

적합성 피드백을 통해 결정된 가중치를 갖는 시각적 특성에 기반을 둔 이미지 검색 모델

(A Image Retrieval Model Based on Weighted Visual
Features Determined by Relevance Feedback)

송지영[†] 김우철[†] 김승우[†] 박상현^{**}

(Ji-Young Song) (Woo-Cheol Kim) (Seung-Woo Kim) (Sanghyun Park)

요약 디지털 이미지의 양이 증가함에 따라 원하는 이미지를 정확하고 빠르게 찾을 수 있는 방법의 필요성이 증가하고 있다. 이미지 검색 방법으로는 이미지의 색상이나 명암과 같은 시각적 특성을 검색 조건으로 이용하는 내용 기반 검색과 이미지를 설명하는 키워드를 검색 조건으로 이용하는 키워드 기반 검색이 있다. 하지만 이러한 방법만으로는 사용자가 원하는 이미지를 정확하게 찾기 힘들다는 문제점이 제기되어 왔다. 따라서 최근에는 검색 도중 사용자의 응답을 받아 사용자의 요구를 파악함으로써 향상된 검색 결과를 제공하는 적합성 피드백에 대한 연구가 많이 진행되고 있다. 하지만 적합성 피드백을 이용하는 방법들도 원하는 결과를 얻기 위해서는 여러 번의 피드백을 필요로 하고 질의 수행이 완료된 후에는 얻어진 피드백 정보를 재사용하지 못한다는 단점이 있다. 따라서 본 논문에서는 이미지에 키워드를 연결한 후 사용자의 피드백 정보를 반영하여 키워드의 신뢰도를 조절함으로써 키워드 기반 이미지 검색의 정확도를 높일 수 있는 모델을 제안한다. 제안된 모델에서는 사용자로부터 피드백을 받은 이미지뿐만 아니라 긍정적 피드백을 받은 이미지들이 공통적으로 가지는 시각적 특성과 유사한 시각적 특성을 가지는 다른 이미지들 까지도 키워드의 신뢰도를 조정함으로써 좀 더 빠른 시간 내에 검색 결과의 정확도를 높이도록 한다. 제안한 방법의 정확성을 검증하기 위한 실험 결과에 따르면, 같은 횟수의 피드백을 받으면서도 재현율과 정확률은 빠른 증가를 보이는 것으로 나타났다.

키워드 : 이미지 검색, 내용 기반 검색, 적합성 피드백, 멀티미디어 데이터베이스

Abstract Increasing amount of digital images requires more accurate and faster way of image retrieval. So far, image retrieval method includes content-based retrieval and keyword based retrieval, the former utilizing visual features such as color and brightness and the latter utilizing keywords which describe the image. However, the effectiveness of these methods as to providing the exact images the user wanted has been under question. Hence, many researchers have been working on relevance feedback, a process in which responses from the user are given as a feedback during the retrieval session in order to define user's need and provide improved result. Yet, the methods which have employed relevance feedback also have drawbacks since several feedbacks are necessary to have appropriate result and the feedback information can not be reused. In this paper, a novel retrieval model has been proposed which annotates an image with a keyword and modifies the confidence level of the keyword in response to the user's feedback. In the proposed model, not only the images which have received positive feedback but also the other images with the visual features similar to the features used to distinguish the positive image are subjected to confidence modification. This enables modifying large amount of images with only a few feedbacks ultimately leading to faster and more accurate retrieval result. An experiment has been performed to verify the effectiveness of the proposed model and the result has demonstrated rapid increase in recall and precision while receiving the same number of feedbacks.

Key words : Image retrieval, content based image retrieval, relevance feedback, multimedia database

[†] 학생회원 : 연세대학교 컴퓨터과학과
jysong@cs.yonsei.ac.kr
twelvepp@cs.yonsei.ac.kr
kimsw@cs.yonsei.ac.kr

^{**} 종신회원 : 연세대학교 컴퓨터과학과 교수
sanghyun@cs.yonsei.ac.kr
논문접수 : 2006년 11월 14일
심사완료 : 2007년 3월 20일

1. 서론

현재 인터넷에서 사용되는 정보는 텍스트 기반에서 이미지, 오디오 등을 이용한 멀티미디어 기반으로 점차 확장되고 있다. 특히 디지털 카메라, 스캐너와 같이 아날로그 데이터를 디지털로 변환할 수 있는 장치의 발달과 인터넷 대역폭의 증가는 인터넷에 포함된 멀티미디어 정보의 증가를 더욱 가속화시키고 있다. 따라서 텍스트 검색을 위주로 하던 기존의 인터넷 검색 시스템들은 이미지, 음악, 비디오 등의 멀티미디어 정보까지도 검색할 수 있도록 검색 엔진을 확장하려 하고 있다. 특히 여러 멀티미디어 정보 중에서 가장 많은 부분을 차지하고 있는 이미지 정보에 대한 효과적인 검색 기술이 요구되고 있으며, 이에 대한 많은 연구들이 진행되고 있다.

기존의 이미지 검색 기술은 대부분 이미지를 저장하는 파일의 이름이나 이미지의 내용을 표현하는 키워드를 검색 조건으로 사용하는 키워드 기반의 검색(Keyword-based search) 방법을 사용한다. 그러나 키워드 기반의 검색 방법은 이미지에 키워드가 연결되어 있지 않거나 이미지에 연결된 키워드가 해당 이미지를 제대로 표현하지 못한다면 검색 결과의 정확도가 매우 낮다는 단점을 가지고 있다. 이러한 단점을 극복하기 위하여 각 이미지의 키워드를 사람이 직접 입력할 수도 있지만 이미지 수의 증가 속도를 고려한다면 이는 좋은 해결책이 될 수 없다.

이미지 검색을 위한 다른 방법으로 이미지의 색상 분포나 명암 같은 시각적 특징을 검색 조건으로 사용하는 내용 기반의 검색(Content-based search)이 있다. 그러나 이 방법은 컴퓨터에 의해서 계산되는 시각적 특징이 사람이 인지하는 시각적 특징과 다를 수 있다는 단점을 가지고 있다[1,2]. 또한, 검색 조건으로서 이미지를 제시해야 하기 때문에 텍스트 기반의 검색 엔진에 익숙한 사용자가 사용하기 어렵다는 단점도 가지고 있다.

결국 내용 기반 검색 방법과 키워드 기반 검색 방법 모두 만족할 만한 검색의 정확도를 제공하지 못한다는 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해서 이미지 검색에 적합성 피드백(Relevance feedback)을 적용하는 연구가 최근 많이 진행되고 있다. 적합성 피드백이란 초기 검색 결과에서 사용자가 몇 개의 예시 답안을 피드백 정보로 시스템에게 알려주면 그 정보를 바탕으로 질의를 수정하여 좀 더 정확한 검색 결과를 얻는 방법이다[3,4].

적합성 피드백을 내용 기반 검색에 적용하는 기존의 연구들은 검색 결과 중 사용자가 긍정적 이미지(Positive image)로 선택한 것들의 공통적인 시각적 특징을 이용해 이미지의 유사성을 비교함으로써 검색 결과의 정확도를 향상시킨다. 그러나 원하는 검색 결과를 얻기

위해서는 피드백 과정을 여러 번 반복해야 하며, 질의 수행이 완료된 후에는 이렇게 얻어진 피드백 정보를 재사용하지 못한다는 단점이 있다[3,4].

적합성 피드백을 키워드 기반 검색에 적용하는 기존의 연구들은 피드백 정보를 이용하여 이미지와 키워드 간 연결의 신뢰도(Confidence)를 자동적으로 조정함으로써 검색 결과의 정확도를 향상시킨다. 즉, 긍정적 이미지에 연결된 키워드의 신뢰도는 높이고 부정적 이미지(Negative image)에 연결된 키워드의 신뢰도는 낮춤으로써, 피드백이 반복됨에 따라 키워드가 이미지를 좀 더 정확하게 표현하도록 하는 방식을 사용한다[5-7]. 그러나 이 방식에서는 피드백을 받은 이미지에 대해서만 키워드의 신뢰도를 조정하기 때문에 전체 이미지의 키워드를 올바르게 조정하기 위해서는 많은 시간이 소요된다는 단점이 있다.

본 논문에서는 적합성 피드백을 키워드 기반 검색에 적용할 때 발생하는 문제점을 해결하기 위해서 키워드 기반 검색 방법에 내용 기반 검색 방법을 결합한 모델을 제안한다. 제안하는 방안은 피드백으로 받은 이미지들의 키워드뿐 아니라 긍정적 피드백을 받은 이미지들을 구분하는데 사용된 시각적 특성과 유사한 시각적 특성을 갖는 다른 이미지들의 키워드까지도 신뢰도 조정의 대상이 되도록 한다. 이를 통해 적은 피드백으로도 많은 이미지의 키워드 신뢰도를 조정할 수 있게 되므로, 궁극적으로는 좀 더 빠른 시간 내에 검색 결과의 정확도를 높일 수 있다는 장점을 가지게 된다.

본 논문의 구성은 다음과 같다. 먼저 2장에서는 이미지 검색에 대한 기존 연구를 살펴본다. 다음으로 3장에서는 본 논문에서 제안하는 검색 모델을 단계별로 설명하며, 4장에서는 실험을 통하여 제안 모델의 성능을 평가한다. 마지막으로 5장에서는 본 논문의 내용을 요약하고 앞으로의 연구 방향을 제시한다.

2. 이미지 검색을 위한 기존 연구

이미지 검색 방법은 크게 내용 기반 검색과 키워드 기반 검색으로 구분된다. 내용 기반 검색은 이미지로부터 색상, 무늬, 윤곽과 같은 시각적 특징을 추출하여 검색 조건으로 사용한다[1,8-10]. 대표적인 내용 기반 검색 시스템으로 QBIC[8], VisualSEEK[9], Virage[10] 등이 있다. QBIC는 색상, 무늬, 예제 이미지, 스케치 등을 질의로 사용하여 이미지를 검색하며, 주로 대용량 이미지 데이터베이스나 비디오 데이터베이스에서 사용된다. VisualSEEK는 사람이 이미지를 인식할 때 가장 중요하게 사용하는 시각적 특징인 색상들의 상하 및 좌우 위치 관계를 이용하여 이미지를 검색한다. Virage는 색상 배치, 무늬, 객체의 외곽선 등을 이미지의 시각적 특

정으로 사용하며, 얼굴 인식 및 안구 인식과 같은 특정 분야에서도 사용된다.

검색의 정확도를 높이기 위해서 대부분의 내용 기반 검색 시스템들은 여러 개의 시각적 특징을 조합하여 사용하지만, 컴퓨터 프로그램에 의해서 계산되는 시각적 특징이 사람이 인지하는 시각적 특징과 다를 수 있으므로 정확도 향상에는 한계가 있다. 또한, 검색 조건으로서 이미지를 제시해야 하기 때문에 텍스트 기반의 검색 엔진에 익숙한 사용자가 사용하기 어렵다는 단점도 가지고 있다[1].

키워드 기반 검색은 이미지의 내용을 표현하는 키워드를 검색 조건으로 사용한다. 따라서 이미지를 표현하는 키워드의 정확도가 높은 경우에는 검색 결과의 정확도를 보장할 수 있다. 키워드의 정확도를 높이기 위한 가장 좋은 방법은 사람이 직접 이미지의 키워드를 입력하는 것이다. 그러나 이미지 데이터베이스의 크기를 고려한다면 사람이 일일이 이미지의 키워드를 입력한다는 것은 거의 불가능하다. 따라서 자동 키워드 연결에 대한 연구들이 최근 활발하게 진행되고 있다[11-13].

Cheng 등은 영역 단위의 클러스터링을 이용하여 새로운 이미지에 자동으로 키워드를 연결하였다[11]. 즉, 새로운 이미지가 주어지면 이미지를 영역 단위로 나눈 후, 각 영역과 가장 유사한 클러스터를 찾아서 그 클러스터에 연결된 키워드를 해당 영역에 할당하는 방법을 사용하였다. Jeon 등은 키워드의 의미적 계층 관계를 이용하여 키워드의 신뢰도를 조정하는 방법을 제안하였다[12]. 예를 들면, 고양이와 강아지는 모두 애완동물이므로, 어떤 이미지에 고양이라는 키워드와 강아지라는 키워드가 모두 연결되어 있다면 그 이미지에 애완동물이라는 키워드를 높은 신뢰도로 연결하였다. Feng 등은 하나의 이미지에 대해서 서로 다른 두 가지 방법을 각각 적용하여 두 개의 키워드 집합을 구한 후, 두 집합에 모두 속한 키워드들은 높은 신뢰도를 할당하고 하나의 집합에만 속한 키워드들은 낮은 신뢰도를 부여하는 방법을 제안하였다[13].

기본적으로 위의 자동 키워드 연결 방법들은 프로그램에 의해 계산되는 시각적 특징의 유사성을 이용하여 키워드를 다른 이미지에 전파(Propagate)하는 방식을 사용한다. 이와 같은 방식을 사용하여 모든 이미지에 키워드를 연결한 후에는 키워드 기반 검색을 수행할 수 있게 되므로 내용 기반 검색에 비해 검색 속도를 향상시킬 수 있지만 검색의 정확도가 낮다는 문제는 여전히 해결하지 못한다. 따라서 검색의 정확도를 높이기 위해서 이미지 검색에 적합성 피드백을 적용하는 연구[3,4]가 최근 많이 진행되고 있다. 적합성 피드백이란 초기 검색 결과의 정확도를 사용자에게 평가받아 사용자가

원하는 데이터의 특성을 파악하여 좀 더 정확한 검색 결과를 제공해 주는 방법이다. 이러한 적합성 피드백은 내용 기반 검색과 키워드 기반 검색 모두에 적용될 수 있다.

내용 기반 검색에서는 두 가지 방식으로 적합성 피드백을 적용한다. 첫 번째 방식은 사용자의 피드백 정보를 이용하여 긍정적 이미지들과는 비슷하고 부정적 이미지들과는 다르도록 질의 이미지를 표현하는 정보를 수정하여 다시 데이터베이스를 검색하는 것이다. 두 번째 방식은 여러 시각적 특징 중 분별력이 높은 것을 찾아 더 높은 가중치를 부여함으로써 유사도 계산의 과정에서 사용자의 피드백 정보를 반영하는 것이다[2-4].

키워드 기반 검색에서는 사용자의 피드백 정보를 이용하여 키워드를 추가 혹은 삭제하거나[7], 키워드의 신뢰도 값을 변경한다[6,11]. 내용 기반 검색과 같이 즉각적으로 향상된 이차 검색 결과를 얻을 수는 없지만 피드백 정보가 누락되면 잘못된 연결된 키워드의 신뢰도는 낮아지고 제대로 연결된 키워드의 신뢰도는 높아지게 된다. 또한 피드백 과정에 의해 이미지로부터 키워드가 제거되기도 하고 새로운 키워드가 추가되기도 한다. 따라서 피드백 정보가 충분히 누락되면 이미지 검색의 정확도가 향상된다. 그러나 이 방식은 사용자로부터 피드백을 받은 이미지에 대해서만 키워드의 신뢰도를 조정하기 때문에 전체적인 검색의 정확도를 높이기 위해서는 많은 피드백 과정이 필요하다는 단점을 가지고 있다.

본 논문에서는 적합성 피드백을 키워드 기반 검색에 적용할 때 발생하는 문제점을 해결하기 위하여 피드백을 받은 이미지들의 키워드뿐 아니라 긍정적 이미지들을 구분하는데 사용된 시각적 특성과 유사한 시각적 특성을 갖는 다른 이미지들을 피드백 적용 범위에 포함하여 신뢰도 조정의 대상이 되도록 한다. 따라서 제시하는 모델은 기존의 모델보다 좀 더 빠른 시간 내에 검색 결과의 정확도를 높일 수 있다는 장점을 가지게 된다.

3. 피드백 확장 방법에 기반한 이미지 검색 모델

이번 장에서는 본 논문에서 제안하는 이미지 검색 모델을 설명한다. 이미지 검색 모델은 크게 1) 이미지들을 이미지 데이터베이스에 수집하는 이미지 수집 부분과 2) 이미지 데이터베이스를 검색하여 사용자가 원하는 이미지를 찾아주는 이미지 검색 부분으로 나눌 수 있다.

이미지를 수집하는 부분은 이미지 검색에 필요한 메타 데이터를 생성하는 과정이며 검색에 필요한 메타 데이터에 따라서 다양한 메타 데이터 수집 방법을 사용한다. 즉, 키워드 기반의 검색을 위해서는 이미지를 표현하는 키워드를 이미지에 연결하는 키워드 연결 과정이 필요하고, 내용 기반의 검색을 위해서는 색상이나 모양

같은 저 레벨(low-level)의 시각적 특성(visual feature)을 이미지로부터 추출하여 저장하는 과정이 필요하다. 본 논문에서는 키워드 기반 검색 방식과 내용 기반 검색 방식을 결합한 모델을 사용하므로 이미지로부터 키워드 및 시각적 특성을 모두 추출해야 하며, 이런 작업을 담당하는 부분을 그림 1.1과 같이 '키워드 및 시각적 특성 연결 모듈'이라 한다.

이미지를 검색하는 부분은 3개의 모듈로 구성되어 있다. 먼저 그림 1.2의 '키워드 기반 이미지 검색 모듈'은 이미지 데이터베이스를 검색하여 사용자가 제시한 질의 키워드를 하나라도 포함하고 있는 이미지들을 검색하는 모듈이며, 이 모듈을 통해 검색된 결과를 '1차 검색 결과'라고 한다. 이러한 1차 검색 결과는 이미지에 연결된 키워드가 이미지의 내용을 제대로 표현하지 못하는 경우에 정확도가 낮다는 문제점을 가지고 있다. 따라서 검색 결과의 정확도를 높이기 위해서 그림 1.3의 '이미지 재정렬 모듈'을 이용해 사용자의 피드백 정보를 입력받아 이를 분석한 후 1차 검색 결과를 재정렬한다. 마지막으로 그림 1.4의 '키워드 신뢰도 조절 모듈'은 사용자로부터 긍정적 피드백을 받은 이미지들의 키워드 신뢰도는 높이고 부정적 피드백을 받은 이미지들의 키워드 신뢰도는 낮추는 역할을 수행한다. 또한, 긍정적 피드백을 받은 이미지들을 구분하는데 사용된 시각적 특성과 유사한 시각적 특성을 갖는 다른 이미지들의 키워드 신뢰도도 함께 높임으로써 적은 피드백으로도 많은 이미지의 키워드 신뢰도가 조정되도록 한다.

3장의 나머지 부분에서는 위의 과정을 따라가면서 각 과정을 자세히 설명하도록 한다. 3.1절에서는 키워드 기반 검색 및 내용 기반 검색을 위해 이미지에 부가적으로 저장해야 하는 이미지의 속성에 대해서 정의한다. 3.2절에서는 이러한 이미지 속성의 초기 값을 생성하는 자동 키워드 연결에 대해서 설명한다. 계속해서 3.3절에서는 키워드 기반 이미지 검색 과정을 설명하고 3.4절에

서는 적합성 피드백을 적용하여 향상된 2차 검색 결과를 제공하는 과정을 설명한다. 마지막으로 3.5절에서는 키워드의 신뢰도를 조절하여 향상된 이미지 검색 결과를 제공하는 방법을 설명한다.

3.1 이미지와 이미지의 속성

본 논문에서 제안하는 방식으로 이미지를 검색하기 위해서는 이미지 데이터에 1) 키워드 기반 검색을 위한 키워드와 2) 키워드에 사용자의 적합성 피드백 결과를 반영하기 위한 키워드의 신뢰도, 3) 내용 기반 검색을 위한 저 레벨의 시각적 특성의 3가지 정보를 추가해야 한다.

$$I = (F, (K, C))$$

정의 1. 이미지 I

- F는 색상, 무늬, 질감과 같이 이미지에서 추출한 저 레벨의 시각적 특성들의 집합으로서 각 이미지는 n개의 특성에 대한 정보를 가지고 있다. 이미지 I의 j번째 특성을 I.Feature_j라고 하면 F는 {I.Feature_j | 1 ≤ j ≤ n}으로 표현된다.
- K는 이미지의 내용을 표현하는 키워드의 집합으로서 각 이미지에는 최대 m개의 서로 다른 키워드가 연결되어 있다. 이미지 I에 연결되어 있는 j번째 키워드를 I.Keyword_j라고 하고 이미지가 가질 수 있는 전체 키워드의 집합을 W라고 하면 K는 {I.Keyword_j | 1 ≤ j ≤ m, I.Keyword_j ∈ W}로 표현된다.
- C는 이미지에 연결된 키워드가 이미지의 내용을 얼마나 정확하게 반영하는지를 나타내는 신뢰도의 집합이다. 이미지 I에 연결되어 있는 j번째 키워드의 신뢰도는 I.Confidence_j로 표현하며 최소값으로 MINCONF, 최대값으로 MAXCONF를 갖는다. 또한, 전체 키워드의 집합 W에 속한 임의의 키워드 kw가 이미지 I에서 가지는 신뢰도를 I.Confidence_{kw}로 표시한다. 만일 키워드 kw가 이미지 I에 연결되어 있지 않다면 I.Confidence_{kw}의 값은 0이 된다.

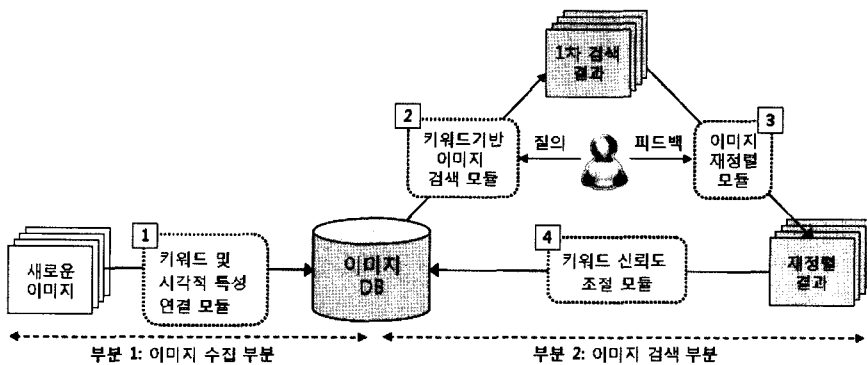


그림 1 제안하는 이미지 검색 모델의 전체 구조도

3.2 자동 키워드 연결

자동으로 키워드를 이미지에 연결하는 방법에는 여러 가지가 있다. 예를 들어 1) 이미지가 저장되어 있는 웹 페이지, 책 등과 같은 환경의 분석을 통해서 얻는 방법, 2) 내용 기반 검색을 이용하는 방법 등이 존재한다. 본 논문에서는 다른 방법에 비해서 상대적으로 높은 정확도를 가지고 있으며 이미지를 얻을 수 있는 환경에 영향을 받지 않는 내용 기반 검색을 통한 자동 키워드 연결 방식을 이용한다.

내용 기반 검색을 이용한 자동 키워드 연결 방식에서는 사람에 의해 미리 키워드가 연결되어져 있는 이미지의 집합인 트레이닝 집합(Training set) T가 필요하다. 트레이닝 집합 T를 이용해 새로운 이미지 I_{new} 에 자동으로 키워드를 연결하는 과정은 다음과 같다. 먼저, 새로운 이미지 I_{new} 와 트레이닝 집합 T에 속한 각 이미지 I_i 사이의 시각적 특성의 유사도 $FeatureSimilarity(I_{new}, I_i)$ 를 계산한다. 이 때 $FeatureSimilarity(I_{new}, I_i)$ 는 0부터 1사이의 값을 갖으며 1에 가까울수록 두 이미지의 시각적 특성이 비슷하다는 것을 나타낸다. 다음으로, 전체 키워드의 집합 W에 속한 각 키워드 kw가 새로운 이미지 I_{new} 에서 가지는 신뢰도 $I_{new}.Confidence_{kw}$ 를 아래 식 (1)을 이용하여 계산한다. 이 때 모든 이미지는 최대 m개의 키워드만을 가질 수 있으므로 신뢰도를 기준으로 상위 m개의 키워드를 선택하여 새로운 이미지 I_{new} 의 키워드 집합으로 설정한다. 만약 선택된 m개의 키워드 중에서 신뢰도가 MINCONF보다 작은 것이 있다면 해당 키워드를 I_{new} 의 키워드 집합에서 삭제한다.

$$I_{new}.Confidence_{kw} = \frac{\sum_{I_i \in T} (Feature.Similarity(I_{new}, I_i) \times I_i.Confidence_{kw})}{\sum_{I_i \in T} Feature.Similarity(I_{new}, I_i)}$$

식 (1). 전체 키워드의 집합 W에 속한 키워드 kw가 새로운 이미지 I_{new} 에서 가지는 신뢰도

3.3 키워드를 이용한 이미지 검색

3.2절에서 설명한 자동 키워드 연결 방법을 사용하여 데이터베이스에 저장된 모든 이미지에 키워드들을 연결한 후에는 키워드 기반의 이미지 검색을 수행할 수 있게 된다. 사용자가 하나 이상의 질의 키워드를 검색 조건으로 입력하면, 먼저 데이터베이스에 저장된 이미지 중에서 하나 이상의 질의 키워드를 포함한 것들을 검색한다. 다음으로, 검색된 이미지들을 질의 키워드 집합에 대한 신뢰도 합을 기준으로 정렬한 후 사용자에게 보여준다. 질의 키워드의 집합을 Q라고 하면 Q에 대한 이미지 I_i 의 신뢰도 합 $ConfidenceSum(Q, I_i)$ 은 아래의 식

(2)를 이용하여 계산한다.

$$ConfidenceSum(Q, I_i) = \sum_{kw \in Q} I_i.Confidence_{kw}$$

식 (2). 질의 키워드의 집합 Q에 대한 이미지 I_i 의 신뢰도 합

3.4 적합성 피드백

이미지에 연결된 키워드의 정확도가 높으면 3.3절의 과정만으로도 사용자가 원하는 이미지를 얻을 수 있다. 하지만 일반적으로 자동 키워드 연결 방식을 통한 키워드 연결은 낮은 정확도를 보인다. 따라서 본 논문에서는 적합성 피드백의 정보를 누적하여 이미지에 연결된 키워드의 정확도를 높이는 방법을 사용한다.

적합성 피드백은 질의 결과로 주어진 이미지 집합의 순위를 사용자의 피드백을 적용해 재정렬하는 것이다. 사용자의 피드백은 찾고자 하는 이미지를 나타내는 긍정적 이미지와 질의와 관련성이 적은 이미지를 나타내는 부정적 이미지로 구분된다. 이러한 피드백 정보를 분석해서 질의와 이미지간의 유사도를 재계산하는 과정을 통해서 이미지 검색의 정확도를 향상시킬 수 있다.

3.4.1 피드백에 사용된 시각적 특성

본 논문에서는 질의 키워드에 따라 결과로 주어진 이미지 중에서 긍정적 이미지와 부정적 이미지를 찾는데 사용되는 이미지의 특성이 다르다는 것을 이용해 사용자의 피드백을 분석하고 그 결과를 키워드의 신뢰도에 반영하고자 한다. 예를 들어 '숲'이라는 키워드를 질의로 사용하는 경우 사용자는 모양이나 무늬보다는 색상이라는 특성을 통해서 이미지가 긍정적 이미지인지 부정적 이미지인지 판단할 것이다. 따라서 긍정적 이미지의 색상과 유사한 색상을 가지고 있는 이미지들을 찾아서 이미지에 연결되어 있는 '숲'이라는 키워드의 신뢰도를 높임으로써 이미지 검색 결과의 정확도를 높일 수 있다.

사용자의 피드백을 키워드의 신뢰도에 반영하기 위해서는 먼저 피드백 내용을 분석하여 어떠한 시각적 특성이 긍정적 이미지와 부정적 이미지를 구분하는데 중요하게 사용되었는가를 판단해야 한다. 이를 위해 본 논문에서는 시각적 특성의 분별력(Discrimination power)이라는 개념을 정의하여 사용한다. 시각적 특성 $Feature_j$ 의 분별력이란 1차 검색 결과로 주어진 이미지들을 $Feature_j$ 를 기준으로 정렬하였을 때 긍정적 이미지들의 순위와 부정적 이미지들의 순위가 얼마나 차이가 나는가를 정량적으로 나타낸 것이다. 즉, 긍정적 이미지들이 가지는 시각적 특성 $Feature_j$ 의 평균값과 1차 검색 결과에 포함된 각 이미지가 가지는 $Feature_j$ 값의 차이를 계산한 후, 두 값의 차이가 작은 순으로 1차 검색 결과를 정렬한 다음에 아래의 식 (3)을 이용하여 $Feature_j$ 의 분

별력 DP_j 를 계산한다.

$$DP_j = \frac{Po_j + Ne_j}{N_p + N_n}$$

식 (3). 시각적 특성 $Feature_j$ 의 분별력

위의 식 (3)에서 N_p 와 N_n 은 각각 사용자에 의해 선택된 긍정적 이미지의 수와 부정적 이미지의 수를 나타낸다. 또한 Po_j 와 Ne_j 는 긍정적 이미지들이 가지는 $Feature_j$ 의 평균값을 기준으로 1차 검색 결과를 재정렬할 때 각각 상위 N_p 순위에 포함된 긍정적 이미지의 수와 하위 N_n 순위에 포함된 부정적 이미지의 수를 나타낸다. 이때 DP_j 가 1에 가까울수록 $Feature_j$ 가 긍정적 이미지와 부정적 이미지를 잘 구분함을 의미한다. 위의 예를 다시 살펴보면, '숲'이라는 키워드를 질의로 사용하는 경우에는 색상을 나타내는 시각적 특성이 모양이나 무늬를 나타내는 시각적 특성에 비하여 상대적으로 높은 분별력을 가지게 됨을 알 수 있다.

3.4.2 이미지의 재정렬

위와 같은 방법으로 각 시각적 특성 $Feature_j$ 의 분별력을 계산한 후 이를 이용하여 1차 검색 결과로 주어진 이미지들을 재정렬하여 사용자에게 보여준다. 이를 위해 먼저 각각의 시각적 특성에 대해 긍정적 이미지들이 가지는 평균값을 이용하여 가상의 질의 이미지 I_{avg} 를 작성한다. 이 때, 가상의 질의 이미지 I_{avg} 의 j 번째 시각적 특성은 긍정적 이미지들의 가지는 j 번째 시각적 특성의 평균값이 된다. 다음으로 1차 검색 결과에 포함된 각 이미지 I_i 와 가상의 질의 이미지 I_{avg} 사이의 유사도를 함수 $WeightedFeatureSimilarity(I_{avg}, I_i, \{DP_1, DP_2, \dots, DP_n\})$ 를 이용하여 계산한다. 이 유사도 함수는 시각적 특성의 분별력을 이용하여 해당 특성의 가중치를 아래의 식 (4)와 같이 구한 후 이를 두 이미지의 유사도 계산에 반영한다. 식 (4)에서 n 은 전체 시각적 특성의 수를 나타내며, DP_j 와 w_j 는 각각 j 번째 시각적 특성의 분별력과 가중치를 나타낸다.

$$w_j = \frac{DP_j}{\sum_{k=1}^n DP_k}$$

식 (4). j 번째 시각적 특성의 가중치

시각적 특성의 가중치를 반영한 위의 유사도 함수를 기반으로 1차 검색 결과에 포함된 이미지들을 재정렬하면 사용자가 긍정적으로 평가한 이미지의 시각적 특성과 비슷한 시각적 특성을 갖는 이미지들은 상위에 배치되고 상이한 시각적 특성을 갖는 이미지들은 하위에 위

치하게 된다. 이를 통해 사용자는 질의 키워드를 포함하고 있는 이미지 중에서 자신이 긍정적인 피드백으로 주었던 이미지들과 비슷한 이미지들을 검색 결과로 얻게 된다.

3.5 키워드의 신뢰도 조정

본 논문에서 제안한 이미지 검색 시스템에서는 사용자의 피드백 정보를 이용하여 1차 검색 결과에 포함된 이미지들의 키워드 신뢰도를 조정하는 피드백 확장 방법을 적용한다. 먼저, 긍정적 이미지에 대해서는 질의로 주어진 각 키워드의 신뢰도를 일정 단위 증가시킨다. 만약 증가된 신뢰도가 $MAXCONF$ 를 초과할 경우에는 신뢰도를 $MAXCONF$ 로 설정한다. 만약 긍정적 이미지에 질의 키워드가 연결되어 있지 않은 경우에는 그 키워드를 $MINCONF$ 의 신뢰도로 해당 이미지에 연결한다. 다음으로, 부정적 이미지에 대해서는 질의로 주어진 각 키워드의 신뢰도를 일정 단위 감소시킨다. 만약 감소된 신뢰도가 $MINCONF$ 보다 작아지는 경우에는 그 키워드를 해당 이미지로부터 삭제한다. 마지막으로 사용자의 피드백을 받지 않은 이미지이지만 재정렬 후에 상위에 위치하는 일정 개수의 이미지들을 선택하여 키워드의 신뢰도를 증가시킨다. 이렇게 추가적으로 선택되는 이미지들은 본 논문에서는 '추가 이미지'라고 호칭한다. 추가 이미지들은 사용자의 직접적인 피드백의 결과가 아니기 때문에 사용자가 원하지 않는 결과일 수도 있으므로, 긍정적 이미지보다는 키워드의 신뢰도를 작게 증가시킨다. 이와 같이 추가 이미지들을 선택하여 키워드의 신뢰도를 조절하는 것을 '피드백 확장'이라고 한다.

피드백 확장 방법을 사용하면 이미지에 연결된 키워드의 신뢰도를 빠르게 조절하는 것이 가능해진다. 하지만 피드백 확장에 의한 신뢰도 조절이 항상 이미지 검색 성능을 향상시키지는 못한다. 예를 들어, 키워드 kw 가 연결된 이미지의 집합에서 정확하게 연결된 키워드들은 모두 $MAXCONF$ 의 신뢰도를 갖고 정확하게 연결되지 않은 키워드들은 $MINCONF$ 와 $MAXCONF$ 사이의 신뢰도를 갖는다고 하자. 이런 경우 피드백 확장 방법을 적용하면 정확하게 연결된 키워드의 신뢰도는 더 이상 오르지 못한 채 잘못된 키워드의 신뢰도만 높아지기 때문에 검색의 정확도가 떨어지게 된다.

이러한 문제점을 해결하게 위해서 본 논문에서는 '키워드별 선택적 피드백 확장' 방법을 사용한다. 즉, 각각의 질의 키워드 kw 에 대하여, 추가 이미지 상에서 kw 가 가지는 신뢰도의 평균을 구한 후 그 값이 정해진 임계값을 초과하지 않는 경우에만 kw 에 대한 추가 이미지의 신뢰도를 높이도록 한다. 만약 추가 이미지 상에서 kw 가 가지는 신뢰도의 평균값이 정해진 임계값을 초과하는 경우에는 kw 에 대한 신뢰도 조정이 충분히 이루

어졌다고 판단하여 kw에 대한 피드백 확장을 적용하지 않는다.

4. 실험을 통한 성능 평가

이 장에서는 다양한 조건 하에서 이미지 검색 결과에 대한 피드백을 받음에 따라 이미지들이 가지고 있는 키워드의 신뢰도가 변하는 정도를 확인한다. 이 실험을 통하여 본 논문이 제안하는 적합성 피드백과 피드백 확장 방법이 검색 결과의 정확도 면에서 기존의 적합성 피드백 방법보다 우수함을 보인다.

4.1 실험 환경

본 실험에서는 CalTech에서 수행된 이미지 연구[14]에서 사용된 이미지 셋을 선택하여 검색의 대상으로 사용하였다. CalTech 이미지 셋은 약 9,281장의 이미지로 이루어져 있으며 각 이미지마다 하나의 객체를 분명히 나타내고 있어 이미지의 내용을 잘 표현하는 시각적 특성을 가지고 있다. 또한, 이미지가 포함하는 객체에 대한 올바른 키워드가 연결되어 있기 때문에 실험 이미지로 사용하였다.

이미지의 시각적 특성은 MPEG-7 XM 소프트웨어 [15]를 사용하여 추출하였다. 실험에는 이미지의 색상, 형태, 무늬를 가장 잘 반영하는 5개의 시각적 특성(Color Layout, Color Structure, Homogeneous Texture, Edge Histogram, Region Shape)을 사용하였다. 시각적 특성의 유사도는 MPEG-7 XM 소프트웨어에 구현되어 있는 유사도 함수를 사용하였다. 실험을 위해 Windows XP 운영체제에서 동작하는, 1GB의 메모리와 80GB의 하드디스크를 가지고 있는 펜티엄 4 2.8GHz의 컴퓨터를 사용하였다.

자동 키워드 연결 과정에 사용되는 트레이닝 집합의 크기는 전체 이미지 수의 약 4%인 360장으로 결정하였다. 자동 키워드 연결과정에서 트레이닝 집합을 이용하여 다른 이미지들에게 자동 키워드를 연결하였고, 트레이닝 집합을 포함한 전체 9,281장의 이미지에 대한 데이터베이스를 구성하여 검색의 대상으로서 실험에 사용하였다.

성능평가는 다음의 두 가지 방식을 대상으로 한다. 하나는 단순히 피드백을 받은 이미지들의 신뢰도를 조절하는 NaiveFeedback 방식으로 Y. Lu 등의 연구[6]에 기반을 둔 방식이다. 다른 하나는 피드백 확장 방법을 적용하여 더 많은 이미지의 신뢰도를 조절하는 ExtendedFeedback 방식으로 본 논문에서 제안하는 기법이다.

4.2 용어 정의

이 장에서는 실험에 대한 이해를 돕기 위하여 실험과 관련된 용어를 설명한다.

실험의 진행에 따라서 바뀌는 이미지 데이터베이스의 정확성을 측정하기 위해서는 피드백 과정 중간에 사

이 개입하여 이미지 검색 결과를 확인하고 피드백을 주어야 한다. 하지만 검색 결과의 모든 이미지와 각 이미지가 가지는 키워드에 대한 신뢰도의 조절 상태를 하나하나 직접 확인하여 정확성을 측정하는 것은 거의 불가능하다. 이러한 문제를 해결하기 위해서 사용자가 직접 정확성을 평가하는 대신 이미지 별로 미리 정의된 정확한 키워드와 적합성 피드백을 통해 변화하는 키워드의 신뢰도를 비교하여 시스템 스스로 정확성을 평가하는 방법을 사용한다. 이를 위해 데이터베이스의 모든 이미지에 미리 올바른 키워드를 사용자가 수작업으로 연결하였으며, 이렇게 연결된 키워드들을 수동(Manual) 키워드라고 정의한다. 수동 키워드는 이미지의 내용을 표현하는 정확한 키워드이기 때문에 신뢰도는 언제나 MAXCONF 값을 가지고 있으며, 검색을 통해 연결된 자동 키워드와는 별도로 관리된다. 시스템은 전체 데이터베이스가 가지는 자동 키워드들의 신뢰도의 총합과 수동 키워드들의 신뢰도의 총합을 비교하여 전체 데이터베이스의 정확성을 스스로 평가하게 된다.

이 장에서는 이미지에 연결된 키워드가 자동 키워드와 수동 키워드의 두 종류로 구분되므로 각각 kw^A 와 kw^M 로 표현한다. 자동 키워드는 시스템이 이미지 검색에 사용하고 피드백 과정을 통해서 변화 시키며, 수동 키워드는 시스템을 통해 변화되는 정확성을 측정하기 위해서만 사용된다. 또한 이미지 I에 갖고 있는 수동 키워드 kw^M 의 신뢰도는 $I_i.Confidence_{kw^M}$ 으로 표기하며 자동 키워드 kw^A 의 신뢰도는 $I_i.Confidence_{kw^A}$ 로 표기한다.

데이터베이스에 있는 모든 이미지는 자동 키워드와 수동 키워드를 가지고 있기 때문에 같은 키워드 kw 가 이미지 I에 연결되어 있더라도 키워드는 자동 키워드로 연결되어 있을 수도 있고 수동 키워드로서 연결되어 있을 수도 있다. 이와 같이 서로 다른 특징의 키워드를 갖고 있는 이미지들을 구분하기 위하여 본 논문은 이미지 집합을 따로 정의한다. 키워드 kw 를 자동 키워드로 가지고 있는 이미지의 집합은 $Set^A(kw)$ 로 정의하며, 수동 키워드로 가지고 있는 이미지의 집합은 $Set^M(kw)$ 로 정의한다.

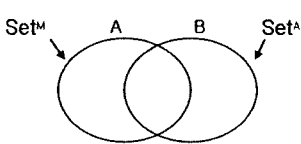
4.3 재현율(Recall)과 정확률(Precision)의 정의

본 장에서는 데이터베이스의 정확성을 측정하기 위한 지표로 사용되는 재현율과 정확률을 설명한다.

데이터를 검색할 때 데이터베이스에 있는 질의와 관련된 정확한 데이터들 중 검색된 정확한 데이터의 비율을 재현율이라고 한다. 이것은 시스템이 정확한 데이터를 검색해 내는 능력을 나타낸다. 한편, 질의를 처리하여 검색된 데이터 중 정확한 데이터의 비율을 정확률이라고 한다.

그림 2에 정확률과 재현율을 도표로 표현하였다. 집합

A는 질의가 주어졌을 때 데이터베이스에 있는 질의와 실제로 관련된 모든 이미지들을 나타내고 집합 B는 질의에 의해 검색된 모든 이미지를 나타낸다. 즉, 집합 A는 질의 키워드 kw를 수동 키워드로서 가지고 있는 집합 $Set^M(kw)$ 이며, 집합 B는 수동 키워드로서 가지고 있는 집합 $Set^A(kw)$ 이다. 이 때 검색 결과 중 사용자가 원하는 정확한 이미지들은 두 집합이 교차하는 부분이며 $Set^A(kw)$ 와 $Set^M(kw)$ 의 교집합으로 표현할 수 있다. 재현율을 높이려면 질의와 관련된 이미지를 많이 찾을 수 있도록 검색 결과의 수를 늘려야 하고, 정확성을 높이려면 두 집합이 교차하는 $A \cap B$ 를 늘리면서 원하지 않은 결과인 B-A를 줄여야 한다.



A : 시스템이 검색해야 하는 정확한 이미지들
B : 시스템이 실제로 검색한 이미지들

그림 2 재현율과 정확률

기존 이미지 검색 연구에서는 재현율을 측정하기 위해 주어진 질의의 결과로 얻어진 상위 검색 결과에서 찾아내지 못한 이미지를 포함한 전체 정확한 이미지에 대한 찾아낸 정확한 이미지의 비율을 이용하였다. 그리고 정확률을 측정하기 위해서는 찾아낸 이미지에 대한 찾아낸 정확한 이미지의 비율을 계산하였다. 그러나 본 연구는 1) 내용 기반 검색이 아닌 키워드 기반 검색방법을 사용하고, 2) 각 키워드가 신뢰도 값을 가지고 있기 때문에 동일한 순위에 온 이미지라도 신뢰도의 값이 다를 수 있다. 따라서 기존의 연구에서 사용한 정확성 측정 방법을 그대로 사용하기에는 무리가 있기 때문에 키워드 신뢰도를 함께 고려한 재현율과 정확률의 측정 방법을 제시한다.

한 키워드의 재현율은 수동 키워드들의 신뢰도 합과 정확하게 연결된 자동 키워드의 신뢰도 합 간의 비율로 정의할 수 있다. 이 값은 키워드가 이미지에 얼마나 정확하게 연결되었으며, 연결된 키워드가 이미지의 내용을 얼마나 잘 표현하는지를 나타내는 기준이 되며 식 (5)와 같이 나타낼 수 있다.

$$Recall(kw) = \frac{\sum_{I_i \in \{Set^M(kw) \cap Set^A(kw)\}} (I_i \cdot Confidence_{kw^A})}{\sum_{I_i \in Set^M(kw)} (I_i \cdot Confidence_{kw^M})}$$

식 (5). 키워드 kw의 재현율

정확률은 자동 키워드들의 신뢰도 합에 대한 정확한 자동 키워드의 신뢰도 합의 비로 표현할 수 있으며 식 (6)과 같이 정의한다. 식 (6)의 값은 키워드 kw가 자동으로 연결된 이미지들 중 정확하게 연결된 이미지의 신뢰도 비율을 측정하여, 실제 검색되는 이미지 중 정확한 이미지가 어느 정도인가를 나타낸다. 정확한 검색 결과를 얻기 위해서는 정확한 자동 키워드 신뢰도 합은 크고 부정확한 자동 키워드 신뢰도 합은 작아야 한다.

$$Precision(kw) = \frac{\sum_{I_i \in \{Set^M(kw) \cap Set^A(kw)\}} (I_i \cdot Confidence_{kw^A})}{\sum_{I_i \in Set^A(kw)} (I_i \cdot Confidence_{kw^M})}$$

식 (6). 키워드 kw의 정확률

위의 방법을 사용하여 각 단어의 재현율과 정확률을 구한다면 데이터베이스 전체의 재현율과 정확률을 구할 수 있다. 데이터베이스 전체의 재현율과 정확률은 시스템에서 사용되는 모든 키워드의 재현율 평균과 정확률 평균으로 구하며 식 (7), 식 (8)과 같이 표현한다. 이때, 시스템에서 정의된 모든 키워드의 집합을 W라고 한다. 집합의 크기는 |W|로 표현한다.

$$Recall = \frac{\sum_{kw \in W} Recall(kw)}{|W|}$$

식 (7). 데이터베이스에 있는 모든 키워드에 대한 재현율

$$Precision = \frac{\sum_{kw \in W} Precision(kw)}{|W|}$$

식 (8). 데이터베이스에 있는 모든 키워드에 대한 정확률

4.4 파라미터 결정

본 실험을 하기 전에 실험의 결과에 영향을 미치는 두 개의 파라미터 $Threshold^{Size}$ 와 $Threshold^{EXTFB}$ 를 결정하기 위한 실험을 수행한다. 파라미터 $Threshold^{Size}$ 는 피드백 확장 방법에서 선택되는 추가 이미지의 개수를 결정하며 $Threshold^{EXTFB}$ 는 키워드별 선택적 피드백 확장 방법에서 피드백 확장의 사용여부를 결정한다.

실험을 위해 7,000번의 피드백을 받는 동안 재현율, 정확률의 변화를 조사하였으며 피드백 정보로 N_p 는 3, N_n 는 3으로 하였다. 키워드의 신뢰도는 MINCONF인 0 부터 MAXCONF인 5까지의 값을 갖는다. 두 방법의 성능차이를 나타내기 위하여 식 (9)와 같이 Extended-Feedback과 NaiveFeedback의 측정값 차이를 구한 후 NaiveFeedback에 대한 백분율로 나타내었다.

Difference of Measures =

$$\frac{\text{ExtendedFeedback} - \text{NaiveFeedback}}{\text{NaiveFeedback}} \times 100$$

식 (9). 두 방법의 측정값 차이를 백분율로 나타내기 위한 식

4.4.1 Threshold^{Size} 파라미터의 결정

Threshold^{Size}의 값이 클수록 많은 이미지를 선택하므로 많은 키워드의 신뢰도가 오르며 따라서 재현율을 낮게 빠르게 오른다. 하지만 추가 이미지의 수가 많아질수록 정확하지 않은 이미지도 결과에 포함될 확률이 높아지므로 정확률은 느리게 오르는 단점이 있다. 따라서 재현율과 정확률 모두 만족할 만한 결과를 얻을 수 있는 파라미터 값을 결정해야한다. 그림 3은 각각 Threshold^{Size} 값을 전체 피드백 이미지 수의 0.5배수부터 1.5배수까지 변화시켰을 때 검색 결과의 재현율과 정확률을 측정한 그래프이다. x축은 피드백의 횟수를 나타내며, y축은 각각 재현율과 정확률의 성능차이를 나타낸다.

그림 3(a)를 보면 Threshold^{Size}가 높을수록 재현율이 NaiveFeedback보다 빠르게 증가하는 것을 볼 수 있다. 하지만 그림 3(b)를 보면 Threshold^{Size}가 높을수록 정확률이 초기에는 빠르게 증가하지만 1,500번 정도의 피드백이 지나면 정확률이 빠르게 낮아지고 7,000번의 피드백을 받은 이후에는 큰 성능의 차이가 없는 것을 볼 수 있다. Threshold^{Size}가 0.5, 1.0일 경우는 거의 같은 수준을 보이고 있으며 1.5일 경우는 정확률이 다른 두 경우에 비해 낮은 값을 갖는다. 파라미터 값이 1일 때 정확률이 전체적으로 가장 좋은 성능을 보였으며 재현율에서도 높은 성능 비교 값을 나타내기 때문에 본 논문은 이후 Threshold^{Size}로 1을 사용한다.

4.4.2 Threshold^{EXTFB} 파라미터의 결정

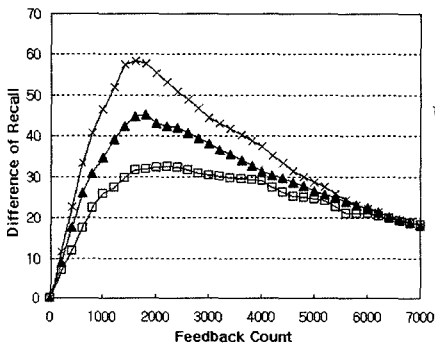
다음은 키워드별 선택적 피드백 확장 방법에서 사용하는 Threshold^{EXTFB}를 결정하기 위한 실험이다. Threshold^{EXTFB}는 키워드 kw의 평균 신뢰도가 MAXCONF 값에 비교하여 얼마만큼의 값을 갖고 있나 계산하여 신뢰도 조정이 충분히 이루어 졌는가 여부를 판단하는 값으로 사용한다. 만약 MAXCONF 값이 5이고 Threshold^{EXTFB}이 0.8이라 하면 MAXCONF의 0.8배인 4를 피드백 확장을 결정하는 기준으로 사용한다.

실험을 위하여 Threshold^{EXTFB}를 0.6부터 0.9까지 변화시켜 가면서 재현율과 정확률을 비교하였다. Threshold^{EXTFB}의 값이 낮을수록 시스템은 피드백 확장 방법의 사용을 일찍 그만두게 되므로 피드백 확장 방법의 단점인 정확률의 느린 증가를 줄일 수 있다. 하지만 정확한 키워드의 신뢰도를 조절할 수 있는 기회가 줄어들기 때문에 재현율이 느리게 오른다. 따라서 재현율과 정확률 모두 높은 값을 가질 수 있는 파라미터 값을 결정해야 한다.

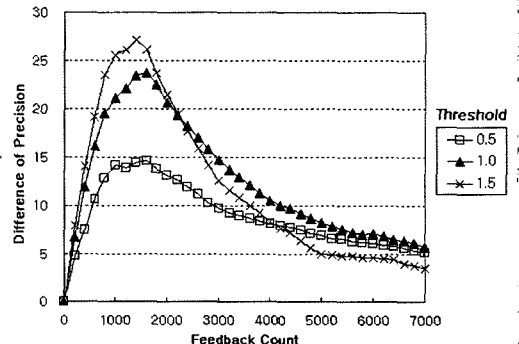
그림 4(a)를 보면 7,000번의 피드백 이후 재현율은 Threshold^{EXTFB}이 0.9일 때 가장 큰 성능 차이를 보였으며 0.7과 0.8일 경우 0.9만큼은 아니지만 좋은 성능 차이를 보였다. 또 그림 4(b)를 보면 정확률은 대체적으로 Threshold^{EXTFB}이 높을수록 성능차이가 적었다. 이것은 Threshold^{EXTFB}가 높으면 키워드의 신뢰도가 MAXCONF값에 이른 후에도 확장 피드백 작업을 하게 되서 부정확한 이미지의 키워드가 선택되어 신뢰도가 오르기 때문이다. 실험 결과 그림 4(b)와 같이 정확률은 Threshold^{EXTFB}이 0.7일 때 가장 높다. 따라서 본 논문은 높은 재현율의 성능 차이를 보이면서 가장 좋은 정확률 성능 차이를 보이는 0.7을 Threshold^{EXTFB}으로 사용한다.

4.5 성능 평가를 위한 실험

실험 1: 피드백의 횟수에 따른 재현율, 정확률 비교



(a) Recall



(b) Precision

그림 3 Threshold^{Size}의 결정 실험

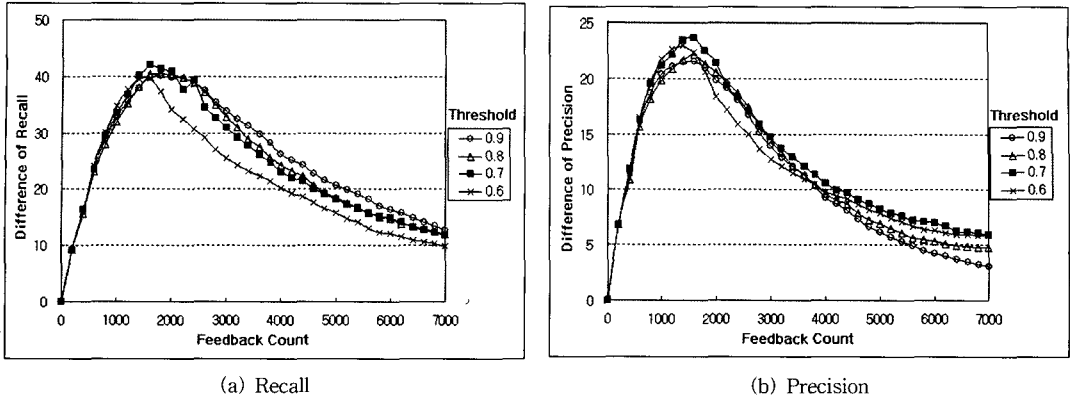


그림 4 Threshold^{EXTFB}의 결정 실험

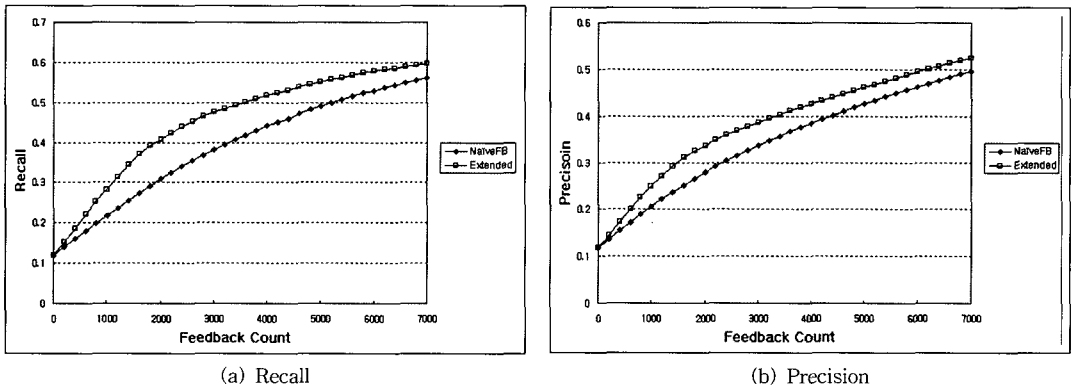


그림 5 피드백의 증가에 따른 재현율, 정확률

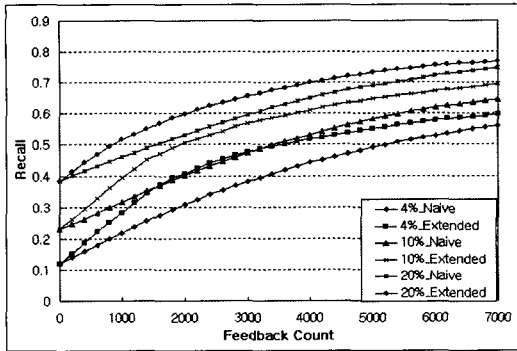
실험 1에서는 예비 실험에서 결정된 파라미터 값을 이용하여 ExtendedFeedback과 NaiveFeedback의 재현율과 정확률을 비교하여 평가한다. 그림 5는 피드백 횟수를 증가시키면서 재현율, 정확률을 측정 한 그래프이다. 이 때, 한 번의 피드백마다 6장의 이미지를 선택하였으며 각각 N^p 는 3, N_n 는 3로 하였다.

실험 결과에 의하면 피드백 횟수가 증가할수록 재현율과 정확률은 오르며 피드백 횟수가 같을 경우에는 ExtendedFeedback의 측정값이 NaiveFeedback보다 높은 값을 갖고 있다. 그림 5(a)를 보면 피드백 횟수가 1,000번에서 24,000번 사이일 경우는 ExtendedFeedback의 재현율은 NaiveFeedback보다 약 30%, 7,000번의 피드백을 받은 이후에는 약 7% 정도 더 높은 값을 갖고 있었다. 이것은 확장 피드백 방법을 사용하여 정확한 키워드의 신뢰도가 더욱 빠르게 조정되는 것을 나타낸다. 또한 그림 5(b)를 보면 피드백 800번에서 2,000번 사이에서 ExtendedFeedback의 정확률은 NaiveFeedback보다 약 20%정도 높고 7,000번의 피드백을 받은 이후는 약 6% 높다. 이것은 ExtendedFeedback에 의한 추가

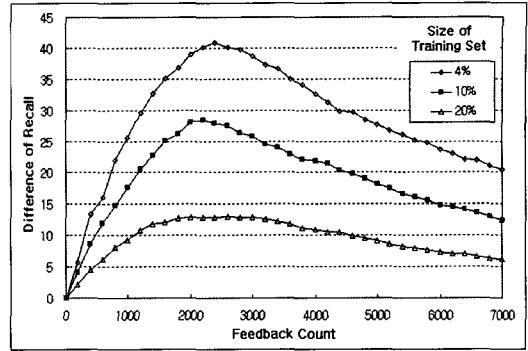
이미지를 선택할 때 정확한 키워드의 신뢰도가 증가함을 나타낸다.

실험 2: 트레이닝 집합의 크기변화에 따른 성능 평가
 트레이닝 집합의 크기가 클수록 자동 키워드는 정확하게 연결된다. 이것은 자동 키워드 연결 단계에서 유사한 특성을 지니는 정확한 이미지의 수가 많아지기 때문이다. 초기 자동 키워드의 상태에 따른 피드백의 반영 효과를 알아보기 위하여 트레이닝 집합의 크기를 각각 40, 100, 200으로 설정한 후 실험을 하였다. 40은 시스템에 사용되는 키워드를 지닌 이미지를 하나씩 선택하여 만들 수 있는 최소 트레이닝 집합의 크기이다. 100과 200은 일반적인 실험에서 사용하는 전체 이미지 수의 10%와 20%이다.

그림 6(a)와 그림 7(a)는 각 트레이닝 집합의 크기 변화에 따른 재현율과 정확률을 나타낸 것이며 그림 6(b)와 그림 7(b)는 재현율의 성능비교와 정확률의 성능비교를 나타낸 그래프이다. 그림 6(a)와 그림 7(a)를 보면 같은 서로 다른 세 가지 크기의 트레이닝 집합에서 ExtendedFeedback이 NaiveFeedback보다 좋은 값을 갖

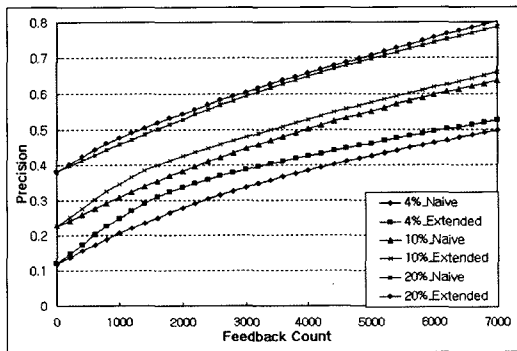


(a) Recall

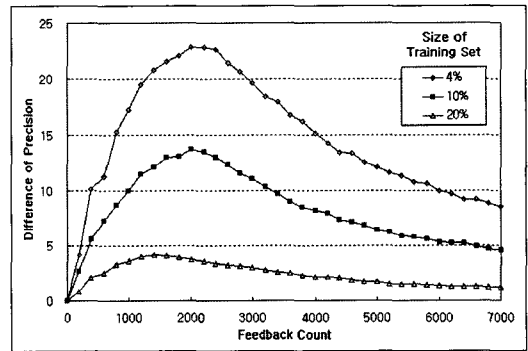


(b) Recall 비교

그림 6 트레이닝 집합 크기에 따른 재현율과 두 방법의 성능 차이



(a) Precision



(b) Precision 비교

그림 7 트레이닝 집합 크기에 따른 정확률과 두 방법의 성능 차이

으며, 그림 6(b)와 그림 7(b)에서 볼 수 있는 것과 같이 트레이닝 집합의 크기가 작을수록 두 방법 간의 큰 성능 차이를 갖는다. 이것은 트레이닝 집합의 크기가 작아 초기 데이터베이스의 키워드가 정확하지 않은 상태일수록 ExtendedFeedback에 의한 신뢰도 조정이 빠르게 진행되어 더 정확한 검색 성능을 가지고 있음을 나타낸다. 따라서 제안하는 피드백 확장 방법은 초기 데이터베이스에 있는 이미지들의 키워드 연결이 충분히 높은 신뢰도를 가지지 않는 환경에서도 좋은 성능을 보인다.

실험 3: 피드백 사이즈에 따른 실험

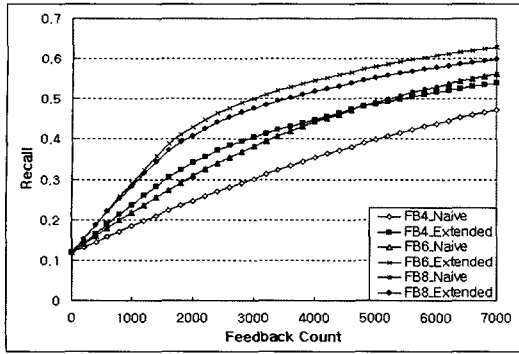
이미지의 신뢰도를 조절하기 위한 가장 이상적인 방법은 검색되는 모든 이미지에 피드백을 주는 것이다. 하지만 키워드에 의해 검색되는 1차 검색 이미지의 수가 많기 때문에 모든 이미지에 피드백 정보를 주는 것은 불가능하다. 일반적으로 사용자는 1차 결과에서 몇 개의 이미지만 피드백 정보로 시스템에 넘겨준다.

사용자에게 받는 피드백 크기에 따라 제안 방식의 성능 향상 정도를 알아보기 위하여 피드백 크기를 바꾸어 가며 성능 차이를 측정하였다. 피드백 크기를 4에서 8가

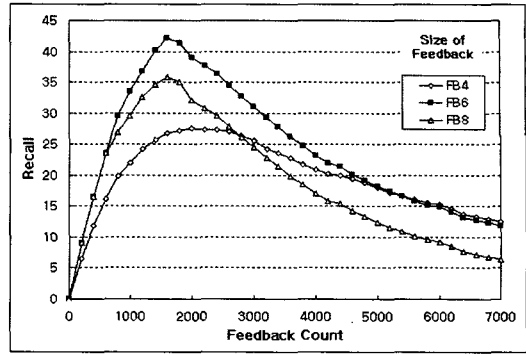
지 변화 시켜가면서 실험하였으며, 이때 N_p , N_n 는 피드백 크기의 반으로 하였다. 이때 사용하는 피드백 크기는 각각 전체 이미지 수의 0.4%, 0.6%, 0.8%인 4, 6, 8을 사용하였다.

그림 8, 9에 피드백 크기의 변화에 따른 결과를 나타내었다. 그림 8(a)와 그림 9(a)를 보면 같은 피드백 정보를 받은 경우 ExtendedFeedback이 NaiveFeedback보다 재현율과 정확률 모두 더 높은 값을 갖는다. 두 방법의 성능차이를 표현한 그림 8(b)와 그림 9(b)를 보면 피드백의 크기가 작을수록 두 방법의 측정값은 큰 성능 차이를 나타낸다.

이 결과에 따르면 1차 결과에서 남기는 피드백의 크기가 작을수록 ExtendedFeedback이 NaiveFeedback에 비해 좋은 성능을 보임을 알 수 있다. 1,000번의 피드백이 진행된 이후의 실험 결과를 보면 피드백을 4개씩 주었을 때 피드백을 8개 주었을 때보다 재현율은 약 15%, 정확률은 약 8%정도 높은 값을 갖는다. 사용자에게 받은 피드백의 크기가 작아도 ExtendedFeedback의 사용으로 인한 추가 이미지의 신뢰도 조정이 빠르게 이루어지고 있음을 나타낸다.

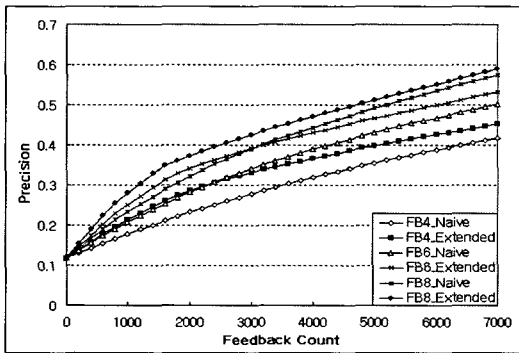


(a) Recall

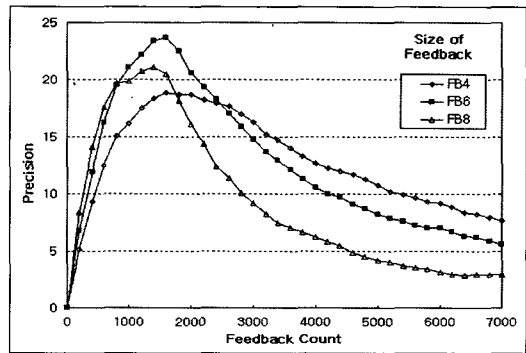


(b) Recall 비교

그림 8 피드백 크기에 따른 재현율과 두 방법의 차이



(a) Precision



(b) Precision 비교

그림 9 피드백 크기에 따른 정확률과 두 방법의 차이

5. 결론

본 논문에서는 적합성 피드백을 키워드 기반 검색에 적용하는 방식에 내용 기반 검색을 결합하는 모델을 제안하였다. 제안된 모델에서는 피드백을 받은 이미지들의 키워드뿐 아니라 긍정적 피드백을 받은 이미지들을 구분하는데 사용된 시각적 특성과 유사한 시각적 특성을 갖는 다른 이미지들의 키워드까지도 신뢰도 조정의 대상이 되도록 한다. 이를 통해 작은 피드백으로도 많은 이미지의 키워드 신뢰도를 조절할 수 있게 되므로, 궁극적으로는 좀 더 빠른 시간 내에 검색 결과의 정확도를 높일 수 있다는 장점을 가지게 된다.

본 논문의 주요 공헌을 요약하면 다음과 같다. 첫째, 키워드 기반 검색, 적합성 피드백, 내용 기반 검색을 모두 사용하여 빠른 시간 내에 이미지 검색 결과의 정확도를 높일 수 있는 새로운 검색 모델을 제안하였다. 둘째, 적합성 피드백 과정에 의해 선택된 긍정적 혹은 부정적 이미지들로부터 내용 기반 검색을 위한 최적의 가중치 유사도 함수를 자동적으로 결정할 수 있는 방안을 제시하였다.

본 논문에서 제안한 방법의 우수성을 검증하기 위하여 실험을 통한 성능 평가를 수행하였다. 실험 결과에 따르면 단순히 피드백을 받은 이미지만 신뢰도를 조정하는 방법에 비해서 본 논문에서 제안한 추가 이미지까지 신뢰도를 조정하는 피드백 확장 방법이 같은 횟수의 피드백을 했을 때 최고 40%까지의 재현율 증가를 보였고 정확률도 최고 20%의 증가를 보여 좋은 검색 결과를 제공하였다.

향후 연구로는 피드백에서 받은 이미지의 시각적 특성과 질의 키워드의 관련성을 조사하여 1차 검색 이미지를 재정렬할 때 긍정적 이미지와 부정적 이미지를 더 잘 구분 지을 수 있는 가중치 유사도 기법을 연구할 계획이다. 각 키워드를 이용하여 이미지를 검색할 때 다 사용자로부터 얻은 피드백정보를 저장, 분석하여 키워드와 이미지의 시각적 특성사이의 상관도(Correlation)를 추론할 수 있는 방법을 제시한다면 내용기반 검색과 키워드기반 검색을 모두 적용 가능한 이미지 검색 모델을 제시할 수 있을 것이다.

참고 문헌

- [1] S. Deb and Y. Zhang, "An Overview of CBIR Techniques," In Proc. the 18th International Conference on Advanced Information Networking and Applications, 2004.
- [2] J. Zachar, S. S. Iyengar, J. Barhen, "Content Based Image Retrieval and Information Theory: A General Approach," Journal of the American Society for Information Science and Technology, 2001.
- [3] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance Feedback: A Power Tool for Interactive Content-Based Image Retrieval," IEEE Transactions on Systems, Man, and Cybernetics, Vol. 8, pp. 644-655, 1998.
- [4] X. Zhu and T. S. Huang, "Relevance Feedback in Image Retrieval: A Comprehensive Review," Multimedia Systems, Vol. 8(6), pp. 536-544, 2003.
- [5] O. Marques and N. Barman, "Semi-Automatic Semantic Annotation of Images Using Machine Learning Techniques," In Proc. International Semantic Web Conference, 2003.
- [6] Y. Lu, C. Hu, X. Zhu, H. J. Zhang, and Q. Yang, "A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems," In Proc. ACM Multimedia, 2000.
- [7] W. Liu, S. Dumais, Y. Sun, H. Zhang, M. Czerwinski, and B. Field, "Semi-Automatic Image Annotation," In Proc. Human Computer Interaction, pp. 326-333, 2001.
- [8] M. Flinker, H. Samhey, W. Niblack et al., "Query by Image and Video Content: The QBIC System," IEEE Computer, Vol. 28, pp. 23-32, 1995.
- [9] J. R. Smith and S. F. Chang, "VisualSeek: A Fully Automated Content-Based Image Query System," In Proc. ACM International Conference on Multimedia, pp. 87-93, 1996.
- [10] J. R. Bach, C. Fuller, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. Shu, "The Virage Image Search Engine: An Open Framework for Image Management," In Proc. Storage and Retrieval for Still Image and Video Databases IV, SPIE, 1996.
- [11] P. J. Cheng and L. F. Chien, "Effective Image Annotation for Search Using Multi-Level Semantics," In Proc. International Conference of Asian Digital Libraries, 2003.
- [12] J. Jeon, V. Lavrenko, and R. Manmatha, "Automatic Image Annotation and Retrieval Using Cross-Media Relevance Models," In Proc. the 26th Annual International ACM SIGIR Conference, 2003.
- [13] H. Feng and T. Chua, "A Learning Based Approach for Annotating Large Online Image Collection," In Proc. 10th International Multimedia Modeling Conference, 2004.
- [14] L. Fei-Fei, R. Fergus and P. Perona, "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories," IEEE. CVPR 2004, Workshop on Generative-ModelBased Vision, 2004.
- [15] MPEG-7 visual part of the eXperimentation Model (version 9.0) January 2001, ISO/IEC JTC1/SC29/WG11 N3914.



송지영

2005년 2월 연세대학교 컴퓨터과학과 졸업(학사). 2007년 2월 연세대학교 컴퓨터과학과 대학원 졸업(석사). 관심분야는 이미지 검색, 멀티미디어 데이터베이스, 데이터 마이닝



김우철

1997년~2003년 B.S. 연세대학교 컴퓨터과학과. 2004년~2006년 M.S. 연세대학교 컴퓨터과학과(석사). 2006년~현재 연세대학교 컴퓨터과학과(박사). 관심분야는 바이오인포메틱스, LBS, 데이터베이스 보안, 멀티미디어 데이터베이스 등



김승우

2005년 2월 연세대학교 컴퓨터과학과 졸업(학사). 2005년 3월~현재 연세대학교 컴퓨터과학과 석사과정. 관심분야는 데이터 마이닝, 데이터베이스 보안 등

박상현

정보과학회논문지 : 데이터베이스
제 34 권 제 1 호 참조