

예제 기반 동영상 검색을 위한 MPEG-7 저작도구 구현

박주현*, 김주영*, 낭종호*, 김경수**, 정병희**, 하명환**
* 서강대학교 컴퓨터학과, ** KBS 기술 연구소
parkjh@mlneptune.sogang.ac.kr

An Implementation of MPEG-7 Authoring Tool for Content Based Video Retrieval by Example

Joo Hyoun Park*, Joo Young Kim*, Jong Ho Nang*,
Gyung Su Kim**, Byung Hee Jung**, Myung Hwan Ha**
* Dept. of Computer Science, Sogang University
** KBS Technical Research Institute

요 약

예제 기반 동영상 검색은 동영상 자체가 검색의 대상이 되는 방법으로써 저급 수준 정보를 추출하여 비교 검색을 수행한다. 멀티미디어 데이터 기술 표준인 MPEG-7 에는 검색에 사용할 수 있는 저급 수준 정보와 비교방법이 기술 되어 있으며, 표준을 준수함으로써 호환 가능한 검색 시스템을 구성할 수 있다. 본 논문에서는 MPEG-7 기반 동영상 검색 시스템을 구성하기 위해 필요한 MPEG-7 구조를 제안하고, 제안된 구조에 맞추어 MPEG-7 인스턴스를 생성하는 저작도구를 설계하고 구현하였다. 저작도구는 MPEG-7 에서 제시하는 visual/audio 기술자중에서 동영상 검색에 효과적으로 쓰일 수 있는 기술자인 Dominant Color, Color Structure, Homogeneous Texture 등 10 가지의 기술자를 자동으로 추출하여 표현할 수 있으며, 오브젝트 단위의 검색이 가능하도록 자동 / 반 자동 오브젝트 추출 기능을 제공한다. 또한 추출된 기술자에 대한 저작자의 이해를 돕기 위해, 각각의 기술자마다 서로 다른 시각화 방법을 제안하고 이를 구현하였다. 동영상에 대한 저작 결과는 검색에 이용될 수 있도록 MPEG-7 표준에 맞는 XML 형식의 문서 파일로 생성한다.

Keyword : Authoring Tool, MPEG-7, Content Based Video Retrieval

1. 서론

컴퓨터 처리 능력의 향상과 인터넷과 같은 컴퓨터 네트워크 성능의 빠른 성장은 컴퓨터에서의 정보 표현 수단을 텍스트 위주에서 동영상 과 같은 멀티미디어 위주로 바꾸어 놓았다. 이는 멀티미디어 데이터의 양적 팽창을 야기하였으며, 이들을 효율적이고 체계적으로 관리할 수 있는 시스템인 디지털 비디오 라이브러리 (Digital Video Library : DVL)에 대한 연구를 활발하게 하였다. 본 논문에서 구현한 저작도구의 목적인 예제 기반 동영상 검색은 대량의 동영상 데이터를 관리하는 DVL 시스템이 갖추어야 할 가장 중요한 요소중에 하나이다. 예제 기반 동영상 검색은 동영상 데이터 자체를 비교하여 검색을 하기 때문에 동영

상에서 어떠한 저급 수준 정보를 추출하여 사용할 것이며, 추출된 정보를 어떻게 비교할 것인지가 매우 중요한 연구 과제가 된다. 이러한 문제에 관해서는 이미 많은 연구가 이루어졌으며, 멀티미디어 기술 표준인 MPEG-7 visual 과 audio [1,2]에서는 검색에 사용할 수 있는 저급 수준 정보와 각 저급 수준 정보에 대한 비교 방법을 제안하고 있다. MPEG-7 을 기반으로 한 동영상 검색을 하기 위해서는 동영상에서 MPEG-7 표준에 근거한 저급 수준 정보를 추출하여 MPEG-7 문서를 생성하는 저작도구가 필요하다. 본 논문에서는 동영상을 분석하여 XML 형태의 MPEG-7 문서를 생성하고, 추출된 저급 수준 정보를 저작자가 쉽게 분석할 수 있도록 저급 수준 정보 시각화를 제공하는 저작도구를 설계 및 구현하였다.

2. 연구 배경

MPEG-7 은 총 7 개의 파트로 구성되어 방대한 양의 오디오/비디오 기술 표준을 제공하고 있다. 본 장에서는 저작도구를 설계하기 위해서 필요한 MPEG-7 의 구성 요소와 객체 추출에 관하여 설명한다.

2.1 MDS(Multimedia Description Schemes)

DS 는 기술자(Descriptor)들과 DS 들간의 관계를 정의함으로써 MPEG-7 기술 기능을 확장 해주는 기능을 갖고 있다[4]. 특히, MDS[3]는 멀티미디어 관련 정보를 다양하게 기술할 수 있도록 100 여개가 넘는 정보들을 기술하고 있으며, 실제 적용에 있어서는 정의되어 있는 수많은 DS 들 가운데서 필요한 DS 들로만 이루어진 스키마를 재구성해야 한다. 본 저작도구에서 사용할 대표적인 DS 는 Segment Ds 이다.

Segment 는 MPEG-7 에서 정의하고 있는 멀티미디어 콘텐츠의 기본단위이며, 세그먼트를 기술하는 기본적인 Segment DS 와 Segment 의 특징 별로 기술할 수 있는 다양한 DS 가 존재한다. 이러한 DS 들 중에서 기술하고자 하는 Segment 의 특징을 나타낼 수 있는 DS 를 선택하여 사용하여야 한다.

2.2 객체 추출

연속된 여러 프레임에 걸쳐 공간적, 시간적인 평가 척도 아래 연관된 영역들의 집합을 객체라고 할 수 있다. <그림 1>은 일반적인 객체 추출 방법 [5]을 보여준다. 투사 및 색 영역 분리(Projection and Segmentation) 모듈은 현재 프레임의 공간적인 영역을 분리한다. 영역을 분리하기 위해서 이전 프레임의 움직임 정보와 공간적인 색 분리

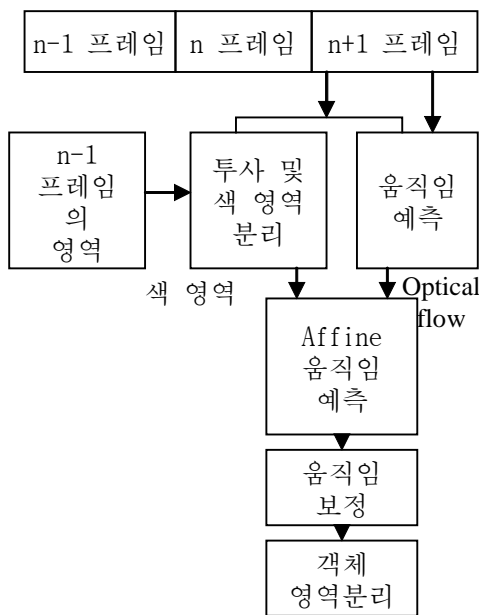


그림 1 객체 추출 방법

(Color Segmentation)방법을 이용한다. 하지만 공

간적인 영역 분리 방법만을 사용해서는 각 영역이 객체의 영역인지 배경의 영역인지를 구분하지 못한다. 이러한 이유로 시간적인 분리 (Temporal Segmentation)작업이 이루어져야 하며 그 방법은 다음과 같다. 현재 프레임과 다음 프레임과의 사이에서 Optical Flow 를 구하고, 이 정보를 바탕으로 나누어진 영역들의 움직임을 예측할 수 있다. 하나의 영역 안에 있는 모션 벡터들은 일정한 방향으로 하여 노이즈를 없애주고, 마지막으로 움직임이 일치하는 인접 영역들끼리 합쳐서 오브젝트를 추출하게 된다. 하지만 카메라 움직임이나 화면 자체의 노이즈, 또는 다중 객체 존재 시, 객체 추출의 정확도는 현저하게 떨어지게 된다. 따라서 최근에는 사용자와의 상호 작용을 고려한 반 자동 객체 추출에 관한 연구[6,7]가 활발히 이루어지고 있다. 이러한 방법은 최초 객체 설정 과정에서 사용자의 도움을 받음으로써 객체 추출의 정확성을 향상시킬 수 있다. 본 논문에서도 최초 프레임의 객체를 사용자가 지정하게 하는 반 자동 객체 추출 방법을 사용하였다.

3. 예제 기반 동영상 검색을 위한 MPEG-7 저작도구 설계

본 장에서는 예제 기반 동영상 검색을 위한 MPEG-7 저작도구를 설계하기 위해서 필요한 스키마 설계와 저급 수준 정보 추출, 그리고 각 저급 수준 정보의 시각화 방법에 관하여 설명한다.

3.1 예제 기반 동영상 검색을 위한 MPEG-7 스키마 설계

예제 기반 동영상 검색을 위한 동영상 구조는 MDS[3]을 이용하여 구성할 수 있으며, 저급 수준 정보는 MPEG-7 Visual/Audio[1,2]에서 제안된 기술자중 동영상 검색에 이용할 수 있는 기술자들을 선택적으로 사용하였다. 먼저, 동영상 구조에 관한 스키마는 다음과 같은 방법으로 설계하였다. 동영상의 구조적인 특징을 살펴보면, 동영상은 각각의 독립된 내용을 가지고 있는 여러 개의 신(Scene)으로 구성되어 있으며, 하나의 신은 화면 전환에 따른 여러 개의 샷(Shot)으로 나눌 수 있다. 또한 샷에서는 샷을 대표하는 키 프레임(Key Frame)이 존재한다. 동영상을 MPEG-7 을 이용하여 기술할 때에는 이러한 동영상에 대한 구조적인 특징을 기술할 수 있어야 한다. 동영상은 오디오와 비디오 정보가 함께 포함되어 있으므로, 전체 동영상에 해당하는 세그먼트의 기술은 오디오와 비디오 정보를 기술할 수 있는 AudioVisualSegment DS 를 이용하도록 한다. 전체 동영상은 여러 개의 신으로 구성되어 있는데 신 또한 오디오와 비디오 정보를 포함하고 있으므로 마찬가지로 AudioVisualSegment DS 를 사용하며, 전체 동영상 세그먼트와는 시간적으로 연결되어 있으므로 TemporalComposition 을 이용하여 시간 관계를 기술한다. 샷 또한 신 안에 시간적으로 연결되어 있기 때문

에 AudioVisualSegment DS 와 TemporalDecomposition 를 이용하여 기술하도록 한다. 한편 샷은 동영상 검색 단위로 사용되기 때문에 샷을 기술하는데에는 오디오 정보와 비디오 정보를 세분화하여 기술하도록 한다. 오디오 정보는 AudioSegment DS 를 이용하며, 비디오 정보는 VideoSegment DS 를 이용하도록 한다. 샷을 대표하는 키 프레임은 한장의 영상이나 비디오 프레임을 기술할 수 있는 StillRegion DS 를 이용하여 기술한다. 객체 단위의 검색을 하기 위해서 객체를 위한 기술도 필요하게 되는데, 객체는 키 프레임에 포함되어 있다고 할 수 있으므로 프레임 안의 영역으로 볼 수 있으며, StillRegion DS 를 이용하여 기술할 수 있다. 키 프레임과 객체는 공간적으로 연결되어 있으므로 SpatialDecomposition 을 이용하여 두 세그먼트 간의 관계를 기술한다. 전술한 스키마를 그림으로 표현하면 <그림 2>와 같다.

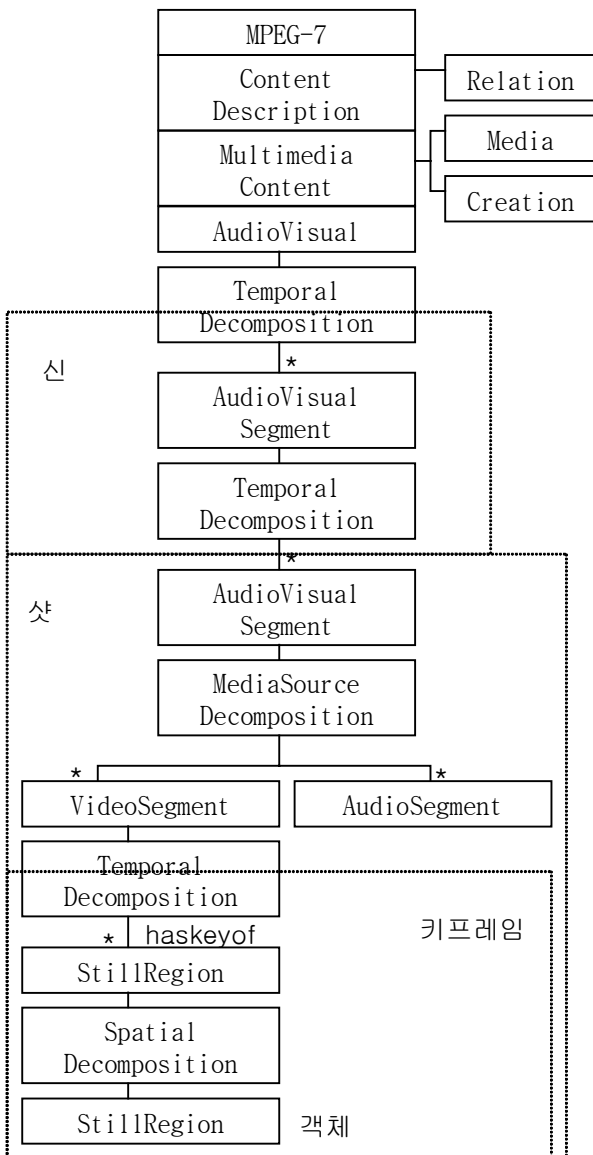


그림 2 스키마 구조

3.2 저급 수준 정보의 시각화 방법

본 논문에서 사용한 MPEG-7 의 저급 수준 정보는 다음과 같다.

- Dominant Color 는 영상의 임의의 영역 안에 많이 포함되어 있는 색을 표현하는 기술자인데, 화소(pixel)간의 색상을 비교하여 차이 값이 임계 값보다 작으면 화소를 병합해 나가며 전체 화면에 많이 포함되어 있는 색상을 찾는 방법을 이용하여 추출한다. 최대 5 개의 대표 색상으로 표현하며 <그림 3>와 같이 막대 그래프로 시각화하여 표현 할 수 있다.

Dominant Color



그림 3 Dominant Color 시각화

- Color Structure는 색상의 뭉쳐진 정도를 히스토그램으로 나타낸 기술자로, 8x8 크기의 윈도우 안에 포함되어 있는 색상의 빈의 값을 증가시켜가면서 영상 전체에 대해 256개의 빈의 히스토그램을 구한다. 각 빈의 대표 색상을 구할 수 있으며 <그림 4>와 같이 가로축을 256개의 빈으로 하고, 세로축을 히스토그램 값으로 한 그래프로 표현할 수 있다.

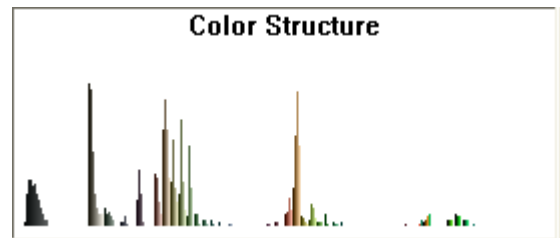


그림 4 Color Structure 시각화

- GoF/GoP Color는 각 프레임에 대해 Scalable Color를 구한 후, 결과로 나온 계수(Coefficient)에 대해 평균을 구한다. Scalable Color는 HSV 형태의 색상 값을 Haar 변환을 통해 인코딩 후 BitPlane형태로 저장한다.
- Homogeneous Texture는 재질을 표현하는 30개의 주파수에 대한 평균, 표준 편차, 에너지, 에너지의 편차를 수치로써 표현하는 기술자이다.
- Edge Histogram은 영상을 4x4 영역으로 나누어, 각 영역에 가로, 세로, 45°, 135°, 무정형의 에지 성분이 얼마나 포함되어 있는지를 히스토그램으로 표현하는 기술자이다. 각 영역에서 가장 많은 분포를 가지는 에지 성분을 <그림 5>와 같이 표현하여 시각화 할 수 있다.

Edge Histogram



그림 5 Edge Histogram 시각화

- Region Shape는 각도와 반지름 성분이 각각 다른 35개의 영역을 기본 함수로 표현하여 이 함수들이 포함되는 정도로써 기술하려는 영역의 화소의 분포를 기술한다. 35개의 기본 성분의 히스토그램은 ART(Angular Radial Transform) 함수를 통하여 35개의 값으로 나타난다. Region Shape와 Contour Shape는 객체의 모양을 표현 하기 위한 것이므로, 선택된 객체의 외곽선을 표현해 줌으로써 시각화 할 수 있다.



그림 6 Shape 시각화

- Motion Activity는 프레임 별로 Block Matching을 구하여 움직임의 정도를 표현하는 기술자이다. 움직임 정보를 좀 더 세분화 하여, 전체 움직임 정도와 방향성을 표현할 수 있으며, 각 Block의 움직임 정도를 공간적, 시간적으로 표현하고 있다. 이 중 시각화에 사용한 Motion Activity정보는 Intensity이며, 샷 간의 변화 정도를 표현할 수 있도록 <그림 7>와 같이 타임라인에 배열하여 표현하였다.



그림 7 Motion Activity 시각화 방법

- Motion Trajectory 는 일정 시간 동안 동영상 내의 객체의 움직임을 2 차원 혹은 3 차원 좌표로 기술하는 기술자이다. 본 논문에서는 2 차원 기술 방법을 사용하였으며, 객체의 중심 이동 좌표를 사용하여 시각화 하였다.

Motion Trajectory



그림 8 Motion Trajectory 시각화

- Region Locator는 영역의 공간적인 위치를 표현할 수 있는 기술자이다. 이 기술자가 자료형(Data Type)으로 가질 수 있는 Poly 기술자는 임의의 도형을 좌표로써 기술할 수 있는데, 이 Poly 기술자를 이용하여 객체가 키 프레임에서 차지하는 공간적 위치 정보를 화소의 좌표로 기술할 수 있다.
- WaveForm 기술자는 진폭(Amplitude) 값을 10ms인 단위로 나누고 1초에 해당되는 100개의 단위에 대해 한 개의 평균(Mean)값을 구한 후, 총 시간만큼 평균값들을 나열한다. 오디오 정보는 일정 시간 동안의 값을 이용하므로, WaveForm 기술자는 샷에 대한 오디오 저급 수준 내용 정보가 된다.

이러한 Visual 와 Audio 기술자들은 비디오의 저급 수준 정보로써 샷, 키 프레임과 객체에 대해 기술이 되어야 한다. XM 에 따르면, GoF/GoP 기술자를 제외한 색상이나 재질의 기술자들은 한 영상에 대해, 외형 기술자는 영상 중에서도 객체와 같은 일부 영역에 마스크 처리된 부분에 대해 기술 및 추출하고 있다. 움직임의 Motion Activity 와 색상의 GoF/GoP Color 기술자는 동영상과 같은 연속된 프레임에 대해서 기술 및 추출하며, Motion Trajectory 기술자는 동영상에 포함된 객체의 움직임 좌표에 대해서 기술하고 있다. 그렇기 때문에 기술하려고 하는 세그먼트에 맞추어 저급 수준 내용 정보를 사용해야 한다. 이러한 점에서, GoF/GoP 와 Motion Activity, Motion Trajectory 기술자는 샷에 대한 정보이며, 나머지 색상과 재질의 기술자들은 키 프레임에 대한 정보로 분류할 수 있다. 또한 외형 기술자들과 Region Locator 기술자는 한 프레임 안의 부분 영역으로 볼 수 있는 객체에 대한 정보이며, 객체는 동영상 검색 시에 질의 단위가 될 수 있으므로 외형 정보뿐만 아니라 키 프레임에서 포함되어 있는 저급 수준 내용 정보도 포함될 필요가 있다. 따라서 객체에 대한 저급 수준 내용 정보로써, GoF/GoP Color 를 제외한 색상, 재질 기술자들도 기술 및 추출을 하였다.

4 예제기반 동영상 검색을 위한 MPEG-7 저작 도구 구현

4.1 사용자 인터페이스

저작도구의 전체 모습은 <그림 9>와 같다. 좌측 상단에는 동영상을 재생할 수 있는 재생 윈도우가 있으며, 그 아래에는 신 단위로 저급 수준

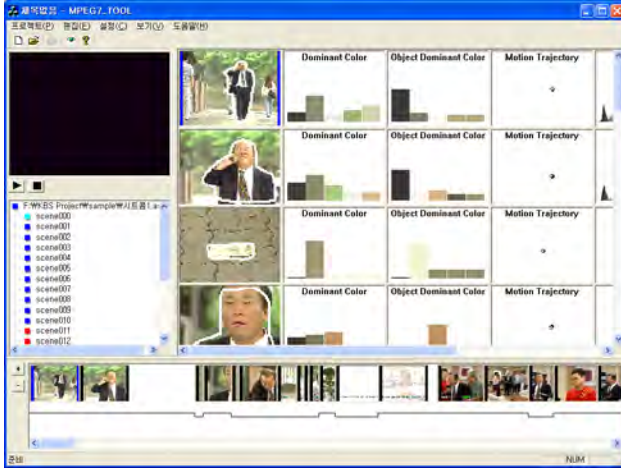


그림 9 MPEG-7 저작 도구

정보를 추출하거나 시각화하기 편하도록 신으로 구성된 트리 뷰가 있다. 중앙에는 현재 선택된 신에 속한 샷에 대한 저급 수준 정보가 행 단위로 보여주는 뷰가 있다. 하단에는 키 프레임과 Motion Activity 를 시간 축으로 배열한 타임라인 뷰가 있다. 저급 수준 정보 뷰는 트리 뷰에서 선택된 신 또는 타임라인 뷰에서 선택된 샷이 포함된 신의 샷들을 행 단위로 표현하는 뷰이다. 좌측 열부터 키 프레임, Dominant Color, Object Dominant Color, Motion Trajectory, Color Structure, Edge Histogram 을 시각화하여 표현하고 있으며, 객체를 추출하지 않았을 경우 3, 4 번째 열은 공란으로 표현된다. 또한 키 프레임을 선택하여 객체 추출을 할 수 있다. 객체 추출도구는 독립된 프로그램으로 존재하며 <그림 10>과 같은



그림 10 객체 추출 도구

인터페이스를 갖는다. 객체는 한번 이상 추출할 수 있으며, 최신의 추출 정보가 이전의 추출 정보를 대체

체하게 된다. 타임 라인 뷰는 동영상 전체 샷을 시간 축 기반으로 키 프레임을 나열 한 것이다. 타임라인에서 샷을 선택하게 되면 선택된 샷이 포함되어 있는 신을 바탕으로 신 트리 뷰, 저급 수준 정보 뷰와 동기화하여 각각의 뷰를 갱신하게 된다. 키 프레임 밑에서 5 단계로 표현되는 Motion Activity 그래프가 그려져 있으며, 키 프레임과 같이 동기화 하여 움직인다.

4.2 객체 추출 도구의 구현

본 논문에서 사용한 객체 추출 방법은 관심 있는 영역을 객체로 지정하여 질의 시 사용할 수 있도록 하는데 목적이 있다. 따라서 반 자동 객체 추출 방법을 사용하였으며 그 방법은 <그림 11>과 같다. 입력 동영상의 첫 프레임에서 저작자가 영역을 지정해 주면 정확한 객체 추출을 위해서 외곽선 자동 수정 과정을 거치게 되는데, 본 논문에서는 snake 알고리즘[8]을 사용하였다. 외곽선 자동 수정 후, 설정된 외곽선에 포함되어 있는 영역을 판단하면서 첫 프레임에 대한 객체 추출 작업을 완료하게 된다. 영역을 판단하기 위해서 사용한 방법은 Mean Shift 알고리즘[9]을 사용한 공간

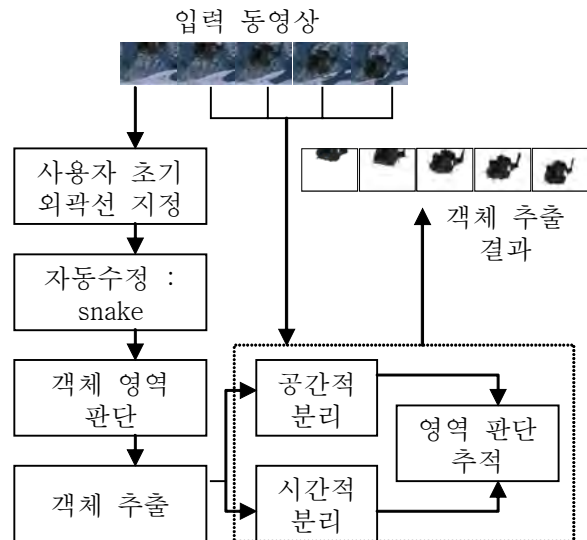


그림 11 객체 추출 및 추적 구성도

적 분리 방법이다. 공간적 분리 후, 같은 색상을 갖는 화소들 간의 병합과정을 통해 영역 분리를 한다. 영역 분리 과정을 거치면 초기 외곽선 안에 포함 여부를 조사하여 객체 포함 여부를 판단한다. 그 후에는 프레임 별로 공간적 분리와 시간적 분리를 하면서 초기 객체에 대응되는 영역들을 추적해 나간다. 전 프레임에서 개체로 판단된 영역들과 현재 프레임에서 공간적 분리를 한 결과로 나온 영역들, 그리고 두 프레임 사이의 밝기 차를 이용하여 영역을 비교한다. 전 프레임에서의 어떤 영역에 밝기 차가 있다면 움직임이 있다는 것을

의미하므로 현 프레임에서의 영역들 가운데에서 같은 영역으로 추정되는 영역을 찾는다. 영역 비교 방법은 영역의 크기와 위치 정보를 바탕으로 하게 된다. 비교하려는 두 영역의 크기의 차와 위치의 차가 임계 값보다 작다면 같은 영역으로 판단하여 현 프레임의 해당 영역은 객체에 포함되는 영역으로 간주하게 된다.

본 연구에서는 기존의 객체 추출 연구에서의 영역 추적 방법들에 비해 비교적 간단한 방법을 사용하였는데, 이는 본 연구가 MPEG-7 을 바탕으로 한 저급 수준 내용 정보를 이용하여 검색 및 비교를 수행 하기 때문이다. MPEG-7 을 바탕으로 하여 객체를 질의로 사용할 때 사용되는 저급 수준 내용 정보는 거의 대부분이 샷 안에서 한 장의 키 프레임에 기술된 객체의 정보이다. 그러한 정보에는 색상, 질감, 외형 정보가 해당된다. 움직임에서 사용되는 정보는 객체의 움직임 좌표를 표현하는 Motion Trajectory 만 해당되며, 이는 색상, 질감, 외형 정보에 비해 중요도가 떨어지므로 객체 추적의 정확도 보다는 시간상의 효율을 고려하는 것이 바람직하다. <그림 12>는 (a)의 가장 왼쪽 프레임에서 사용자가 헬리콥터를 지정해 준 후, 자동 수정을 거쳐 객체 추적을 한 결과를 (b)에서 보여준다.



(a) 입력 동영상



(b) 객체 추적 결과

그림 12 객체 추적 결과

4.3 저작

동영상에 대한 최초 저작일 경우에는 인덱싱 도구를 통해 동영상을 인덱싱 한 후에 인덱싱 결과를 사용하여 새로운 MPEG-7 인스턴스를 생성한다. 이후의 저작을 위해서는 MPEG-7 인스턴스와 동영상 파일만을 가지고도 가능하며 MPEG-7 인스턴스 파일 열기를 통해 초기화 된다. 이와 같이 초기화 된 상태에서 신 단위의 저작이 가능하게 된다. 저작도구는 자동으로 선택된 신에 속해 있는 각 샷에 대해서 저급 수준 정보를 추출하게 되며, 추출된 저급 수준 정보에 대해서는 하나의 샷에 대한 저작이 끝날 때마다 MPEG-7 인스턴스에 저장되어 저작 결과를 유지하게 된다.

5 결론

멀티미디어 기술 표준인 MPEG-7 은 다양한 동영상 기술 방법을 제안하고 있으며, 이를 이용하여 효율적인 동영상 검색 시스템을 구성할 수 있

다. 하지만 이러한 MPEG-7 을 사용한 응용을 하기 위해서는 표준에 따른 동영상 분석 도구가 필요하게 된다. 본 논문에서는 예제 기반 동영상 검색을 하는데 적합하도록 MPEG-7 스키마 구조를 제안하였으며, 제안된 스키마에 맞추어 동영상을 분석/저작하여 MPEG-7 문서를 생성하는 저작도구를 설계하고 구현하였다.

본 저작도구는 MPEG-7 표준을 따름으로 해서 다른 시스템과의 호환성을 보장할 수 있으며, 각 기술자들에 대한 시각화를 제공함으로써 저작자가 동영상의 저급 수준 정보를 분석하기 쉽도록 하였다. 또한 객체 단위와 샷 단위의 기술자를 따로 기술함으로써 생성된 MPEG-7 문서를 샷 단위 검색 뿐만 아니라 객체 단위의 검색에도 이용할 수 있도록 하였다.

본 저작도구의 개선 방향은 크게 두 가지로 생각해 볼 수 있다. 첫 번째는 고정된 스키마가 아닌 가변적인 스키마를 사용할 수 있도록 하여 확장 성을 향상시키는 것이고, 두 번째는 저작자가 고급 수준 정보를 입력할 수 있도록 하여 저급 수준 검색과 고급 수준 검색을 병행할 수 있도록 하는 것이다. 이러한 두 가지 개선점을 보완한다면 보다 나은 동영상 검색을 위한 MPEG-7 저작도구를 구현할 수 있을 것이다.

6 참고 문헌

1. ISO/IEC JTC1/SC29/WG11 *Information Technology -Multimedia Content Description Interface-Part3: Visual*, 2001.
2. ISO/IEC JTC1/SC29/WG11 *Information Technology -Multimedia Content Description Interface-Part4: Audio*, 2001.
3. ISO/IEC JTC1/SC29/WG11/N3966, Text of 15938-5 FCD *Information Technology-Multimedia Content Description Interface-Part5:Multimedia Description Schemes*, 2001.
4. B.S.Manjunath, P.Salembier, and T.Sokora, *Introduction to MPEG-7*, John Wiley & Sons, 2002.
5. S.F.Chang, W.Chen, H.J.Meng, H.Sundaram and D.Zhong, "VideoQ:An Automated Content Based Video Search System using Visual Cues," *Proceedings of ACM Multimedia 97*, 1997.
6. S.Cooray, N.O'Connor, S.Marlow, N.Murphy, T.Curran. , "Semi-Automatic Video Object Segmentation Using Recursive Shortest Spanning Tree and Binary Partition Tree," *Workshop on Image Analysis for Multimedia Interactive Services*, 2001.
7. C.Gu, M.C Lee , "Semantic video object tracking using region-based classification," *Proceedings of International Conference on Image Processing*, 1998.
8. M.Kass, A.Witkin and D.Terzopoulos, "Snake : Active Contour Models," *International Journal of Computer Vision*, 1988.
9. D.Comanicu, P.Meer, "Robust analysis of feature spaces: color image segmentation," *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997.