

인간의 인지도에 근거한 질의를 통한 영상 검색의 성능 향상

(Performance Improvement of Image Retrieval System by Presenting Query based on Human Perception)

유 헌 우 [†] 장 동 식 ^{**} 오 근 태 ^{***}
(Hun-Woo Yoo) (Dong-Sik Jang) (Geun-Tae Oh)

요 약 영상간의 유사도는 일반적으로 영상으로부터 추출한 특징벡터간의 벡터공간상의 거리를 계산해서 판단한다. 그러나 이러한 특징벡터가 유사도 계산을 위한 하나의 방법이지만 항상 인간의 유사도 개념을 충실히 반영하지는 않는다. 그러므로 현존하는 대부분의 영상검색시스템들은 각 특징간의 중요도를 선정하여 유사도에 반영하는 방법을 사용하고 있다. 본 논문에서는 영상검색을 위한 새로운 초기 가중치 설정과 갱신 알고리즘을 제안한다. 이를 위해서 먼저 데이터베이스 영상을 인간의 인지도 판단에 의해 그룹화 한 후, 내부질의를 외부질의를 수행하고, 검색된 영상중 유사한 영상이 어느 그룹에 속하는지 알아내어 각 영상별로 유사도 계산에 필요한 최적 특징 가중치를 계산한다. 2000개의 영상데이터에 대한 실험을 통해서 제안된 알고리즘의 우수성을 보인다.

키워드 : 영상검색, 최적 가중치, 인간의 인지도, 그룹화

Abstract Image similarity is often decided by computing the distance between two feature vectors. Unfortunately, the feature vector cannot always reflect the notion of similarity in human perception. Therefore, most current image retrieval systems use weights measuring the importance of each feature. In this paper new initial weight selection and update rules are proposed for image retrieval purpose. In order to obtain the purpose, database images are first divided into groups based on human perception and, inner and outer query are performed, and, then, optimal feature weights for each database images are computed through searching the group where the result images among retrieved images are belong. Experimental results on 2000 images show the performance of proposed algorithm.

Key words : image retrieval, optimal feature weight, human perception, grouping

1. 서 론

인터넷 기술의 발전으로 가정에서도 손쉽게 외부의 정보를 습득할 수 있는 시대가 도래하였다. 이러한 인터넷 기술의 근간 중에는 멀티미디어 데이터의 내용

(Contents)을 생성, 전송, 검색하기 위한 관련기술이 하나의 부류로 자리잡고 있다. 본 논문에서는 이러한 멀티미디어 데이터 중에서 영상에 초점을 맞추어 검색해 내는 알고리즘을 제안한다.

영상을 검색하는 초기의 방법은 사용자가 찾고자 하는 영상에 대한 키워드를 제시하면 검색시스템이 해당 키워드를 포함하는 영상을 검색해 보여준다. 예를 들어 숲속에 있는 호랑이를 나타내는 영상을 찾고자 할 때 키워드 "숲속 호랑이" 등을 입력하면 데이터베이스에서 관련 영상을 검색해 낸다. 이와 같은 텍스트기반검색(Text-based retrieval)[1, 2]은 정확도 면에서 어느 정도 신뢰할 수 있지만, 색인자가 일일이 모든 데이터베이스 영상에 대해 색인 키워드를 작성

본 논문은 2002년도 학술진흥재단의 지원에 의해서 연구되었음
(KRF-2002-005-H20002)

[†] 정 회 원 : 연세대학교 인지과학연구소 교수
paulyh@yonsei.ac.kr

^{**} 정 회 원 : 고려대학교 산업시스템정보공학과 교수
jang@korea.ac.kr

^{***} 비 회 원 : 수원대학교 산업정보공학과 교수
gtoh@mail.suwon.ac.kr

논문접수 : 2002년 5월 22일

심사완료 : 2002년 11월 19일

해야 하고, 때로는 텍스트만을 가지고는 표현하기 어려운 영상들도 있다. 또한 영상에 따라서 색인자와 검색자의 관점이 틀리기 때문에 추후 검색할 때는 원하는 영상을 찾기가 어려운 점도 있다. 따라서 이러한 단점을 극복하기 위해서 텍스트 외에 영상 자체의 내용(Contents)을 추출/색인하여 객관적으로 검색하는 방법이 연구되고 있다. 이 방법은 컴퓨터 비전기술을 이용하여 영상을 자체적으로 분석하고 색상, 질감, 모양, 또는 영상내의 객체위치의 상호관계를 자동적으로 추출하여 검색에 이용한다.

초기의 내용기반 검색은 영상의 색상, 질감, 모양 등의 정보를 추출하여 이러한 정보가 비슷한 영상을 순서대로 검색해 내는 방법을 이용하였다. 영상의 내용에 기반한 검색은 모든 영상을 작업자가 일일이 색인 하는 부담을 줄일 수 있고, 관점의 불일치로 인한 검색의 실패를 어느 정도 줄일 수 있어서 텍스트기반 방식의 보완차원에서 하나의 대안으로 해석되고 있다. 그러나 이 방법이 객관적이고 손쉽게 영상을 검색해 낼 수 있지만 아직까지는 텍스트 방식에 의한 방법보다는 검색의 정확도면에서 많이 뒤지고 있다. 그 이유는 영상의 내용, 즉, 색상, 질감, 모양 등의 정보는 저차원의 영상특징에 불과하고 인간이 영상을 이해하는 고차원적인 특징을 표현하지는 않는다[3, 4, 5]. 예를 들어, 파란 하늘에 새가 날고있는 영상과 비행기가 떠있는 영상은 인간의 고차원적인 인지개념에서는 두 영상이 전혀 다른 영상임을 판단할 수 있지만, 저차원적인 특징에 근거해 판단할 때 만일 새와 비행기가 어느 정도 유사한 색상이라면 전혀 다른 영상이라고 구별 짓기가 어려워진다. 따라서 이와 같은 내용기반 검색의 한계를 극복하기 위하여 검색과정 중에 인간의 인지능력과 경험을 포함시키는 유사도 피드백(Relevance feedback) 방법에 대한 연구가 최근에 활발히 진행되고 있다. 이 방법은 사용자가 예제 영상을 사용하여 질의를 하게 되면 예제 영상의 색상, 질감, 모양 등의 정보가 유사한 영상을 데이터베이스에서 검색해 보여주게 된다. 그 후에는 검색된 영상들에 대해 인간의 유사성 판단에 의해 각 영상들을 유사하다(Relevant), 유사하지 않다(Non-relevant) 등의 사용자의 판단을 체크(Check)하고 재차 검색을 하게 되면 내부적인 알고리즘에 의해 2차 검색 때는 질의가 다시 정제되어(Query refining) 보다 인간의 인지도면에서 유사한 영상들을 얻게 된다. 이와 같은 체크 과정과 질의 과정을 여러 번 반복해서 사용자가 원하는 많은 유사한 영상들을 얻는 방법이다.

유사도 피드백에 관한 연구는 크게 질의점 이동(query point movement)방법, 가중치갱신(reweighting)방법, 그리고 통계적 기법이나 신경망 이론을 이용하는 방법으로 요약될 수 있다. 질의점 이동방법은 초기의 검색결과를 얻은 후, 질의와 유사한 영상으로 질의를 이동하거나 혹은 질의와 유사하지 않은 영상의 경우는 멀어지게 하는 기법으로 이상적인 질의를 찾는 방법이다[6, 7, 8, 9]. 가중치 갱신방법은 각각의 영상에서 n 차원 특징벡터를 추출하여 n 차원 공간에서 하나의 점으로 표현한 후, 초기의 검색결과로 질의와 유사하다고 생각되는 영상들이 어떤 축을 따라 넓게 분포되어 있으면(분산이 크면) 해당 축은 입력질의와 유사하지 않다고 판단할 수 있으므로 해당 축에 낮은 가중치를 주고 반대의 경우는 높은 가중치를 주는 방법이다[4, 5, 10, 11, 12, 13, 14]. 통계적 기법이나 신경망 이론을 이용하는 방법은 찾고자 하는 목표영상을 확률모형에 의해 예측을 하거나 패턴인식에서 많이 사용하는 신경망을 통해서 영상을 분류하는 방법이다[3, 6, 12, 15, 16, 17, 18].

2. 유사도 피드백

초기의 영상검색방법들은 영상의 기본 특징인 색상, 질감, 모양 등의 정보를 추출하고 이 정보간에 상호 중요도에 따라 고정된 가중치를 주어 유사도를 비교 검색하였다. 이 때 각 특징간에는 최대값과 최소값의 범위가 틀리므로 정규화 과정을 거쳐게 된다. 각 특징의 정규화를 끝내면 주어진 데이터베이스에 대해 경험적으로 가중치를 선정하여 유사도 계산에 사용한다. 일반적인 검색시스템에서는 영상간의 유사도 판단시 모양보다는 색상에 더 좌우가 되므로 색상에 보다 가중치를 둘 수 있지만, 트레이드마크 검색 같은 분야에서는 색상보다는 모양이 더욱 중요하므로 모양에 더 가중치를 둘 수 있다. 그러나 어느 특징에 어느 정도의 가중치를 주어야 하는지의 기준이 없기 때문에 정확한 검색을 하는 데에 어려움이 있다. 본 단원에서는 주어진 데이터베이스에 대해 각 영상을 가지고 질의할 때 영상마다 사용된 특징간의 가중치를 최적으로 선택하는 방법을 제시한다. 본 논문에서는 간략한 설명을 위해 색상과 질감특징만이 사용되었다고 가정한다. 그러나 기타 다른 특징들의 사용도 손쉽게 확장 가능하다.

2.1 초기 가중치 선정(데이터베이스내의 영상)

각 영상마다 최적의 가중치는 그 최적의 가중치를 가지고 질의했을 때 유사한 영상이 가장 많이 나오게

되는 것을 의미한다. 데이터베이스, 그리고 색상과 질감특징이 주어졌다는 가정 하에서, 인간의 인지도에 의한 적절한 초기 가중치를 아래와 같이 계산한다. N 개의 데이터베이스 영상을 $X=(x_1, x_2, x_3, \dots, x_N)$ 이라고 하고 이 영상들을 인간의 인지에 의해 N 개로 군집화한 집합을 $X=(C_1, C_2, C_3, \dots, C_M)$ 이라 하자. I 번째 군집군 C_I 가 K 개의 영상으로 이루어질 경우 $C_I=(x_{I1}, x_{I2}, x_{I3}, \dots, x_{IK})$ 로 표시하자.

Step 1: N 개의 데이터베이스 영상 집합 X 를 인간의 인지에 의해 군집화 한다(Perceptual grouping). 여기서 군집화는 데이터베이스를 구축하는 색인자의 주관에 의한 군집화를 의미하고 중복(overlap)을 허용한다(예를들어, 한 개의 영상에 하늘과 바다와 배가 동시에 존재한다면 이 영상은 각각 3개의 군집군에 공통적으로 속할 수 있다).

Step 2: 데이터베이스에서 질의 영상 x_Q 를 선택한다 [x_Q 가 군집화 C_Q 에 속한다고 가정하자. 혹은 x_Q 가 중복 영상이어서 Step 1의 예와 같이 하늘, 바다, 배(순서대로 군집화 C_Q, C_R, C_S 에 속한다고 가정하자)].

Step 3: x_Q 영상과 유사한 영상을 색상특징을 가지고 계산해서 유사도 순서대로 L 개를 추출한다. 추출된 L 개의 영상 중에 군집군 C_Q 에도 공통적으로 포함되어 있는 영상의 개수 v (중복영상일 경우 C_Q 에 공통적으로 포함되어 있는 영상 수 $+C_R$ 에 공통적으로 포함되어 있는 영상 수 $+C_S$ 에 공통적으로 포함되어 있는 영상 수 $= v$)를 구한다(v 가 클수록 색상특징이 중요함을 알 수 있다).

Step 4: x_Q 영상과 유사한 영상을 질감특징을 가지고 계산해서 유사도 순서대로 L 개를 추출한다. 추출된 L 개의 영상 중에 군집군 C_Q 에도 공통적으로 포함되어 있는 영상의 개수 w (중복영상일 경우 C_Q 에 공통적으로 포함되어 있는 영상 수 $+C_R$ 에 공통적으로 포함되어 있는 영상 수 $+C_S$ 에 공통적으로 포함되어 있는 영상 수 $= w$)를 구한다(w 가 클수록 질감특징이 중요함을 알 수 있다).

Step 5: 질의영상 x_Q 에 대해 색상과 질감의 가중치를 상대적인 비율을 고려하여 각각 $v/(v+w)$ 와 $w/(v+w)$ 로 결정한다.

Step 6: x_Q 를 제외한 데이터베이스내의 모든 영상에 대해 Step2-Step 5의 과정을 반복한다.

위와 같은 과정을 통해서 주어진 데이터베이스와 주어진 특징 내에서 데이터베이스에 있는 각 영상들에 대해 색상특징과 질감특징의 적절한 초기 가중치

를 얻을 수 있다. 예로써 만약 데이터베이스가 $X=(A,B,C,D,E,F,G,H,I,J,K,L,M)$ 의 13개의 영상을 가지고 있고, 이러한 X 를 인간의 인지도에 의해 2개의 그룹 $C_1=(A,B,C,D,E,F)$, $C_2=(G,H,I,J,K,L,M)$ 으로 나누었다고 가정하자. 질의 영상을 그룹 C_1 에만 속해 있는 E 로 선정하면 영상 E 의 색상과 질감 특징치의 최적 가중치의 계산은 아래와 같은 방법을 통해서 구한다. E 를 질의 영상으로 색상특징만을 이용하여 유사도 계산시 처음 L (5개로 정하면)개가 순서대로 $\{E,A,B,C,I\}$ 이고, 질감특징만을 이용하여 유사도 계산시 처음 L (5개로 정하면)개가 순서대로 $\{E,B,I,J,L\}$ 이 나왔다고 가정하면, E 가 속한 그룹 C_I 과 색상질의 결과인 $\{E,A,B,C,I\}$ 과의 중복된 영상은 $\{E,A,B,C\}$ 의 4개 ($v=4$)가 되고, 질감질의 결과인 $\{E,B,I,J,L\}$ 와의 중복된 영상은 $\{E,B\}$ 의 2개 ($w=2$)개가 되므로 색상의 가중치는 $4/(4+2)$ 이고 질감의 가중치는 $2/(4+2)$ 이다. 따라서 영상 E 를 가지고 질의를 할 때 각각 색상 가중치 0.67, 질감가중치 0.33을 사용하여 계산하면 된다. 상기의 과정은 그림 1에 간략히 나타나 있다.

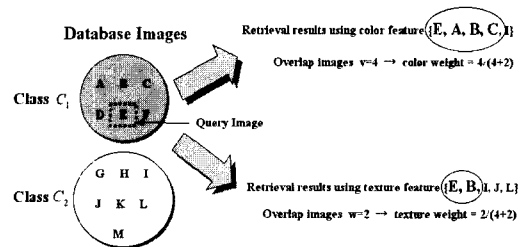


그림 1 적절한 초기 가중치 계산방법

만약 데이터베이스가 $X=(A,B,C,D,E,F,G,H,I,J,K,L,M)$ 중에서 인지도에 의해 4개의 그룹 $C_1 = \{A,B,C,D\}$, $C_2 = \{A,E,F,G\}$, $C_3 = \{H,I,J\}$, $C_4 = \{K,L,M\}$ 으로 공통적으로 속한 A 를 질의로 삼을 경우로 나누는 상태에서 2개의 그룹 C_1 과 C_2 에 가중치 계산은 아래와 같다. 색상과 질감을 이용한 처음 5개의 검색결과가 순서대로 $\{A,B,E,G,H\}$ 와 $\{A,B,C,J,K\}$ 이면 색상과 질감을 이용한 중복영상은 그룹 C_1, C_2 를 모두 고려해서 각각 $\{A,B,E,G\}$ 와 $\{A,B,C\}$ 이므로 색상 가중치는 $4/(4+3)$ 이고 질감가중치는 $3/(4+3)$ 이 된다.

상기의 방법은 데이터베이스 영상을 미리 사용자가 분류함으로써 비효율적인 면이 있을 수 있지만 인간의 인지도에 의한 분류는 보다 높은 검색성능을 얻을 수 있는 장점이 있다. 유사한 방법으로 캘리포니아 대학[19]에서도

유사영상들을 미리 군집화 한 후 텍스트와 내용(특징벡터)을 색인하고 추후 검색시에는 먼저 유사한 군집군을 텍스트로 찾은 후 각 군집군에서 영상의 저차원 특징벡터의 유사성을 가지고 최종 검색 결과를 얻는 2단계 검색방법이 사용되었다.

2.2 초기 가중치 선정(데이터베이스 외부의 영상)

데이터베이스에 있지 않는 외부의 영상을 질의로 한 경우에는 이 영상과 데이터베이스의 영상들간의 색상과 질감에 대한 어떠한 상관정보도 없기 때문에 어떠한 초기 가중치를 가지고 유사도를 계산할지를 결정해야 한다. 이전 단락에서는 주어진 데이터베이스내의 모든 영상들에 대한 질의를 오프라인에서 수행한 후 검색결과를 가지고 적절한 가중치를 선택할 수 있어서 결과로 유사한 영상을 얻는 방법이였지만, 이제 데이터베이스 외부의 영상을 질의로 할 경우는 온라인 상에서 실행되어야 하므로 인지도에 의해 적절한 특징들의 가중치를 얻는 것이 쉽지 않다. 그러나 이전 단락에서 데이터베이스영상을 M개의 군집군으로 나누어 놓았으므로 각 군집군의 중심값(평균)과 분산(variance)를 이용해서 질의와 가장 근접한 군집군을 찾고 해당 군집군의 멤버(해당군집군에 속한 영상)들의 색상 가중치의 평균과 질감 가중치의 평균을 질의 영상의 색상과 질감에 대한 초기 가중치로 삼는다.

2.3 가중치 갱신

단락 2.2에서 설명한 방법을 통해 얻은 질의 x_0 의 초기의 색상 가중치를 w_c , 초기의 질감 가중치를 w_l 라 하면, 이 가중치를 가지고 질의하여 L개의 결과를 얻는다. 각 결과 영상은 데이터베이스 내에서 인지도에 의해 군집화 방법을 통해서 검색된 것이므로 단순히 군집화를 하지 않고 검색한 결과보다 좋은 성능을 나타낼 것이다. 그러나 초기의 w_c 와 w_l 는 아직 질의와 데이터베이스간의 완전한 인지도에 의한 가중치가 아니기 때문에 보다 정확한 결과를 얻기 위해서 아래의 방법을 통해 가중치를 갱신한다.

Step 1: L개의 초기 검색결과를 인간의 인지도에 의해 질의 영상과 유사하면 '+'를 체크(check)하고 유사하지 않거나 판단하기가 어려우면 체크를 하지 않는다.

Step 2: Step 1에서 '+'로 체크한 영상들의 군집군을 확인해서 해당 군집군의 가중치를 증가시킨다. 예를 들어 '+'체크 영상이 3개고 각각 군집군 C_1, C_2, C_3 에 소속해 있으면 w_{c1}, w_{c2}, w_{c3} 를 1씩 증가시킨다.

Step 3: '+'체크 영상의 색상 가중치 평균 w_c^{avg} 와 질감 가중치 평균 w_l^{avg} 를 구한 후 최초 질의 x_0 의 색상 가중치 w_c , 질감가중치 w_l 와 비교해서 차이만큼(혹은 아래의 식 (1), (2)와 같이 차이의 반만큼)을 증감하

여 갱신된 새로운 가중치 w_c, w_l 를 얻는다.

$$w_c = w_c + \frac{|w_c - w_c^{avg}|}{2}, \text{ if } w_c \leq w_c^{avg} \tag{1}$$

$$w_c = w_c - \frac{|w_c - w_c^{avg}|}{2}, \text{ otherwise} \tag{2}$$

(w_l 도 동일하게 적용)

Step 4: Step 3에서 구한 w_c, w_l 를 가지고 다시 질의하여 L개의 검색결과를 얻고 Step 1 - Step 4의 과정을 계속 반복한다(반복중단 조건은 더 이상의 검색결과 영상이 변하지 않거나 사용자가 일정횟수 반복 후에 멈출 수 있다).

예로써 만약 데이터베이스가 $X = \{A, B, C, D, E, F, G, H, I, J, K, L, M\}$ 의 13개의 영상을 가지고 있고, 이러한 X를 인간의 인지도에 의해 2개의 그룹 $C_1 = \{A, B, C, D, E, F\}$, $C_2 = \{G, H, I, J, K, L, M\}$ 으로 나누었다고 가정하자. 질의 영상을 데이터베이스 X에 있지 않는 외부영상 P로 선정하면 군집군 C_1, C_2 중 영상데이터의 색상과 질감의 중심값과 가장 가까운 군집군을 찾고 만약 거리가 같으면 분산값이 큰 군집군을 찾는 방법으로 P와 가장 유사한 군집군을 찾는다. 여기서 가깝다는 말은 벡터공간상에서 유클리디안 거리가 가장 작음을 의미하고 색상과 질감의 거리비율을 동일하게 하여서 계산한다. 만약 C_1 이라고 가정하면 C_1 내의 6개의 영상에 대해 색소가중치의 평균 w_c 와 질감가중치의 평균 w_l 를 외부영상 P의 초기 가중치로 삼는다. 각각 $w_c = 4.2$, $w_l = 3.1$ 로 가정하자. 이 w_c 와 w_l 를 가지고 질의 영상 P에 대한 유사영상을 전체 데이터베이스에서 계산했을 때 처음 L(5개로 정하면)개가 순서대로 {C, E, K, M, L}이 검출되고 이중 C(C_1 에 속함), K, M(모두 C_2 에 속함)이 질의와 유사하여 '+'체크를 했다면, 색소가중치 w_c 와 질감가중치 w_l 의 갱신은 아래와 같은 방법을 통해서 구한다. 영상 {C, K, M}의 색상평균 가중치가 5.4, 질감평균 가중치가 2.1로 계산되었다고 한다면 질의 영상 P의 새롭게 갱신된 가중치는 색상의 경우는 $4.2 + (|4.2 - 5.4|)/2$ 가 되고 질감의 경우는 $3.1 + (|3.1 - 2.1|)/2$ 이 된다. 이 새로운 가중치를 가지고 다시 질의를 수행하면 1차 질의 결과보다 향상된 결과를 얻을 수 있다. 상기의 과정은 그림 2에 간략히 나타나 있다.

2.4 특징값과 유사도 함수

본 논문에서는 영상을 표현하는 특징 값으로 색상과 질감을 이용하였다. 색상특징값으로는 우리의 이전 연구에서 사용했던 것처럼 영상의 국부적 위치로부터 색상(hue), 채도(saturation), 명도(value)를 이용했다[10, 20]. 먼저 영상을 3×3의 영역으로 분할한 후 각각의 영역에

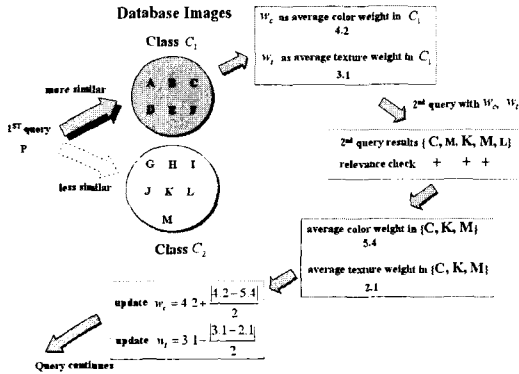


그림 2 외부영상질의시 가중치 갱신방법

서 가장 중요한 색상, 명도, 채도값을 얻는다. 이러한 특징값은 RGB영상이 주어지면 HSV영상으로 변환하고 노이즈에 강인하고 데이터 양을 줄이기 위해 색상, 채도, 명도를 각각 18, 3, 3개로 양자화 한 후 명도축에 있는 5개의 그레이 값을 합산하여 얻는다. 즉 모든 색상은 이러한 167개의 색상(= $18 \times 3 \times 3 + 5$) 중 하나로 표시된다.

질감특정도 색상과 마찬가지로 국부적인 3×3 영역에서 얻는다. 먼저 각 영역에서 그레이레벨 코어커런스 행렬을 구한 후 이로부터 5개의 유효한 특징값(2차모멘트, 콘트라스트, 상관계수, 분산, 엔트로피)을 사용한다[20].

상기의 과정을 통해서 각 영역에서 색상과 질감특징을 얻으면 영상간의 유사도는 이러한 색상과 질감특징값의 유클리디안 거리의 차를 합산하여 계산한다[20].

3. 실험

제안된 알고리즘의 적합성을 보이기 위해서 데이터베이스 내부에 있는 영상과 외부영상을 가지고 질의를 실시하였다. 실험에서 사용된 영상은 총 2000장이고 이것을 인간의 인지도에 의해 24개로 그룹화하여 데이터베이스에 색인 하였다.

3.1 내부질의

내부질의를 위해 사용한 영상은 각 그룹에서 5개를 발채하여 총 120(=24×5)의 영상을 사용하였고, 알고리즘의 성능을 보이기 위해 정확도(Precision)과 회수율(Recall)을 계산하였다. 정확도와 회수율을 계산할 때는 질의 영상과 데이터베이스영상간에 유사도 함수에 의해 유사도를 계산한 후 가장 유사한 처음 10개의 결과 영상을 얻는 방법을 모두 120개의 영상에 대해 실시한 후 결과를 평균한 것이다. 그림 3은 색상특징만을 가지고

검색한 결과, 질감특징만을 가지고 검색한 결과, 색상과 질감의 가중치를 각각 1:1과 3:1로 주어진 상태에서의 결과를 보여준다.

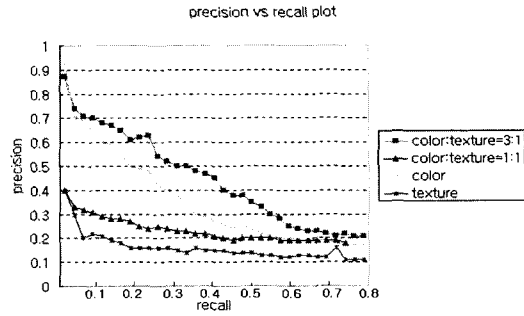


그림 3 평균정확도와 평균회수율 곡선(색상과 질감특징을 각각 하나씩 사용하거나, 1:1의 가중치를 결합해서 사용했을 경우, 3:1의 가중치를 결합해서 사용했을 경우)

그림을 보면 일반적인 자연영상에 대해서는 인간의 유사도 판단시 색상이 질감보다 더 중요하다는 것을 알 수 있다. 색상과 질감의 비율을 1:1로 했을 경우는 질감보다는 좋지않은 색상이 질감보다는 오히려 떨어짐을 알 수 있다. 이것은 아직 각각의 특징에 대해 좋은 가중치를 선정하지 못했기 때문이다. 그러나 색상에 보다 가중치를 주어서 3:1의 가중치를 가지고 검색했을 경우는 보다 향상된 성능을 얻을 수 있었다.

그림 4는 제안된 알고리즘을 적용하여 데이터베이스내의 질의 영상들의 가중치를 가지고 유사도를 계산한 결과이다. 그림 3의 3:1가중치와 비교하여 보다 향상된 결과를 보여준다. 이 결과는 데이터베이스 색인시 미리 영상들에 대해 특징 별로 최적의 가중치를 계산해 놓았기 때문이다.

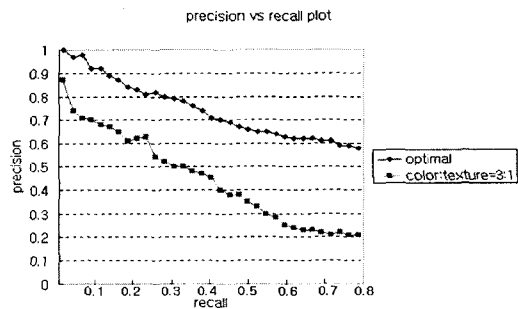


그림 4 평균정확도와 평균회수율 곡선(색상:질감=3:1의 가중치와 최적가중치를 사용했을 경우)

3.2 외부질의

데이터베이스 외부에 있는 영상들을 질의로 사용할 경우 색상과 질감에 대해서 데이터베이스 내부 영상과는 어떠한 상관관계가 없기 때문에 최적가중치를 찾을 수가 없다. 따라서 먼저 질의가 시작되면 질의와 유사한 데이터베이스내의 25개 그룹가운데서 특징들의 중심값과 분산을 고려하여 가장 유사한 그룹을 찾는다. 그 후 그 그룹내의 영상의 색상과 질감가중치의 평균을 최초가중치로 삼는다. 이 초기 가중치를 가지고 여러 번 질의를 반복하면 보다 정확한 검색결과를 얻을 수 있다.

실험에 사용된 외부영상은 총 120개이며 이 각각의 영상은 인지도 면에서 25개의 그룹 중에 하나에 속한 것들로 택했다. 그림 5는 이 120개의 영상에 대한 정확도와 회수를 곡선이다. 그림에서 보듯이 최초의 결과보다 반복횟수가 늘어남에 따라 좋은 결과를 얻을 수 있음을 알 수 있다. 그림 6은 한 가지 예로 '치타' 영상을 질의로 하여 유사도를 계산했을 때 최초 검색 시에는 6개의 유사한 영상을 얻었지만 3번의 반복 후에는 8개의 유사한 영상을 얻을 수 있음을 보여준다.

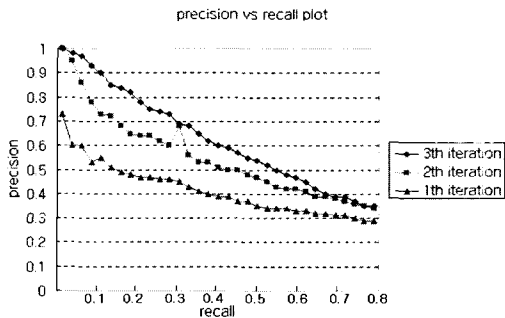


그림 5 질의 횟수의 증가에 따른 평균정확도와 평균회수율 곡선



(a) 최초질의의 결과



(b) 3번의 반복질의의 후 결과

그림 6 '치타' 질의 영상검색결과

제안된 알고리즘을 Y. Rui[4]의 논문과 비교하였다. 유사한 조건에서 비교하기 위해서 Rui의 알고리즘에는 제안된 알고리즘에서 사용된 것과 같은 동일한 색상과 질감정보 2개만을 이용하였고, 유사도 비교방법은 유클리디안 거리를 이용하였다. 피드백의 반복횟수는 3회로 제안하였다. 그림 7은 비교 결과를 나타낸다. 그림에서 보듯이 내부질의는 Rui의 방법보다 우수함을 알 수 있고, 외부질의의 경우도 근소하지만 우수한 성능을 나타내고 있다. 이러한 이유는 데이터베이스 색인시 미리 영상들에 대한 특징 별로 적절한 가중치를 계산해 놓았기 때문이다.

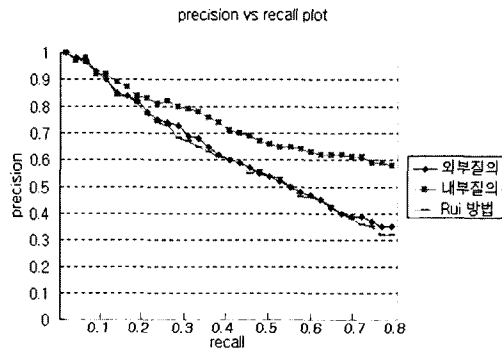


그림 7 평균정확도와 평균회수율 곡선(외부질의, 내부질의와 Rui의 방법을 사용했을 경우)

4. 결론

영상검색은 멀티미디어 시대의 중요한 부분을 차지한다. 본 논문에서는 영상의 내용에 기반한 검색에서 사용되는 특징들간의 최적가중치 선정 알고리즘을 제안했다. 사용된 특징으로는 HSV 색상모델과 코-오커런스 행렬로부터의 5가지 질감특징들을 영상의 국부적 위치로부터 추출하여 이용했고 유사도 검색시 유클리디안 거리를 이용하여 검색하였다. 질의 방법으로는 데이터베이스 내부의 영상에 의한 질의와 외부영상에 의한 질의를 사용하였다.

최적가중치를 선정하기 위하여 먼저 데이터베이스 색인시 색인된 영상들을 인간의 인지도에 의해 그룹화 하고 (Perceptual grouping), 각 영상들의 내부 질의를 통해서 최적의 가중치를 계산하여 색인 하였다. 실험결과 계산된 최적의 가중치를 가지고 검색했을 경우 각각 색상 혹은 질감 정보만을 이용하거나, 색상과 질감이 1:1 과 3:1의 가중치를 가지고 검색했을 경우보다 좋은 검색 결과를 얻을 수 있었다. 이 결과는 미리 각 영상들에 대

한 최적의 가중치를 계산해 놓았기 때문이다.

데이터베이스 외부에 있는 영상들을 질의로 사용할 경우는 먼저 가장 유사한 그룹을 찾고 그 그룹내의 영상들의 색상과 질감가중치를 평균하여 최초 가중치로 삼았다. 이후 제안된 가중치 갱신알고리즘에 의해서 여러 번 반복질의를 수행함에 따라 보다 향상된 검색결과를 얻을 수 있었다.

그러나 본 논문은 아직 제안된 알고리즘의 가능성을 보여줄 뿐이다. 보다 많은 영상(>2000)들에 대한 실험 결과가 필요하고, 잘 알려진 안정된 알고리즘과의 보다 폭넓은 비교도 필요할 것으로 생각된다. 비록 오프라인(off-line)이지만 영상을 미리 군집화 하는 것도 부담스러운 면이 될 수 있다. 제안된 알고리즘이 보다 효과적이기 위해서는 사용되는 특징벡터의 중요성은 말할 필요도 없다. 따라서 인간에 유사한 인지능력을 표현할 수 있는 특징을 찾아내는 것도 지속적인 연구과제로 생각된다.

참 고 문 헌

- [1] N.S. Chang and K.S. Fu, "Query-by pictorial-example," *IEEE Trans. Software Eng.* vol SE-6, no. 6, pp. 519-524, 1980.
- [2] S.-K. Chang, C.W. Yan, D.C. Dimitroff, and T. Arndt, "An intelligent image database system," *IEEE Trans. Software Eng.* vol 14, no 5, pp. 681-688, 1988.
- [3] M. Li, Z. Chen, H.J. Zhang, "Statistical correlation analysis in image retrieval," *Pattern Recognition*, 2002 (To appear).
- [4] Y. Rui, T.S. Huang, M. Ortega, S. Mehrotra, 1988, "Relevance feedback : A power tool in interactive content-based image retrieval," *IEEE Trans. Circuits and Systems Video Technology* vol 8, no 5, pp. 644-655, 1998.
- [5] S. Aksoy, R.M. Haralick, F.A. Cheikh, M. Gabbouj, "A weighted distance approach to relevance feedback," *Proc. IAPR Int. Conf. Pattern Recognition*, pp. 812-815, 2000.
- [6] I.J. Cox, M.L. Miller, T.P. Minka, T.V. Papathomas, P.N. Yianilos, "The Bayesian image retrieval system, PicHunter : theory, implementation and psychophysical experiments," *IEEE Trans. Image Process.* vol 9, no 1, pp. 20-37, 2000.
- [7] S. Sclaroff, L. Taycher, M. La Cascia, "Image Rover: a content-based image browser for the World Wide Web," *Proc. Workshop on Content-Based Access of Image and Video Libraries*, pp. 2-9, 1997.
- [8] M.E.J. Wood, N.W. Campbell, B.T. Thomas, "Iterative refinement by relevance feedback in content-based digital image retrieval," *Proc. Sixth ACM Multimedia Conference*, pp. 13-20, 1998.
- [9] N. Vasconcelos, A. Lippman, "Learning from user feedback in image retrieval systems," *Proc. Neural Information Processing Systems* 12, pp. 977-983, 1999.
- [10] H.W. Yoo, D.S. Jang, S.H. Jung, J.H. Park, and K.S. Song, "Visual Information Retrieval via Content-based Approach," *Pattern Recognition*, vol 35, no. 3, pp. 749-769, 2002.
- [11] K. M. Lee and W. N. Street, "Incremental feature weight learning and its application to a shape-based query system," *Pattern Recognition Letters*, 2002 (To appear).
- [12] Y. Ishikawa, R. Subramanya, C. Faloutsos, "MindReader: query databases through multiple examples," *Proc. 24th Int. Conf. Very Large Databases*, pp. 218-227, 1998.
- [13] Y. Lu, C. Hu, X. Zhu, H.J. Zhang, Q. Yang, "A unified framework for semantics and feature based relevance feedback in image retrieval systems," *Proc. Eighth ACM Multimedia Conference*, pp. 31-38, 2000.
- [14] R. Schettini, G. Ciocca, I. Gagliardi, "Content-based color image retrieval with relevance feedback," *Proc. Int. Conf. Image Processing*, pp. 75-79, 1999.
- [15] C.S. Lee, W.Y. Ma, H.J. Zhang, "Information embedding based on user's relevance feedback for image retrieval," *Proc. SPIE, vol. 3846 (Multi-media Storage and Archiving Systems IV)*, pp. 294-304, 1999.
- [16] J. Laaksonen, M. Koskela, S. Laakso, and E. Oja, "Self-Organising Maps as a Relevance Feedback Technique in Content-Based Image Retrieval," *Pattern Analysis & Applications* vol 4, pp. 140-152, 2001.
- [17] T.P. Minka and R.W. Picard, "Interactive learning with a 'society of models'," *Pattern Recognition*, vol. 30, pp. 565-587, 1997.
- [18] J. Peng, B. Bhaunu, S. Qing, "Probabilistic feature

relevance learning for content-based image retrieval," *Computer Vision and Image Understanding* vol 75 no 1/2, pp. 150-164, 1999.

- [19] S. Newsam, B. Sumengen, and B.S. Manjunath, "Category-Based Image Retrieval," *Proc. Int. Conf. Image Processing, Special Session on Multimedia Indexing, Browsing and Retrieval*, vol. 3, pp. 596-599, 2001.
- [20] H.W. Yoo, S.H. Jung, D.S. Jang, and Y.K. Na, "Extraction of Major Object Features Using VQ Clustering for Content-based Image Retrieval," *Pattern Recognition*, vol 35, no. 5, pp. 1115-1126, 2002.



유 현 우

1966년 12월 24일생. 1992년 인하대학교 전기공학과 졸업, 동대학 전기공학 석사(1994). 고려대학교 산업시스템정보공학 박사(2001). 현재 연세대학교 인지과학연구소 연구교수. 관심분야는 컴퓨터비전, 멀티미디어시스템, 제어이론



장 동 식

1956년 12월 16일생. 1979년 고려대학교 산업공학과 졸업. 텍사스 주립대 산업공학 석사(1985). 텍사스 A&M 산업공학 박사(1988). 1989~현재 고려대학교 산업시스템정보공학과 교수, 고려대학교 관리처장. 관심분야는 컴퓨터비전, 멀티미디어시스템, 로봇비전



오 근 태

1957년 11월 13일생. 1980년 고려대학교 산업공학과 졸업. KAIST 산업공학 석사(1982). KAIST 산업공학 박사(1987) 현재 수원대학교 산업정보공학과 교수. 관심분야는 SCM, 확률모델링