

동영상 이미지의 특징정보 분석 시스템 설계 및 구현

최 내 원* 지 정 규**

Design and Implementation of the Feature Information Parsing System for Video Image

Nae-won Choi*, Jeong-gyu Jee**

요 약

컴퓨터 응용기술의 급속한 발전으로 인해 동영상 정보는 인터넷 및 사회전반의 다양한 분야에서 활용되고, 그 수가 기하급수적으로 증가되고 있다. 동영상 정보 분석 시스템은 기본적으로 텍스트를 기반으로 하기 때문에, 동영상 정보가 가지는 애매성을 표현하기 곤란하며, 주석 작성에 따르는 과도한 작업부담 및 객관성 결여 등의 문제점을 가지고 있다. 본 논문에서는 대용량의 동영상 정보를 효율적으로 분석하기 위해 동영상 이미지의 분할영역에서 색상정보와 모양정보를 이용한 방법을 제안하고자 한다. 색상 정보를 추출하기 위해서는 기존의 RGB 방식에서 HSI방식으로 색상변환 하여 대표색상과 매칭 되는 특징 정보를 사용한다. 그리고 모양정보는 물체의 윤곽선에 해당하는 화소들만을 대상으로 연산을 수행하는 향상된 불변 모멘트(IMI)를 이용한다.

Abstract

Due to the fast development in computer application technologies, a video is now being more widely used than ever in many areas. The current information analyzing systems are basically built to process text-based data. Thus, it has little bits problems when it needs to correctly represent the ambiguity of a video, when it has to process a large amount of comments, or when it lacks the objectivity that the jobs require. We would like to purpose the method that is capable of analyze a large amount of video efficiently. To extract the color, we translate the color from RGB to HSI and use the infomation that matches with the representative colors. To extract the shape information, we use improved moment invariants(IMI) so that we can solve many problems of histogram intersection.

* 숭실대학교 컴퓨터학과

** 한국학술진흥재단

I. 서론

급속한 디지털 처리의 발달과 함께 동영상에 대한 검색도 다양화되고 있으나 활용적인 측면에서는 성과가 미미한 실정이다. 특히 문자기반방식을 이용한 방대한 동영상 데이터를 관리하고 보존하는데 있어서 검색 및 질의 방법과 시각 정보, 그리고 이러한 부분들의 문자정보를 연결하는 것은 아직도 초기수준밖에는 미치지 못한다[1].

동영상에서 시각 정보를 위한 표현을 얻기 위해서는 내용분석이 요구된다. 기존의 이미지에 몇 가지 내용분석 방법을 가지고 접근을 시도한다. 이러한 방법은 칼라, 텍스처, 모양을 비교하여 분석을 하게 된다. 연속된 비디오에서 내용 분석은 샷이라는 유사한 공간적인 세그먼트를 통해 처음에 연속성에서 분리하게 되고, 그때 그러한 세그먼트는 하나 또는 몇 개의 프레임으로 요약되고, 시각이나 음성적인 특징을 가지는 샷 사이에서의 관계성을 마지막으로 결정하게 된다[2].

지금 까지 연구되어진 동영상 내용분석 및 내용검색의 목적은 방대한 동영상 정보의 자동적인 분류 및 검색을 가능케 하고, 기존의 문자기반(SQL)방식을 해결하며 주관적인 해석을 배제하여 동영상 데이터에 대한 일관적인 접근이 가능하도록 하는 것이다. 본 연구에서는 기존의 동영상 분석기법을 기반으로 이미지 특징기법에 색상과 모양정보를 사용하여 분석내용의 정확도와 효율성을 향상시키기 위하여 동영상 이미지의 특징정보 분석 시스템을 설계하고 구현하는 것이 목적이다.

II. 동영상 이미지의 내용 표현

2.1 색상 표현

색상을 표현하는 칼라모델은 사용되는 용도에 따라 다양하다. 가장 일반적으로 사용되는 하드웨어 기준 모델들

은 칼라 모니터와 많은 종류의 색채 비디오 카메라를 위한 RGB모델, 칼라 프린터를 위한 CMY모델, 칼라 텔레비전 방송의 표준인 YIQ모델, 이미지 처리를 위해 가장 많이 사용되는 HSI모델, HSV모델 등이 있다.

RGB 모델은 빨강, 녹색, 파랑의 원색 당 하나씩 모두 3개의 독립적인 이미지 평면으로 구성되어 있는데, RGB 모니터에 보내진 이 세 이미지는 인광체 스크린에서 결합되어 복합색 이미지를 만든다. 따라서 이미지 처리를 위한 RGB 모델의 사용은 이미지 자체 값들이 3가지 색상 평면으로 자연스럽게 표현될 때 적당하다. HSI 모델의 장점은 첫째, 밝기 성분 I가 이미지에서 색 정보로부터 분리되어 있다. 둘째, 색상과 채도 성분은 사람이 색을 인지하는 방식과 같은 관계를 가지고 있다

이러한 특성은 HSI 모델이 인간 시각 시스템의 색채 감지 특성에 기초한 이미지 처리 알고리즘 개발을 위한 이상적인 도구가 되게 한다.

2.2 색상특징 정보

색상특징 정보를 표현하기 위하여 Swain이 제안한 색상 히스토그램[3]을 많이 사용한다.

장점으로는 전체적인 이미지의 성질을 대표할 수 있고 알고리즘이 간단하며, 물체의 회전이나 작은 이동 등과 같은 기하학적인 변형에는 강건한 특징이 있다. 그러나 빛의 밝기와 이미지 내의 물체 크기에 민감하고, 전혀 다른 이미지도 같은 색상 분포를 가질 수 있는 단점이 있다.

칼라 특징정보는 R,G,B로 표현되는 칼라 이미지의 최대 히스토그램을 좌표로 표현할 수 있다. 각 색상 값에 대해 히스토그램은 식(1)과 같이 생성된다[4].

$$h(i) = \frac{n(i)}{n} \quad (1)$$

식(1)에서 이미지의 전체 픽셀수가 n , 특정 칼라값이 i , i 칼라값을 갖는 픽셀의 합을 $n(i)$ 로 표현한다. 구한 각 칼라에 대한 히스토그램의 최대점을 좌표로 표현하면 R,G,B 각각에 대해 3개의 좌표가 얻어지는데 그 좌표값을 이미지의 키 값 $K_c = (X_0, Y_0, X_1, Y_1, X_2, Y_2)$ 으로 사용하였다[4][5][6].

2.3 모양특징 정보

모양정보를 이용하는 방법은 윤곽선 정보를 추출하여

매칭에 이용하거나 불변 모멘트를 계산하여 대표값으로 사용한다. 이미지 내에 지배적 객체를 색상 특징정보와 윤곽선 정보를 사용하여 얻어 낼 수 있다. 모양의 유사성을 비교할 때, 대수적 모멘트(algebraic moment)나 인수곡선(parameter curve)보다 더 인간의 모양 감지에 가까운 측정법을 제공하기 때문에 누적 전환각도(cumulative turning angle)를 사용한다. 이러한 값들은 회전이나 이동에도 변하지 않으며, 불룩하거나 불룩하지 않은 다면체에 대하여 변하지 않고 상대적으로 계산이 쉽다.

Ⅲ. 동영상 이미지 특징 분석 시스템의 설계

3.1 시스템의 구성

동영상 이미지 특징분석 시스템 모델은 (그림 1)과 같이 대표 프레임 추출, 특징정보 추출, 정보분류, 정보출력 모듈로 구성된다.

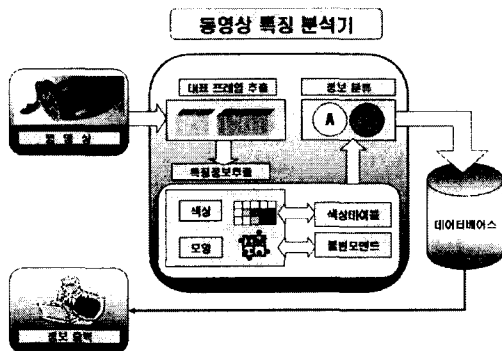


그림 1. 동영상 이미지 특징분석 시스템 모델

본 시스템은 멀티미디어 검색시스템을 서비스하기 위한 특징량 중 색상정보와 모양정보를 추출하기 위한 시스템이다. 먼저 입력되는 동영상에 대한 대표 프레임을 추출하고 추출된 이미지에 대하여 부분영역을 추출한다. 그리고 부분영역의 각 색상정보와 모양정보는 클래스별로 분류하여 데이터베이스에 저장하게 된다.

3.2 대표 프레임 추출

에지 영상을 사용한 검출 방법은 영상의 에지를 구하기 위하여 Canny, Laplacian, Range, Sobel 등을 오퍼레이터로 사용하는데 이것은 시간이 많이 소요되는 작업으로 에지 영상을 이용한 검출 방법이 여러 대표 프레임 검출 방법 중에서 가장 시간이 오래 걸리게 한다.

대표 프레임 추출 과정에서는 수많은 데이터 프레임들을 처리해야 하므로 처리 속도가 시스템 성능에 중요한 영향을 미친다. 따라서 처리 속도를 빠르게 하는 것이 중요한 문제가 된다. 특히 색상이나 모양 특징을 찾기 위하여 대표 프레임을 추출하는 경우, 몇 초간 지속되므로 매우 정확한 장면전환 검출은 요구되지 않는다.

따라서 본 연구에서는 RGB 칼라 히스토그램을 이용한 검출 방법을 사용하였다. RGB 칼라모델은 빨강(red), 녹색(green), 파랑(blue) 등 요소의 조합으로 칼라를 표현하는 방법으로 대부분의 컴퓨터 그래픽스 시스템들이 사용하는 칼라 모델로서 구현하기가 간편하다. RGB 칼라 히스토그램을 이용한 검출방법은 원 영상에 대하여 바로 처리를 하므로 빠른 처리 속도를 가지고 있으며 또한 칼라 정보를 이용하므로 대표 프레임 검출률도 우수하다[7].

대표 프레임 추출과정은 다음과 같다.

- 단계 1 : 첫 번째 프레임에 대하여 RGB 각각의 히스토그램을 구한다.
- 단계 2 : 두 번째 프레임에 대하여 RGB 각각의 히스토그램을 구한다.
- 단계 3 : 프레임 간의 RGB 각각의 히스토그램 차이의 누적값을 구한다.
- 단계 4 : RGB 히스토그램 중 하나라도 누적값이 임계치(threshold) 이상이면 그 프레임을 컷으로 검출한다.

3.3 특징정보 추출

특징정보 추출을 위해서 부분영역을 먼저 구분하는데, 부분영역 추출은 이미지를 구성 성분이나 물체로 구분하여 나누는 것이다. 이 구분의 세분화 범위는 해결하고자 하는 문제에 따라 다르다. 즉 세그멘테이션은 관심의 대상 물체가 분리되어졌을 때 멈추어야 한다. 세그멘테이션의 이상적인 결과는 사람이 인식하는 의미 있는 물체를 추출해 내는 것이지만 현재까지는 사람이 보기에 의미 있는 객체추출의 알고리즘은 알려진 바 없다.

따라서 부분영역 추출은 세그멘테이션 방법을 사용하고, 구체적으로 영역성장(region growing) 및 영역결합(region merging) 알고리즘을 적용한다. 기본적으로 칼라 이미지에 적용되는 것을 전제로 하기 때문에 식(2)에 따라 영역 성장의 임계값을 설정한다.

$$\sum |H - H_s|^2 + |S - S_s|^2 + |I - I_s|^2 \leq T \quad (2)$$

색상특징의 가중치를 달리함에 따라 추출되는 부분영역의 질이 달라질 수도 있다. 본 H, I의 승수는 2로 설정하였으며, S의 승수는 2~9까지 변경하면서 영역이 분할되는 결과의 변화를 기대할 수 있다.

① 모양 특징정보 추출

윤곽선을 추출하기 위해 기울기 연산자를 사용하여 크기와 방향을 갖는 벡터값으로 표현하는 방법, 소벨 연산법, 라플라시안 2차 미분법 등 많은 이론이 있지만 정확한 윤곽선을 찾기는 어렵다.

본 논문에서는 부분영역의 경계선을 추출하기 위해 템플레이트 매칭에 의한 윤곽선 추출 방법을 사용한다. 템플레이트 매칭방법은 마스터 패턴을 준비해 두고 경계선과 일치하는 패턴을 선택하여 패턴을 적용하게 된다.

모멘트는 어떤 축을 기준으로 한 값들의 분포의 척도이다. 이미지 처리에서는 2차원의 모멘트를 이용하여 그 레이 스케일 값들의 분포를 기술하는데 사용된다. 모멘트는 스칼라 양이고 패턴인식이나 패턴해석에서 모양 인식과 식별을 위한 특징으로 사용된다.

불변 모멘트에서 이미지 $f(i, j)$ 의 $(p+q)$ 차 모멘트 m_{pq} 는 다음과 같이 정의된다.

$$m_{pq} = \sum_i \sum_j i^p j^q f(i, j) \quad (3)$$

이진 이미지인 경우 $f(i, j)$ 는 대상 물체 내에서 1, 배경에서 0이 될 것이다. 식(3)에서 이진 이미지의 0차 모멘트 m_{00} 는 $f(i, j)$ 의

총합이 되므로 면적과 동일하다. 또 1차 모멘트 m_{10}, m_{01} 을 m_{00} 으로 정규화 하면 무게중심의 좌표를 얻을 수 있다. 즉

$$i_G = \frac{m_{01}}{m_{00}}, j_G = \frac{m_{10}}{m_{00}} \quad (4)$$

중심 근처에 있는 모멘트 u_{pq} 를 특히, 중심 모멘트(central moment)라고 하며 다음과 같이 나타낼 수 있다.

$$u_{pq} = \sum_i \sum_j (i - i_G)^p (j - j_G)^q f(i, j) \quad (5)$$

또 2차 모멘트는 관성 모멘트라고 불린다. 이미지 $f(i, j)$ 의 2차 중심 모멘트는 식에서 $p=0, q=2$ 또는 $p=2, q=0$ 로 하여 구하여지므로 다음과 같은 식이 된다.

$$u_f = \sum_i \sum_j \{ (j - j_G)^2 + (i - i_G)^2 \} f(i, j) \\ = u_{02} + u_{20} \quad (6)$$

식(6)에서 중심을 원점으로 하는 경우에는 x축 및 y축에 대하여 2차 모멘트를 각각 u_{02} 및 u_{20} 으로 하면 원점 근처에 있는 2차 모멘트 u_2 는 다음과 같이 정의할 수 있다.

$$u_2 = u_{02} + u_{20} = \sum_i \sum_j j^2 f(i, j) \\ + \sum_i \sum_j i^2 f(i, j) \quad (7)$$

이러한 모멘트는 Hu가 제안하였는데, 이는 모양 경계선과 그것과 관련된 내부 전체를 포함하여 모든 화소에 대해 계산해야 하므로 상당히 많은 계산량이 요구되는 단점이 있다[6].

이것을 개선하기 위하여 모든 화소에 대한 연산을 행하지 않고 템플레이트 매칭에 의해 검출된 윤곽선에 해당하는 화소들만 연산하는 향상된 불변 모멘트(Improved Moment Invariants)를 사용한다. 이것은 연산량을 현저히 감소시켜준다.

수정된 $(p+q)$ 차 모멘트 m_{pq} 는 다음과 같이 정의된다.

$$m_{pq} = \sum_{(x,y) \in C} (x - \bar{x})^p (y - \bar{y})^q \quad (8)$$

그리고 차수 $p+q \leq 3$ 인 중심 모멘트는 다음과 같다.

$$\begin{aligned} \mu_{00} &= m_{00}, \mu_{10} = 0, \mu_{01} = 0 \\ \mu_{11} &= m_{11} - \bar{x}\bar{y}m_{10}, \\ \mu_{20} &= m_{20} - 3\bar{x}\bar{y}m_{10}, \mu_{02} = \bar{y}m_{01}, \\ \mu_{30} &= m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2m_{10}, \\ \mu_{12} &= m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}^2m_{10}, \\ \mu_{21} &= m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}^2m_{01}, \\ \mu_{03} &= m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2m_{01} \end{aligned} \quad (9)$$

식(9)를 축적변화에 불변하도록 정규화 하기 위해서 식(10)을 사용한다.

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{p+q+1}}, (p+q=2,3,\dots) \quad (10)$$

식(9)의 값으로부터 7개의 불변 모멘트 값을 식(10)와 같이 구하여 식(11)에와 같은 7개의 대표값을 계산하여 사용한다.

$$\begin{aligned} \Phi_1 &= \eta_{20} + \eta_{02} \\ \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \\ \Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\ \Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \\ &\quad - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03}) \\ &\quad \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \Phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ &\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \end{aligned}$$

$$\begin{aligned} &- (\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \\ &\times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (11)$$

② 색상 특징정보 추출

색상정보의 특징값을 추출하는 과정을 살펴보면 입력 되는 이미지는 대부분의 프로그램에서 장치 독립적인 이미지 파일인 비트맵 이미지이다. 비트맵 파일의 픽셀 구성은 RGB로 표현되며, 이로부터 특징값 추출에 사용되는 값인 HSI로 변환이 필요하다[이동근 99].

기본적으로 영상에서 사용하는 색상인 RGB 보다는 사람의 눈이 인식하는 색과 유사하며, 각 축이 시각적으로 독립적인 HSI 즉, 색조(Hue), 채도(Saturation), 명암(Intensity)을 이용하는 것이 더 좋은 결과를 산출한다. 이는 RGB 모델은 RGB의 비율에 의해 색을 결정하는 수학적 모델이며, R,G,B 각각의 채널에 대한 서로의 간섭이 많기 때문에 색상 값을 정확히 알기 어려우나 HSI 모델은 직관적 모델이기 때문에 색상 값을 정확히 알 수 있으며, 또한 H, S, I 각 축이 독립적이기 때문에 각각에 대한 값을 쉽게 구할 수 있는 장점이 있기 때문이다.

본 논문에서는 RGB 세 가지의 기본 색에 의해 표현되어 있는 입력 영상의 각 화소의 색상을 HSI 모델로 변환시킨다. RGB 모델에서 HSI 모델로의 변환 과정은 식(12)와 같다.

$$\begin{aligned} I &= 0.3R + 0.59G + 0.11 \\ V_1 &= R - I = 0.7R - 0.59G - 0.11B \\ V_2 &= B - I = -0.3R - 0.59G + 0.89B \\ H &= \tan^{-1}\left(-\frac{V_1}{V_2}\right), S = \sqrt{V_1^2 + V_2^2} \end{aligned} \quad (12)$$

색상 변환식을 사용하여 각 채널에 대한 색조, 명도, 채도에 대한 값을 구한다. H는 0~180°의 값이 나오는데, B>G 일 경우 H = 360° - H 한다. S는 0~1 사이의 값을 얻게 되는데, H와 S는 다시 0~255 사이의 값을 갖도록 정규화 한다.

위에서 구한 HSI 값으로 각 채널의 평균과 표준편차 왜도를 식(13)과 같이 구한다.

$$\begin{aligned}
 E_i &= \frac{1}{M \cdot N} \sum_{j=1}^C P_{ij} \\
 S_i &= \sqrt{\frac{1}{M \cdot N} \sum_{j=1}^C (P_{ij} - E_i)^2} \\
 \sigma_i &= \sqrt[3]{\frac{1}{M \cdot N} \sum_{j=1}^C (P_{ij} - E_i)^3} \quad (13)
 \end{aligned}$$

여기서 P_{ij} 는 입력 이미지 $M \times N$ 의 j 번째 채널에 대한 i 번째 화소의 값을, E_i 는 각 채널에 대한 평균, S_i 는 표준편차, σ_i 는 왜도를 C 는 HSI의 각 채널을 나타낸다. 따라서 색상 특징정보는 $S_k(E_j, S_j, \sigma_j)$ 으로 표현되고, S_k 는 전체 이미지 또는 k 번째 부분영역의 색상 특징정보이다.

3.4 정보분류

동영상에서 히스토그램에 의해 대표 프레임으로 추출된 이미지는 그 자신의 프레임 번호를 기억하고, 이미지의 특징값 추출과정을 거쳐 색상정보와 모양정보에 대한 특징값을 각각 가지게 된다. 앞에서 하나의 이미지에서 추출된 다수의 색상 및 영역 특징정보에 따라 다수의 클래스에 할당되고 검색시에도 질의 이미지는 다수의 클래스로 매핑되는데 각 특징정보가 해당하는 모든 클래스의 이미지를 검색하여 유사도를 계산하게 된다. 주의할 점은 하나의 이미지가 하나의 클래스로 매핑되는 것이 아니라 여러개의 클래스를 갖는다는 것이다.

동영상에서 추출된 이미지의 프레임 번호와 이미지의 부분영역별 색상특징 정보와 모양특징 정보에 해당하는 특징값에 따라 데이터베이스에 저장할 하게 된다. 색상정보에 대한 분류 기준은 각 채널에 대해 추출된 색상정보인 평균, 표준편차, 왜도가 사용된다. 따라서 9개의 특징정보에 따라 9개의 클래스로 분류된다.

모양정보에 대한 분류 기준은 부분영역의 윤곽특징, 유한 요소 모델 등으로 표현될 수 있다. 윤곽정보는 부분영역의 모양 유사성에 대한 주관적인 해석이 가능하므로 어느 한 가지 기준으로 클래스화 하는 것은 상당히 어렵다. 따라서 자연 사진과 같은 특정 도메인에 적합한 템플릿을 정의하여 사용하는 것으로 실용화되고 있다. 템플릿은 질의 이미지의 윤곽정보와 비교하여 가장 유사한 템플릿의 그룹으로 분류된다. 템플릿을 비교할 때 이미지 전체를 픽셀 단위로 계산하므로 많은 계산량이 소요된다.

본 논문에서는 템플릿에 의한 분류는 사용되지 않고, 불변 모멘트 정보의 스칼라 값의 분포와 색상정보의 색조, 채도, 명도 값을 사용한다.

IV. 구현 및 성능 평가

4.1 구현 환경

시스템의 구현에 사용되는 하드웨어로는 RAM 512MB이고, OS는 윈도우즈 XP Professional 이 탑재된 1.5GHz 펜티엄4 PC가 사용되었다. 성능평가를 위해 저사양의 RAM 64MB와 윈도우즈 98이 탑재된 350MHz 펜티엄II PC도 사용되었다.

동영상 이미지의 특징정보 추출 및 데이터베이스 구축은 Visual C++ 6.0(Service Pack 5)를 사용하여 프로그래밍 하였다. 데이터베이스는 ODBC보다 접근 속도가 빠른 DAO를 사용하여 구축하였다.

시스템에서 사용되는 동영상은 용량에 제한되는 것을 방지하기 위해 100MB이상 되는 동영상 10여종, 50MB급 10여종, 10MB급 10여종을 선택하여 실험하였고, 실험의 일관성 유지를 위해 상영시간 3분이내의 CF 동영상으로 선택하여 실험하였다.

4.2 결과 및 성능 평가

(그림 2)는 동영상에서 히스토그램을 사용하여 각 채널별 값들의 임계값이 넘는 대표 프레임을 추출하는 추출기이다.

(그림 3)은 추출된 대표 프레임에서 RGB를 변환하여 HSI로 색상특징 정보를 추출하고 템플릿 매칭 방법으로 추출된 윤곽선정보를 이용하여 7개의 대표값을 계산한다.

〈표 1〉의 (a)는 이미지의 색상특징을 기존의 RGB방식을 적용하여 추출한 결과이다. 이미지에서 전체적인 색상톤은 색상간섭에 의하여 단일색상으로 표현이 되고 있다. 반면에 HSI를 사용하는 (b)는 같은 이미지에 대하여 각각 다른 값을 유도하기가 쉬움을 알 수 있으며 평균값을 적용할 경우 전체 이미지의 색상톤을 알아내기 쉽게 되어있다.

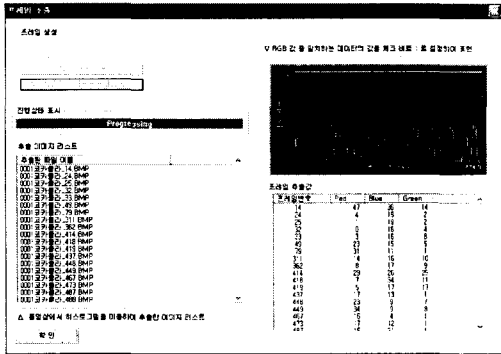


그림 2. 동영상 프레임 추출

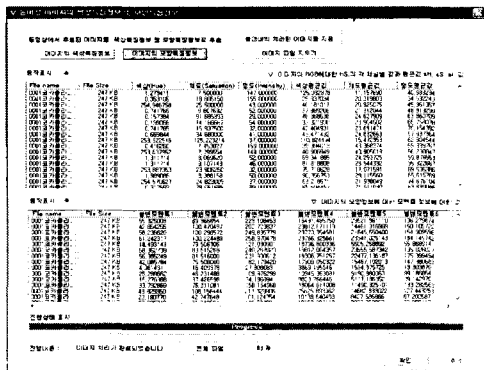


그림 3. 색상특징 정보와 모양특징 정보 추출

〈표 2〉의 (a)는 Hu가 제안하는 불변모멘트에 대한 연산결과이다. (a)는 모양 경계선의 내부 전체를 포함하여 모든 화소에 대해 계산해야 하므로 상당히 많은 계산량이 요구되는 단점이 있고, 뚜렷한 윤곽선에 대하여 1, 2차 모멘트를 구하게 되므로 또 다른 윤곽선에 대하여 연산하는 것은 불가능하다.

(b)는 모든 화소에 대한 연산을 행하지 않고 템플레이트 매칭에 의해 검출된 윤곽선에 해당하는 화소들만 연산하는 향상된 불변 모멘트를 사용하여 연산한 결과이다. 윤곽선 구분이 (a)에 비하여 미세한 부분까지 구별을 할 수 있음을 보인다.

〈표 1〉 (a) 이미지의 RGB 적용 결과

색상특징	코카콜라1.b mp	코카콜라2.b mp	코카콜라3.b mp	코카콜라4. bmp
Red	0	255	255	0
Green	20	1	100	15
Blue	150	50	50	60

(b) 이미지의 HSI와 평균 적용 결과

색상특징	코카콜라1.b mp	코카콜라2.b mp	코카콜라3.b mp	코카콜라4.b mp
색조	0.3531	254.9466	254.9465	0.1990
채도	18.6051	9.8076	91.6853	14.1666
명도	155	52	54	42
색조 평균	29.9374	20.9250	37.8892	33.3219
채도 평균	15.4723	43.9050	28.1154	14.1946
명도 평균	69.5353	50.1073	80.4412	28.7853

〈표 2〉 (a) 불변모멘트 적용 결과

모양특징	코카콜라1.b mp	코카콜라2.b mp	코카콜라3.b mp	코카콜라4.b mp
1차 모멘트	42.8542	61.2423	4.9614	30.1424
2차 모멘트	19306.751	10945.353	10917.903	1268.4458
	3	3	2	

(b) 향상된 불변모멘트 적용 결과

모양특징	코카콜라1.b mp	코카콜라2.b mp	코카콜라3.b mp	코카콜라4.b mp
불변모멘트1	42.8542	61.2423	4.9614	30.1424
불변모멘트2	81.7404	76.3110	108.1584	42.7476
불변모멘트3	249.8353	231.9306	74.1868	58.6083
불변모멘트4	19306.7513	10945.3533	10917.9032	1268.4458
불변모멘트5	8427.5868	10195.0255	280.7870	222.5573
불변모멘트6	135.0260	87.2025	1.1678	13.3889
불변모멘트7	290.0953	200.1345	40.7034	58.0039

V. 결론

최근에 동영상 검색과 저장방법에 대한 연구가 활발히 진행 중에 있는데, 지금까지 연구되어진 동영상 내용분석 및 내용검색의 목적은 방대한 동영상 정보의 자동적인 분류 및 검색을 가능케 하고, 기존의 문자기반방식을 해결함은 물론 주관적인 해석을 배제하여 동영상 데이터에 대한 일관적인 접근이 가능하도록 하는 것이다. 동영상은 연속적인 이미지의 집합과 마찬가지로 동영상 검색한다는 것은 이미지에서 정보를 추출한다는 것과 같다.

동영상 이미지의 색상과 모양정보를 이용하여 데이터

베이스에 저장하고 색상 질의와 모양 이미지와의 비교를 통하여 이미지 특성을 얼마만큼 정확하고 효율적으로 처리 할 수 있는지를 연구하였다. 이미지의 색상정보 추출은 색상변환을 통하여 색상값의 간섭을 줄여 보다 정확한 값을 추출하였다. 모양정보는 템플레이트 매칭을 이용한 윤곽선과 향상된 불변모멘트의 영역정보를 사용하여 이미지에서 표현하고자 하는 특징적 부분을 보다 효과적으로 추출 할 수 있었다. 제안 시스템은 동영상 검색 부분에서 대용량의 자료를 관리하고 액세스할 수 있는 하나의 방법이 될 수 있으며, 동영상이 아닌 이미지 검색 시스템에서도 텍스트 저장방식에서 벗어난 자동화 방법을 구축하는데 도움이 될 수 있다.

향후 연구과제로는 이미지를 단 몇 개의 특징으로만 분류하기에는 방대한 특징들을 포함하고 있으므로 이미지의 객체를 좀더 세분화하는 작업이 필요하고, 잡음이나 조명의 변화에 따른 색상의 변화에 대한 균일성 보정에 관한 연구가 필요하다.

참고문헌

- [1] A.M Tekalp, "Digital Video Processing", Prentice Hall, 1993.
- [2] H. Zhang, A. Kankanhalli and S. W. Smoliar, "Automatic Partitioning of Full-motion Video "Multimedia System, Vol. 1, No. 1, pp. 10 - 28, 1993.
- [3] N. Ikonomakis K.NPlataniotis, M.Zervakis, A.N. Venetsanopoulos, "Region Growing and Region Merging Image Segmentation", IEEE DSP 97. p299-302, 1997.
- [4] 이동근 "색상 및 영역특징 기반 이미지검색 시스템", 숭실대학교 석사학위논문, 1999.
- [5] 김철원, 최기호, "칼라 지정을 이용한 내용기반 화상검색 시스템 구현," 정보처리 논문지 제4권 제4호, 1997.
- [6] 김봉기, "멀티미디어 데이터베이스를 위한 2 단계 내용기반 영상 검색 기법", 숭실대학교 박사학위

논문, 1998.

- [7] Y. Wu, and D. Suter, "A Comparison of Methods for Scene Change Detection in Noisy Image Sequence." Proc. of the First International conference on Visual Information Systems, Melbourne, Australia, pp. 459 - 468, 1996.
- [8] M. Kilger, "A Shadow Handler in a Video-based Real Time Traffic Monitoring System", IEEE Workshop on Applications of Computer Vision, pp.1060-1066, 1992.
- [9] A. Hanjalic, R. L. Lagendijk, and J. Biemond. " A New Key-Frame Allocation Method for Representing Stored Video-Streams", Proc. of the First International Workshop on Image Databases and Multimedia Search, Armsterdam, The Netherlands, pp. 67 - 74, 1996.



최 내 원

1980년 광운대학교 전자계산학과(공학사)
 1988년 숭실대학교 전자계산학과(공학석사)
 1996년~숭실대학교 컴퓨터학과 박사과정
 1983년~현재 명지전문대학 교수
 관심분야 : 멀티미디어, 데이터베이스, 영상처리



지 정 규

1987년 서울산업대학교 전자계산학과(공학사)
 1989년 숭실대학교 전자계산학과(공학석사)
 1998년 숭실대학교 전자계산학과(공학박사)
 1978년~1996년 (주)삼호, 서울시설관리공단 전산실
 1996년~현재 한국학술진흥재단 학술정보팀
 관심분야 : 데이터베이스, 멀티미디어, 디지털도서관