영역기반 검색환경을 위한 FRIP 시스템

고병철 °, 변혜란
연세대학교 컴퓨터과학과

FRIP System For Region-based Image Retrieval

ByongChul Ko, Hyeran Byun
Dept. of Computer Science, Yonsei University

요 약

본 논문에서는 영역기반 검색환경을 제공하는 FRIP(Finding Region in the Pictures) 시스템을 소개한다. FRIP 시스템은 영역 기반 검색환경을 제공하기 위해서, 우선적으로 영상을 분할하고, 각 분할된 영역으로부터 색상, 질감, 크기, 모양, 위치 정보와 같은 최적의 특징 벡터들을 추출하여 색인화 시킨다. 그런 뒤에, 사용자가 검색하고자 하는 영역과 검색 영상 수 k를 입력하면, 유사성 측정 식에 의해 가장 유사한 k만큼의 영상을 우선 순위 형태로 사용자에게 보여주게 된다. 본 시스템에서는 영상을 분할하기 위해서 기본적인 RGB 색상계를 확장(Scaling) 및 이동(Shifting) 알고리즘을 통해 영상의 대비 정도가 향상된 새로운 색상계로 변환 시키고, 원형 필터를 설계하여, 영역 안에 포함된 의미 없는 작은 영역을 제거하도록 하였다. 그리고 이렇게 분할된 각 영역들로부터, 본 시스템에서 제안하는 모양 기술자인 MRS(Modified Radius-based Signature)를 포함하여 5가지의 최적의 특징 벡터들을 전처리 단계에서 데이터베이스에 색인으로 저장하고 유사성 측정을 위한 수치로 사용하였다.

1. 서 론

영상 기술의 급속한 성장과 더불어, 내용기반 영상 검색은 전자 도서관, 영상처리, 멀티미디어 데이터베이스 시스템 등과 같은 응용 분야에서 연구자들의 큰 관심을 끌고 있다. 더 나아가, 최근 빠른 하드웨어 기술의 성장으로 인해, 대용량의 멀티미디어 데이터를 빠른 시간 안에 저장하고 검색하는 기술이 가능해 졌다 [1].

현재, '영상에 의한 질의(Query-by-image)'는 가장 대표적인 내용기반 영상검색 기술의 하나다. 하지만, 만약 영상 검색 시스템이 단지 영상의 전역적 특성만을 이용한다면, 영상내의 포함된 지역적인 특성이 무시되어 많은 잘못된 결과를 나타낼 수 있다. 따라서, 최근의 내용기반 영상검색은 객체 또는 영역단위의 영상검색에 초점이 맞춰 지고 있다.

2. 관련연구

현재 몇몇 시스템에서 영상 데이터베이스로부터 색상, 질감, 모양 성분들을 이용하여 객체 또는 영역 단위로 검색하는 연구가 진행되어 오고 있다 (예, QBIC, Netra, VisualSEEk, Blobworld). 각 시스템은 기본적으로 다음과 같은 몇 가지의 공통된 특징을 가지고 있다.

(1) 영상은 몇 개의 영역들로 분할된다. (2) 분할된 각 영역들로부터 최적의 특징 벡터들이 추출된다. (3) 특징 벡터들의 집합은 데이터베이스에 색인으로 저장된다.

질의 과정에서, 사용자가 스케치한 영역이나, 질의 영상으로부터 분할된 영역들로부터 특징 벡터들이 추출되고, 이 특징 벡터들과 데이터베이스 색인과의 비교(matching) 작업이 이루어지게 된다.

2.1 FRIP 시스템의 개요

*FRIP(Finding Region In the Pictures)*은 기본적으로 3가지의 검색환경을 제공한다. 첫 번째는 'Flexible subblock'에 기반한 영상검색 환경이고[2], 두 번째 '얼굴에 의한 영상검색' 환경이다[3]. 마지막으로 세 번째는 '영역에 의한 영상검색' 환경이다. 본 논문에서는 FRIP 시스템의 3번째 질의 환경인 '영역에 의한 검색'에 초점을 맞추고 있다.

'영역에 의한 검색' 환경을 위해 FRIP 시스템에서는 기본적으로 영상을 확장(Scaling) 및 이동(Shifting)된 색상 영역으로 변환하고, 본 시스템에서 제안하는 원형 필터(Circular filter)를 이용하여 영상을 분할한다. 다음으로, 분할된 영역으

로부터, 색상, 질감, 크기, 모양, 위치 정보들을 추출하여, 전처리 과정에서 데이터베이스에 색인으로 저장한다. 그런 뒤에 질의 과정에서 사용자가 영상을 분할하고, 분할 된 영역중 한 영역을 선택하여 질의하면, 미리 저장된 색인 값과 질의 영상에서 추출된 특징 벡터와의 유사성 측정에 의한 가장 유사한 영역만을 k-nearest neighbor search에 의해 사용자에게 보여주게 된다.

3. 영상 분할 알고리즘

본 시스템에서는 영상을 의미 있는 단위로 분할하고 분할 시간을 단축시키기 위해 기존의 영상 분할과 다른 기법을 사용한다. 영상 분할 과정은 크게 다음의 두 단계를 거친다.

(1) 색상 변형 & 원형 필터와 영역 병합(region merge) 방법을 이용한 1-단계 분할.

(2) 영역 표시(region labeling) 방법과 영역 병합을 이용한 반복-단계 분할.

3.1 색상 좌표계의 변형

일반적으로 컬러 영상은 채도와 명도 성분을 변경시키기 위해서 RGB 색상으로부터 YIQ 색상으로 변경된다. 변경 후에 명도와 채도 성분을 변형시키고 다시 RGB 색상으로 변경시킨다. 하지만, 이러한 작업은 계산상으로 매우 비효율적이다. 따라서, 본 시스템에서는 영상 분할을 위한 전처리 단계로 영상의 질을 향상시키기 위해서 확장과 이동 색상 좌표를 위한 하이브리드 기술 [4]을 적용하였다.

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = k\frac{L'}{L}\begin{bmatrix} R \\ G \\ B \end{bmatrix} + (1-k)\left(\begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} L'-L \\ L'-L \\ L'-L \end{bmatrix}\right) \quad (1)$$

$R'G'B'$: RGB 색상으로 부터 변형된 색상 좌표계

L' : 원 명도(L)값 으로 부터 $Contrast-stretch$된 명도 값

k : 확장($scaling$) 및 이동($shifting$)을 위한 $factor$

식 (1)에서 명도 성분(L)은 LHS 색상 좌표계로 부터 다음과 같은 수식에 의해 얻어진다

$$L=0.299R+0.587G+0.114B.$$

본 시스템에서는 L'을 contrast-stretch된 명도 값으로 정의한다. 따라서, 대비 정도가 좋은 영상(화소 값이 한쪽에 치우치지 않고 고르게 분포)은 그대로 유지하고, 대비 정도가 좋지 못한 영상은 수식 (1)에 확장 및 이동된 형태의 영상으로 변환된다. 본 시스템에서는 실험에 의해 채도 값을 확장시키는 것이 명도 값을 이동시키는 것보다 영상 분할을 위해서 중요하다는 것을 발견하였다. 따라서, 확장 및 이동에 필요한 k값을 0.7로 정하여 사용하였다. 식(1)의 하이브리드 기술은 색상의 확장과 이동을 통해서 채도 성분이 과다하게 증가하거나 감소하는 문제를 막아준다 [4]. 또한, 영상 분할을 위해 RGB 색상을 다른 색상계로 변환하지 않음으로써, 본 시스템에서는 향상(enhance)된 영상을 얻을 수 있었을 뿐만 아니라 색상 변환에 필요한 계산 시간도 단축시킬 수 있었다.

3.2 원형 필터를 이용한 1단계 영상 분할

변형된 색상 좌표계를 이용하여, 다음과 같은 단계를 거쳐 초기 분할된 영상을 얻을 수 있다.

① 영상의 양자화 과정 (Quantization)

② 영상의 평균화 작업 (Averaging filtering)

③ 영상의 메디안 필터링 작업 (Median filtering)

④ 원형 필터 작업(Circular filtering)

⑤ 영역 병합 과정(Region merging)

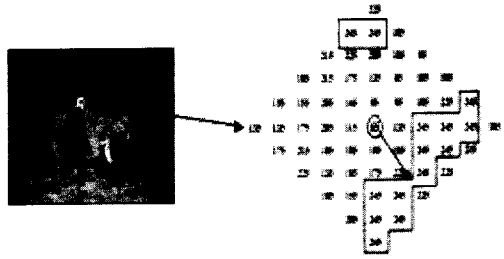

그림 1 . 원형 필터 과정

3.3 영역 표시 방법과 병합 방법을 이용한 반복-단계 분할

1-단계 과정을 통해 분할된 영역들은 변화된 색상 원형 필터를 사용하였다 하더라도 아직 과다하게 영역이 분할된 상태다. 또한, 경우에 따라서, 어떤 영역들은 서로 의미적으로 다른 영역임에도 불구하고, 유사한 속성을 갖기 때문에 한 개의 영역으로 선언될 수 있다. 이러한 문제점을 해결하기 위해 본 시스템에서는 영역을 구분하여 표시(region-labeling)하는 connected-component 알고리즘을 사용하였다. 본 시스템에서는 모양 기술(description)의 저장공간을 효과적으로 관리하기 위해 최대 분할 될 수 있는 영역의 수를 30으로 한정하였다.

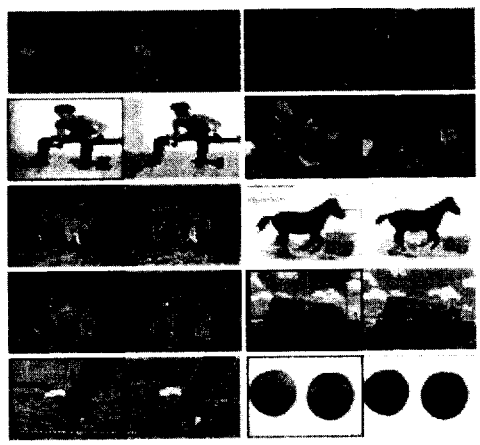

그림 2 . 영상 분할 결과

4. 특징 벡터 추출

영상을 영역 단위로 분할 한 뒤에, 각 영역들로부터 영역을 표현할 수 있는 특징 값들을 추출하는 과정이 필요하다. 본 시스템에서는 유사성 측정을 위해 각 영역들로부터 다음과 같은 5가지의 특징 벡터를 추출하였다.

① 색상 정보: 각 영역으로부터 RGB 평균 색상 성분
② 질감 정보: 쌍직교 웨이블릿 변환을 통한 수직 수평 방향 성분
③ 크기 정보: 영역에 포함된 총 화소 수
④ 모양 정보: 본 시스템에서 제안하는 변형된 반경 기반 시그네쳐와(Modified Radius-based Signature) 이심율(Eccentricity)
⑤ 위치 정보: 영역의 중심 좌표

그림 3 . (a) 원 영상 (b) 영역의 경계선 영상 (c) FRIP 모양 기술 방법에 의한 polygon

마지막으로, 이렇게 각 영역들로부터 구해진 12개의 특징 값들은 색인화 되어 데이터베이스에 저장된다. 최종적으로 저장되는 특징 값들은 표 1과 같다.

특징 번호	기술내용(description)	특징 번호	기술 내용(description)
1	영역 번호	7	무게중심의 y-좌표(Cy)
2	평균 Red 값(Ar)	8	장축(R_{max})
3	평균 Green 값(Ag)	9	단축(R_{min})
4	평균 Blue 값(Ab)	10	시그네쳐(Si, i=1,...,12)
5	영역의 총 화소 수(NP)	11	영역의 y-방향성분 크기(Yd)
6	무게중심의 x-좌표(Cx)	12	영역의 x 방향성분 크기(Xd)

표 1. 영역에 대한 기본적인 12개의 특징 값

5. 유사성 측정

실제적인 정합(matching) 작업은, 데이터베이스에 저장된 영역들 중에서 질의 영상과 가장 유사한 특징을 가지는 k개의 영상(k-nearest neighbor)을 보여 주도록 설계되었다. 유사성 측정 과정을 살펴 보면, 우선, 사용자는 질의 하고자 하는 영상을 분할한다. 이 때 특징 값들은 분할과 동시에 자동으로 추출된다. 그런 뒤에 질의 조건을 선택한다. 최종적으로 사용자가 입력한 질의 조건 즉, (1) 크기를 고려할 것인가? (scale-care/don't care) (2) 모양을 고려 할 것인가? (shape care/don't care) (3) 위치를 고려할 것인가? (location care/don't care) 에 따라서 전체적인 정합 점수(matching score)가 계산된다. 여기서 (1)(2)(3)의 각 조건들은 선형결합(linear combination)에 의해 한 개의 통합된 값, 즉 최종 점수(score)로 나타내어지게 되고, 사용자가 입력한 k만큼의 가장 유사한 영역들을 결과로 보여준다.

6. 실험 결과 및 평가

본 시스템에서는 웹과 Corel Photo CD에서 수집한 자연 영상, 그래픽 영상, 그림 영상 등을 포함하는 2,600개의 영상 데이터베이스를 사용하였다. 본 시스템은 Visual C++ 6.0 언어를 사용하여 구현되었고, 오프라인(off-line) 시스템이다. 본 시스템에 대한 보다 자세한 내용 및 검색 결과는 http://vip.yonsei.ac.kr/Frip 에서 확인 할 수 있다. 현재, 영역기반 검색 방법에 대한 성능 측정을 위한 특별한 기준이 없으므로, '해, 독수리, 호랑이' 와 같은 특정 물체(object)에 대해서 검색을 실시하고, precision에 의해 검색 결과를 평가하였다 (그림 4).

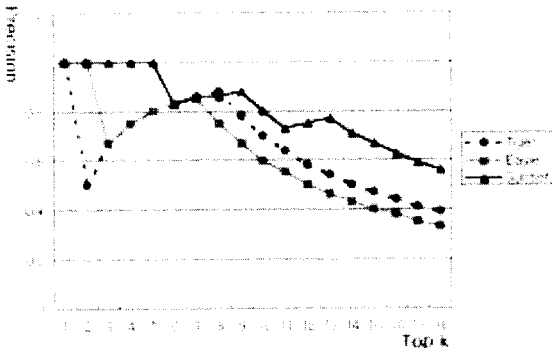

그림 4. FRIP 시스템의 성능 평가 (K-18)

6.1 향후 연구 방향

현재 FRIP 시스템의 유사성 측정과정은 단순히 특징 값들의 차를 구하고 그 값들을 사용자 질의 조건에 따라서 선형 조합하는 형태를 띄고 있다. 하지만, 앞으로의 연구에서는 사용자 피드백(feedback)을 통한 확률적인 학습 과정을 통해, 사용자가 본 시스템과 상호작용 함으로써 보다 정확하고 의미 있는 능동적인 검색환경을 제공할 것이다. 또한, 시스템 환경에 익숙하지 못한 초보자들이나 어린이, 노인들을 위해 동작 인식(Gesture Recognition)과 같은 새로운 개념의 내용기반 검색 환경도 제공할 것이다.

3. 참고문헌

[1] Moghaddamzadeh, N. Bourbakis, A fuzzy region growing approach for segmentation of color images, Pattern Recognition. Vol. 30, No. 6, pp.867 881,1997.

[2] ByoungChul Ko, Hae-Sung Lee, Hyeran Byun, "A Flexible subblock based Image Retrieval Scheme," CVPRIP '2000 , Feb. 27 ,2000 Atlantic city,USA

[3] 고병철, 이해성, 변혜란,"효율적인 브라우징 환경을 위한 비디오 색인", 정보과학회 논문지: 소프트웨어 및 응용 제 27권 제1호,2000, pp 74-83.

[4] Christopher C. Yang, etc., Efficient Luminance and Saturation Processing Techniques for Color Images, Journal of visual comm.