Journal of Korea Multimedia Society Vol. 8, No. 4, April 2005(pp. 450-460)

향상된 불변모멘트를 이용한 동영상 이미지의 특징정보 분석에 관한 연구

이창수†, 전문석††

요 약

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

A Study on Feature Information Parsing of Video Image Using Improved Moment Invariant

Chang-Soo Lee†, Moon-seog Jun††

ABSTRACT

Today, multimedia information is used on the internet and various social areas by rapid development of computer and communication technology. Therefor, the usage is growing dramatically. Multimedia information analysis system is basically based on text. So, there are many difficult problems like expressing ambiguity of multimedia information, excessive burden of works in appending notes and a lack of objectivity. In this study, we suggest a method which uses color and shape information of multimedia image partitions efficiently analyze a large amount of multimedia information. Partitions use field growth and union method. To extract color information, we use distinctive information which matches with a representative color from converting process from RGB(Red Green Blue) to HSI(Hue Saturation Intensity). Also, we use IMI(Improved Moment Invariants) which target to only outline pixels of an object and execute computing as shape information.

Key words: RGB, HSI, IMI, Video Image(비디오 이미지), Key Frame(키 프레임)

1. 서 론

최근의 동영상 정보 검색기술은 방대한 양의 디지털 정보의 수준에 도달하였으나 이에 비해 낮은 사용률을 보인다. 동시에 디지털 저장 매체의 실행 비율은 꾸준히 증가하고 있다. 디지털 정보를 간단하고 낮은 비용으로 얻을 수 있는 것뿐만이 아니라 정보를 저장하고 모으는데 점점 더 쉽게 만들어 내고 한계성을 줄여가고 있다. 그러므로 전문가나 소비자의 환경에서 디지털 처리는 빠르게 성장하고 있다[1].

디지털 라이브러리의 방대한 동영상 데이터를 관리·보존하는데 있어서 현재는 문자기반(SQL)방식

※ 교신저자(Corresponding Author) : 이창수, 주소 : 서울특별시 동작구 상도5동(156-743), 전화 : 02)826-6526, FAX : 02)817-6158, E-mail : powerofmicro@yahoo.co.kr
접수일 : 2004년 8월 5일, 완료일 : 2004년 11월 29일
† 준회원, 숭실대학교 대학원 컴퓨터 공학과 박사과정
(E-mail : powerofmicro@yahoo.co.kr)
†† 숭실대학교 정보과학대학 정교수

을 이용한다. 그러나 검색방법, 질의 방법과 시각 정보(이미지, 그래프, 비디오), 그리고 이러한 부분들의 문자정보를 연결하는 것은 여전히 초기수준밖에는 미치지 못한다[2].

동영상에서 시각 정보를 위한 표현을 얻기 위해서는 내용분석이 요구된다. 기존의 이미지에 몇 가지 내용분석 방법을 가지고 접근을 시도한다. 이러한 방법은 칼라, 텍스쳐, 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[3].

지금 까지 연구되어진 동영상 내용분석 및 내용검색의 목적은 방대한 동영상 정보의 자동적인 분류 및 검색을 가능케 하고, 기존의 문자기반(SQL)방식을 해결하고 또한 주관적인 해석을 배제하여 동영상 데이터에 대한 일관적인 접근이 가능하도록 하는 것이다. 본 연구에서는 기존의 동영상 분석기법을 채택하고 이미지 특징기법에 색상과 모양정보를 사용하여 분석내용의 정확도와 효율성을 향상시키는데 목적이 있다.

본 논문에서 다루는 연구의 내용은 동영상 특징정보의 효율적인 분석을 위한 "동영상 이미지의 특징정보 분석 시스템"의 설계 및 구현에 관한 것이다.

본 논문에서는 동영상에서 컷을 검출하는 단계와 검출된 컷을 이용하여 이미지 내의 특징 정보를 추출하는 두 단계로 분리되어 진행한다. 첫째, 컷을 검출하는 방법으로는 기존의 방법인 히스토그램 차이를 이용한 방법을 적용한다. RGB를 이용한 히스토그램 방법은 기존의 방법 중 구현이 쉽고 컷 검출률을 높일 수 있다. 둘째, 검출된 컷 이미지를 사용해서 특징 추출 기법으로 색상 특징 정보와 모양 특징 정보를 사용하여 이미지 내의 특징을 추출하게 된다.

색상특징 정보는 지역 색상 분포 특성을 고려하기 위하여 입력 영상을 동일한 크기의 다섯 개의 부 영역으로 분할한 후, 분할된 각 부 영역에 대하여 대수적 불변 모멘트 값을 구한다. 모양특징 정보 추출은 템플레이트 매칭 방법을 사용하여 윤곽선을 추출하고, 물체의 윤곽선에 해당하는 화소들만을 대상으로 연산을 수행하는 IMI(Improved Moment Invariants) 처리를 통해 일곱 개의 계수 값을 구한다.

2. 동영상 프레임 검출

2.1 화소 단위의 검출 방법

비디오에서 컷을 검출하는 가장 단순한 방법이다. 화소 단위의 검출 방법은 연속된 두 프레임에서 상응하는 화소 사이의 특징차를 구하여, 그 차이가 임계값 이상이 되면 컷으로 검출하는 방법으로 주로 화소의 농도값을 특징으로 사용한다. 대표적인 방법으로는 Zhang 등이 사용한 방법이 있다[3]. Zhang 등은 연속된 프레임에서 같은 좌표에 위치한 화소간 명암값의 차이가 일정크기보다 큰 화소들의 개수가 임계치 이상인 프레임을 컷으로 검출하였다. 이 방법에서는 카메라의 움직임과 잡영에 의한 영향을 줄이기 위하여 3×3 평균 필터를 사용하였다.

2.2 부분 영역 단위의 검출 방법

부분 영역 단위의 검출 방법은 한 장의 프레임을 몇 개의 부분 영역으로 나누고 부분 영역별로 특징을 구한 후, 연속된 두 프레임에서 상응하는 부분 영역간의 특징차가 임계값 이상이 되는 프레임을 컷으로 검출하는 방법이다.

프레임 전체에 대하여 하나의 히스토그램을 사용하는 검출 방법은 영상 전체의 변화에는 민감한 반면에, 지역적인 움직임에는 영향을 받지 않기 때문에 점진적인 장면 전환을 검출하는 데는 적합하지 않다. 이와는 달리 부분 영역별 히스토그램을 사용한 검출 방법은 각 부분 영역의 변화를 검출할 수 있기 때문에, 전체적인 히스토그램을 사용한 방법에 비하여 보다 상세한 분석을 할 수 있다는 장점을 갖는다[4].

이 방법에서는 하나의 프레임을 객체를 단위로 하는 몇 개의 영역으로 분할하고, 각 영역 별로 농도 히스토그램을 구하여, 그 차의 누적값이 임계값 이상이 되는 프레임을 컷으로 검출한다.

2.3 프레임 단위의 검출 방법

가장 일반적으로 사용되는 프레임 단위의 검출 방법에서는 한 장의 프레임에 대하여 대표적인 특징값을 구하고, 각 프레임간의 특징값의 차이가 임계값 이상이 되는 프레임을 컷으로 검출하는 방법이다. 프레임 단위의 검출 방법에서는 주로 농도 히스토그램, 칼라 히스토그램 등을 특징값으로 사용하며, 이 외에

차영상의 히스토그램이나 에지 영상 등을 사용한 검출 방법이 있다.

농도 히스토그램을 사용한 장면 전환 검출 방법은 영상의 농도값 만을 특징으로 사용하기 때문에 서로 다른 장면의 프레임이라 하더라도, 농도 히스토그램의 분포가 비슷하면 같은 장면으로 분류 되는 문제점이 있다. 이러한 문제점을 해결하기 위해 사용하는 칼라 히스토그램 검출방법은 칼라 모델에 따라 다른 결과를 나타낸다. 따라서 어떤 칼라모델을 적용하는가가 성능에 중요한 영향을 미친다. 검출에 적용되는 대표적인 칼라 모델로는 RGB, HSI, YUV 등이 있다.

Hanjalic 등은 YUV 모델을 사용하였는데 YUV 모델을 사용하여 칼라 히스토그램을 구하고 Y, U, V 각각에 대한 칼라 히스토그램 차이의 누적 값이 임계값 이상이 되는 프레임을 컷으로 검출한다[4,5].

3. 동영상 이미지의 특징정보

3.1 색상 표현

색상을 표현하는 칼라모델은 사용되는 용도에 따라 다양하다. 가장 일반적으로 사용되는 하드웨어 기준 모델들은 칼라 모니터와 많은 종류의 색채 비디오 카메라를 위한 RGB모델, 칼라 프린터를 위한 CMY 모델, 칼라 텔레비전 방송의 표준인 YIQ모델, 이미지 처리를 위해 가장 많이 사용되는 HSI모델, HSV모델 등이 있다.

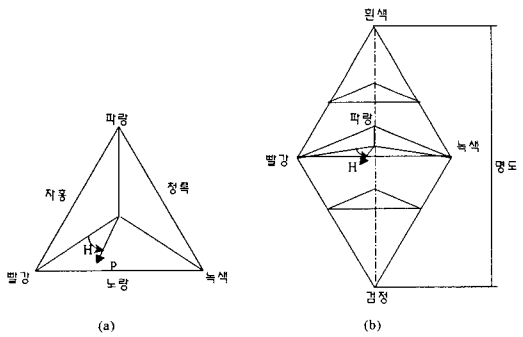

그림 1. 색상 공간

RGB 모델은 그림 1의 (a)에서 보는 바와 같이 원색 당 하나씩 모두 3개의 독립적인 이미지 평면으로 구성되어있다. 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.

3.2 색상특징 정보

색상특징 정보를 표현하기 위하여 Swain이 제안한 색상 히스토그램을 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[8].

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

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[8,9].

3.2 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. 그림 2는 이러한 값들은 회전이나 이동에도 변하지 않으며, 볼록하거나 볼록하지 않은 다면체에 대하여 변하지 않고 상대적으로 계산이 쉽다.

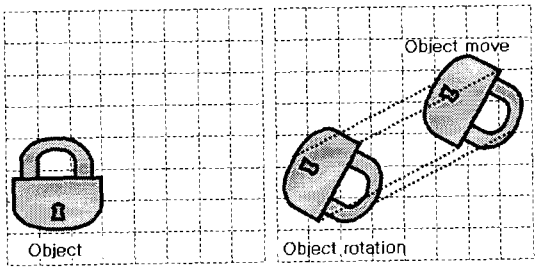

그림 2. 불변모멘트의 회전과 이동

4. 동영상 이미지 특징 분석 시스템

그림 3은 동영상 이미지 특징분석 시스템 모델을 보인다. 분석하고자 하는 동영상 데이터로부터 대표 프레임 추출 후 특징정보를 추출하고 정보 분류 단계를 거쳐 각 클래스별로 특징량을 저장하게 된다. 동영상 대표 프레임 추출 및 특징정보 추출 과정은 다음과 같다.

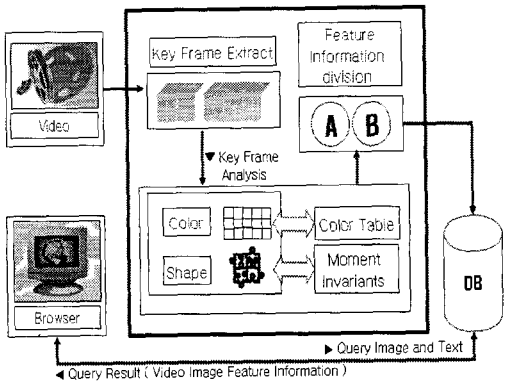

그림 3. 시스템 구성도

단계 1 : 대표 프레임 추출 - 칼라 히스토그램
단계 2 : 특징정보 추출
- 추출된 프레임 이미지에 대한 부분영역 추출
- 색상정보 및 모양정보 추출
단계 3 : 정보분류
- 색상과 모양정보에 대한 유사클래스검색 및 분류, 저장
단계 4 : 정보출력 - 동영상 검색

멀티미디어 검색시스템을 서비스하기 위한 특징량 중 색상정보와 모양정보를 추출하기 위한 시스템이다.

본 시스템은 입력되는 동영상에 대한 대표 프레임을 추출하고 추출된 이미지에 대하여 부분영역을 추출한다. 부분영역의 각 색상정보와 모양정보는 클래스별로 분류하여 데이터베이스에 저장하게 된다.

4.1 칼라 히스토그램을 이용한 대표 프레임 검출

대표 프레임을 검출하기 위하여 칼라 히스토그램을 사용하였다. 기존의 키 프레임 검출 방법 중에서 히스토그램을 사용한 검출 방법은 비교적 정확하게 컷을 검출하기 때문에 동영상 이미지의 모양,색상정보를 얻기 위해서 기존의 방식을 사용한다.

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.

따라서 본 논문에서는 대표 프레임 검출의 정확도와 속도를 고려하여 가장 적당한 방법으로서 칼라 히스토그램을 사용한 검출 방법을 적용하였다[5].

칼라 히스토그램을 사용한 장면 전환 검출 방법은 적용하는 칼라 모델에 따라서 RGB, YUV, HSI의 히스토그램을 이용한다. 실험 결과 어떤 칼라 모델을 적용하던지 대표 프레임 검출의 정확도는 크게 다르지 않았으나 YUV, HSI 칼라 모델을 적용하는 경우 영상을 대상 모델로 변환시키는데 많은 시간이 소요됐다. 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 적색(red), 녹색(green), 파랑색(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. 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.

4.2 템플레이트 매칭에 의한 윤곽선 추출

윤곽선을 추출하는 방법은 많은 이론이 있지만 정확한 윤곽선을 찾기는 어렵다. 1, 2차 미분을 통해 기울기가 변하는 점을 윤곽선으로 검출하는 방법이 가장 기본적인 원리라 할 수 있고, 그것을 응용한 방법들도 많이 제안되어 있다. 기울기 연산자를 사용하여 크기와 방향을 갖는 벡터값으로 표현하는 방법, 소벨(Sobel) 연산법, 라플라시안 2차 미분법 등이 그 예이다. 보다 정확한 윤곽선을 구하기 위하여 끊어진 윤곽선을 연결하는 방법 또는 윤곽선 화소들을 의미있는 경계로 결속하기 위해 고안된 다른 경계 검출 방법을 거치는 경우도 많이 사용한다.

본 논문에서는 부분영역의 경계선을 추출하기 위해 템플레이트 매칭에 의한 윤곽선 추출 방법을 사용한다. 그림 4와 같이 마스터 패턴을 준비해 두고 경계선과 일치하는 패턴을 선택적으로 선택하여 패턴을 적용하여 효과적으로 윤곽선을 추출한다.

제안하는 향상된 불변모멘트의 윤곽선 추출방법은 템플레이트 매칭 방법을 적용하여 기존의 모든 화소를 검사하여 윤곽선을 추출하는 불변모멘트의 연산량을 줄이는 데 효과적으로 사용한다.

4.3 이미지에 대한 부분영역(segment) 추출

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

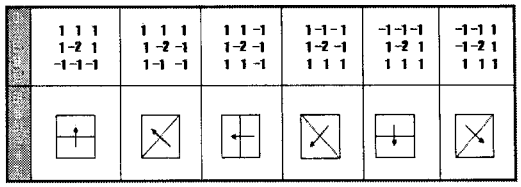

그림 4. 템플레이트 매칭

한다. 세그멘테이션의 이상적인 결과는 사람이 인식하는 의미 있는 물체를 추출해 내는 것이지만 현재까지는 사람이 보기에 의미 있는 객체추출의 알고리즘은 알려진 바 없다[7].

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

$$\sum |H - H_{segment}|^2 + |S - S_{segment}|^2 + |I - I_{segment}|^2 \le T \quad (2)$$

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.

그림 5는 객체의 윤곽선에서 씨앗점의 영역성장을 통하여 부분영역을 추출하는 과정을 도식화 하였다. 객체의 윤곽선은 4.2절에서 설명한 템플레이트 매칭에 의한 윤곽선 추출 방법에 의하여 객체의 윤곽선을 추출하고 씨앗점을 설정하여 부분영역을 설정한다.

4.4 향상된 불변모멘트

모멘트는 어떤 축을 기준으로 한 값들의 분포의 척도이다. 이미지 처리에서는 2차원의 모멘트를 이용하여 그레이 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.

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

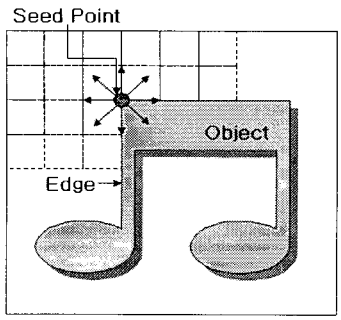

그림 5. 씨앗점과 영역성장

이진 이미지인 경우는 $f(i,j)$는 대상 물체내에서 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 정규화하면 무게중심의 좌표를 얻을 수 있다. 즉

$$i_G = \frac{m_{01}}{m_{00}}, j_G = \frac{m_{01}}{m_{00}} \tag{4}$$

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

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

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

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

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_2는 다음과 같이 정의 할 수 있다

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

식(7)의 2차 모멘트는 u_2는 식(6)에 의하여 $u_{02} + u_{20}$가 성립되므로 각 중심모멘트를 구하여 더한것과 같다.

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

본 논문에서는 이러한 단점을 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 감소시켜 주고 정확한 모양정보의 대표값을 얻을 수 있다.

수정된 (p+q)차 모멘트 m_{pq}는 다음과 같이 정의된다.

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

그리고 차수 p+q≤3인 중심 모멘트는 다음과 같다.

$$\mu_{00} = m_{00}, \mu_{10} = 0, \mu_{01} = 0$$
$$\mu_{11} = m_{11} - \bar{y} m_{10},$$
$$\mu_{20} = m_{20} - \bar{x} m_{10}, \mu_{02} - \bar{y} m_{01},$$
$$\mu_{30} = m_{30} - 3\bar{x} m_{20} + 2\bar{x}^2 m_{10},$$
$$\mu_{12} = m_{12} - 2\bar{y} m_{11} - \bar{x} m_{02} + 2\overline{y^2} m_{10},$$
$$\mu_{21} = m_{21} - 2\bar{x} m_{11} - \bar{y} m_{20} + 2\overline{x^2} m_{01},$$
$$\mu_{03} = m_{03} - 3\bar{y} m_{03} + 2\overline{y^2} m_{01} \tag{9}$$

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

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

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

$$\Phi_1 = \eta_{20} + \eta_{02}$$
$$\Phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$
$$\Phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2$$
$$\Phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2$$
$$\Phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2$$

$$-3(\eta_{21}+\eta_{03})^2]+(3\eta_{21}-\eta_{03})\times(\eta_{21}+\eta_{03})$$
$$\times[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2]$$
$$\Phi_6=(\eta_{20}-\eta_{02})[(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2]$$
$$+4\eta_{11}(\eta_{30}+\eta_{12})(\eta_{21}+\eta_{03})$$
$$\Phi_7=(3\eta_{21}-\eta_{03})(\eta_{30}+\eta_{12})[(\eta_{30}+\eta_{12})^2$$
$$-(\eta_{21}+\eta_{03})^2]+(3\eta_{12}-\eta_{30})(\eta_{21}+\eta_{03})$$
$$\times[3(\eta_{30}+\eta_{12})^2-(\eta_{21}+\eta_{03})^2] \quad (11)$$

4.5 부분 영역에 대한 색상 특징정보 추출

색상정보의 특징값을 추출하는 과정을 살펴보면 입력되는 이미지는 대부분의 프로그램에서 장치 독립적인 이미지 파일인 비트맵 이미지이다. 비트맵 파일의 픽셀 구성은 RGB로 표현되며, 이로부터 특징값 추출에 사용되는 값인 HSI로 변환이 필요하다[8].

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 논문에서도 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.

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

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.

정규화된 값은 각 채널의 평균과 표준편차 왜도를 식(13)와 같이 구한다.

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

여기서 P_{ij}는 입력 이미지 M×N의 j번째 채널에 대한 i번째 화소의 값을, E_i는 각 채널에 대한 평균, S_i는 표준편차, σ_i는 왜도를 C는 HSI의 각 채널을 나타낸다. 따라서 색상 특징정보는 $S_k(E_j, S_j, \sigma_j)$으로 표현되고, S_k는 전체 이미지 또는 k번째 부분영역의 대표 색상 특징정보이다.

4.6 이미지의 특징값 분류 및 저장

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 향상된 불변 모멘트에 의한 대표값을 이용한 모양특징정보에 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 저장을 하게 된다. 색상 정보에 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, 명도값을 사용한다.

5. 실험 및 성능 평가

5.1 실험 환경

시스템의 구현에 사용되는 하드웨어로는 RAM 512MB이고, OS는 윈도우즈 XP Professional 이 탑재된 1.7GHz 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 3분 이내의 CF 동영상으로 선택하여 실험하였다.

5.2 실험 결과 및 성능 평가

그림 6는 제안하는 시스템의 메인화면이고, 그림 7는 동영상에서 히스토그램을 사용하여 각 채널별 값들의 임계값이 넘는 대표 프레임을 추출하는 추출기이다.

그림 8은 추출된 대표 프레임에서 RGB를 변환하여 HSI로 색상특징 정보를 추출하고 템플레이트 매칭 방법으로 추출된 윤곽선정보를 이용하여 7개의 대표값을 계산한다.

그림 9의 (a)는 모양정보의 대표값과 색상값을 추출하여 회전 및 이동에 대한 특징값을 저장하고 동영상에서 대표프레임으로 추출되는 이미지와의 유사성을 비교하기 위하여 인위적으로 생성한 이미지이다. (b)는 샘플이미지에서 향상된 불변모멘트의 7차

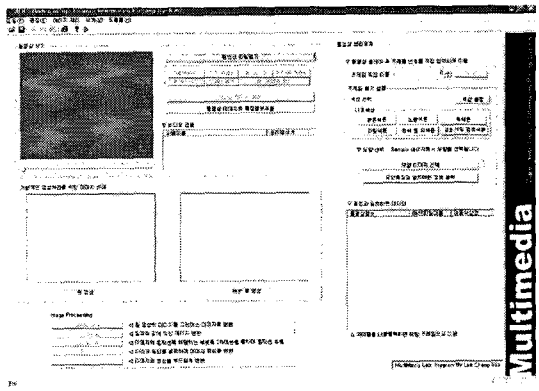

그림 6. 메인 화면

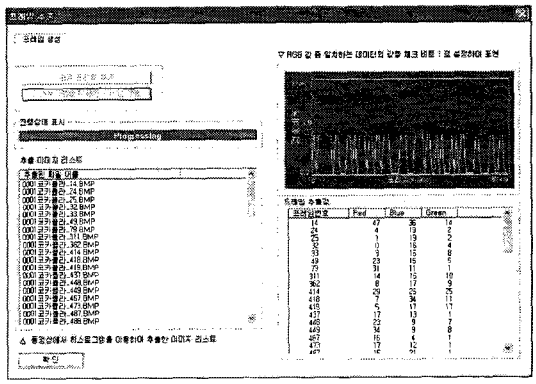

그림 7. 동영상 프레임 추출

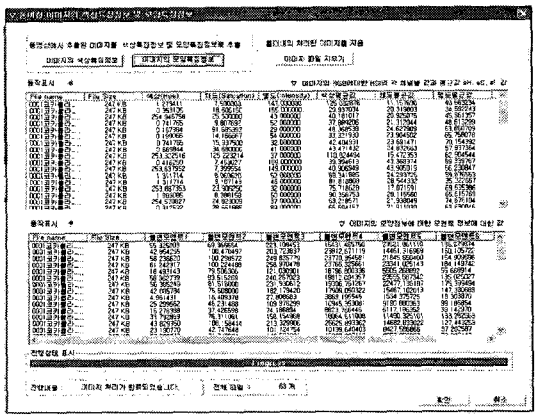

그림 8. 색상특징 정보와 모양특징 정보 추출

모멘트 대표값과 색상정보, 영역정보의 특징값을 추출한 그림이다.

표 1의 (a)는 이미지의 색상특징을 기존의 RGB방식을 적용하여 추출한 결과이다. 이미지에서 전체적인 색상톤은 색상간섭에 의하여 단일색상으로 표현

(a) 샘플이미지 (b) 샘플이미지의 특징값 추출

그림 9 동영상 이미지의 특징정보와 유사도를 비교하기 위한 샘플이미지와 특징값 추출

표 1. (a) 이미지의 RGB적용 결과

색상특징	코카콜라 1.bmp	코카콜라 2.bmp	코카콜라 3.bmp	코카콜라 4.bmp
Red	0	255	255	0
Green	20	1	100	15
Blue	150	50	50	60

(b) 이미지의 HSI와 평균 적용 결과

색상특징	코카콜라 1.bmp	코카콜라 2.bmp	코카콜라 3.bmp	코카콜라 4.bmp
색조(Hue)	0.3531	254.9466	254.9465	0.1990
채도	18.6051	9.8076	91.6853	14.1666
명도	155	52	54	42
채널평균값	29.9374	20.9250	37.8892	33.3219
표준편차	15.4723	43.9050	28.1154	14.1946
왜도	69.5353	50.1073	80.4412	28.7853

표 2. (a) 불변모멘트 적용 결과

모양특징	코카콜라 1.bmp	코카콜라 2.bmp	코카콜라 3.bmp	코카콜라 4.bmp
1차 모멘트	42.8542	61.2423	4.9614	30.1424
2차 모멘트	19306.7513	10945.3533	10917.9032	1268.4458

(b) 향상된 불변모멘트 적용 결과

모양특징	코카콜라 1.bmp	코카콜라 2.bmp	코카콜라 3.bmp	코카콜라 4.bmp
1차 모멘트	42.8542	61.2423	4.9614	30.1424
2차 모멘트	81.7404	76.3110	108.1584	42.7476
3차 모멘트	249.8353	231.9306	74.1868	58.6083
4차 모멘트	19306.7513	10945.3533	10917.9032	1268.4458
5차 모멘트	8427.5868	10195.0255	280.7870	222.5573
6차 모멘트	135.0260	87.2025	1.1678	13.3889
7차 모멘트	290.0953	200.1345	40.7034	58.0039

이 되고 있다. 반면에 HSI의 표준편차, 평균 및 왜도 값을 사용하는 (b)는 같은 이미지에 대하여 각각 다른 값을 유도하기가 쉬움을 알 수 있으며 평균값을 적용할 경우 전체 이미지의 색상톤을 알아내기 쉽게 되어있다.

표 2의 (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 모멘트만을 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.

그림 10은 질의 이미지로 샘플이미지의 특징정보(모양특징, 색상특징)를 사용하여 데이터베이스에 저장되어 있는 동영상 이미지 중 대표프레임으로 추출된 이미지와 비교하였을 때 우선순위가 높은 순으로 나열한 결과이다.

샘플이미지의 7차 모멘트 정보와 색상의 대표값을 데이터베이스의 동영상 이미지 특징정보와 비교하였을 경우, 샘플이미지의 모멘트 정보와 색상값이 가장 유사한 값인 코크1.AVI의 대표프레임 229번의 이미지가 가장 유사하다. 모멘트는 회전, 이동, 축소,

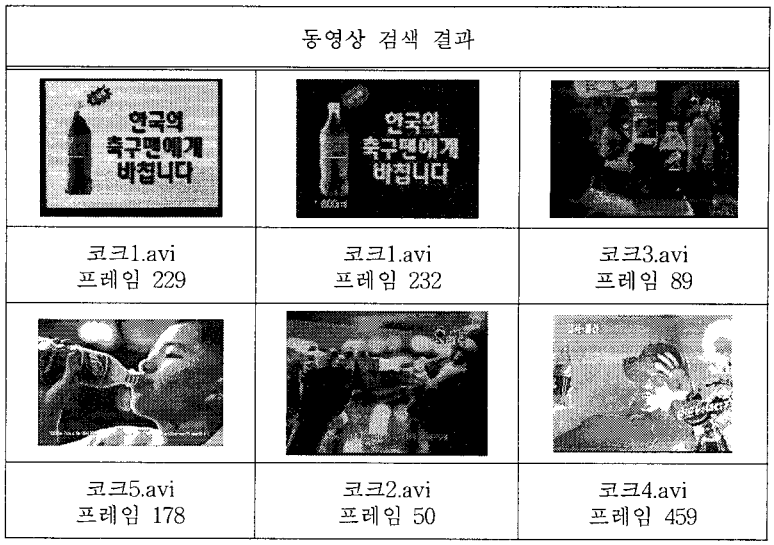

그림 10. 샘플이미지의 특징값을 사용한 동영상 검색결과

확대에 대하여 강인성을 가지고 있지만 샘플이미지와 유사한 크기와 위치를 가지는 이미지가 가장 유사하게 된다.

6. 결 론

동영상에서 시각 정보를 위한 표현을 얻기 위해서는 내용분석이 요구된다. 최근에 동영상 검색과 저장방법에 대한 연구가 활발히 진행 중에 있다. 지금 까지 연구되어진 동영상 내용분석 및 내용검색의 목적은 방대한 동영상 정보의 자동적인 분류 및 검색을 가능케 하고, 기존의 문자기반(SQL)방식을 해결하고, 또한 주관적인 해석을 배제하여 동영상 데이터에 대한 일관적인 접근이 가능하도록 하는 것이다. 동영상은 연속적인 이미지의 집합과 마찬가지이다. 따라서 동영상을 검색한다는 것은 이미지에서 정보를 추출한다는 것과 같다. 이미지 정보 추출은 칼라, 텍스쳐, 모양 등 여러 가지 속성을 추출하는 것이다. 동영상 이미지의 색상과 모양정보를 이용하여 데이터베이스에 저장하고 색상 질의와 모양이미지와의 비교를 통하여 이미지 특성을 얼마만큼 정확하고 효율적으로 처리 할 수 있는지를 실험하였다. 또한 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 표현하는데 보다 효과적으로 추출 할 수 있었다.

제안한 방법의 기대효과는 동영상 검색 부분에서 대용량의 자료를 관리하고 엑세스 할 수 있는 하나의 방법이 될 수 있으며, 동영상이 아닌 이미지 검색 시스템에서도 텍스트 저장방식에서 벗어난 자동화 방법을 구축하는데 도움이 될 수 있다.

참 고 문 헌

[1] P.R.Hill, C.N.Canagarajah, D.R.Bull, "Image segmentation using a texture gradient based watershed transform", IEEE Transactions on Image Processing, VOL.12 NO.12 pp.1618~1633, DECEMBER 2003.

[2] J.L.Starck, F.Murtagh, E.J.Candes, D.L.Donoho, "Gray and color image contrast enhancement by the curvelet transform", IEEE Transactions on Image Processing, VOL. 12 NO.06 pp.0706~0717, JUNE 2003.

[3] H.Zhang, A.Kankanhalli, S.W.Smoliar, "Automatic Parti-tioning of Full-motion Video", Multimedia System, Vol.1, No.1, pp.10-28, 1993.

[4] A.Hanjalic, R.L.Lagendijk, 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, Amsterdam of The Netherlands, pp.67-74, 1996.

[5] Y.Wu, D.Suter, "A Comparison of Methods for Scene Change Detection in Noisy Image Sequence." Proc of the Fist International conference on Visual Information Systems, Melbourne, Australia, pp.459-468, 1996.

[6] N. Ikonomakis K.NPlataniotis, M.Zervakis, A.N. Venetsanopulos, "Region Growing and Region Merging Image Segmentation", IEEE DSP 97. p299-302, 1997.

[7] S.Y.Wan, W.E.Higgins, "Symmetric region growing", International Conference on Image Processing 2000, VOL.12 NO.09 pp.1007~1015, SEPTEMBER 2003.

[8] 이동근 "색상 및 영역특징 기반 이미지검색 시스템", 숭실대학교 석사학위논문, 서울, 1999.

[9] 김철원, 최기호, "칼라 지정을 이용한 내용기반 화상검색 시스템 구현," 정보처리 논문지, 제4권, 제4호, 1997.

이 창 수

1999년 2월 한서대학교 컴퓨터과학과(학사)
2002년 2월 숭실대학교 컴퓨터공학과(석사)
2002년 3월~현재 숭실대학교 컴퓨터공학과 박사과정

관심분야 : 영상처리, 멀티미디어, 네트워크보안

전 문 석

1980년 숭실대학교 컴퓨터공학과(학사)
1986년 University of Maryland 전산과(석사)
1989년 University of Maryland 전산과 (박사)
1989년 Morgan State University 전산수학과 조교수
1989년~1991년 New Mexico State University 부설 Physical Science Lab. 책임연구원
1991년~현재 숭실대학교 정보과학대학 정교수
관심분야 : 네트워크보안, 암호학, 컴퓨터알고리즘