

論文2002-39CI-4-4

영상 검색을 위한 적합성 피드백의 개선

(Improvement of Relevance Feedback for Image Retrieval)

尹帥征*, 朴東權*, 元致善*

(Su-Jung Yoon, Dong-kwon Park, and Chee-sun Won)

요약

본 논문에서는, 확률적 방법과 질의 위치 이동 방법을 융합하여 검색 성능을 향상시키는 영상검색 방법을 제안한다. 제안한 알고리즘은, 질의 영상과 데이터베이스 영상 사이의 유사도 계산에서, 확률적 방법의 유사도와 질의 위치 이동 방법의 유사도를 융합한다. 본 논문에서 이용된 확률적 방법은 부정적 예제들을 다루기에 적합하다. 반면에, 질의 위치 이동 방법은 긍정적예제의 통계적인 특성을 다룬다. 본 논문의 목적은 이러한 두 방법을 융합함으로써, 각각의 방법이 가지고 있는 단점을 극복하는 것이다. 실험결과는 제안한 방법이 확률적 방법과 질의 위치 이동 방법을 각각 적용한 경우보다 더 나은 성능을 나타낸다는 것을 보여준다.

Abstract

In this paper, we present an image retrieval method for improving retrieval performance by fusion of probabilistic method and query point movement. In the proposed algorithm, the similarity for probabilistic method and the similarity for query point movement are fused in the computation of the similarity between a query image and database image. The probabilistic method used in this paper is suitable for handling negative examples. On the other hand, query point movement deals with the statistical property of positive examples. Combining these two methods, our goal is to overcome their shortcoming. Experimental results show that the proposed method yields better performances over the probabilistic method and query point movement, respectively.

Key Words : Image indexing and retrieval, relevance feedback

I. 서론

적합성 피드백은 마치 사용자가 검색 시스템의 일부처럼 사용되어 시스템과 사용자 사이의 상호작용을 함으로써 내용기반 영상검색에서 드러난 단점을 줄이기 위한 방법이다. 사람이 느끼는 의미론적인 방법을 고

수준의 레벨이라고 하고 컴퓨터에 의해서 처리 가능한 비주얼 한 특징(Feature)을 저 수준의 레벨이라고 하면, 저 수준의 특징을 이용해서 검색된 정보를 사용자가 고 수준의 레벨에서 선별하고, 이 선별된 정보가 다시 저 수준의 특징의 갱신에 사용되어, 사용자에게 새로운 정보를 제공한다. 이러한 과정의 반복이 적합성 피드백의 기본 개념이다. 여기에는 한가지 가정이 들어 있다. 사용자의 의미론적인 선택을 좀더 낮은 수준의 특징이 반영할 수 있어야 한다는 것이다. 사용자에게 의해서 피드백 되는 데이터는 크게 두 가지로 구분될 수 있다. 긍정적 예제의 집합(Positive examples)과 부정적 예제의 집합(Negative examples)이다. 긍정적 예제의 집합은 사용자가 판단할 때 자신이 찾고자 하는 영상과 관련이 있는 영상들의 집합을 나타내고 부정적 예제의

* 正會員, 東國大學校 電子工學科

(Department of Electronic Engineering, Dongguk University)

※ "이 연구는 2001학년도 동국대학교 연구원 지원에 의하여 이루어졌음"

接受日字:2001年10月16日, 수정완료일:2002年5月17日

집합은 관련이 없는 영상들의 집합을 나타낸다. 긍정적 예제의 집합을 처리하는 부분을 긍정적 피드백(Positive feedback)이라고 하고 부정적 예제의 집합을 처리하는 부분을 부정적 피드백(Negative feedback)이라고 한다. 많은 기존의 적합성 피드백 시스템에서는 긍정적 피드백과 부정적 피드백을 내부적으로 같은 과정으로 처리하였다. 대개, 질의 위치 이동이나 재가중치 적용 중 한 가지 방법을, 혹은 두 가지 방법을 함께 사용해서 처리하였는데, 이 방법들은 예제들의 관련성을 이용해서 영상의 특징을 구성하는 빈(Bin)을 갱신하는 방법이다. 그런데, 긍정적 예제의 집합을 구성하는 예제는 사용자가 판단하는 의미론의 관점에서 볼 때 찾고자 하는 영상과 유사한 영상들을 선택함으로써 비슷한 의미를 갖는 경우가 많다. 또한, 시스템에 이용된 특징이 영상과 관련된 내용을 충분히 반영한다면 빈의 관점에서도 유사성을 가지고 있는 경우가 대부분이며, 그 특징에 의해서 생성되는 벡터공간 내에서는 무리를 이루는 경우가 많다. 그러나, 부정적 예제의 집합을 구성하는 예제는 대개 의미론적인 관련성이 거의 없다. 따라서, 빈의 관점에서도 유사성이 거의 없으며, 벡터공간에서는 독립적, 혹은 고립되어 단독으로 존재하는 경우가 많다. 따라서, 통계적인 특성을 가지고 긍정적예제의 집합을 다루는 긍정적 피드백을 사용한다면, 부정적 피드백은 부정적 예제의 공통된 특성을 고려하지 않는 다른 방법을 고려해 보아야 한다.^[1] 본 논문에서는 기존의 영상검색에 적용되는 적합성 피드백에서 소홀이 취급되던 부정적 피드백을 처리하는 최근에 제시된 방법과 그 단점을 보완해 줄 수 있는 질의위치이동방법을 융합함으로써, 검색 효율을 향상 시키는 방법을 제안한다. 본 논문의 구성은 다음과 같다. 2장에서는 적합성 피드백의 소개와 적합성 피드백에 쓰이는 방법들을 알아보고, 3장에서는 제안한 방법에 대해서 설명한다. 4장에서는 제안한 방법과 기존의 방법을 이용한 결과를 비교하고, 5장에서 결론을 맺는다.

II. 적합성 피드백

적합성 피드백은 기본적으로 벡터 스페이스 모델을 기본으로 하며 여러 가지 방법이 개발되었다. 그 중에 영상검색과 관련해서 두 가지 주요한 접근 방법이 있다.^[2] 그 중 하나는 질의 위치 이동(Query point movement)이고 다른 하나는 재가중치 적용(Re-

weighting)이다. 본 논문에서는 이 두 방법 중 질의 위치 이동방법을 이용하므로 질의 위치 이동방법만 설명한다.

1. 질의 위치 이동 (Query point movement)

질의 위치 이동은 사용자에게 의해서 피드백된 정보의 통계적인 특성을 이용해서 이상적인 질의 위치(Ideal query point)에 접근하는 방법으로 벡터들의 크기와 방향을 고려해서 공통적인 의미를 갖는 벡터를 찾는 것이다. 그림 1 은 질의 위치 이동을 설명하고 있다. X와 Y의 두 특징요소에 대해 반복 검색 횟수를 K라고 정의할 때 원의 내부는 탐색범위를 나타내고 검은 점들은 사용자가 K번 결과에서 관련 있다고 표시한 영상들의 특징요소의 위치를 나타낸다. 관련 영상들의 특징정보를 계산해서 K+1번 검색에 이용한다. 우측의 그림에서 탐색범위의 중심이 이동하는 것을 볼 수 있다. 이처럼 사용자와 시스템 사이의 상호작용을 반복해서 탐색범위의 중심이 사용자가 원하는 결과를 가장 잘 반영할 수 있는 이상적인 질의 위치에 접근해가는 방법이 질의 위치 이동이다.

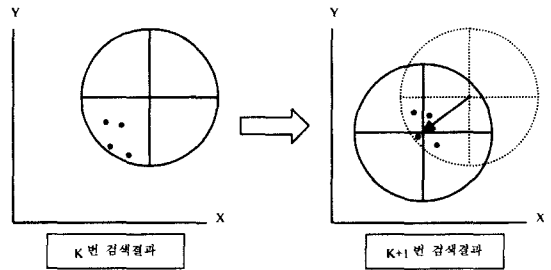


그림 1. 질의 위치 이동
Fig. 1. Query Point Movement.

일반적으로는 질의 위치 이동은 벡터공간에서 질의와 관련 있는 집합의 평균을 많이 사용한다. 이는 Standard Rocchio의 방식을 따른 것으로 벡터공간에의 검색에서 다음 번 질의 위치를 결정하는 방법으로 널리 알려져 있다. 관련식은 식 (1)과 같다.^[3]

$$Q' = \alpha Q + \beta \frac{1}{n_p} \sum_{x \in E_p} x - \gamma \frac{1}{n_n} \sum_{x \in E_n} x \quad (1)$$

Q'은 새로운 질의 위치를 나타내고, Q는 초기 질의 위치를 나타낸다. E_p는 검색결과 중에서 긍정적 예제의 집합을 나타내고, E_n은 검색결과 중에서 부정적 예제의 집합을 나타낸다. n_p는 E_p의 원소의 개수

를 나타내고, n_n 은 E_N 의 원소의 개수를 나타낸다. α, β, γ 는 각각 적절한 가중치를 나타낸다. 위의 수식에서 두 번째 항은 $\beta\mu_p$ 로 나타낼 수 있다. μ_p 는 E_p 의 평균을 의미한다. 반면에, 세 번째 항은 $\gamma\mu_n$ 으로 나타낼 수 있다. μ_n 은 E_N 의 평균을 의미한다. 문자와 음향에 관련된 정보검색에서는 $\alpha=1.0$, $\beta=0.75$, $\gamma=0.25$ 로 고정한 경우에 더 좋은 결과를 보였다.^[4] 그러나, 영상 검색에 위의 가중치를 그대로 적용하기에는 무리가 있다. 그 이유는, 문자는 사용자가 관련 있는 문자와 관련 없는 문자를 질의로 선택하였을 때, 문서 내에서 선택한 문자의 출현빈도를 계산하여 검색하면 되지만, 영상검색은 데이터베이스 내에 같은 영상이 존재하지 않는 경우가 대부분이므로 출현빈도를 계산할 수가 없고, 질의와 완전히 같은 영상을 찾는 것이 목적이 아니라, 유사한 영상을 찾는 것이 목적이기 때문이다. 따라서, 영상검색과 관련해서는 다른 가중치 접근 방법이 제시되었다.^[5] 그것은 $\alpha=0$, $\beta=1$, $\gamma=0$ 으로 하는 것으로 새로운 질의의 위치를 관련 있는 집합의 평균으로 하는 것이다. 이 방법은 영상검색에서 좋은 결과를 보이고 있다. 그 이유는, 적합성 피드백을 사용하면, E_p 의 관점에서 검색과정을 반복하여 결과를 보았을 때, 결과에 유사한 영상이 반복되어 나타나고 그 반복되는 유사한 영상이 새로운 질의로 선택되므로 이전검색의 데이터를 고려하지 않아도 되므로 α 를 고려하지 않은 것이다. 그리고, 대개 관련 없는 영상의 집합은 공통적인 성질이 없고 고립적이며, 독립적인 경우가 많다. 따라서, 식(1)과 같은 통계적 특성을 적용하는 방법을 그대로 관련 없는 영상의 집합에 적용하면, 검색결과에 대해서 긍정적인 영향을 크게 주지 못하고, 오히려 검색 효율이 저하될 수도 있다. 따라서, γ 를 고려의 대상에서 제외한 것이다. 그러나, 관련 없는 영상의 집합이 의미가 없는 것은 아니다. 관련 없는 영상의 집합에 대해서 적절하게 다른 방법을 이용하면, 검색 효율을 높일 수 있다.

2. 부정적 피드백을 다루기 위한 방법

영상 검색 시스템에 도입된 적합성 피드백은 질의 위치 이동 방법과 재가중치 적용방법이 가장 많이 적용되는 방법이지만, 그 외에도 몇 가지 종류가 있다. 그 중에 한가지는 확률적 모델을 이용하는 것이다. 영상검색에 적용되는 적합성 피드백은 유사한 영상을 찾는다는 관점에서 볼 때, 미지의 데이터를 이미 알고 있는

데이터와 비교해서 분류하는 클러스터링(Clustering)과 관련이 깊다. 클러스터링과 관련되어 많이 사용되는 알고리즘으로 K-NN(K-nearest neighbor)알고리즘이 있다. 이는 입력된 데이터가 어떠한 특성을 갖는 지를 알 수 없을 때, 그 데이터와 가장 가깝게 있는 K개의 이미 알고있는 데이터를 조사해서, 입력된 데이터의 특성을 추정하는 방법이다. 예를 들어서, 허리둘레와 가슴둘레로 성별을 구별하는 문제가 있다고 가정한다. 성별을 모르는 어떤 사람의 허리둘레와 가슴둘레의 데이터가 입력되었을 때, 그 데이터에 가장 가까운 K개의 데이터를 조사한다. 그 결과가 남성이 많으면, 입력된 데이터는 남성으로, 여성이 많으면, 입력된 데이터를 여성으로 판단하는 것이다. 이러한 방법을 영상검색에 적용할 수 있다. 즉, 이미 알고 있는 데이터로 사용자가 선택한 긍정적 예제의 집합과 부정적 예제의 집합을 이용하고, 입력된 데이터는 데이터베이스내의 영상의 정보를 이용하는 것이다. 만약, 시스템에 사용된 특징이 충분히 영상의 특성을 반영할 수 있다면, 이 특징을 사용하여 구성된 공간 내에서 사용자가 긍정적 예제로 선택한 데이터와 가까운 데이터는 사용자가 찾으려고 하는 영상과 유사한 영상일 가능성이 크고, 사용자가 부정적 예제로 선택한 데이터와 가까운 데이터는 사용자가 검색과정에서 배제하고자 하는 영상과 유사한 영상일 가능성이 크다. 따라서, 이런 경우에, 긍정적 예제에 가까운 영상 데이터는 유사도 측정에서 거리를 줄여서 결과에서 상위에 위치하게 해야 하고, 부정적 예제에 가까운 영상 데이터는 거리를 늘여서 결과에서 하위에 위치하게 해야 한다. 이때, 긍정적 예제와 부정적 예제를 각각 나누어서 생각하고, K를 1로 하면, 입력된 데이터에서 가장 가까운 긍정적 예제와 부정적 예제의 거리를 계산할 수 있다. 본 논문에서는 Brunell과 Mich가 제안한 방법^[6]을 이용하겠다. 입력된 데이터와 긍정적 예제와의 최소거리와 부정적 예제와의 최소거리를 이용하여, 두 거리 사이의 비율(Rate)을 구할 수 있으며, 이러한 비율을 이용하면 사용자의 의도를 좀더 잘 반영하는 거리 계산이 가능하다. 내용을 살펴보면 식 (2)(3)(4)와 같다.

$$D(I, E) = \min_{I' \in E} d(I, I') \quad (2)$$

$$D'(I, E_p, E_n) = D(I, E_p) \left[\frac{D(I, E_p)}{D(I, E_n)} \right]^\gamma \quad (3)$$

$$d(I, I') = \sum_{j=1}^J |F(j) - F'(j)| \quad (4)$$

d 는 두 영상(I 와 I')사이의 벡터공간에서의 거리를 나타낸다. F 는 영상 I 의 내용을 나타내는 특징이고, 특징 F 는 J 개의 bin으로 구성되어 있다고 정의한다. 따라서, $F(j)$ 는 j 번째 bin의 값을 나타낸다. E 는 사용자가 선택한 예제의 집합을 나타낸다. E_p 는 긍정적 예제의 집합을 나타내고, E_N 은 부정적 예제의 집합을 나타낸다. 따라서, $D(I, E_p)$ 는 영상 I 와 긍정적 예제의 집합을 이루는 원소들 사이의 최소거리를, $D(I, E_N)$ 는 영상 I 와 부정적 예제의 집합을 이루는 원소들 사이의 최소거리를 의미한다. D' 은 영상 I 와 긍정적 예제의 집합, 부정적 예제의 집합을 모두 고려한 거리를 나타낸다. γ 는 D' 의 크기를 조절하는 요소이며, 양의 값을 갖는다. 만약, $D(I, E_N)$ 을 고려하지 않는다면, 식 (4)에서 $[(D(I, E_p)/D(I, E_N))^\gamma]$ 는 적합한 의미를 갖지 못하게 되므로, 고려의 대상에서 제외하면, $D'(I, E_p, E_N) = D(I, E_p)$ 가 된다. 표 1은 $D(I, E_p)$ 과 $D(I, E_N)$ 을 비교하였을 때, 식(4)에 의한 $D'(I, E_p, E_N)$ 의 변화를 나타낸다.

표 1. 영상 I 와 예제와의 최소거리에 따른 D' 의 변화

Table 1. The variance of D' by minimum distance between image I and example

$D(I, E_p)$ 와 $D(I, E_N)$ 의 비교	$D'(I, E_p, E_N)$ 의 변화
$D(I, E_p) \approx D(I, E_N)$	$D'(I, E_p, E_N) \approx D(I, E_p)$
$D(I, E_p) < D(I, E_N)$	$D'(I, E_p, E_N) < D(I, E_p)$
$D(I, E_p) > D(I, E_N)$	$D'(I, E_p, E_N) > D(I, E_p)$

- (1) $D(I, E_p)$ 과 $D(I, E_N)$ 이 거의 같다면, 최종거리는 긍정적 예제의 집합을 이루는 원소들과의 최소거리와 비슷하게 된다는 것으로, 순위 결정에서 부정적 예제는 특별한 영향을 주지 않는다는 것을 의미한다.
- (2) $D(I, E_p)$ 이 $D(I, E_N)$ 보다 작다면, D' 에 의한 순위 결정에서 $D(I, E_p)$ 과 $D(I, E_N)$ 이 거의 같은 경우보다 더 작은 거리를 갖게 된다는 것이다. 따라서, 현재 영상 I 가 사용자에게 보여지는 검색결과에 포함될 가능성이 커진다.

(3) $D(I, E_p)$ 이 $D(I, E_N)$ 보다 크다면, D' 에 의한 순위 결정에서 $D(I, E_p)$ 과 $D(I, E_N)$ 이 거의 같은 경우보다 더 큰 거리를 갖게 된다는 것이다. 따라서, 현재 영상 I 가 사용자에게 보여지는 검색결과에서 배제될 가능성이 커진다.

[6]에서는 긍정적 예제들을 클러스터링 하는 방법을 사용한다. 그 이유는 검색 시스템의 응답속도를 빠르게 하고, 사용자가 선택한 예제가 하나의 무리를 이루지 않고 여러 무리를 이루는 경우에도 검색의 결과를 가능한 한 좋게 하기 위해서 이다. 하지만, 대개의 경우, 사용자는 검색과정에서 자신이 찾고자 하는 영상과 관련된 방향으로 예제를 선택한다. 따라서, 사용자가 선택한 긍정적 예제들은 단일한 클러스터로 볼 수 있으며, 그러므로, 제안한 방법에서는 긍정적 예제를 클러스터링 하는 방법을 사용하지 않고, 3장에서 제안한 방법을 적용하였다.

III. 제안한 방법

식 (4)방법은 사용자가 선택한 예제들을 각각 이용함으로써, 긍정적 피드백과 부정적 피드백을 하나의 방법으로 함께 사용할 수 있다는 장점이 있다. 반면에, 각각의 예제 가운데 최소거리를 갖는 2개(긍정적 예제 1개, 부정적 예제 1개)만을 다루기 때문에 예제들의 통계적인 특성은 고려하지 않고 있다. 질의 위치 이동 방법은 예제의 통계적인 특성을 잘 반영하고 있는 반면, 부정

표 2. 질의 위치 이동 방법과 식(4)방법의 각각의 장단점

Table 2. The characteristic between query point movement and the eq.(4) method.

	질의 위치 이동 방법	식(4) 방법
장점	긍정적 예제의 집합의 통계적 특성을 이용한다.	부정적 피드백을 긍정적 피드백과 함께 다룰 수 있다.
단점	부정적 피드백을 긍정적 피드백과 같은 방법으로 계산하는 것은 적합하지 않다.	긍정적 예제의 통계적 특성을 이용하지 않는다.

적 피드백을 긍정적 피드백과 같은 방법으로 다루기에 부적합한 면이 있다. 따라서, 위의 두 가지 방법은 상호 보완적인 면을 가지고 있다는 것을 알 수 있다. 표 2에 각각의 장단점을 정리 하였다.

그러므로, 질의 위치 이동 방법과 식 (4)방법을 융합해서 사용하는 것은, 영상 검색에서 긍정적 예제들의 통계적인 특성을 반영하면서, 부정적 예제를 이용하는 적합한 방법을 제시함으로써, 검색의 효율을 높일 수 있다. 이러한 내용을 나타낸 것이 그림 2와 그림 3이다. 점선원은 예제로부터의 일정거리를 나타낸다. 그림 2에서 긍정적 예제들의 평균은 영상 I와 I'에 대해서 같은 거리를 갖는다($d_1=d_2$). 따라서, 질의 위치 이동 방법만을 이용했을 때 두 영상은 검색 결과에서 같은 순위를 갖게 된다. 그러나, 식 (4)방법을 융합해서 사용하면, 영상 I가 I'보다 긍정적 예제에 보다 더 가깝고, 따라서 사용자의 의도를 더 잘 반영하는 영상이라는 것을 알 수 있다. 융합된 결과에서는 영상 I에서 긍정적 예제들의 평균까지의 거리, 그리고, 영상 I에서 가장 작은 거리를 갖는 긍정적 예제까지의 거리를 합산하여 계산하므로, 결과적으로 영상 I가 영상 I'보다 검색 결과에서 더 높은 순위를 갖게 된다.

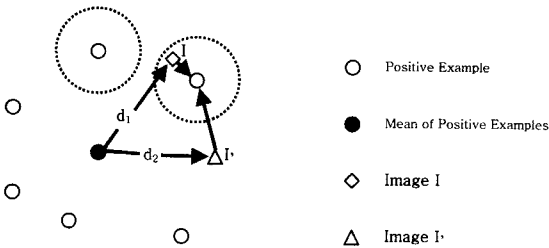


그림 2. 제안한 방법에 의한 질의 위치 이동 방법의 향상
Fig. 2. Enhancement of query point movement method by using the proposed method.

그림 3도 같은 방법으로 해석될 수 있다. 식 (4)방법에서 영상 I와 I'이 가장 가까운 긍정적 예제와 같은 거리 값을 갖는다고 하면($d_3=d_4$), 식 (4)방법만을 이용하였을 때 이 두 영상은 같은 순위를 갖게 된다. 하지만, 질의 위치 이동 방법을 융합하여 이용하면, 긍정적 예제의 평균에 가까운 영상 I가 영상 I'보다 융합된 결과에서 작은 거리를 갖게 되어, 더 높은 순위를 갖게 된다.

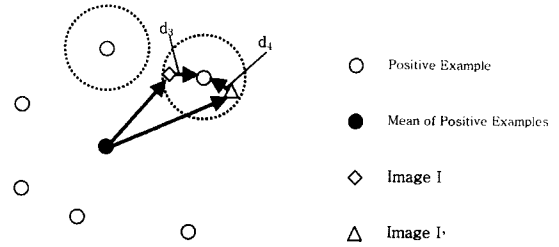


그림 3. 제안한 방법에 의한 식(4)방법의 향상
Fig. 3. Enhancement of the eq.(4) method by using the proposed method.

검색과정에 필요한 거리를 계산하기 위해서 다음과 같은 방법을 이용하였다. 제안한 방법에서는 질의 위치 이동 방법에서 다음 번 질의 위치를 구할 때, Standard Rocchio의 방식을 이용하였으며 $\alpha=0, \beta=1, \gamma=0$ 로 하여 긍정적 예제의 평균을 계산하였다. 이러한 내용을 식 (5)에 표시하였다. E_p 는 사용자가 피드백한 긍정적 예제의 개수를 나타내고, Q' 은 새로운 질의 위치를 나타낸다. Z 는 데이터베이스 내에서 하나의 영상 정보의 특징을 나타내는 데이터이다. Q' 에서 Z 까지의 거리 (D_{mean})는 식 (6)과 같다. 거리 계산에는 L_1 메트릭(Metric)을 이용하였다.

$$Q' = \frac{1}{|E_p|} \sum_{x \in E_p} x \tag{5}$$

$$D_{mean} = d(Q', Z) \tag{6}$$

식 (4)방법의 거리 계산은 식 (7)(8)을 이용하였다. 긍정적 예제의 집합과 데이터베이스내의 하나의 영상을 나타내는 특징과의 최소거리를 $MinPe$ 로, 부정적 예제의 집합과 데이터베이스내의 하나의 영상을 나타내는 특징과의 최소거리를 $MinNe$ 로 정의한다. E_p 와 E_N 의 최소 상대 거리(D_{min})의 계산은 식 (9)와 같다.

$$MinPe = D(Z, E_p) = \min_{I' \in E_p} d(Z, I') \tag{7}$$

$$MinNe = D(Z, E_N) = \min_{I' \in E_N} d(Z, I') \tag{8}$$

$$D_{min} = \left[\frac{MinPe}{MinNe} \right]^r \tag{9}$$

순위를 결정하기 위한 최종적으로 계산되는 거리(D_{final})는 식 (10)과 같이 각각 계산된 두 거리를 단순히 더하여 유사도 계산에 이용하면 된다.

$$D_{final} = D_{mean} + D_{min} \quad (10)$$

기존의 많은 적합성 피드백을 적용한 논문에서는 질의 위치 이동 방법을 사용할 때, 가중치 방법도 함께 사용하였다. 하지만, 제안한 방법에서는 D_{mean} 을 구할 때 가중치를 사용하지 않았다. 그것은, 제안한 방법은 검색에 사용되는 유사도를 구할 때, 질의 위치 이동 방법으로 구한 유사도와 식 (4)의 방법으로 구한 유사도를 융합해서 이용한다. 그런데, D_{min} 를 계산할 때, 사용자가 선택한 예제를 각각 다루고 있기 때문에 bin의 관련성을 고려하지 않는다. 그러므로, 질의 위치 이동 방법의 유사도를 나타내는 D_{mean} 을 계산할 때만 가중치를 적용하는 것은 적합하지 않으며, 두 방법 모두 같은 거리 계산 방식을 이용하는 것이 적합하다. 그래서, 질의 위치 이동 방법을 계산할 때, 가중치를 고려하지 않았다. 이러한 내용을 검증하기 위해서, 가중치를 고려한 방법과 고려하지 않은 방법을 비교해 보았다. 그림 4는 질의 위치 이동 방법을 적용할 때 가중치를 적용한 결과(Weight,^[7] Weight(1/std)^[2])와 적용하지 않는 결과(No Weight)를 나타낸다. 여기서 std는 표준편차(σ)를 나타낸다. 가중치 적용방법은 다음 두가지 방법을 각각 적용해 보았다.

1. Weight방법 : $w_i = \left(\sum_j^{nb} w_j \right)^{-1} w_i'$

$$w_i' = 1 - \left(\frac{\sigma_i^2 - \sigma_{min}^2}{\sigma_{max}^2 - \sigma_{min}^2} \right)$$

2. Weight(1/std) 방법 : $w_i = \left(\sum_j^{nb} \sigma_j^{-1} \right)^{-1} \frac{1}{\sigma_i}$

사용된 기호의 의미는 다음과 같다.

w_i (Wight of i^{th} bin), nb (Number of bins), $i(i^{th}$ bin), (Standard deviation)

ANMRR(Average normalized modified retrieval rank)은 검색된 영상들의 순위를 이용해서 검색성능을 나타내는 한 방법이다. 하나의 질의는 검색결과로부터 수정된 검색 순위의 정규화한 값을 나타내는 NMRR(Normalized modified retrieval rank)값을 갖는다. 따라서, 실험에 사용된 질의 영상의 모든 NMRR값을 더하

고, 그 총합을 질의 영상 개수로 나눈값이 ANMRR이다. 즉, NMRR의 평균값이다. ANMRR은 0에서 1사이 값을 갖으며, 낮은 값을 갖을수록 결과가 우수하다.^[8] 검색 횟수는 1이 초기검색을 의미하며, 2는 적합성 피드백을 한번 적용하고, 4는 적합성 피드백을 세 번 적용하였다는 것을 의미한다. 그림 4에서 알 수 있듯이, 질의 위치 이동 방법을 적용할 때, 가중치를 이용하지 않은 방법이 가중치를 적용한 두 방법보다 나은 결과를 보이고 있고, 가중치를 적용한 두 방법은 거의 비슷한 결과를 나타내고 있다.

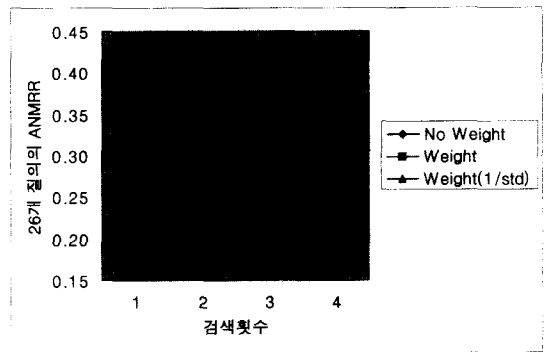


그림 4. 제안한 방법에서 질의위치이동방식에 가중치를 적용하였을 때와 적용하지 않았을 때의 비교
Fig. 4. Comparison of query point movement method with and without weights.

적합성 피드백을 적용한 검색은 여러 번의 반복과정을 하면서, 질의 정보를 갱신한다. 매번 반복되는 검색과정마다, 긍정적 예제와 부정적 예제가 선택된다. 선택된 예제들에 대한 정보는 계속 누적될 필요가 있다. 긍정적 예제의 집합은 반복되는 검색과정에서 계속적으로 보여지는 경향이 있다. 따라서, 중복되는 데이터가 존재하므로, 긍정적 예제의 집합은 그 크기가 증가하는 비율이 상대적으로 높지 않다는 것을 알 수 있다. 또한, 식 (1)을 이용할 때, α 를 고려하지 않으면, 긍정적 예제는 데이터를 축적하지 않아도 된다. 반면에, 부정적 예제는 매번 반복과정마다 바뀌는 경향이 있으므로, 부정적 예제의 집합의 크기의 증가율은 상대적으로 매우 높은 편이다. 반복되는 검색과정에서 누적되는 부정적 예제의 집합 다루기 위한 두가지 방법을 고려해 볼 수 있다. 첫 번째 방법은, 반복되는 검색과정에서 선택된 E_N 을 모두 저장하고, 이렇게 누적된 E_N 을 검색과정에 이용하는 방법이다. 시스템을 동시에 사용하는 사용자

가 많을 경우에도, 사용자마다 누적된 E_N 을 저장하면 되므로 필요한 메모리의 크기는 크지 않다. 하지만, 이 방법은 $MinNe$ 를 찾기 위해서 데이터베이스내의 하나의 영상(Z)과 E_N 을 비교하는 시간이, 누적되는 E_N 의 개수가 많아짐에 따라 증가한다는 문제가 있다. 두 번째 방법은, 식 (8)에서 알 수 있듯이, 부정적 예제의 집합을 나타내는 E_N 의 원소가 아무리 많다고 해도, 필요한 최소값은 단 하나뿐이라는 사실을 이용한 방법이다. 데이터베이스내의 영상 개수 만큼의 $MinNe$ 를 저장하기 위한 메모리공간을 할당한다. 이렇게 하면, 이전검색 과정까지의 누적된 와 데이터베이스내의 각각의 영상에 대한 $MinNe$ 를 구한 후 저장할 수 있다. 따라서, 현재검색에서 사용자가 선택한 E_N 에 대해서만 $MinNe$ 를 구한 후 저장된 $MinNe$ 와 비교하여 더 작은 값을 $MinNe$ 로 선택하고 이 값을 다시 저장한다. 이러한 과

정을 반복하면, 첫 번째 방법에서 누적된 모든 E_N 을 이용한 계산결과와 같은 결과를 얻을 수 있다. 할당된 공간에 저장된 $MinNe$ 값은 사용자가 새로운 질의로 검색을 할 때 초기화 된다. 이 방법은 $MinNe$ 를 찾기 위한 계산시간의 관점에서 볼 때, 현재검색에서 선택된 개수만 고려하고, 이전검색과정까지의 누적된 E_N 의 개수는 고려하지 않아도 되므로, 첫번째 방법에 비해서, 응답속도가 빠르다는 장점이 있다. 하지만, 동시에 사용자가 많아지면, 각각의 사용자를 위해 필요한 메모리공간의 크기가 첫번째 방법보다 많이 요구된다는 단점이 있다. 제안한 방식에서는 두 번째 방법을 이용하여 실험하였다.

그림 5는 제안한 방식의 검색 순서도를 나타낸다.

성능평가를 위해서 재현율(Recall)과 ANMRR을 이용하였다. 재현율은 데이터베이스내의 적합자료 중 탐색한 적합자료의 비율을 나타낸다.

IV. 실험결과

제안한 방법에서는 영상의 특징을 나타내기 위해서 EHD(Edge histogram descriptor)를 이용하였다.^[9] EHD는 80개의 로컬 bin(Local bin)으로 구성되고, 검색 시에는 글로벌(Global)과 세미글로벌(Semi-global)성분이 로컬 bin으로부터 재구성되며, 전체적으로 150개의 bin으로 확장되어 이용된다.^[10] 실험영상으로는 11639개의 자연영상을 이용하였다. 이 11639개의 자연영상에는 질의 영상 51개와 각각의 질의 영상과 유사한 영상이 3개부터 33개까지 존재한다. 나머지 영상들은 무작위로 포함된 영상들이다. 그림 6은 실험에 사용된 질의 영상 26개를 나타낸다. 26개의 질의 영상은 51개의 질의영상 가운데, 초기검색에서 유사영상을 모두 찾거나, 유사영상을 하나도 찾지 못해서 적합성 피드백을 적용할 필요가 없는 영상들을 제외한 질의영상을 이용하였다. 결과를 평가하기 위해서 재현율을 사용하였다. 식 (9)의 값으로는 2를 주었다.

그림 7은 26개 질의 영상의 초기 검색 결과에, 적합성 피드백을 1회만 적용하였을때의 결과를 재현율의 평균으로 나타내었다. 초기검색은 적합성 피드백을 사용하기 전의 결과를 나타내고, (PE)는 선택과정에서 긍정적 예제만을 선택한 경우를 의미하며, (PE,NE)는 긍정적 예제와 부정적 예제를 함께 선택한 경우를 나타낸다. QPM은 질의 위치이동 방법을 이용하였다는 것

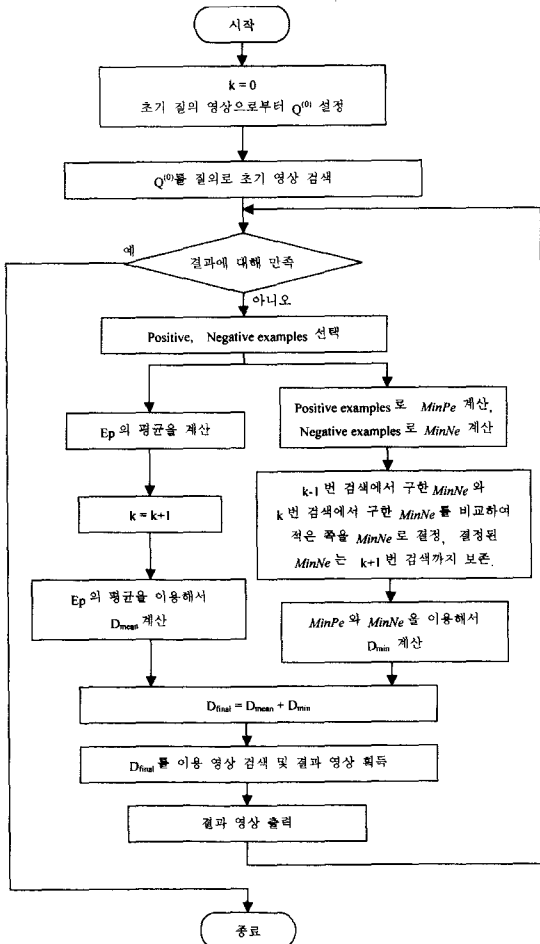


그림 5. 제안한 방식의 순서도
Fig. 5. Flowchart of the proposed method.

을 나타내고, 제안한 방법은 Proposed Method로 표시하였다. 따라서, QPM(PE)는 질의 위치 이동 방법만을 사용하고 긍정적 예제만을 선택하였을 때의 결과를 나타내고, Proposed Method(PE)는 제안한 방법을 적용하고 긍정적 예제만을 선택하였을 때의 결과를 표시한다. Proposed Method(PE,NE)는 제안한 방법을 적용하고 긍정적 예제와 부정적 예제를 함께 선택하였을 때의 결과를 나타낸다. 긍정적 예제만을 선택한 경우에, 제안한 방법이 질의 위치 이동방법만을 이용한 방법보다 우수함을 알 수 있다. 또한, 제안한 방법에 한해서는 긍정적 예제와 부정적 예제를 함께 선택하였을 때의 결과가 우수함을 알 수 있다. 이것은 부정적 예제를 검색에 이용되는 것이 향상된 결과를 가져올 수 있고, 그 방법으로는 부정적 예제는 각각 독립적으로 다루어지는 것이 적합하다는 것을 의미한다.

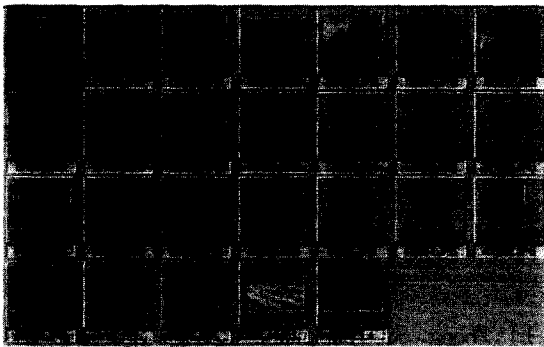


그림 6. 실험에 사용된 26개의 질의영상
Fig. 6. 26 Query images for experiments.

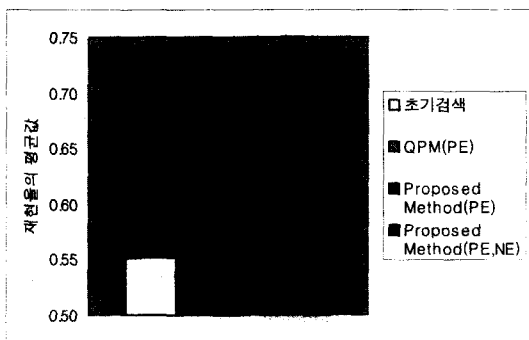
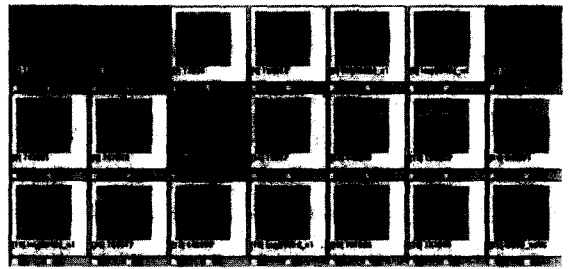


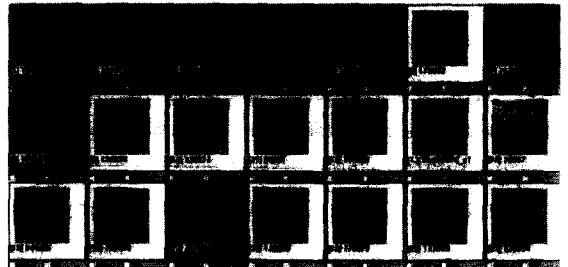
그림 7. 각각의 방식에 따른 검색결과의 재현율의 평균값
Fig. 7. Average recall values for each method.

그림 8은 제안한 방법의 반복적인 적용에 의한 검색 결과의 한 예를 나타내었다. 그림 8의 (a)(b)(c)에서 최

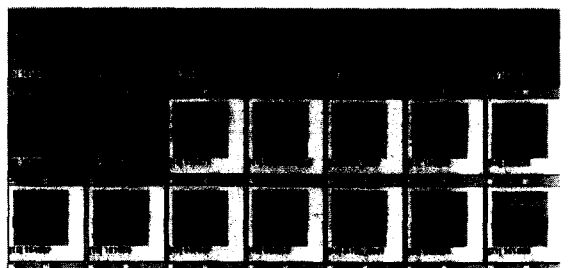
상단 가장 좌측의 영상이 가장 높은 순위를 갖는 영상이고, 우측 하단으로 갈수록 낮은 순위를 갖는다. (a)에서 가장 좌측 상단의 테두리가 어두운 색을 갖는 영상은 첫번째 순위 영상이면서, 동시에 초기검색에서 질의로 사용된 영상이다.(b)(c)의 결과에서 알 수 있듯이 검색을 반복함에 따라 유사한 영상이 빠르게 상위에 랭크되는 것을 볼 수 있다. 구체적으로 11639개의 실험영상 가운데 실험에 사용된 질의영상과 유사한 영상이 질의영상을 포함하여 9개가 존재하는데, (c)의 결과를



(a) 초기검색 결과
(a) First Retrieval Result



(b) 제안한 방법 1회 적용 결과
(b) Retrieval result after applying the proposed method once



(c) 제안한 방법 2회 적용 결과
(c) Retrieval result after applying the proposed method two times

그림 8. 실험 결과의 예
Fig. 8. Experimental results for each step of the proposed relevance feedback.

보면 9개를 모두 찾았을 뿐만 아니라, 모두 상위에 검색된 것을 볼 수 있다. 실험에 사용된 나머지 질의 영상에 대한 결과도 이와 비슷한 결과를 얻는 것을 확인할 수 있었다.

그림 8은 26개 질의 가운데 단지 하나의 질의에 대한 실험결과를 나타낸 것이다. 그러므로, 26개의 질의 영상을 이용하여 각각의 질의 영상에 대한 NMRR값을 구한 후, NMRR 값을 모두 더한 총합을 질의 영상의 개수로 나눈 ANMRR값을 이용해서 26개 질의 영상에 대한 결과를 나타낼 수 있다. 그림 9는 26개의 질의를 이용하고 적합성 피드백을 3회 적용할 때, 매번 적용할 때 마다 구한 ANMRR 결과이다. 검색 횟수는 1이 초기검색을 의미하며, 2는 적합성 피드백을 한번 적용하고, 4는 적합성 피드백을 세 번 적용하였다는 것을 의미한다. 식 (4) Method는 확률적 방법을 이용한 결과를 의미하고, Proposed Method는 제안한 방법의 결과를 나타낸다. QPM(PE), (PE,NE)는 앞서 재현율을 이용한 시험과 같은 의미를 갖는다. ANMRR은 검색순위를 나타내므로 낮은 값을 갖을수록 좋은 결과를 의미한다. 따라서, 제안한 방법이 질의 위치 이동 방법 혹은, 식 (4)방법만을 이용한 경우보다 우수하다는 것을 알 수 있다. 또한, 적합성 피드백을 반복해서 이용할 경우, 부정적 예제를 긍정적 예제와 함께 고려하여 검색하였을 경우가 긍정적 예제만을 고려한 경우보다 결과가 우수하며, 적합성 피드백의 적용횟수가 높아짐에 따라, 더욱 그 차이가 나는 것을 알 수 있다. 적합성 피드백 알고리즘의 우수함을 나타내는 기준은 사용자가 원하는 영상과 유사한 영상을 얼마나 반복 횟수를 적게 하고 데이터베이스에서 찾을 수 있는가로 측정할 수 있다.

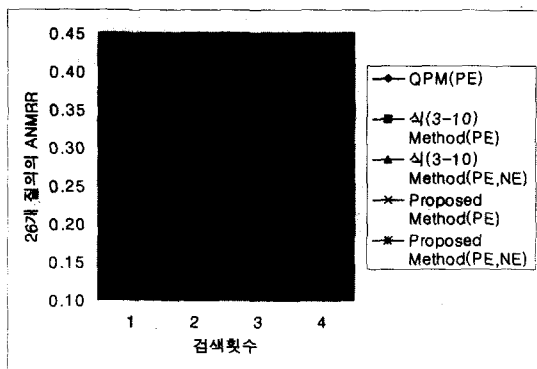


그림 9. 26개 질의영상에 의한 검색결과 ANMRR
Fig. 9. ANMRR of retrieval result by 26 query images.

ANMRR의 변화를 보면, 제안한 방법이 초기검색에서 처음 적합성 피드백을 적용하였을 때의 ANMRR값이, 다른 방법과 비교하여, 가장 낮은 값을 갖는 것을 알 수 있다. 이것은 검색과정을 여러 번 반복하지 않고도 좋은 결과를 얻을 가능성이 높다는 것을 나타내며, 따라서, 제안한 방법의 우수성을 입증한다.

V. 결 론

본 논문에서는 최근에 제시된 부정적 피드백을 효과적으로 다룰 수 있는 확률적 방법과 질의 위치 이동 방법을 융합함으로써 검색효율을 향상시킬 수 있는 방법을 제안하였다. 제안한 방법은 유사도 계산에서, 각각의 부정적 예제와 유사한 영상을 효과적으로 검색결과에서 배제할 수 있는 확률적 방법과 긍정적 예제의 집합의 통계적 특성을 이용하는 질의 위치 이동 방법을 이용하여 각각의 거리를 계산한 후 계산된 두 거리를 가산하여 검색에 이용하였다. 두 방법이 상호 보완적인 면을 가지고 있기 때문에, 두 방법을 융합한 방법이 각각의 방식을 이용한 결과보다 우수한 결과를 나타낸다는 것을 실험결과로 확인할 수 있었다.

참 고 문 헌

- [1] Zhong Su, Hongjiang Zhang, Shaoping Ma, "Using Bayesian Classifier in Relevant Feedback of Image Retrieval," in *Proc IEEE Int. Conf.*, 2000.
- [2] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in *Proc IEEE int. Conf. on Image Proc.*, 1997.
- [3] J.J.Rocchio: "Relevance Feedback in Information Retrieval," In *G.Salton ed, The SMART Retrieval System - Experiments in Automatic Document Processing*, pp. 313~323, Prentice Hall, Englewood Cliffs, N.J., 1971.
- [4] P. Schauble. *Multimedia Information Retrieval: Content-Based Information Retrieval from Large Text and Audio Databases*, Kluwer, Boston, 1997.

- [5] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "Mindreader: Query databases through multiple examples," in *Proc. Of the 24th VLDB Conference*, New York, 1998.
- [6] Roberto Brunelli, Ornella Mich, "Image Retrieval by Examples," *IEEE Transactions on Multimedia*, Vol. 2, No. 3, pp. 164~171, September 2000.
- [7] Y. Rui, T. S. Huang "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," *IEEE Circuits and systems for Video technology*, Vol. 8, No. 5, 1998.
- [8] "Color and Texture Core Experiments," *ISO/IEC/JTCl/SC29/WG11* : document no. N3090, Maui, Dec. 1999.
- [9] "CD 15938-3 Multimedia Content Description Interface - Part 3 Visual," *ISO/IEC JTCl/SC29/WG11/W3703*, La Baula, Oct. 2000.
- [10] D.K. Park, Y.S. Jeon, C.S. Won, S-J Park, "Efficient use of local edge histogram descriptor," *Int. Workshop on Standards, Interoperability and Practices, ACM*, pp. 52~54, Marina del Rey, CA, Nov. 4, 2000.

저 자 소 개

尹 帥 征(正會員)

1999년 동국대학교 학사, 2002년 동국대학교 석사. <주관심분야: 영상 검색>

朴 東 權(正會員)

1996년 동국대학교 학사. 1998년 동국대학교 석사. 1998년 ~ 현재 동국대학교 박사과정. <주관심분야: 영상분할, 영상검색, 비디오 브라우징, 객체추적>

元 致 善(正會員)

1982년 고려대학교 학사. 1986년 University of Massachusetts/Amherst, 석사. 1990년 University of Massachusetts/Amherst 박사. 1989년~1992년 금성사 가전연구소(현, LG전자 멀티미디어 연구소) 선임연구원. 1992년~현재 동국대학교 전자전기공학부 교수. <주관심분야: 영상분할, 영상검색, 비디오 브라우징, 디지털 워터마킹>