

색상과 질감정보의 적응적 가중치 기법을 이용한 내용기반 영상검색

황춘화[○], 김계영^{**}, 최형일^{*}

^{○*}승실대학교 미디어공학과

^{**}승실대학교 컴퓨터공학과

e-mail: chunhua0305@gmail.com, gykim11@ssu.ac.kr, hic@ssu.ac.kr

Content-based Image Retrieval using adaptive weight of Color and texture information

Chun-Hua Huang[○], Gye-Young Kim^{**}, Hyung-Il Choi^{*}

^{○*}Dept. of Media, Soongsil University

^{**}Dept. of Computer Science, Soongsil University

● 요약 ●

본 논문에서는 영상들의 특징들을 추출하여 특징 값들의 비교를 통하여 질의 영상의 유사 영상을 검색하는 방법을 제안한다. 제안하는 방법은 입력 영상들의 색상 히스토그램으로 색상 특징 값들을 추출하고 질감 정보인 에지 정보와 이웃화소간의 공간 관계를 분석하여 질감 특징 값들을 추출하여 저장한 후 질의 이미지의 색상과 질감 특징들을 구하여 비교를 통하여 유사도를 분석하고 결과 영상을 보여준다. 또한 색상과 질감을 혼합하여 사용할 때 적응적으로 가중치를 부여함으로써 가중치가 적합하지 않아 발생하는 오 검출될 현상을 피할 수 있게 되었다. 실험을 통하여 기존의 방법과의 성능을 비교분석하였고 본 방법의 우수성을 입증하였다.

키워드: 영상검색(image retrieval), 특징 추출(feature extraction), 가중치 설정(fix weight)

1. 서론

멀티미디어 데이터베이스에서 디지털 영상 및 비디오 데이터들이 차지하는 비중이 증가됨에 따라 많은 사람들이 멀티미디어 데이터베이스에 접근하여 정보를 검색하고 있다. 특히 IPTV 등과 같은 정보 공간의 확산을 통한 영상 유통과 소비 시스템이 급격히 변화하면서 영상의 양적 성장 및 접근성 향상에 따른 영상 분류 이슈가 부상하고 있다. 그러나 정보량이 늘어날수록 원하는 데이터의 검색은 더욱 어려워져 멀티미디어 정보의 검색, 저장, 관리 기술에 대한 요구가 증가되고 있다. 1990년대 주를 이룬 텍스트 기반의 분류 기술이 가진 키워드 선정의 주관성과 키워드와 영상을 연결시키는 고비용 문제로 인해 영상검색은 내용 및 의미 특징 기반의 영상 분류 기술 트렌드로 전환 되었다. 즉 영상 및 비디오의 효율적인 검색을 위하여 멀티미디어가 가지고 있는 속성을 효율적으로 추출하고 이를 이용한 데이터베이스의 구축, 색인, 그리고 검색을 수행하는 새로운 기술들의 개발이 요구되고 있다. 이를 기반으로 영상정보로 구성된 멀티미디어 데이터에서 전체 이미지나 객체의 색상, 질감, 모양 등의 내용을 이용하여 원하는 유형의 이미지 정보를 추출하여 저장하거나 검색하는 내용 기반의 검색 시스템(content-based retrieval system)이 활발히 연구되고 있다. 때문에 더욱 빠르고 효율적인 검색 및 색인 기술들은 항상 필요하

고 있고 많은 연구들이 진행 되어 왔고 지금도 진행 중에 있다.

이와 관련된 연구를 보면 아래와 같은 몇 가지가 있다.

C.S Won, D.K Park, S.J Park[1]에서는 에지 정보를 이용한 방법을 사용하였는데 영상을 블록 단위로 나눠서 각 블록에서 강한 에지를 추출하여 지역에지 히스토그램 및 전역 에지 히스토그램을 구하여 비교함으로써 유사한 이미지를 추출한다. 그러나 이 방법으로는 물체의 대체적인 공간 위치정보는 어느 정도 추출되지만 물체의 내부 속성이나 색상 정보가 부족하기 때문에 단순한 이미지내의 객체를 추출하는 데는 적합하지만 자연영상에서의 영상 검색에는 좋은 결과를 얻지 못한다는 단점이 있다.

Robert[2]는 이웃화소간의 공간 픽셀들 간의 의존성을 나타내는 것으로서 패턴의 반복적인 정도를 나타낸다. 즉 거칠거나 부드러운 정도 등 질감의 특징을 나타냄으로써 물체 내부의 속성을 표현한다. 그러나 이 특징으로는 단순히 물체 특징의 판단에 사용되기에 적합하고 물체의 분별이나 검색에는 효과적이지 못하다. 왜냐하면 비슷한 질감을 가진 물체가 많기에 위치정보나 공간 정보, 색상정보 등 중요한 정보들이 뚜렷하지 않기 때문이다.

본 논문에서는 위에서 존재하는 문제점들을 보완하여 보다 강한 영상검색 방법을 제안한다.

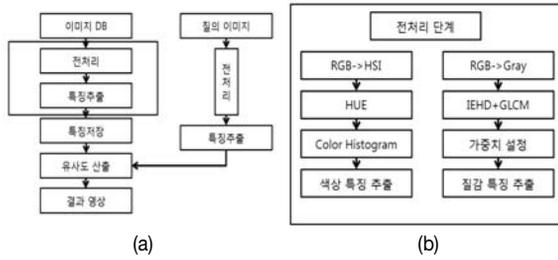


그림 1. 시스템 구조
Fig. 1. System Architecture

그림 1은 본 논문에서 제안하는 시스템 구성도이다. 주로 전처리 단계, 특징 추출 및 저장 단계, 유사도 산출 및 영상검색 등 3가지 단계를 거친다. 특징 추출에는 색상특징 추출과 질감 특징 추출이 있는데 색상은 RGB 영상을 HSI색상공간으로 변환하여 색상정보를 추출하여 색상의 히스토그램을 구하고 저장한다. 질감 정보로는 EHD[1]방법의 문제점을 개선한 IEHD방법과 GLCM[2]방법을 이용한 다중 질감 방법으로 특징들을 추출하고 이 특징들을 데이터베이스에 저장한다. 질의 이미지가 들어오면 질의 이미지의 특징들을 추출하고 DB내의 특징들과의 비교를 통하여 가장 유사한 영상들을 사용자에게 보여준다.

II. 색상 및 질감 특징 추출

1. 색상 특징 추출

본 논문에서는 영상처리에서 가장 많이 쓰고 있는 HSI색상 모델을 사용하여 색상(H)채도(S), 명도(I)값을 추출해서 색상(H)값만을 사용하였다. 왜냐하면 색상은 조명이나 밝기에 덜 민감하기에 서로 다른 밝기거나 조명 값에서도 잘 사용될 수 있기 때문이다. 또한 오직 하나의 색상 값으로만 연산하였기 때문에 모든 채널을 사용했을 때보다 연산이 줄어들고 연산시간도 단축된다는 장점이 있게 되었다.

HSI의 색상(H)은 0-360사이의 값을 갖게 되는데 이번 논문에서는 색상 정보에서 대표적인 색상60개를 추출하여 60개의 색상(H)정보만 비교를 진행한다. 즉 질의 이미지의 360개의 색상 히스토그램을 구해서 히스토그램을 내림차순으로 정렬 한 다음 그중 히스토그램의 빈도수가 가장 많은 색상 정보 60개를 추출하여 그 영상의 대표 색상(HUE)로 정한다. 그 다음 이미지 데이터베이스에서 비교하는 영상들에서 질의 이미지에 대응되는 60개의 색상 히스토그램을 추출한 후 두 영상의 대표색상들 사이의 유사도를 계산한다.

대부분의 영상에서 보면 물체는 주로 가운데 위치하여 있다. 영상에서 배경보다 물체의 정보가 더 중요하기 때문에 이번 논문에서는 물체에 더 큰 가중치를 두고 배경에 더 작은 가중치를 두는 방법을 사용한다. [3]

색상특징을 이용하여 유사도를 구하는 공식을 보면 아래와 같다.

$$Dis_{co} = w_1 \times Dis_{object} + w_2 \times D_{background}$$

$$Dis_{object,background} = \sum (\frac{Min(H(i),H(j))}{Max(H(i),H(j))})^2 (H(i),H(j) > 0)$$
(1)

$-H(i),H(j)$ 는 질의 이미지와 비교이미지의 60개의 대표 색상이고 $w_1 = 0.7, w_2 = 0.3$ 으로 설정하였다.[3]

본 논문에서는 기존의 방법에서 변형된 유사도제공[6]방법으로 비교되는 영상들의 차이를 강조하였다.

그러나 색상 히스토그램 방법은 한 이미지 내의 어떤 색상이 얼마만큼 존재하는지의 정도를 나타낼 뿐 이미지 내의 색상의 위치를 표현하지 못하기 때문에 위치 정보가 사라진다. 때문에 이번 논문에서는 질감 정보를 추가하여 위치 정보도 함께 사용하여 성능을 높이고자 한다.

2. 질감 특징 추출

질감은 영상에서 동일한 패턴을 가지는 영역의 특성으로 객체의 표면이 가지는 거칠 정도나 밝기 변화 정도를 나타낸다.

2.1 에지 히스토그램

본 절에서는 질감 특징 중의 하나인 에지 히스토그램에 대해서 설명한다.

EHD(Edge Histogram Descriptor)[1]는 이미지의 히스토그램을 구하여 전역 에지 히스토그램과 지역 에지 히스토그램으로 이미지 사이의 유사도를 측정하는 방법이다. EHD는 일반적으로 5가지 에지를 표현한다. 즉 $0^\circ, 45^\circ, 90^\circ, 135^\circ$ 및 비방향성 에지이다.

EHD 알고리즘은 이미지를 16개의 서브 이미지로 나누고 각 서브 이미지를 8개의 블록으로 나눴으로써 모두 128개의 블록을 구한다. 그리고 각 블록에서 5개의 에지 중 가장 강한 에지를 구하고 $80(16 \times 5 = 80)$ 개의 지역에지 히스토그램과 1개의 전역에지 히스토그램을 구해서 유사도를 구한다. 그 수식을 보면 아래의 식(2)와 같다.

$$D(A, B) = \sum_{i=0}^{70} |LocalA[i] - LocalB[i]| + 5 \times \sum_{i=0}^4 |GlobalA[i] - GlobalB[i]|$$
(2)

아래 그림2에서의 경우 이미지 1과 이미지 2의 히스토그램 거리의 차이도 120이고 이미지 1과 이미지3의 히스토그램 거리의 차이도 120이다. 하지만 이미지 2와 이미지 3의 에지 히스토그램이 서로 다르기 때문에 이는 서로 다른 이미지이다. 즉 한 이미지에 대하여 같은 거리 값이 나오는 이미지가 여러 개 나올 수 있고 이 유사도 측정 방법을 사용하면 이들은 모두 같은 유사도를 가지기 때문에 분류가 제대로 되지 않는다. 이는 해결되어야 할 문제이다.

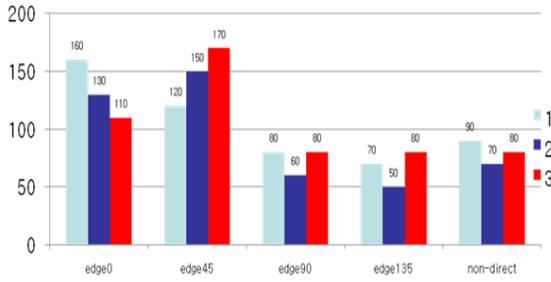


그림 2. 다른 히스토그램의 같은 거리 값
Fig. 2 same distance of different histogram

이번 논문에서는 이와 같은 오 분류된 경우를 해결하기 위하여 아래와 같은 방법을 제안한다.

2.2 개선된 에지 히스토그램(IEHD)

두 영상의 유사도를 구하기 위하여 두 히스토그램의 차이 값의 누적을 사용하는 대신 이번 논문에서는 두 히스토그램 각각의 유사한 정도를 나타내는 수식을 사용한다.

$$Dis_{cht} = \sum_{i=0}^{79} \frac{(LocalA[i]) \wedge (LocalB[i])}{\max((LocalA[i]), (LocalB[i]))} + \sum_{i=0}^4 \frac{(GlobalA[i]) \wedge (GlobalB[i])}{\max((GlobalA[i]), (GlobalB[i]))} \quad (if, \max \neq 0) \quad (4)$$

이 수식을 그림 2에 적용하면 수식(2)을 통하여서는 서로 같은 값이 나오지만 (3)을 통하여서는 1과 2의 유사도는 3.85이고 1과 3의 유사도는 4.16이라는 서로 다른 결과가 나오기 때문에 누적 값이 같은 결과가 나오는 것을 피할 수 있다.

2.3 Gray Level Co-occurrence Matrix(GLCM)

통계적 분석방법의 GLCM[2]은 그레이 레벨의 값으로 질감 특징을 추출하는 방법으로 거리와 방향이 일치하는 픽셀 쌍이 얼마나 자주 나타나는가를 표시한다. 즉 빈도수를 갖는 매트릭스로서 질감 패턴상의 그레이 레벨이 반복되는 출현 회수로서 이 매트릭스를 이용해서 공간상의 픽셀 간 의존성을 분석할 수 있다. 에지의 정보로만은 모든 화소들 사이의 상관 관계를 표현하기 힘들기 때문에 GLCM을 이용하여 좀 더 구체적인 부분의 질감 특징들을 구함으로써 성능을 높일 수 있게 되었다.

본 논문에서는 [2]의 28개의 특징벡터 중 Energy, Contrast, Correlation, Variation, Inverse Difference Moment, Entropy 등 6개의 특징벡터를 사용하여 모두 4개의 방향에서의 24개의 특징 벡터를 사용한다.

24개의 특징 벡터로 GLCM 특징들의 유사도를 구하는 수식은 아래와 같다.

$$Dis_{glcm} = \sum_i \frac{\min(f_i(q), f_i(d))}{\max(f_i(q), f_i(d))} \quad (i = 24) \quad (5)$$

여기서 $f_i(q), f_i(d)$ 는 각각 질의 이미지와 테스트 이미지의 특징 값을 가리킨다.

2.4 적응적 가중치 설정 방법

본 논문에서는 색상과 질감 특징을 혼합하여 사용할 때 각각의 특징들에 가중치를 부여하여 유사도를 산출하는 방법을 사용한다. 이때 가중치는 유사도를 구함에 있어서 아주 중요한 작용을 하기 때문에 가중치를 정확하게 설정하는 것이 아주 필요하다.

이번 논문에서는 질감 특징 중의 하나인 GLCM의 $\sum_i \sum_j p(i, j)^2$ 을 임계값으로 하는 방법을 제안한다. 즉 이 값은 질감의 에너지로서 질감 정보가 많으면 에너지 값이 크고 질감 정보가 적으면 에너지 값도 따라서 작아진다. 때문에 이 값으로 영상의 질감 정보의 유무를 판단할 수 있다.

$$similarity = w_1 \times Dis_{co} + w_2 \times Dis_{cht} + w_3 \times Dis_{glcm} \quad (6)$$

식 (6)은 이번 논문에서 영상을 검색하는 유사도 공식이다. 자동 가중치를 설정할 때 임계값을 $T_1, T_2 (T_1 > T_2)$ 로 설정하고 아래와 같이 세 가지 경우를 고려한다.

첫 번째는 질감 특징 값이 중요한 경우이다.

$$f_1 \geq T_1, w_1 = 0, w_2 = 0.1, w_3 = 0.9 \quad (7)$$

두 번째는 색상 특징 값이 중요한 경우이다.

$$f_1 \leq T_2, w_1 = 1, w_2 = 0, w_3 = 0 \quad (8)$$

세 번째는 색상과 질감이 모두 중요한 작용을 하는 경우이다.

$$T_2 < f_1 < T_1, w_1 = 0.3, w_2 = 0.1, w_3 = 0.6 \quad (9)$$

위와 같은 방법으로 가중치를 각각 구하여 식(6)에 대입하여 유사도를 구할 수 있다.

III. 실험 결과 및 결론

본 실험에서는 AMD Dual Core 2.70Ghz CPU, 3GB 메모리 컴퓨터를 사용하였으며 운영체제는 Microsoft Windows 7, 사용 툴은 Microsoft Visual Studio 2008이다. 실험 영상은 500여장의 256×384 해상도 크기로 정규화 된 영상[4][5]으로 구성되었고 그중에서 임의의 150여장의 이미지를 질의 이미지로 사용하여 정확도를 산출하였다.

실험을 통하여 이번 논문에서는 임계값을 각각 $T_1 = 0.08, T_2 = 0.03$ 으로 정하였을 때 가장 좋은 결과를 나타냈다.



그림 3. 결과 영상
Fig. 3 result image

표 1. 성능 평가 비교 결과
Table 1. performance estimation result

	말	꽃	버스	공룡	사람	평균
Color	91,8	37,2	79,8	55,8	84,4	69,8
Texture	65,2	92,6	74,6	98,8	57,8	77,8
Fixed weight	90,2	87,4	88	98	82,4	89,2
제안한 방법	92,6	93,2	91,8	98,8	85,6	92,4

그림 3은 제안한 방법을 사용했을 때의 결과 영상의 예를 보여 준다. 그림에서 봤을 때 상위 10개의 결과 영상은 모두 같은 부류의 영상으로 정확하게 검출했고 높은 정확율도 보인다. 또한 표1은 정확율의 성능 평가인데 성능 평가표에서 다른 방법들과의 비교를 통하여 제안하는 방법으로 언제나 가장 좋은 결과를 얻을 수 있다는 것을 알 수 있다.

본 논문은 색상 특징과 질감특징을 추출하여 적응적으로 가중치를 설정함으로써 기존의 방법들보다 좋은 성능을 갖고 있다. 그러나 색상과 질감정보가 명확하지 않은 영상에서는 검색이 힘들다는 단점이 있으므로 앞으로 모양에 대한 특징도 병합하면 더 좋은 결과를 얻을 수 있을 것이다.

참고문헌

- [1] C.S Won, D.K Park, S.J Park "Efficient Use of MPEG-7 Edge Histogram Descriptor", ETRI Journal, Volume 24, Number 1, February 2002.
- [2] Robert M.Haralick, K.Shanmugam, Its'Hak Dinstein "Textural Features for Image Classification", IEEE Transactions on Systems, Man And Cybernetics, Vol. SMC-3, No.6, November 1973, page 610~621
- [3] J.B Lee, G.H Lee, C Choi, S.M Cho, J.C Park, "Image Retrieval using Shape by Edge Feature and Texture and Color", Korean Multimedia Institute, Spring Science Presentation Papers, 2002
- [4] J. Li, J. Z. Wang, "Automatic linguistic indexing of pictures by a statistical modeling approach," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 9, pp. 1075~1088, 2003
- [5] J. Z. Wang, J. Li, G. Wiederhold, "SIMPLiCity: Semantics-sensitive Integrated Matching for Picture Libraries," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol 23, no.9, pp. 947~963, 2001
- [6] S.Y Shin, Y. W Rhee, "New Abrupt/Gradual Scene Change Detection", Korean Sea information Communication Conference paper, Vol. 13, No. 11,2009