

## 적합성 피드백을 이용한 멀티미디어 검색

이팔진\*, 윤보현\*\*, 최은화\*\*\*

\*초당대학교 컴퓨터과학과

\*\*한국전자통신연구원 휴먼정보처리부 지식처리팀

\*\*\*공주대학교 응용수학과

e-mail : [pjlee@chodang.ac.kr](mailto:pjlee@chodang.ac.kr), [ybh@etri.re.kr](mailto:ybh@etri.re.kr), [euchoi@kongju.ac.kr](mailto:euchoi@kongju.ac.kr)

### Multimedia Retrieval using Relevance Feedback

Pal-Jin Lee\*, Bo-Hyun Yun\*\*, Eun-Ha Choi\*\*\*

\*Dept. of Computer Science, Chodang University

\*\*Knowledge Processing Team, Dept. of Human Information Processing, ETRI

\*\*\*Dept. of Applied Mathematics, Kongju National University

#### 요 약

본 문서에서는 사용자 적합성 피드백을 적용한 멀티미디어검색 기법을 제안한다. 적합성 피드백은 멀티미디어검색에 있어 사용자가 요구하는 정보를 반영할 수 있어 영상의 검색 효율을 높일 수 있다. 이 실험에서는 긍정적 피드백과 부정적 피드백을 함께 사용하였다. 실험결과, 적합성 피드백을 이용하면 적은 횟수의 반복검색으로 우수한 결과를 얻을 수 있음을 알 수 있다.

#### 1. 서론

영상은 그 영상의 특성을 표시하는 좀더 낮은 레벨의 특징벡터로 나타내어 질 수 있다. 특징벡터 가운데 비주얼 한 특성을 갖는 벡터들이 있다. 예를 들어, color, shape, texture등이 바로 그것이다. 이러한 특징벡터는 비주얼 한 특징이 있기 때문에 영상의 검색에 쓰이는데 유리한 면을 가지고 있다. 특징벡터를 이용해서 원하는 영상을 찾고자 하는 시도가 90년도 초에 많이 제기되었다. 이듬하여 멀티미디어 검색 시스템(CBIR)이다[1,2,5,6,7].

CBIR은 영상의 비주얼 한 특징벡터가 사람의 개입이 없이도 추출이 가능하며, 또한 항상 같은 특징을 유지하고, 컴퓨터를 이용해서 고속의 처리가 가능하다는 이점을 가지고 있다. 위와 같은 이점으로 인하여 대용량의 영상데이터를 고속으로 탐색해서 원하는 결과를 얻는 것이 가능해졌다. 하지만, CBIR에는 근본적인 두 가지 문제가 있었다. 그 중 한가지는 사람이 영상을 바라볼 때 사용하는

의미론적인 접근 방식과 비주얼 특징벡터사이의 차이점이다.

예를 들어서, 단순한 Color나 Shape을 갖는 영상이 있다고 가정하자. 특징벡터를 이용해서 컴퓨터로 검색된 이러한 영상에 대해서 사람도 같은 의미를 생각할 수 있다. 하지만, 항상 그런 것은 아니다. 오히려 사람이 느끼기에 전혀 다른 의미를 갖는 두 영상이 비주얼 특징벡터에서는 매우 유사한 의미를 갖는 경우가 상당수 있다. 두 번째 문제는 사람에 따른, 혹은 동일한 사람이라도 그 사람의 현재의 입장에 따른 의미해석의 문제이다. 사람은 개개인의 경험, 지식, 성장배경등에 따라서, 동일한 특징벡터에 의해서 검색된 영상에 대해서도 다른 의미를 부여할 수 있다. 또한, 그 사람의 관점이 어떠한 특정한 특징벡터에 더 강하게 집중될 수도 있다. 예를 들면, 같은 영상이라고 해도, 어떤 사람은 Color부분에 중점을 둘 수 있고, 또 다른 사람은 Shape이나 Texture에 중점을 둘 수도 있다.

두 번째 문제에 대한 대안책으로 CBIR에서는 사용자가 특징벡터를 선택하고 그 특징벡터에 대해서 가중치를 주는 방법이 제안되었다. 그러나, 이 방법은 사용되는 시스템을 잘 알지 못하는 일반적인 사용자에게는 적합하지 않으며, 또한, 대개 일반적인 사용자는 가중치를 얼마나 주어야 하는지를 알지 못한다는 단점이 있다. 그래서 영상처리에 도입된 개념이 적합성 피드백이다. 이 장에서는 적합성 피드백의 개념을 살펴보고 제안된 방법을 설명하겠다.

## 2. 적합성 피드백의 개념

적합성 피드백은 사용자가 마치 시스템의 일부처럼 사용되어 시스템과 사용자의 상호작용으로 위의 언급한 두 가지 문제를 해결하고자 하는 방법이다. 사람이 느끼는 의미론적인 방법을 고 수준의 레벨이라고 보고 컴퓨터에 의해서 처리 가능한 비주얼 특징벡터를 저 수준의 레벨이라고 보면, 저 수준의 특징벡터를 이용해서 검색된 정보를 사용자가 고 수준의 레벨에서 선별하고, 이 선별된 정보가 다시 저 수준의 특징벡터의 갱신에 사용되어, 사용자에게 새로운 정보를 제공한다. 이러한 과정의 반복이 적합성 피드백의 기본 개념이다[3,4]. 물론, 여기에는 한가지 가정이 들어 있다. 사용자의 의미론적인 선택을 좀더 낮은 수준의 특징벡터가 반영할 수 있어야 한다는 것이다. 적합성 피드백은 기본적으로 벡터 스페이스 모델을 기본으로 하며 두 가지 주요한 접근방법이 있다. 그 중 하나는 query point movement이고 다른 하나는 re-weighting이다.

사용자에 의해서 feed-back되는 데이터는 크게 두 가지로 구분될 수 있다. 긍정적 예제와 부정적 예제다. 긍정적 예제는 사용자가 판단할 때 자신이 찾고자 하는 영상과 관련이 있는 영상들의 집합을 나타내고 부정적 예제는 관련이 없는 영상들의 집합을 나타낸다. 긍정적 예제를 처리하는 부분을 긍정적 피드백이라고 하고 부정적 예제를 처리하는 부분을 부정적 피드백이라고 한다. 많은 기존의

적합성 피드백 시스템에서는 긍정적 피드백과 부정적 피드백을 내부적으로 같은 과정에 의해서 처리하였다.

대개 위에서 언급한 query point movement나 re-weighting, 혹은 두 가지 방법을 함께 사용해서 처리하였는데, 이 방법들은 관련성을 이용해서 특징벡터를 갱신하는 방법이다. 그런데, 긍정적 예제는 의미론의 관점에서, 또한 특징벡터의 관점에서도 관련성을 가지고 있는 경우가 대부분이며, 벡터공간 내에서 무리를 이루는 경우가 많지만, 부정적 예제는 대개 관련성이 거의 없거나, 심지어는 독립적으로, 혹은 고립되어 단독으로 존재하는 경우가 많다. 따라서, 같은 방법으로 긍정적 예제와 부정적 예제를 다루는 것은 적절하지 못하며 두 방법, 즉, 긍정적 피드백과 부정적 피드백은 각각 독립적인 방법으로 다루어질 필요성이 있다.

## 3. 제안한 적합성 피드백 방법

제안한 방법에서는 에지 히스토그램 디스크립터(Edge Histogram Descriptor)를 이용하였다. 에지 히스토그램 디스크립터는 80개의 로컬빈으로 구성되며, 검색시에는 글로벌과 세미글로벌성분이 로컬빈으로부터 재구성되며 전체적으로 150개의 빈으로 확장되어 이용된다. 위에서 언급한 query point movement방법을 이용하기 위해서 빈들의 평균을 구하고 re-weighting을 위해서 분산을 구한다[3,4]. 식 (3-1)에서  $J^{(k)}$ 는 k번 검색에서 사용자가 feed-back한 긍정적 예제의 개수를 나타내고, I는 에지 히스토그램 디스크립터를 구성하는 빈의 개수를 나타낸다.  $\overline{X_j^{(k)}}(i)$ 는 k번 검색에서 j영상의 i번 빈값을 의미하며,  $\overline{Q^{(k)}}(i)$ 는 이전 검색에서 사용되었던 갱신된 질의 정보를 나타낸다.  $\overline{X_j^{(k)}}(i)$ 와  $\overline{Q^{(k)}}(i)$ 에 의해서 갱신될 k+1번째의 쿼리 특징벡터  $\overline{Q^{(k+1)}}$ 는 다음과 같다.

$$\overline{Q}^{(k+1)}(i) = \frac{1}{J^{(k)}+1} (\overline{Q}^{(k)}(i) + \sum_{j=1}^{J^{(k)}} \overline{X}_j^{(k)}(i)), \quad i=1, \dots, I \quad (3-1)$$

$V^{(k)}(i)$  를 k번 검색에서  $J^{(k)}$  개의 긍정적 예제의 i번째 빈값의 분산을 의미한다고 하면 다음의 식에 의해 계산할 수 있다.

$$V^{(k)}(i) = \frac{1}{J^{(k)}} \sum_{j=1}^{J^{(k)}} (\overline{X}_j^{(k)}(i) - \overline{M}^{(k)}(i))^2, \quad i=1, \dots, I \quad (3-2)$$

여기서,

$$\overline{M}^{(k)}(i) = \frac{1}{J^{(k)}+1} (\overline{Q}^{(k)}(i) + \sum_{j=1}^{J^{(k)}} \overline{X}_j^{(k)}(i)), \quad i=1, \dots, I$$

$V_{\max}^{(k)}$  는 I개의 분산값에서 최대치를 나타내며,  $V_{\min}^{(k)}$  은 I개의 분산값에서 최소치를 나타낸다.

$$V_{\max}^{(k)} = \max\{V^{(k)}(i)\} \quad i=1, \dots, I \quad (3-3)$$

$$V_{\min}^{(k)} = \min\{V^{(k)}(i)\} \quad i=1, \dots, I \quad (3-4)$$

가중치  $w^{(k+1)}(i)$  는 식 (3-5)로 계산되며, 결과는  $[0, 1]$ 로 정규화된 값을 갖게 된다.

$$w^{(k+1)}(i) = 1 - \frac{V^{(k)}(i) - V_{\min}^{(k)}}{V_{\max}^{(k)} - V_{\min}^{(k)}}, \quad i=1, \dots, I \quad (3-5)$$

유사도는 식 (3-1)의  $\overline{Q}^{(k+1)}(i)$  와 식 (3-5)에서

계산된  $w^{(k+1)}(i)$  를 이용해서 식 (3-6)으로 계산된다.  $z(i)$ 는 데이터베이스내의 저장된 각각의 영상정보의 특징벡터에서 i번째 빈값을 의미한다.

$$d(\overline{Q}^{(k+1)}, \overline{Z}) = \frac{\sum_{i=1}^I w^{(k+1)}(i) |q^{(k+1)}(i) - z(i)|}{\sum_{i=1}^I w^{(k+1)}(i)} \quad (3-6)$$

긍정적 예제와 부정적 예제는 각각 다른 방법으로 다루어지는 것이 바람직하다. 그래서, 제안한 방법에서는 부정적 예제로 Black-list를 작성한다. 작성된 Black-list를 이전의 Black-list와 병합한 후 긍정적 예제를 이용해서 검색된 결과로부터 Black-list와 일치하는 영상을 제거하고 최종적인 결과를 출력하는 방식을 사용하였다. 이 방법은 특히, 긍정적 예제의 개수가 적을 때 효과적인 성능을 나타내었다.

#### 4. 실험결과

제안한 방법의 성능측정을 위해서 자연영상으로 테스트하였다. 자연영상의 개수는 11639를 사용하였고, 자연영상은 Ground truth(GT)을 이용하여 테스트하였다. 실험에 사용한 프로그램에서 중간아래에 보이는 영상이 질의 영상이고, 좌측위부터 우측아래 방향으로 검색순위가 정렬된 결과영상이 표시된다.

실험방법은 질의영상을 이용해서 검색된 결과내에서 GT에 속하는 영상의 개수를 구하고, 그 개수를 GT의 개수로 나누어서 평가하였다. 다시말해서, 찾아야 하는 영상중에 얼마나 찾았는가를 알수있는 Recall(재현율)을 구해서 평가의 기준으로 채택하였다. 질의 영상으로 자연영상을 이용한 결과는 와 같다. 초기검색은 질의 영상을 이용하여 검색한 결과를 그대로 사용하였고,

초기검색결과로부터 사용자가 긍정적 예제만을 선택하고 적합성 피드백을 1회 실행했을때의(즉, 긍정적 피드백만 이용) 결과가 긍정적 피드백 에 나타나 있다. 그리고, 긍정적 예제와 부정적 예제 모두를 사용해서 적합성 피드백을 1회 실행했을때의 결과가 Positive & 부정적 피드백에 나타나 있다.

실험결과, Feedback 방법의 사용이 초기검색된 결과보다 우수하다는 것을 알 수 있다. 또한 긍정적 피드백과 부정적 피드백을 함께 이용한 방법이 긍정적 피드백만을 이용한 방법보다 우수하다는 것을 알았다.

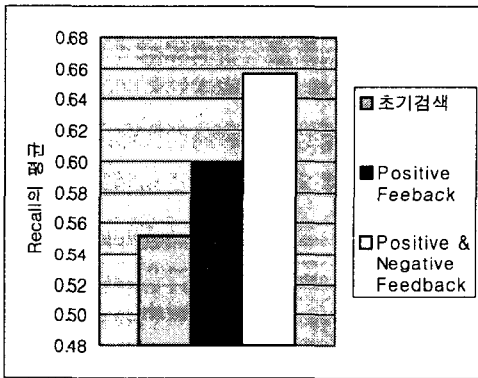


그림 1. 각각의 방식에 따른 지연영상검색결과와 평균

그림1에서 세가지 방법의 실험결과를 보인다. 역시 긍정적 피드백과 부정적 피드백을 함께 사용한 방법이 우수한 결과를 나타냄을 확인할 수 있다.

다음으로, 반복적인 적합성 피드백에 의한 결과를 살펴보겠다. 이 실험에서는 긍정적 피드백과 부정적 피드백을 함께 사용하였다. K는 적합성 피드백을 적용한 반복횟수를 나타낸다. 결과영상들로부터 알 수 있듯이 적합성 피드백을 이용하면 적은 횟수의 반복검색으로 우수한 결과를 얻을 수 있음을 알 수 있다.

## 5. 결론

본 문서에서는 사용자 적합성 피드백을 적용한 멀티미디어검색 기법을 제안한다. 적합성 피드백은 멀티미디어검색에 있어 사용자가 요구하는 정보를 반영할 수 있어 영상의 검색 효율을 높일 수 있다. 이 실험에서는 긍정적 피드백과 부정적 피드백을 함께 사용하였다. 실험결과, 적합성 피드백을 이용하면 적은 횟수의 반복검색으로 우수한 결과를 얻을 수 있음을 알 수 있었다.

## 참고문헌

- [1] ISO/IEC/JTC1/SC29/WG11 : "Overview of the MPEG-7 Standard," MPEG document N3752, La Baule, Oct. 2000.
- [2] ISO/IEC JTC1/SC29/WG11 : "CD 15938-3 MPEG-7 Multimedia Content Description Interface - Part 3 Visual", MPEG document N3703, La Baule, Oct. 2000.
- [3] Eugenio Di Sciascio and Marina Mongiello, "Query by Sketch and 적합성 피드백 for Content-Based Image Retrieval over the Web," Journal of Visual Language and Computing, No. 10, pp. 565-584, 1999.
- [4] Y. Rui, T. S. Huang, and S. Mehrotra, "Content-based image retrieval with 적합성 피드백 in MARS," Proc. IEEE Int. Conf. On Image Proc., Oct. 1997.
- [5] Y. Ishikawa, R. Subramanya, and C. Faloutsos, "Mindreader: Query databases through multiple example," in Proc. Of the 24th VLDB Conference, pp. 218-227, 1998.
- [6] Dong Kwon Park, Yoon Seok Jeon, Chee Sun Won and Soo-Jun Park, "Efficient Use of Local Edge Histogram Descriptor," Proceedings ACM Multimedia 2000 Workshops, pp. 51-54, 2000.
- [7] ISO/IEC/JTC1/SC29/WG11 : "Core Experiment Results for Spatial Intensity Descriptor (CT4)," MPEG document M5374, Maui, Dec. 1999.