retrieval Precision(%)

	Without RF	First RF	Second RF	Third RF
Raw feature	63.87	72.26	74.61	76.55
Normalized by mean	31.64	41.99	48.05	51.37
Normalized by mean*std dev	35.66	45.70	51.37	55.08
Gaussian normalization	51.17	63.28	66.41	67.97

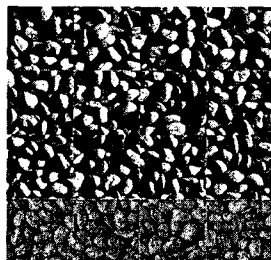
<표 2> INRIA의 Parametric접근 방법에 의한 Retrieval Precision(%)

	Without RF	First RF	Second RF	Third RF
Gaussian normalization	51.17	55.66	59.18	60.74

<표 2>는 INRIA의 Parametric접근 방법에 의한 Retrieval Precision을 나타낸 것이다. 표에서 관련성 귀환의 효과는 <표 1>과 마찬가지로 첫 번째 귀환이 가장 효과가 있는 것으로 보인다. 그러나 제안한 방법에 비하여 복잡하면서 성능은 첫 번째 귀환에서 약 8% 낮은 것으로 나타났다.



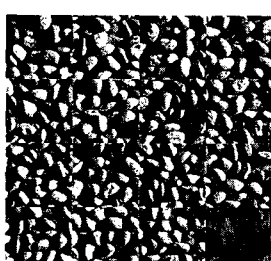
(a) Without RF



(b) First RF

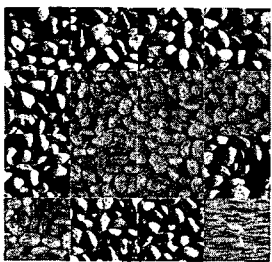


(c) Second RF

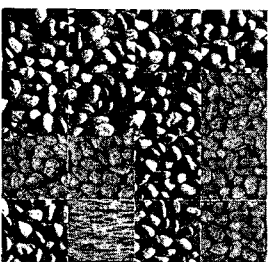


(d) Third RF

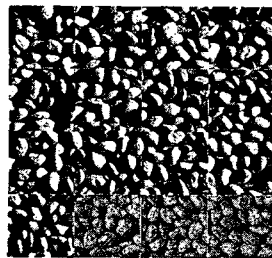
(그림 2) 제안된 방법에 의한 관련성 귀환의 결과의 예



(a) Without RF



(b) First RF



(c) Second RF



(d) Third RF

(그림 3) INRIA 방법에 의한 관련성 귀환의 결과의 예

(그림2)와 (그림 3)은 제안된 방법과 INRIA의 관련성 귀환의 결과의 예를 각각 보인 것이다. 그림들에서 좌측 상단의 첫 번째 그림이 질의 영상이며, 좌우 상하 순으로 유사도에 따라 출력한 것이다.

그림의 예들에서 제안된 방법이 INRIA의 경우보다 각 귀환 단계에서 질의 영상과 유사한 영상들을 더 많이 검색한 것을 확인할 수 있다.

### V. 결 론

본 연구에서는 질감정보를 이용한 영상 데이터베이스 검색 시스템에 간단하면서 효과적인 관련성 귀환을 적용하였다.

먼저 영상 데이터베이스 내에 있는 영상들에 대하여 Gabor Wavelet 변환을 이용하여 질감특징을 추출하고, 이를 바탕으로 다양한 형태의 특징 값들을 구하고 이를 검색에 사용하였다. 또한, 초기 검색결과에 대하여 제안한 간단한 상별 귀환 알고리즘의 관련성 귀환을 영상 검색시스템에 적용하고, 이를 기존의 관련성 귀환을 가진 영상 검색시스템과 비교하였다.

16종류의 512개의 영상으로 구성된 영상 데이터베이스에 대하여 실험한 결과, 전체적으로 관련성 귀환을 적용한 경우가 관련성 귀환을 적용하지 않는 경우에 비하여 Retrieval Precision이 높은 값을 보였다. 그리고 관련성 귀환 중에서는 첫 번째 귀환이 가장 효과가 있는 것으로 나타났다.

또한 제안된 귀환을 가진 검색방법은 INRIA의 방법보다 각 귀환단계에서 약 7 ~ 8%의 높은 검색 효율을 보였다.

### 참고 문헌

- [1] J.P.Eakins, M.E.Graham. 'Content-based Image Retrieval,' Technical Report, University of Northumbria at Newcastle, UK, 1999.
- [2] Yong Rui, Thomas S. Huang and Sharad Mehrotra, Content-based image retrieval with relevance feedback in MARS, Proc. of IEEE Conference on Image Processing 97, vol.2, pp.815-818, Santa Barbara, 1997.
- [3] Yong Rui, Thomas S. Huang and Sharad Mehrotra,, Relevance Feedback: A Power Tool for Interactive Content Based Image Retrieval. IEEE Transactions on Circuits and Video Technology, 8(5): 644-655, 1998.
- [4] C. Nastar, M. Mitschke , C. Meilhac and N. Boujemaa, Surfimage: a flexible content-based image retrieval system, ACM Multimedia 98, Bristol,1998.
- [5] T. Chua, C. Chu and M. Kankanhalli, Relevance feedback techniques for image retrieval using multiple attributes, Proc. of IEEE International Conference on Multimedia Computing and Systems 99, vol.2, pp.890-894, Florence Italy,1999.
- [6] Yong Rui, Thomas S. Huang and Sharad Mehrotra, Relevance feedback techniques in interactive content-based image retrieval, Proc. of IS&T and SPIE Storage and Retrieval of Image and video Database vol.VI, pp.25-36, San Jose, 1998.
- [7] Y. Ishikawa, R. Subramanya and C. Faloutsos, MindReader: Query database through multiple examples, VLD 98 Conference, New York, 1998.
- [8] A.D. Doulamis, Y.S. Avrithis, N.D. Doulamis and S.D. Kollias, Interactive content-based retrieval in video database using fuzzy classification and relevance feedback, Proc. of IEEE International Conference on Multimedia Computing and Systems 99, vol.2, pp.954-958, Florence Italy,1999.
- [9] M.E.J. Wood, N.W. Campbell and B.T. Thomas, Iterative refinement by relevance feedback in content-based digital image retrieval, ACM Multimedia 98- electronic Proceedings,1998.