

영상에 대한 Semantics 축적이 가능한 Relevance Feedback

오상욱[†], 정민교^{**}, 설상훈^{***}

요 약

Relevance Feedback(RF)은 사용자의 인지적 피드백(perceptual feedback)을 사용하는 영상 검색 기법 중의 하나로서, 사용자 피드백을 통해 얻게 되는 적합성 정보(relevance information)를 이용하여 사용자 질의(query)를 점진적으로 구체화하게 된다. 그러나, 기존 RF 기법에서는 이러한 적합성 정보가 매우 유용한 정보임에도 불구하고, 검색이 끝나는 순간 없애버리고 만다. 그래서, 본 논문에서는 사용자의 인지적 피드백 정보를 버리지 않고, 저장하는 새로운 개념의 RF를 제안한다. 새로 제안된 RF는 시간의 흐름에 따라 축적되어 저장된 상위 레벨의 적합성 정보(high-level relevance information)를 하위 레벨의 특징벡터(low-level feature vectors)와 동적으로(dynamically) 결합하여 사용함으로써, 검색의 효율성을 크게 향상시킨다. 제안 방법의 우수성을 입증하기 위해 다양한 실험 결과도 제시한다.

Semantics Accumulation-Enabled Relevance Feedback

Sangwook Oh[†], Min Gyo Chung^{**}, Sanghoon Sull^{***}

ABSTRACT

Relevance Feedback(RF), a method to use perceptual feedback in image retrieval, refines a query by the relevance information from a user. However, the user's feedback information is thrown away as soon as a search session ends. So, this paper proposes an enhanced version of RF, which is designed to accumulate human perceptual responses over time through relevance feedback and to dynamically combine the accumulated high-level relevance information with low-level features to further improve the retrieval effectiveness. Experimental results are presented to prove the potential of the proposed RF.

Key words: Relevance Feedback(적합성 피드백), Semantics Accumulation(시맨틱 축적), CBR(내용기반 검색)

1. 서 론

하위 레벨 특징벡터(low-level feature vectors)를 기반으로 하는 영상 검색 시스템은 몇 가지 문제점들을 가지고 있다. 예를 들어, 고차원적인 인간의 인지적 개념을 하위 레벨 특징벡터만을 사용하여 나타내

는 것 자체가 어려울 뿐만 아니라, 심지어 하위 레벨 특징벡터 값들도 영상의 모양이나 크기, 각도, 색깔의 미묘한 차이에 민감하게 반응하기 때문에 유사한 영상일지라도 다르게 인식되는 문제가 발생하게 된다. 이와 같은 문제점들을 극복하고자 다양하고도 많은 기법들이 제안되었으며, 그 중에서도 Relevance

※ 교신저자(Corresponding Author) : 정민교, 주소 : 서울시 노원구 공릉 2동 126번지(139-774), 전화 : (02)970-5753, FAX : (02)970-5981, E-mail : mchung@swu.ac.kr
접수일 : 2005년 3월 2일, 완료일 : 2005년 6월 10일

[†] 고려대학교 전자컴퓨터공학과 대학원
(E-mail : osu@mpeg.korea.ac.kr)

^{**} 정회원, 서울여자대학교 교수

^{***} 고려대학교 전자컴퓨터공학과 교수
(E-mail : sull@mpeg.korea.ac.kr)

※ 본 연구는 2004학년도 서울여자대학교 교내학술연구비에 의하여 수행되었음.

Feedback(RF)은 고차원적인 인간의 인지적 개념과 하위 레벨 특징벡터를 통합함으로써 영상 검색 시스템의 성능을 향상시키는 기법이다[1-3]. 즉, RF 기법은 질의 영상과 결과 영상의 유사성 정도를 사용자가 직접 입력하게 하고, 이렇게 입력 과정에서 얻어진 사용자의 인지적 개념을 이용하여 하위레벨 특징 벡터에 대한 가중치를 계속적으로 변화시키면서 영상 검색을 수행하게 된다.

기존의 RF는 사용자와 영상 검색 시스템이 상호 작용(interaction) 함으로써 영상 검색 결과의 정확성을 높이는 것이었지만, 사용자 피드백(feedback)으로부터 수집된 중요한 적합성 정보(relevance information)를 검색 종료 후 폐기하는 단점을 갖고 있었다. 그러나, 이렇게 상호작용을 통하여 얻어진 적합성 정보에는 영상에 대한 사용자의 인지적 개념(즉, semantics)이 포함되어 있기 때문에, 영상 자체의 내용을 표현하는 데 있어서 하위 레벨 특징벡터보다 더 직관적이며 훨씬 활용도가 높다는 것을 발견할 수 있었다.

따라서, 본 논문에서는 영상에 대한 semantics를 포함하고 있는 적합성 피드백(relevance feedback) 정보를 버리지 않고 저장하여 재사용함으로써, 영상 검색 시스템의 검색 성능을 향상시키는 새로운 RF 기법을 제안한다. 이렇게 하기 위하여, 제안된 RF 기법에서는 시간의 경과에 따라서 축적된 사용자 피드백 정보를 바탕으로, 영상들 간의 개념적 유사 관계를 표현하는 semantic space라는 새로운 상위 레벨 특징(high level feature)을 도입하여 사용한다. 결국, 새로 제안된 RF는 이렇게 semantic space에 저장된 상위 레벨 특징정보를 하위 레벨 특징벡터와 동적으로(dynamically) 결합하여 사용함으로써 검색의 효율성을 크게 향상시키게 된다.

다양한 실험 결과들은 본 논문에서 제안하고 있는 RF 기법이 기존 RF 기법보다 영상 검색 성능에서 좋은 결과를 나타내고 있음을 보여주고 있으며, 사용자의 인지적 개념을 나타내는 정보가 누적될수록 시스템의 성능이 더욱 향상되고 있음을 보여주고 있다.

본 논문의 구성은 다음과 같다. 제 2절에서는 본 논문에서 사용하는 semantic space에 대한 기본 개념 및 구성 방법을 간단히 설명하고, 상위 레벨 semantics 정보를 하위 레벨 특징벡터와 동적으로 통합하는 방법을 자세히 설명한다. 제 3절은 제안된 방법의 성능을 나타내는 다양한 실험 결과를 포함하

고 있으며, 결과 분석을 통하여 제안된 방법의 우수성을 보여주고 있다. 제 4절에서는 본 논문에 대한 결론을 짧막하게 맺고 있다.

2. 새롭게 제안하는 RF 기법

2.1 Semantic Space

기존의 RF 기법에서는 영상 검색 중에 사용자에게 의하여 생성된 적합성 피드백(relevance feedback) 정보가 생성되는 즉시 소멸되어 없어지지만, 제안된 RF에서는 이런 피드백 정보를 버리지 않고 semantic space라고 불리는 곳에 저장하게 된다. 개념적으로 간단히 말하면, semantic space는 영상과 영상이 내용적으로 얼마나 유사한지를 나타내는 일종의 특징 공간(feature space)이라고 할 수 있다. 실제로는, 하나의 영상 데이터베이스에 대한 semantic space는 $n \times n$ 행렬 $M=(m_{ij})$ 로 나타낼 수 있다. 여기서 n 은 데이터베이스가 가지고 있는 전체 영상의 숫자를 나타내고, m_{ij} 는 현재까지 누적된 질의 영상 i 와 데이터 베이스내의 임의 영상 j 사이의 적합성(relevance) 정도를 나타내는 값이다. 영상 i 와 영상 j 가 내용적 또는 개념적으로 유사하면 유사할수록 m_{ij} 값은 커지게 된다. 앞으로 위의 행렬 $M=(m_{ij})$ 을 본 논문에서는 semantic matrix라고 칭한다.

그림 1은 5개 영상으로 구성된 영상 데이터베이스에 대한 semantic matrix $M=(m_{ij})$ 의 예를 보여 주고 있다. 여기서, 모든 영상 i, j 에 대하여 m_{ij} 의 초기값은 0 값을 갖는다. 그러나, 영상 검색이 수행됨에 따라 사용자 피드백 정보가 점차 누적되어, m_{ij} 는 영상 i, j 간의 적합성 값(relevance values)으로 빠르게 채워지게 된다. 주어진 질의 영상 i 에 대하여, 결과 영상

	영상1	영상2	영상3	영상4	영상5
영상1	0	1	2	1	1
영상2	3	0	1	0	2
영상3	3	1	0	2	4
영상4	1	2	2	0	5
영상5	0	1	2	3	0

그림 1. 5개 영상으로 구성된 영상 데이터베이스에 대한 semantic matrix M 의 예

중 어느 한 영상 j 가 “relevant”하다고 사용자가 마크(mark)하면, m_{ij} 는 시스템이 정한 일정한 값만큼 증가하게 된다. 예를 들어, 그림 1에서 영상 4가 질의 영상이고 영상 5가 결과 영상 중 “relevant”한 영상이라고 마크된 경우에는, m_{45} 의 값이 5에서 $5+\alpha$ 로 변하게 된다. 여기서 α 는 적합성 값의 증액분으로서, α 의 실제 값을 결정하는 방법에는 여러 가지가 존재할 수 있다. 그러나, 본 논문에서는 “relevant”로 마크된 경우에는 $\alpha=1$, 그렇지 않은 경우에는 $\alpha=0$ 과 같은 간단한 규칙을 사용하였다.

Semantic matrix M 을 유지하고 갱신하는 방법에는 비대칭(asymmetric) 갱신방법과 대칭(symmetric) 갱신방법이 있을 수 있다. 비대칭 갱신방법에서는 사용자가 질의 영상 i 에 대해 영상 j 를 “relevant”하다고 선언하면, 행렬 M 의 m_{ij} 값만 새로운 값으로 갱신되고, m_{ji} 값은 변하지 않고 그대로 있게 된다. 반면에, 대칭 갱신방법은 위와 같은 경우 m_{ij} 와 m_{ji} 를 동시에 갱신하게 된다. 이 두 가지 갱신방법에 대한 효율성이나 특성에 대해서는 더 많은 연구가 수행되어야 하겠지만, 비대칭 갱신방법은 영상 i 와 영상 j 중 어느 영상을 질의 영상으로 사용하느냐에 따라서 상이한 검색 결과를 나타내는 경향이 있으나, 대칭 갱신방법은 영상 i 와 영상 j 중 어느 영상이 질의 영상으로 사용되는 상관없이 동일한 검색 결과를 나타낸다. 비대칭 갱신방법이 대칭 갱신방법보다 더 일반적이고, 인간의 인지적 개념에서 미세하게 차이가 나는 영상들을 보다 더 잘 차별화할 수 있기 때문에, 본 논문에서는 비대칭 갱신 방법을 사용하였다.

2.2 상위 레벨 semantics 정보와 하위 레벨 특징벡터의 결합

본 절에서는 semantic space로 표현되는 영상의 상위레벨 정보를 영상의 하위레벨 특징벡터와 RF 프레임워크 안에서 효과적으로 결합하는 방법에 대하여 자세히 설명한다. 설명의 편의를 위해, 영상을 두 개의 하위 레벨 특징(색깔과 텍스처)과 하나의 상위레벨 특징(2.1 절에서 언급한 semantic space)으로 표현할 수 있다고 가정하자. 실제로, 본 논문에서는 색깔은 color correlogram[4]을, 텍스처는 Shim & Choi[5]가 제안한 방법을 사용하여 나타내었다. (그러나, 본 논문에서 제안한 상위레벨 특징은 특정 하위 레벨 특징과 연관되어 있는 것이 아니므로, 위에서 사용한 색깔이나 텍스처 특징 외에 다른 형태의

하위 레벨 특징을 사용해도 검색 성능 향상이라는 동일한 효과를 가져 올 수 있습니다.)

그리고, 다음과 같은 기호와 정의를 아래 설명에서 사용하기로 한다.

- $S_C(i,j)$, $S_T(i,j)$, $S_S(i,j)$ 는 각각 영상 i 와 영상 j 간의 색상(color), 텍스처(texture) 및 semantics 특징에 대한 유사도를 나타낸다.
- 그리고, W_C , W_T , W_S 는 각각 색상, 텍스처 및 semantics 특징에 대한 가중치를 나타낸다.

2.2.1 유사도 계산

질의 영상 i 와 임의의 영상 j 에 대한 전체 유사도 $S(i,j)$ 는 위의 기호와 정의를 이용하여 다음과 같이 나타낼 수 있다.

$$S(i,j) = W_C \frac{S_C(i,j)}{\max_j S_C(i,j)} + W_T \frac{S_T(i,j)}{\max_j S_T(i,j)} + W_S \frac{S_S(i,j)}{\max_j S_S(i,j)}$$

여기서, $S_S(i,j)$ 의 값으로는 semantic matrix M 에서의 m_{ij} 를 사용한다. 위 식에서 $\max_j S_C(i,j)$, $\max_j S_T(i,j)$, $\max_j S_S(i,j)$ 는 각각 질의 영상 i 에 대한 색깔, 텍스처, semantics 특징벡터의 최대 유사도 값을 나타내며, 각 특징의 유사도 값을 정규화 할 목적으로 사용된다.

2.2.2 가중치 갱신

본 논문에서 제시한 RF 기반의 영상 검색 시스템은 사용자가 선택한 질의 영상에 대해서, 내부 검색 알고리즘을 수행하여 그 결과를 사용자에게 제시하게 된다. 그러면, 사용자는 결과 영상을 보고 어떤 영상들이 질의 영상과 “relevant”한지 또는 그렇지 않은지를 판단하게 된다. 이러한 방법을 통해 얻어진 적합성 정보(relevance information)는 영상의 하위레벨 특징의 가중치와 상위 레벨 특징(즉, semantics 특징)의 가중치를 동적으로 갱신하는데 사용된다. 즉, 각 가중치는 현재 사용 중인 특징들의 상대적인 중요도에 비례해서 적절하게 변경되도록 해야 한다.

k 를 사용자가 검색 결과로써 표시하고자 하는 영상의 개수라고 하고, 집합 R 은 전체 유사도 $S(i,j)$ 에 따른 k 개의 가장 유사한 영상 집합이라고 정의하자. 집합 R_C , R_T 및 R_S 도 비슷한 방법으로 정의한다. 즉,

집합 R_f 는 유사도 $S_f(i, j)$ 에 따른 k 개의 가장 유사한 영상 집합으로 정의할 수 있으며, 여기서 f 는 C, T, S 중 어느 하나의 특징을 나타낸다. 그러면, 이제 특징 f 에 대한 새로운 가중치는 아래 알고리즘(Rui[1]에서 제시한 방법과 유사)을 사용하여 동적으로 계산할 수 있다.

1. $W_{sum} = 0$
2. $[C, T, S]$ 에 속한 각 특징 f 마다 아래 과정을 차례로 수행한다.
 - ① $w_f = 0$ 로 초기화한다.
 - ② 집합 R_f 에 속한 각 영상 P 에 대해서, P 가 집합 R 에도 속하면 $W_f = W_f + \alpha$ 와 같이 변경한다. (여기서, α 값은 “relevant”로 마크된 경우에는 1, 그렇지 않은 경우에는 0 값을 각각 갖는다.)
 - ③ $W_{sum} = W_{sum} + W_f$
3. 위 단계 2에서 얻어진 각 가중치를 전체 가중치인 W_{sum} 를 사용하여 정규화 한다. 즉, $W_f = \frac{W_f}{W_{sum}}$.

위 알고리즘을 자세히 살펴보면, 집합 R_f 와 R 의 겹치는 부분이 커질수록, 가중치 W_f 의 값도 커진다는 것을 알 수 있다. 이렇게 갱신된 가중치들은 질의 이미지와 인지적으로 유사한 영상들을 쉽고 빠르게 찾아내는데 사용된다.

본 논문에서 제안된 RF 기반 영상 검색 시스템에서는 검색이 진행됨에 따라, 가중치는 사용자의 필요나 의도에 따라서 동적으로 변하게 된다. 사용자는 질의 영상과 인지적으로 유사한 영상들만을 선택하여 시스템에게 피드백해주기 때문에, 검색 과정이 진행됨에 따라 $S_f(i, j)$ 유사도에 따라 결정되는 영상이 더욱 많이 나타나는 경향을 갖게 된다. 결과적으로, 색상 또는 텍스처 특징 벡터에 대한 가중치는 점점 작아지고, semantics 특징에 대한 가중치는 점점 커지게 된다. 이런 현상은 semantics 특징이 관련 유사 영상을 빠르게 찾는데 매우 중요한 역할을 한다는 것을 의미한다. Semantics 특징이 보여주는 이러한 장점으로 인해, 결과 영상에서 나타나는 false positive 비율이 현저하게 줄어들게 된다.

3. 실험 환경 및 결과

3.1 실험 환경

제안된 방법의 효율성을 입증하기 위해, 그림 2에

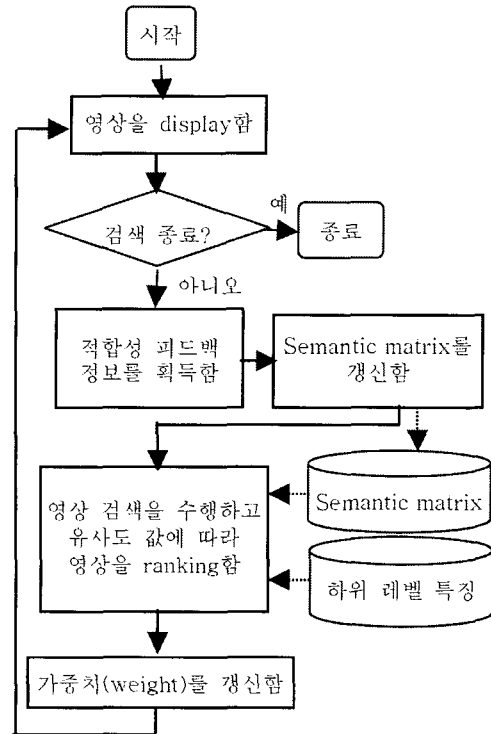


그림 2. 제안된 영상 검색 시스템의 순서도

나타낸 것과 같은 영상 검색 시스템을 구현하였다. 색상, 텍스처 및 semantics 특징들에 대한 초기 가중치는 모두 1/3로 하였다. 또한, semantic matrix M 을 갱신하는 방법으로는 비대칭 갱신기법을 사용하였다. 왜냐하면 비대칭 갱신방법이 대칭 갱신방법보다 더 일반적이고, 인간의 인지적 개념에서 미세하게 차이가 나는 영상들을 보다 더 잘 차별화할 수 있기 때문이다.

실험에서 사용한 이미지 데이터베이스는 2700개의 일반 자연 영상들(natural images)로 구성되었다 (이것은 semantic matrix M 의 차원이 2700×2700 이라는 것을 의미한다). 표 1은 실험에 사용한 영상의 종류를 보여주고 있는데, 그 내용에 따라서 인물, 영화, 동식물, 스포츠 등과 같은 다양한 그룹으로 나뉘어져 있었다. 즉, 전체 이미지 데이터 베이스는 기본 그룹 10개와 각 기본그룹을 조금 더 상세하게 세분한 하위 그룹으로 구성하였는데, 하위그룹은 기본 그룹당 평균 5개 정도를 유지하였다. 각 영상은 표현하고 있는 특징이 확인한 경우 해당 하위 그룹으로 분류하였지만, 그렇지 못한 경우 하위 그룹 중 하나인 기타

표 1. 실험에 사용된 영상 데이터베이스 구성표

그룹 번호	기본 그룹명	하위 그룹명	그룹 번호	기본 그룹명	하위 그룹명
1	사람, 인물	국내연예인, 외국연예인, 스포츠, 정치, 화제의인물, 기타	6	스포츠	야구, 농구, 축구, 골프, 스키, 기타
2	영화, 비디오, 드라마	국내영화, 외국영화, 국내드라마, 외국드라마, 기타	7	자연경치	하늘, 산, 바다, 별자리, 일출/일몰, 기타
3	동식물	강아지, 고양이, 새, 꽃, 나무, 기타	8	문화	회화, 디자인, 건축, 연극, 기타
4	애니메이션	국내만화, 외국만화, 캐릭터, 코믹, 기타	9	생활	야외, 실내, 축제, 결혼, 기타
5	운송수단	자동차, 비행기, 오토바이, 선박, 기타	10	기타	

그룹으로 분류하였다.

인간(유명 인사, 연예인 등), 운송 수단(차, 비행기, 오토바이 등), 동물, 식물, 별자리, 자연 경치(일출, 일몰, 구름, 번개 등), 스포츠, 만화영화 캐릭터 등과 같은 다양한 그룹으로 나뉘어져 있었다. 그리고, semantic matrix M 을 의미 있는 값으로 채우기 위하여, 일정기간 동안 불특정 다수가 본 논문에서 제안한 새로운 영상 검색 시스템을 자유롭게 사용하도록 하였다. 위 과정을 통해 얻은 초기 행렬 M 은 다음과 같은 값의 분포를 갖고 있었다.

- 각 질의영상에 대하여 검색된 결과 영상들 중, 질의영상과 “relevant” 한 것으로 마크된 것들의 총 개수 (즉, $\sum_{i,j} m_{ij}$)는 17823 이었다.
- 행렬 M 에서 행벡터(row vector)가 영벡터(zero vector)가 아닌 행벡터의 수는 271개 이었다. 따라서 한 개의 행벡터당 사용자들로부터 평균 $17823/271=66$ 개의 “relevant” 피드백을 받았다.

이미지 데이터베이스의 크기가 커질수록, semantic matrix M 은 많은 저장 공간을 요구하게 된다. 그러나, 실제로는 행렬 M 의 값들이 대부분 0 값을 갖거나 비어있는 sparse 행렬 형태를 갖고 있기 때문에, 메모리를 적게 사용하면서도 sparse 행렬을 효율적으로 나타내는 기존의 다양한 방법을 사용하면 영상 데이터베이스의 크기 변화에 따른 semantic matrix의 표현이나 관리에는 별다른 문제가 없게 된다.

3.2 실험 결과

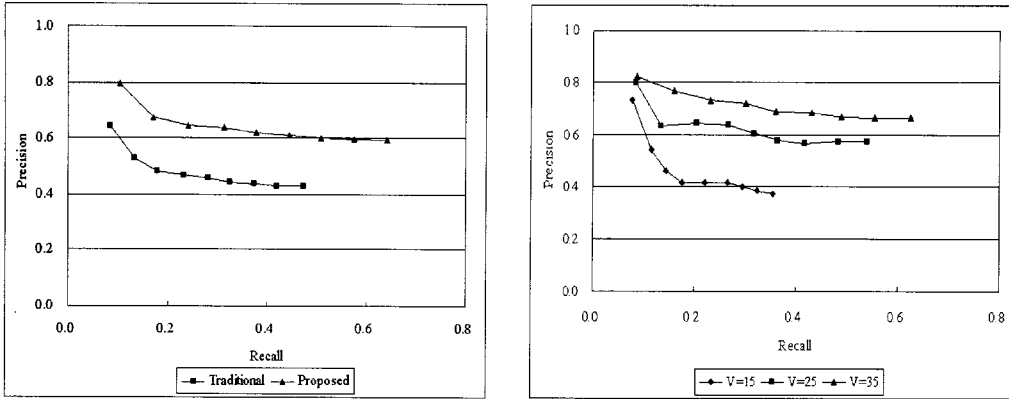
그림 3(a)는 기존의 RF 방법과 제안된 RF 방법의 precision-recall 그래프를 나타낸 것으로, 그래프 내에 표시된 9개의 점들은 결과 영상의 수를 9, 18, 27, ..., 81과 같이 변화시켜 가면서 얻은 것이다. 여기서, 기

존의 방법은 두 개의 하위 레벨 특징 벡터(색깔 및 텍스처)만을 사용했고, 제안된 방법은 두 개의 하위 레벨 특징벡터 뿐만 아니라 상위 레벨 특징인 semantic matrix도 같이 사용하였다. 아래 precision-recall 그래프는 20개의 임의 질의 영상에 대한 검색 결과를 평균하여 작성하였다. 특히, 제안된 방법에 대한 precision-recall 그래프는 3.1절에서 설명한 초기 semantic matrix M 을 사용하여 작성되었다. 결론적으로, 제안된 기법이 기존의 기법보다 약 20% 정도 향상된 검색 성능을 보여주고 있음을 알 수 있었다.

그림 3(b)는 semantic matrix를 사용자가 피드백해주는 적합성 정보들(relevance information)로 점점 채워 나갈 때, 제안된 RF 기법의 precision-recall 그래프가 어떻게 변화하는지를 보여주고 있다. 범례에 있는 기호 V 는 semantic matrix의 하나의 행벡터(row vector)에 포함된 적합성 정보의 평균량을 나타낸다. V 의 값이 커질수록 semantic matrix에 누적된 적합성 정보가 많다는 것을 뜻한다. 예상대로 결과 그래프는 V 의 값이 커질수록 제안된 검색 시스템의 성능이 향상되고 있음을 보여주고 있다.

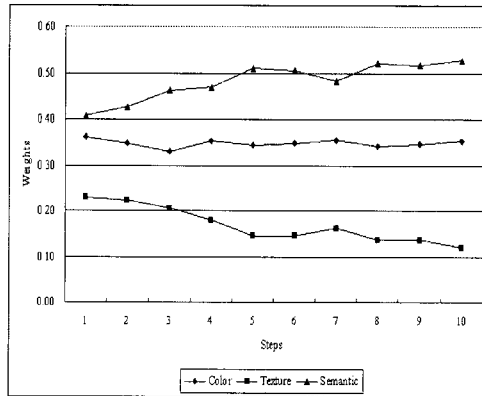
그림 3(c)는 하나의 검색 세션(session)에서 검색이 진행됨에 따라 3개의 특징 벡터에 대한 가중치가 어떻게 변화하는지를 보여주는 그래프이다. 검색의 진행에 따라, semantics 특징에 대한 가중치는 증가하는 경향을 보이지만, 색깔 또는 텍스처 특징에 대한 가중치는 감소하거나 거의 변화가 없는 경향을 보여주고 있다. 이런 결과는 semantics 특징이 사용자가 유사 영상을 빠르게 찾는데 크게 기여하고 있음을 단적으로 보여주는 것이다.

표 2는 기존의 RF 기법과 제안한 RF 기법을 사용하였을 때의 각 그룹별 성능을 비교한 데이터이다.



(a)

(b)



(c)

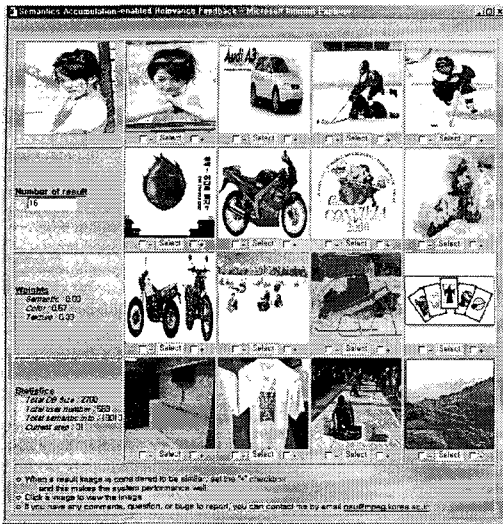
그림 3. (a) 기존 RF 방법과 제안된 RF 방법의 precision-recall 그래프, (b) 사용자 피드백 정보가 semantic matrix에 점진적으로 누적되어갈 때의 precision-recall의 변화 그래프, (c) 하나의 검색 세션 중, 시간 경과에 따라서 각 특징에 대한 가중치가 어떻게 변화하는지를 보여주는 그래프.

표 2. 각 그룹별 기존의 방법과 제안된 방법의 성능 비교표 (%)

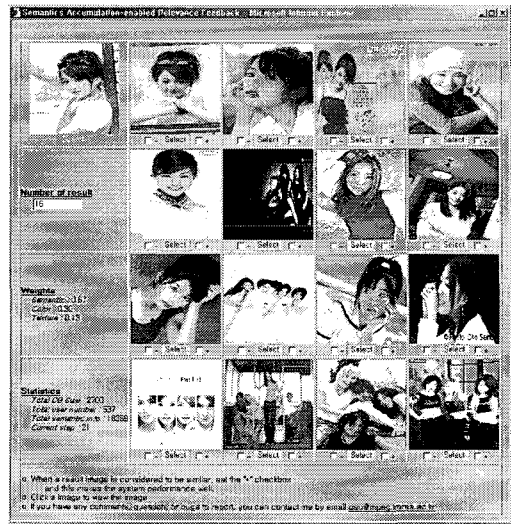
번호	기존 기법		제안된 기법		성능 대비		번호	기존 기법		제안된 기법		성능 대비	
	A1	B1	A2	B2	$(A2-A1)/A1$	$(B2-B1)/B1$		A1	B1	A2	B2	$(A2-A1)/A1$	$(B2-B1)/B1$
그룹1	56	27	89	64	59	66	그룹6	67	43	92	65	45	39
그룹2	48	24	75	51	48	48	그룹7	61	47	64	49	5	4
그룹3	65	41	83	53	32	21	그룹8	49	22	54	25	9	5
그룹4	50	28	63	34	23	11	그룹9	46	19	52	23	11	7
그룹5	62	32	69	37	13	9	평균	56	31	71	44	27	22

검색된 결과에서 해당 그룹에 속하는 영상의 개수를 비교하였다. A는 해당하는 기본 그룹 내 영상을 유사 영상으로 인정한 경우이며, B는 해당하는 하위 그룹 내 영상만을 유사 영상으로 인정한 경우의 검색 성능을 나타낸다. 결과를 보면, 그룹 1은 인물에 관련된

그룹으로 다양한 색깔과 형태의 인물 영상이 존재함으로써, 기존의 RF 기법에서 사용하는 하위 레벨 특징만을 이용하여 검색할 경우 높지 않는 검색 성능을 보이지만, 제안된 RF 기법에서 사용하는 semantic 특징 정보를 사용할 경우 semantic 특징 정보가 축적



(a)



(b)

그림 4. (a) 기존 RF 기법을 이용하였을 때의 검색 결과 예, (b) 제안된 RF 기법을 이용하였을 때의 검색 결과 예.

됨에 따라 높은 성능 향상을 나타내고 있음을 알 수 있다. 그룹 1(인물), 그룹 2(영화) 또는 그룹 6(스포츠) 등과 같이 사용자의 관심이 많은 분류일수록, 제안된 기법이 검색 시스템의 성능 향상에 효과적임을 알 수 있다. 이와 같은 결과는 영상을 바라보는 사용자의 인지적 개념이 일정 부분 통일될 수 있음을 나타내고 있다. 그룹 8(문화)와 같은 경우, 사용자로부터 semantic 특징 정보를 많이 획득하지 못함으로 인하여 성능이 크게 달라지지 않음을 볼 수 있다. 이것은 semantic 특징 정보의 양이 일정량 이상으로 축적되어야 함을 나타내고 있다.

그림 4(a)는 기존 RF 기법만을 이용하였을 때의 검색 결과 예이며, 그림 4(b)는 제안된 RF 기법을 이용하였을 때의 검색 결과 예이다. 화면의 좌측 상단이 질의 이미지이고, 나머지 16개의 이미지가 질의 이미지에 대한 검색 결과 이미지이다. 하위레벨 특징만을 이용하는 기존 RF 방법은 질의 이미지인 연예인과 관계없는 이미지들도 검색 결과로서 나타나지만, 제안한 기법에서는 질의 이미지인 연예인과 관련된 이미지들이 주로 검색 결과로써 나타나게 된다.

4. 결 론

Relevance Feedback(RF)은 사용자의 인지적 피드백(perceptual feedback)을 사용하는 영상 검색 기법 중의 하나이다. 비록 RF가 사용자와의 상호작용

(interaction)을 통한 영상 검색이라는 독특한 개념에 기반을 두고 있지만, 사용자 피드백을 통해 얻게 되는 적합성 정보(relevance information)를 무시하고 사용하지 않는 단점을 갖고 있었다.

그래서, 본 논문에서는 사용자의 인지적 피드백 정보를 저장하여 효과적으로 재사용하는 기능을 갖춘 새로운 개념의 RF를 제안하였다. 다시 말하면, 제안된 RF 기법은 사용자의 피드백을 통해 누적된 적합성 정보를 이용하여 semantic space라고 하는 개념을 도입하였고, semantic space에 저장된 상위 레벨 적합성 정보(high-level relevance information)를 영상의 하위 레벨 특징벡터(low-level feature vectors)와 동적으로 결합하여 사용함으로써 검색의 효율성을 크게 향상시켰다.

실험 결과를 통하여, 제안된 기법의 검색 성능이 기존 방법에 비해 약 20% 정도 크게 향상되었다는 것을 알 수 있었고, 특히, 시간이 지남에 따라 더욱더 그 성능이 좋아짐을 발견하였다. 또한, semantics 특징은 사용자가 유사 영상을 빠르게 찾게 해주거나, 검색결과 of false positive 비율을 낮추는 데 매우 중요한 역할을 하고 있음도 알 수 있었다.

참 고 문 헌

[1] Yong Rui, Thomas S. Huang, Michael Ortega, and Sharad Mehrotra, "Relevance Feedback:

A Power Tool in Interactive Content-Based Image Retrieval," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 8, No. 5, pp. 644-655, Sept. 1998.

- [2] I. J. Cox, M. L. Miller, T. P. Minka, T. V. Papathomas, and P. N. Yianilos, "The Bayesian image retrieval system, PicHunter:theory, implementation, and psychophysical experiments," *IEEE Trans. on Image Processing*, Vol. 9, No. 1, pp. 20-37, Jan. 2000.
- [3] H. Muller, W. Muller, S. Marchand Maillet, and T. Pun, "Strategies for positive and negative relevance feedback in image retrieval," *Proc. of IEEE Conference on Pattern Recognition*, Vol. 1, pp. 1043-1046, Sept. 2000.
- [4] Jing Huang, S. Ravi Kumar, Mandar Mitra, Wei Jing Zhu, and Ramin Zabih. "Image Indexing Using Color Correlograms," *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 762-768, June 1997.
- [5] Seong-O Shim and Tae-Sun Choi, "Edge color histogram for image retrieval," *2002 Int'l Conf. on Image processing*, Vol. 3, pp. 957-960, June 2002.



오 상 욱

1998년 8월 고려대학교 전자공학과 졸업(공학사)
 2000년 8월 고려대학교 전자공학과 대학원 졸업(공학석사)
 2005년 8월 고려대학교 전자공학과 대학원 졸업(공학

박사)

관심분야: 디지털 방송, 영상처리 및 검색, 멀티미디어 통신



정 민 교

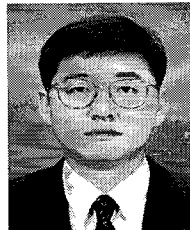
1985년 서울대학교 전자계산기공학과 졸업(공학사)
 1987년 KAIST 컴퓨터학과 졸업(공학석사)
 1996년 University of Iowa 컴퓨터학과 졸업(공학박사)
 1987년~2001년 Korea Telecom

멀티미디어 연구소 (선임연구원)

2001년~2003년 Vivcom Inc. (Founder & Engineering Director)

2003년~현재 서울여자대학교 조교수

관심분야: 컴퓨터 비전, 영상처리/코딩, 멀티미디어 검색, 콘텐츠 보호, 생체 인식



설 상 훈

1981년 서울대학교 전자공학과 졸업(공학사)
 1983년 한국과학기술원 전기및전자공학과 졸업(공학석사)
 1993년 Univ. Of Illinois, Urbana

Champaign 전자공학과

졸업(공학박사)

1993년~1994년 Beckman Institute, Univ. Of Illinois, USA 연구원

1994년~1996년 Flight Deck Branch, NASA, Ames Research Center, USA 연구원

1996년~1997년 IBM Almaden Research Center 연구원

1997년~현재 고려대학교 전자공학과 교수

관심분야: Digital Media, TV and Internet, Video Indexing and Search, Video codec, Home Networking, International Standards (MPEG, TV-Anytime, ATSC)