

論文2001-38CI-1-5

동적 분할 기법을 이용한 비디오 데이터의 대표키 프레임 추출 (A Dynamic Segmentation Method for Representative Key-frame Extraction from Video data)

李淳熙*, 金榮熙**, 柳根鎬***

(Soon Hee Lee, Young Hee Kim, and Keun Ho Ryu)

요 약

시간적 특성을 가진 비디오 자료와 같은 멀티미디어의 자료에 접근하기 위해서는 내용 기반 이미지 검색 기술이 필요하다. 더욱이, 내용 기반 이미지 검색의 기본적인 기술 중의 하나가 대표키 프레임들의 추출이다. 제안된 방법을 구현했을 뿐만 아니라, 대표키 프레임들이 비디오 데이터의 특성에 대한 데이터 분석을 사용하여 추출될 수 있음을 보였다. 또한, 제안된 방법이 정확할 뿐만 아니라 효과적이라는 것을 증명하였다. 제안한 방법은 비디오 데이터 베이스를 위해 색인을 구축하는데 매우 유용하다. 그러므로 제안한 방법이 실제 계에서 비디오 데이터 베이스를 구축하는데 사용되기를 기대한다.

Abstract

To access the multimedia data, such as video data with temporal properties, the content-based image retrieval technique is required. Moreover, one of the basic techniques for content-based image retrieval is an extraction of representative key-frames. Not only did we implement this method, but also by analyzing the video data, we have proven the proposed method to be both effective and accurate. In addition, this method is expected to solve the real world problem of building video databases, as it is very useful in building an index.

I. 서 론

멀티미디어의 디지털 이미지, 동영상, 음성 자료 등과 같은 비정형 자료의 활용은 대용량 저장 장치와 컴퓨터 처리 속도의 개선, 초고속 자료 전송이 가능한 통신

망의 구축, 그리고 비정형 멀티미디어 처리 기술 등에 힘입어 상당한 발전을 이루어 정보 통신 사회에서 다양한 응용 사례와 서비스를 제공하고 있다.

멀티미디어 자료인 동영상은 영상 정보 뿐 아니라 음성 정보와 문자 정보를 비롯한 여러 가지의 정보들을 가지고 있으며, 포함하고 있는 내용이 방대할 뿐 아니라 매우 다양하고 시간적, 복합적, 비정형적인 특징을 지닌다. 이러한 동영상 자료에 대한 검색의 필요성은 방송이나 교육 등 비디오 자료를 매체로 이용하여 정보를 전달하는 분야에서 더욱 증대되고 있다. 또한 인터넷과 디지털 도서관 등을 통하여 산재된 멀티미디어 정보들을 손쉽게 이용할 수 있게 됨에 따라, 원하는 자료를 고속으로 검색할 수 있는 효율적인 시스템이 필요하게 되었다.

그러나 동영상의 방대한 크기로 인하여, 필요한 정보

* 正會員, 案山工科大學 컴퓨터情報科

(Dept. of Computer Information, Ansan college of Technology)

** 正會員, 成均館大學校 大學院 電氣電子컴퓨터工學部
(Dept. of Electrical and Computer Engineering, Sungkyunkwan University)

*** 正會員, 忠北大學校 컴퓨터科學科

(Dept. Computer Science, Chungbuk National Univ.)

接受日字:2000年5月2日, 수정완료일:2000年12月27日

를 기존의 키워드나 색인 기반의 방법으로 검색하기에는 이미 한계점에 도달한 실정이다. 그러므로 동영상의 내용을 효과적으로 표현하고 이를 효율적인 검색을 위하여 다양한 방법의 내용 기반 검색에 대한 연구가 진행되고 있다^[1].

영상자료에 대한 정보 검색의 시도는 이미 1980년대부터 시작되었는데, 초기에는 텍스트 중심의 검색 방법을 이용하였다. 이 방법은 멀티미디어 자료를 검색하는데 제약이 많았고, 검색을 위한 사전 작업의 하나인 색인 작업에 훈련된 전문가가 요구되었기 때문에 시간과 비용 면에서 효율이 떨어졌다. 또한 시간적인 특성을 가지고 있는 자료를 사용자의 요구에 의해 상연해 볼 수 있도록 하는 시간 지원 데이터베이스의 영상화에 관한 연구^[2]와 비디오 자료의 내용을 검색하기 위하여 비디오의 내용을 대표할 수 있는 특징을 추출하고, 이를 기초로 색인을 구성하여 검색할 수 있는 내용 기반 검색에 관한 연구가 있었다^[3,4,5].

비디오 자료를 장면으로 분할하거나 구조적으로 브라우징을 하기 위해서는 기본적인 영상 정보를 사용한다. 이 영상 정보에 대한 내용 기반 검색을 위하여 다음과 같은 핵심 요소 기술들이 요구된다.

첫째, 비디오 파싱은 비정형 적이고 대용량인 비디오 자료를 계층적 구조 형태로 재구성하기 위한 기술이다. 이 기술은 비디오 분할 단계와 비디오 색인 단계로 구성된다. 또한 사용되는 자료의 종류에 따라서 비압축 비디오와 압축 비디오에 대한 기술로 나눌 수 있다.

둘째, 비디오 검색 및 브라우징 기술이 있다. 이는 기존의 데이터베이스는 텍스트로 이루어진 질의어를 사용하여 검색을 수행하지만, 비디오 데이터베이스는 텍스트뿐만 아니라 영상이나 오디오 정보 등으로 구성된 질의어를 사용하며, 특징간 유사도를 사용한 정합 방식을 통하여 접근할 수 있다^[6].

이러한 기술은 비디오 자료를 분할하고 분할된 영역의 내용을 나타낼 수 있는 프레임을 선택하는 것이다. 비디오 자료의 분할은 샷(shot)이 기준이 된다. 이는 인접한 두 프레임간의 유사도의 차이가 동일한 샷 안에서 보다 장면 전환점(cut)에서 훨씬 크다는 사실을 이용한다. 즉 인접한 두 프레임간의 특징 차이를 구한 후 비교 값이 임계값 보다 클 경우에 장면 전환이 이루어졌다고 보는 것이다. 이러한 연구에는 연속된 두 프레임에서 상응하는 화소간의 특징 차이가 임계값을 초과하면 장면 전환이 이루어졌다고 보는 화소간의 차이

비교법^[7], 연속된 프레임들에서 대응하는 일정 영역의 통계 값에 의한 유사율 측정법^[8], 화소의 세기나 색상을 히스토그램으로 표현하여 유사도를 히스토그램 분포 차이 비교^[9,10], 그밖에도 에지 변화 비교^[11], 압축 상관 계수 비교^[12], 그리고 움직임 벡터 비교^[13] 등의 방법이 있다.

그러나 비디오 자료는 특히 자료량이 방대하기 때문에 장면 전환 검출을 통하여 얻은 샷별로 그 내용을 잘 나타낼 수 있는 소수의 키 프레임을 추출하고, 내용 기반의 연산을 이들의 집합에 적용함^[14~17]에 너무 많은 자료를 다루어야 한다. 이러한 점을 고려하여 최근 연구에서는 샷뿐만 아니라 씬(scene)에 관련된 대표키 프레임들을 검출하는 기법들이 등장하고 있다.

이런 방법에는 계승된 특징 차이를 이용한 방법과 내용기반 클러스터링 방법이 있다. 두 프레임간의 특징 차이를 구하는데 있어서, 계승된 특징 차이를 이용한 방법은 두 프레임은 다른 장면을 나타내지만 우연히 특징 차이가 비슷할 수도 있다는 점을 고려하였다. 즉, 두 프레임의 중간 프레임과의 특징 차이와 비교하여 가장 큰 차이 값을 가지는 것을 선택하여 적용하는 기법이다^[15]. 또한 내용 기반 클러스터링 방법은 미리 규정한 일정 크기로 분할하고, 해당 분할 단위의 첫 프레임과 마지막 프레임간의 특징을 구한다. 산출된 특징 값을 차이가 큰 값과 작은 값을 갖는 집단으로 정렬하여, 차이가 적은 집단은 첫 프레임과 마지막 프레임을 대표키 프레임으로 추출한다. 이런 방법을 원하는 수의 대표키 프레임이 될 때까지 반복적으로 적용하는 방법이다.

지금까지의 연구에서는 대표키 프레임을 추출하기 위하여 균등 분할 방법을 사용하였다. 이는 비디오 자료를 필요 이상으로 분할하거나 또는 적게 분할함으로써 대표키 프레임이 중복 추출되거나 또는 추출 대상에서 제외되는 경우가 발생하는 문제점이 있다. 이 연구에서는, 이런 문제점을 개선하여 비디오 분할에 동적 분할 기법을 제안한다. 이 기법은 비디오 자료의 특성을 분석하여 분할 기준을 동적으로 적용하여, 대표키 프레임을 추출하는데 효율성과 정확성을 높일 수 있다.

이 논문의 구성은 다음과 같다. 2장에서는 내용 기반 비디오 검색의 핵심 기술에 관한 기존의 기법들을 알아보고, 이어 3장에서는 비디오 자료의 구조 분석과 대표키 프레임 추출을 위한 분할 기법을, 4장에서는 본 논문에서 제안한 비디오 특성을 고려한 동적 분할 방

식, 마지막으로 5장에서는 결론 및 향후 연구 방향에 관하여 서술한다.

II. 키 프레임 검출 기법

비디오 자료에서, 연속적인 카메라 동작에 의해 촬영된 샷은 비디오 분할의 기본 단위이면서 몇 개의 프레임들로 구성된다. 비디오 자료는 특성상 내용별, 분야별로 평균 프레임 수가 차이가 있다. 또한 하나 이상의 샷들의 집합으로 씬이 구성되는데, 이는 하나의 일관된 이야기를 나타내며, 어떤 의미를 부여 할 수 있는 내용 분류의 기준이 된다.

샷을 이루고 있는 프레임에서, 연속된 두 프레임의 특징 차이가 임계값을 초과하면 장면 전환이 있다고 판단하고, 이를 기준으로 키 프레임을 추출한다. 몇 개의 샷들로 구성된 씬의 내용을 나타낼 수 있는 프레임을 대표키 프레임이라 한다. 즉 연속된 프레임들의 집합에서 첫 번째 프레임과 마지막 프레임의 특징 차이가 비슷하다면, 이는 하나의 씬을 나타내고, 이를 대표하는 프레임이 대표키 프레임이다. 이런 프레임 검출 기법은 크게 키 프레임 검출 기법과 대표키 프레임 검출 기법으로 분류할 수 있다.

1. 샷 경계면 검출 기법

비디오 분할은 장면 전환이나 점진적 장면 전환 효과로 나타나는 샷이나 장면들 사이의 경계를 검출하는 기법이다. 비디오 자료의 연속된 두 프레임의 특징을 측정하여, 그 차이가 임계값 이상일 경우에 샷의 경계 지점으로 검출한다.

내용 기반의 검색을 위한 비디오 자료 처리의 첫 단계는 비디오 자료를 샷 단위로 분할하는 일이다. 비디오 분할에서 샷 경계 검출이란, 샷이 전환하는 경계 프레임을 검출하는 방법으로, 샷이 전환되는 종류는 편집 방법에 따라 급진적인 장면 전환과 점진적인 장면 전환으로 나눌 수 있다^[6].

점진적인 장면 전환은 급진적인 장면 전환의 것에 비하여 두 샷의 프레임간 특징 차이가 뚜렷하지 않기 때문에 샷 경계를 검출하는데 어려움이 따른다. 현재까지의 많은 연구들이 비디오 분할 방법들은 컷을 대상으로 하였지만, 점진적인 장면 전환이나 카메라 동작에 의해서 발생하는 장면 전환을 효과적으로 검출하기 위한 새로운 방법에 대한 연구도 진행되고 있다^[4].

대부분 인접한 두 프레임간의 특징 차이가 동일한 샷 안에서 보다 장면 전환점에서 훨씬 크다는 사실을 이용한다. 이러한 특성을 이용한 가장 간단한 방법은 인접한 두 프레임간의 특징 차이를 구한 후 비교 값이 임계값 보다 클 경우, 장면 전환이 이루어졌다고 보는 것이다.

화소간의 차이 비교 방법은 화면을 구성하는 화소들은 동일한 장면 안에서 변화가 적다는 성질을 이용한다. 연속된 두 프레임에서 상응하는 화소 사이의 특징 차이가 임계값을 초과하면 그 프레임간에는 장면 전환이 있다고 보아 장면 전환이 생긴 프레임을 검출하는 방법으로, 화소간의 명도 값이나 에지 영상의 차이 값을 특징으로 한다.

$$DP_i(k, l) = \begin{cases} 1 & \text{if } |P_i(k, l) - P_{i+1}(k, l)| > t \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

식 (1)에서 P_i 는 i 번째 프레임, (k, l) 은 화소의 위치를 나타낸다. t 는 적당한 임계값이다. 식(1)은 식(2)처럼 화소당 평균 변화율로 표현될 수도 있다. 화소 단위 검출 방법은 잡음이나 물체의 이동에 매우 민감하다는 단점을 갖는다.

$$\frac{\sum_{k,l=1}^{M,N} DP_i(k, l)}{M \times N} \times 100 > t \quad (2)$$

유사율 측정법은 영상 정보 안에서 낱말의 화소를 비교하는 것이 아니라 연속된 프레임들에서 대응하는 일정 영역의 통계 값을 비교하는 방법이다^[6]. 먼저 프레임을 몇 개의 구역으로 나눈다. 그 다음 구역마다 유사율을 계산하여 이 값이 임계값 t 를 초과하는 구역이 많을 때 컷이 있다고 간주한다.

$$\left[\frac{S_i + S_{i+1} + \left(\frac{m_i - m_{i+1}}{2}\right)^2}{S_i \times S_{i+1}} \right] > t \quad (3)$$

여기서 S_i 는 i 번째 프레임의 분산, m_i 는 평균이다. 이 방법은 느린 동작의 영상에서 일어나기 쉬운 오검출에 강하다. 그러나 비교되는 두 구역의 평균과 분산이 같으나 완전히 다른 확률 밀도 함수를 갖는다면, 아무런 장면 변화를 발견하지 못한다.

히스토그램 비교법은 화소의 세기(Y성분)나 색상(Cb, Cr)을 히스토그램으로 표현하여 식 (4)처럼 히스토그램 차로 유사도를 측정한다^[6,10]. 히스토그램 법은

화소 단위 비교보다 카메라 또는 물체의 움직임에 민감하다는 특징을 갖는다.

$$|\sum P_i(k, l) - \sum P_{i+1}(k, l)| > t \quad (4)$$

여기서 $H_i(j)$ 는 i 프레임의 세기 j 의 도수이다. SD_i 가 주어진 임계값 보다 크면 컷이라고 정의한다. 이를 다음처럼 정규화 하여 쓰기도 한다.

$$SD_i = \frac{\sum_{j=1}^c |H_i(j) - H_{i+1}(j)|^2}{H_{i+1}(j)} \quad (5)$$

식 (5)은 두 프레임 사이의 차이를 강조하여 카메라 이동이나 물체 이동에 따른 미세하게 발생한 작은 차이도 크게 한다. 이런 방법들은 선정된 임계값이 카메라 동작보다 낮을 때 오검출의 가능성이 있으며, 너무 높으면 점진적인 씬 전환이 포함된 비디오 열을 거의 검출하지 못한다. 이 검출법은 화소 단위 비교보다 움직이는 물체가 존재하거나 또는 일반적인 샷 안에서의 두 프레임 사이의 섬광이나 선명한 조명이 있을 경우 카메라 브레이크를 검출하는 오류를 발생한다. 부분 영역 단위의 샷 경계 검출 기법은 한 장의 프레임을 겹치는 N 개의 부분 영역으로 분할하고, 각 영역 특성별로 특징 값을 구한 후, 연속된 프레임에서 상응하는 부분 영역간의 특징 차이가 임계값 이상이 되는 프레임을 샷의 경계 프레임으로 검출하는 방법이다. 부분 영역 단위의 샷 경계 검출 기법에서는 주로 히스토그램이나 이동 벡터 등을 특징 값으로 사용한다^[13]. 프레임 단위의 검출 방법은 가장 일반적으로 계산 과정이 간단하고 처리 시간이 짧다는 장점을 갖는 반면, 프레임의 위치 정보를 고려하지 않는다는 단점을 가지고 있다.

2. 키 프레임 추출 기법

키 프레임 추출 기법은 내용 기반 검색에서 대부분 브라우징을 목적으로 제안되었기 때문에, 그 추출 과정에 있어서 프레임간 순차성의 제한을 받으며 휴리스틱한 요소에 상당 부분을 의존한다. 따라서 키 프레임 문제를 좀 더 정량적으로 해결할 수 있는 연구가 필요하다.

기존의 키 프레임 추출 방법 중에서, 가장 간단한 것은 각 샷마다 임의의 한 프레임을 추출하여 이를 키 프레임으로 삼는 것이다. 우선 샷의 첫 프레임을 키 프레임으로 선택하고, 다음 프레임들을 탐색하면서 처음

프레임과 특징 차이가 임계값 이상이면, 그 프레임을 또 다른 키 프레임으로 선택한다. 즉, 마지막으로 선택한 키 프레임과 유사하지 않은 프레임으로 판명이 날 경우, 그 프레임을 대표해 줄 수 있는 키 프레임이 없으므로 이를 또 다른 키 프레임으로 선택하는 것이다. 이와 같은 과정을 샷의 가장 마지막 프레임에 도달할 때까지 순차적으로 반복하여 실행한다. 내용의 변화도가 많지 않은 샷에 대해서는 적은 수의 키 프레임을, 내용의 변화도가 극심한 샷에 대해서는 상대적으로 많은 수의 키 프레임을 추출할 수 있다.

키 프레임이 추출되면, 이들을 순차적으로 나열하여 압축된 형태의 브라우징을 할 수 있다. 그러나 TV 30분 분량의 동영상을 가정할 경우, $30(\text{min}) * 60(\text{sec}/\text{min}) * 30(\text{frame}/\text{sec}) = 54,000$ 개의 프레임을 가지고 있다. 그러나, 5% 비율의 키 프레임 집합으로 대표한다고 할 때 여전히, 2,500개 이상의 키 프레임을 가지게 되므로 효과적인 브라우징을 하기가 어렵다. 이를 위하여 계층적(hierarchical)형태의 브라우징^[17] 기법과 장면 전이 그래프(scene transition graph, STG)^[18] 등의 브라우징 기법들이 있다.

3. 대표키 프레임 추출 기법

동영상은 특히 자료 양이 방대하다. 그러므로 요약하여 대표적으로 보여줄 수 있는 소수의 대표 프레임을 추출하여, 내용 기반의 연산을 이들 대표 집합에만 적용함으로써 빠른 시간 내에 검색 결과를 취득하고자 하는 시도가 있었다^[14,15,17]. 그러나 대부분은 장면 전환 검출을 통하여 얻은 샷 별로, 그 내용을 잘 나타낼 수 있는 소수의 대표키 프레임을 추출하는 방식이다. 비디오의 장면 전환에서 발생하는 샷은, 기존의 연구에서는 샷의 경계 지점을 검출하는데 관심이 모아졌으나, 최근의 연구에서는 샷뿐만 아니라 씬과 관련된 대표키 프레임들을 검출하는 기법들도 등장하고 있다.

1) 계승된 특징 차이를 이용한 기법

계승된 특징 차이를 이용한 기법은, 비디오 분할 과정에서 프레임들 간의 특징 차이를 계승하여 비디오의 주요 프레임들을 짧은 시간 안에 검출할 수 있는 기법이다. 이 방법에서는 사용자가 요구한 대표키 프레임 수 n 의 p 배수로 균등 분할하는 방법을 사용하고 있다. 이 방법은 사용자가 비디오를 미리 상연 해 보고, 그 내용을 분석하여, 이를 근거로 개략적인 씬의 개수를 파악하여 대표키 프레임 수를 요구한다면, 매우 우수한

결과를 기대할 수 있다. 그러나 이는 미리 내용을 파악한 경우에 적용이 가능한 일이며, 내용을 파악하지 못한 모든 비디오에 적용하는데는 한계가 있다. 왜냐하면 비디오 자료는 분야마다 내용의 변화, 카메라 워킹, 편집 스타일 등과 같은 특성이 다르기 때문이다.

이 기법은 계승된 특징 차이를 이용하여 n 개의 대표키 프레임을 검출하는데, 먼저 비디오 전체를 균등 분할하여 m ($m > n$)개의 프레임을 추출한다. 다음으로, 계승된 특징 차이가 가장 작은 두 개의 프레임 F_i 와 F_j 를 선택하고, 이 두 프레임을 해당 프레임 그룹의 중간 프레임인 $F(i+j)/2$ 로 대체한다. 마지막으로 남은 프레임 수가 n 이면 다시 계승된 특징 차이가 가장 작은 두 개의 프레임 선택하여 반복적으로 적용한다^[5].

2) 내용 기반 적응 클러스터링 방식

내용 기반 적응 클러스터링 방식^[9]은 기존의 일반적인 방식처럼 비디오 컷들을 검출하지 않고도 대표키 프레임을 찾는다. 우선, 비디오의 각각의 프레임의 특징 값을 구한다. 그리고 미리 규정한 L 의 크기로 비디오의 프레임 간격을 나눈다. 나누어진 구간마다 각각의 구성 단위 차를 구하는데, 이는 해당 단위의 처음 프레임과 마지막 프레임의 특징 차이를 의미한다. 그 다음 구성 단위마다 갖는 특징 차의 크기로 정렬한 후, 구성 단위 차의 값 r 에 따라 차이가 큰 집단과 작은 두 집단을 형성한다. 이를 토대로 특징 차이가 적은 집단에 있는 각 구성 단위들의 처음 프레임과 마지막 프레임만 남기고 모든 프레임을 삭제한다. 차이가 작은 집단에 있는 각 구성 단위들의 처음과 마지막 프레임과 차이가 큰 집단의 모든 프레임들을 현재 추출된 대표키 프레임으로 간주하는 것이다. 최종적으로 남아 있는 프레임의 수가 대표키 프레임 개수가 원하는 개수보다 적으면 작업을 중지하고, 그렇지 않으면 남아 있는 그룹의 프레임들을 새로운 내용을 갖는 비디오로 간주하여 동적으로 재그룹 하여 프레임을 구간으로 나누는 단계로 돌아간다.

비디오는 내용별, 분야별로 샷당 프레임 수가 다른 특성이 있으며, 어떤 의미를 전달하기 위한 씬을 나타내는 샷의 개수도 해당 비디오마다 다른 특징이 있다. 그러나 위의 두 방법에서는 사용자의 개입이나 미리 정해진 일정길이의 분할 기준을 적용하고 있어서, 모든 대표키 프레임을 효율적으로 추출하지 못하는 문제점이 있다.

이 문제를 해결하기 위하여 비디오의 특성에 따라서 분할 기준을 달리하는 동적 분할 기법을 제안한다.

III. 동적 분할 기법을 적용한 대표키 프레임 추출

1. 비디오 자료의 구조 설명

비디오 자료는 시간적 특성을 가진 프레임들의 집합이 샷이 되고, 이런 샷들을 대표하는 키 프레임이 있으며, 각각 시간적 특성을 가진 샷들의 집합으로 하나의 의미를 전달 할 수 있는 씬을 나타내는 구조적 특징을 가지고 있으며, 이런 씬을 대표하는 대표키 프레임이 있다.

이와 같은 비디오 자료의 시간적, 구조적 특성을 다음과 같이 정의할 수 있다.

【정의 3.1】 비디오 자료(VS)를 구성하는 구조적 성분(*scene, shot, frame*)은 다음과 같이 정의된다.

$$VS = (Oid, Type, a_1, a_2, \dots, a_m)$$

여기에서

Oid : 비디오 단위를 나타내는 식별자

$$Type \in \{scene, shot, frame\} \wedge a_1, \dots, a_m$$

비디오 자료 VS는 객체식별자 Oid를 가지며, $a_1, a_2, a_3, \dots, a_m$ 으로 이루어져 있으며, 이들은 각각 구조적 성분중의 하나인 *scene, shot, frame* 중의 한 형태를 가지며, \wedge 는 AND, \vee 는 OR를 의미한다.

【정의 3.2】 비디오 자료의 구조적 성분인 *scene, shot, frame*의 시간 관계성 정의는 다음과 같다.

$$frame = (VS.Ts = VS.Te)$$

$$shot = (VS.Ts < VS.Te) \wedge (s_frame < e_frame)$$

$$scene = (s_shot < e_shot)$$

여기서

Ts : 비디오 자료의 시작 시간, Te : 비디오 자료의 끝 시간

s_frame : 시작 프레임, e_frame : 끝 프레임

s_shot : 시작 shot, e_shot : 끝 shot

*frame*은 연속된 비디오 자료상의 어느 한 순간 즉 시작 시간과 끝 시간이 같은 하나의 프레임을 나타낸다. *shot*은 비디오 자료 중 시간적으로 어느 시점(Ts)에 시작되어 이보다 늦은 다른 시간(Te)에 종료되며, s_frame이 e_frame보다 앞서는 시간성을 갖는 일련의 *frame*들의 집합이다. *scene*은 *shot*들의 집합으로 구성되며, *shot*들 사이에도 시간적으로 시작 *shot*이 먼저

오고, 나머지 shot들도 시간적 순서에 따라 구성된다.

【정의 3.3】 샷을 대표하는 키 프레임(Key-frame)은 샷을 구성하고 있는 프레임중의 하나이며, 다음과 같이 정의한다.

$$Key-frame = (VS.Ts = VS.Te) \wedge (s_frame \text{ between } e_frame)$$

Key-frame은 비디오 자료상의 한 frame이면서, shot의 시작 프레임과 끝 프레임 사이에 존재하는 프레임이다.

【정의 3.4】 씬을 대표하는 대표키 프레임(Rep-Keyframe)은 씬을 구성하는 프레임중의 하나이며 다음과 같이 정의한다.

$$Rep-Keyframe = Key-frame \wedge (s_shot \text{ between } e_shot)$$

Rep-Keyframe은 씬의 시작 샷과 끝 샷들 사이에 있는 한 개의 Key-frame이다.

비디오 자료는 정의된 것과 같이 시간 및 구조적 특성을 가지고 있다. 하나 이상의 샷들로 구성된 씬은 의미를 전달하는 기본 단위이다. 씬의 의미를 전할 수 있는 대표키 프레임을 추출하여 시간적 순서로 나열하면 전체적인 내용의 흐름을 나타낼 수 있고, 이들 대표키 프레임을 이용하여 비디오 데이터베이스의 색인을 구축하여 내용 기반 검색을 지원한다.

2. 분할 기준 선정

비디오의 분야별 특성을 파악하기 위하여 실험을 통하여 표1로 정리하였다. 표1에서와 같이 분야별로 샷당 평균 프레임 수가 다른 특징이 있었다. 즉 비디오 자료는 분야별로 표현하려는 내용에 따라 편집 기법이 다르기 때문이다. 즉 30초 분량의 상업용 광고의 경우 총 프레임 수가 2455, 샷이 65개로 샷당 평균 프레임 수는 38개 프레임이 된다. 뉴스의 경우는 14분 분량일 때 프레임 수가 25628개, 샷이 210개로 샷당 평균 프레임 수는 122개 프레임이다. 이와 같이 상업용 광고라는 분야 특성상 샷당 프레임 수가 뉴스보다 적은 특성을 가지고 있다. 이런 자료의 분야별 특성을 이용하여 분할 기준을 제시할 수 있으나^[1] 같은 분야의 비디오도 내용별로 샷당 평균 프레임 수 크게 다르다. 따라서 이 연구에서는 비디오 자료에 따른 특성을 고려한 동적 분할 기법을 적용한 대표키 프레임 추출을 제안한다.

표 1은 분야별 샷당 평균 프레임 수를 나타낸 것으

로 이를 위한 기본 자료는 표2에 제시한 비디오 영상을 기초로 하여 계산하였다. 이를 보면 비디오는 분야별로 특징이 있음을 알 수 있을 뿐만 아니라 같은 분야에서도 내용별로 차이가 많이 있음을 알 수 있다.

표 1. 비디오 자료의 특성
Table 1. Property of video data.

Video Sequence	# Frames	# Shots	Min:Sec	Frames per average shot
Commercials	2455	65	1:48	38
Movie clip	12329	187	6:44	66
Music Video	14395	108	7:52	133
News	25628	210	14:00	122

비디오는 분야가 갖는 특성에 의해 샷당 평균 프레임 수가 다르며, 이는 어떤 의미의 구분을 나타내는 씬에 해당하는 영상의 길이도 해당 분야별로 다른 특징을 나타내고 있다. 비디오는 표1에서와 같이 샷당 프레임 수가 각각의 영상 분야별로 특징이 있음을 알 수 있다. 즉, 광고나 뮤직 비디오같이 많은 내용을 짧은 시간에 함축적으로 표현하는 분야는 샷당 프레임 수가 적고, 드라마나 뉴스, 스포츠 등과 같이 그 내용이 설명적이거나 어떤 사실을 보여주는 분야는 샷당 프레임 수가 매우 크다. 이런 분야별 특징을 반영하지 않고 일정한 크기로 분할하여 대표 키 프레임을 추출하는 것은 문제가 있다. 그러므로 대표키 프레임을 추출하기 위한 분할 기준을 비디오의 특징에 맞도록 적용을 해야 한다. 그러나 기존의 대표키 프레임을 추출하는 기법에서는 분할 기준을 일정한 크기로 고정하거나, 사용자가 원하는 대표키 프레임 수의 일정 배수가 되도록 분할 기준을 정하고 있다.

대표키 프레임은 임의의 어떤 장면에 해당하는 씬을 대표하는 키 프레임을 말한다. 즉 씬의 범위 안에서 분할이 이루어지는 것이 가장 효율적이다. 이러한 특성을 배제하고 일정한 길이로 분할하는 것은, 과도한 분할이 되거나 혹은 적은 분할로 대표키 프레임을 찾는데 효율성이나 정확성이 낮아지게 된다. 이런 문제점을 해결하고자 우선 분할 기준을 선정하는 다음과 같은 알고리즘을 제안한다.

단계1) 비디오 자료 중에서 일정량의 자료를 설정한

단계2) 선정된 범위 길이 내의 샷과 프레임 수를 추출한다.

단계3) 추출된 샷당 프레임 수를 산출한다.

단계4) 샷당 가장 큰 프레임 수와 가장 작은 프레임 수를 제거한다.

단계5) 나머지로 샷당 평균 프레임 수를 계산하여 이를 분할 기준인 P 로 사용한다.

설정된 비디오 자료에 대하여 샷의 수와 샷당 프레임 수를 히스토그램 비교법(샷을 추출할 수 있는 다른 방법들도 가능함)을 이용하여 추출한다. 추출된 자료 중에서 샷당 프레임 수가 가장 많은 자료와 적은 자료는 제외시키는데 이는 전체 평균수에 미치는 영향이 매우 크기 때문이다. 이렇게 계산된 샷당 평균 프레임 수는 분할 기준인 P 가 된다.

P 를 구하는 식은

$$P = \frac{\sum_{i=1}^n FS_i - F_{max} - F_{min}}{n-2} \quad (6)$$

식 (6)에서 FS_i 는 i 번째 샷의 프레임 수이며, F_{max} 는 프레임 수가 가장 많은 것, F_{min} 은 프레임 수가 가장 작은 것, n 은 샷 수를 나타낸다.

씬은 1 개 이상의 샷으로 이루어진다. 한 개의 샷으로 구성된 씬의 경우에 분할 기준 P 를 적용하면 샷 단위로 분할되어 적어도 모든 샷을 구분할 수 있게 된다. 이러한 근거로 P 를 분할 기준으로 선택한다.

하나의 동영상 자료 안에서도 도입, 전개, 결말, 즉 내용의 전개상 샷의 길이가 다르다. 따라서 이 논문에서도 하나의 비디오 자료 중에서도 도입, 전개, 결말의 3개 부분으로 나누어서, 분할 기준을 선정하여 적용하였다. P 를 위한 자료의 선정이 일정 분량 이상을 선정하면, 계산 량이 많아지게 되므로 비효율적이 되고, 너무 적으면 정확도가 떨어지게 된다.

3. 동적 분할 기법 이용 키 추출 과정

제안한 알고리즘은 어떤 분야의 비디오라도 자료의 특성을 고려하여 동적 분할 기준을 선정하여 적용할 수 있어 기존의 균등 분할로 인한 문제점을 해결하였으며, 사용자가 대표키 프레임의 수를 지정하는 등의 사용자 개입을 배제 할 수 있어 자동화가 가능하다.

다음은 대표키 프레임을 추출하기 위해서 비디오 자료의 특성이 고려된 동적 분할 기법을 적용한 제안하는 전체 알고리즘이다.

단계1) 동적 분할 기준 알고리즘을 적용하여 분할 기준 P 를 선정

단계2) P 로 비디오 자료를 분할하여 경계지점의 프레임들을 추출하여 대표키 프레임의 후보로 선정한다.

단계3) 후보로 선정된 프레임 중 연속된 두 프레임 F_i 과 F_j 를 선택하고, 두 프레임간의 특징 차이 값은, 선택된 두 프레임과 두 프레임의 중간 프레임과의 각각의 차이 값들 중 가장 큰 값을 선택하고, 이들 값이 임계값 보다 작으면, 두 개의 프레임 F_i 과 F_j 를 선택하고 단계3)으로 가고, 특징 차이가 임계값 보다 큰 두 프레임 F_i 과 F_j 은 대표키 프레임으로 선정된다.

단계4) 두 프레임의 F_i, F_j 를 해당 프레임 집단의 중간 프레임인 $F(i+j)/2$ 로 대체한다.

단계5) 더 이상 특징 차이가 임계값 보다 적은 것이 없으면 멈추고, 그렇지 않으면 단계3)으로 돌아간다.

단계3)에서 연속된 두 프레임과 두 프레임의 중간 프레임과의 특징 차이 값을 비교하는 이유는, 두 프레임은 각기 다른 씬을 나타내지만 우연히도 특징이 비슷하여 히스토그램의 값이 비슷할 수도 있다. 이 경우에는 중간 프레임과의 값의 차이로 인하여, 두 프레임이 각각 다른 씬을 나타내는 대표키 후보에서 탈락되는 것을 방지할 수 있다.

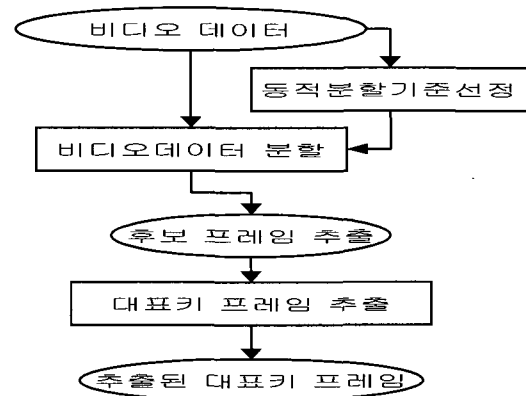


그림 1. 대표키 프레임 추출 과정

Fig. 1. Process of Representative Key-frame Extraction.

그림1은 동적 분할 알고리즘을 적용하는 전체적인 흐름을 나타내고 있다. 영상 자료에서 대표키 프레임을 추출하기 위하여 입력 대상이 되는 비디오 자료의 10%~15%를 선정하였다. 짧은 자료의 경우에 선정된 부분

내의 샷의 수가 적어 오차가 커지므로 15%의 자료에 의해 분할 기준을 계산하였다. 또한 대부분의 비디오 자료는 이야기 도입, 전개, 결말이 있다. 이 논문에서도 이를 적용하여 3개 부분으로 나누어 각각 적용하였다.

선정된 비디오 자료에 대하여 샷의 수와 샷당 프레임 수를 히스토그램 비교법을 이용하여 추출하였다. 샷을 검출할 수 있는 다른 어떤 방법도 가능하지만, 이 논문의 전체적인 알고리즘에는 영향이 없으므로 히스토그램 비교법을 이용하였다. 추출된 자료 중에서 샷당 프레임 수가 가장 많은 자료와 가장 적은 자료는 제외하였다. 이는 전체 평균수에 미치는 영향이 매우 크기 때문이다. 이렇게 계산된 샷당 평균 프레임 수는 분할 기준인 P가 된다. 썸은 최소한 1 개 이상의 샷으로 이루어지므로, 이 기준 P를 적용하면 샷 단위로 분할되며, 대표키 프레임 추출시 중간 프레임과의 비교하는 과정에서 분할 기준 이하의 짧은 썸도 고려하게 된다. P의 크기를 근거 비디오 자료를 분할하고 분할 경계선에 해당 프레임들을 추출한다. 추출된 프레임들 중에서 연속된 두 프레임간의 특징 차이 값은, 두 프레임과 이들의 중간 프레임과의 값들 중 가장 큰 값을 선택한다. 선택한 값이 임계값 보다 적으면, 두 개의 프레임 F_i 과 F_j 를 선택하여, 이들 두 프레임의 F_i , F_j 를 해당 프레임 그룹의 중간 프레임인 $F(i+j)/2$ 로 대치한다. 이와 같은 작업은 계속해서 진행되며, 더 이상 특징 차이가 임계값 보다 적은 것이 없으면 멈추고, 그렇지 않으면 계속해서 연속된 두 프레임간의 특징 차이 값을 비교한다. 두 프레임간의 특징 차이 값이 적다는 것은 두 프레임이 비슷한 장면을 나타내므로, 두 프레임의 중간 프레임을 택하여도 이는 이들을 나타내는 대표키 프레임이 될 수 있기 때문이다. 또한 추출된 프레임 중 연속된 두 프레임의 중간 프레임과의 특징 값을 비교하는 것은 분할된 기준 이하의 아주 짧은 썸이 있을 경우에도 해당 썸의 대표키 프레임을 추출하기 위한 것이다.

IV. 구현 및 평가

비디오 자료의 내용 및 분야별 특성을 고려하여, 동적 분할 기법을 적용하여 대표키 프레임을 추출하는 방법을 평가하기 위하여, 실험에 사용된 시스템 전체의 구조와 설계된 시스템 각각의 구성 요소들을 기술하고, 그 결과를 평가하였다.

1. 시스템 구조

전체 시스템의 구조는 그림2와 같으며, 이 논문에서는 표시된 부분과 같이 비디오 자료의 분석 부분과 대표키 프레임 추출 부분을 다루었다. 비디오 자료의 구조화를 통하여 나온 대표키 프레임은 객체 지향 모델에서 비디오 자료에 관한 내용 기반 검색을 위한 색인을 구성하기 위한 메타 데이터로 이용하게 된다.

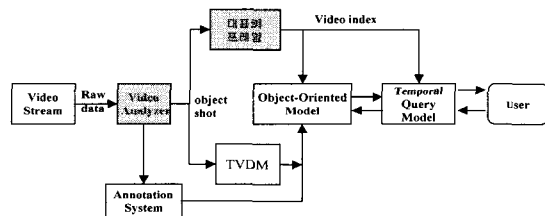


그림 2. TVDM의 전체 시스템 구조
Fig. 2. System Architecture of TVDM.

비디오 자료의 입력은 비디오 분석기(Video Analyzer)에서 메타 데이터를 위한 주석 처리 부분(Annotation System)과 비디오 자료 색인을 위한 대표키 프레임을 추출하기 위한 분석 작업을 한다. 생성된 메타 데이터와 비디오 색인을 이용하여, 비디오의 시간적 특성이 표현된 시간 지원 비디오 데이터 모델(Temporal Video Data Model)은 객체 지향 데이터베이스를 기반으로 한다. 구축된 비디오 자료는 메타 데이터와 비디오 색인과 시간 특성이 반영된 시간 질의 모델(Temporal Query Model)에 의해 사용자(User)가 이용할 수 있다.

2. 구현

구현은 Windows98 환경에서 Visual C++ 6.0을 이용하여 MFC라이브러리와 MCI(Media Control Interface) 기능을 이용하여 AVI, MPG, WAV, MID, RMI 파일을 대상으로 하였다. 그림3은 구현된 프로그램의 수행 과정이다. 실험 자료 중의 하나인 광고에서 대표키 프레임을 추출하여 일부 프레임들을 보여주고 있으며, 대표키 프레임의 프레임 번호를 화면상에 표시하였다.

대표키 프레임을 추출하기 위해서 자료가 있는 드라이브를 선택하고, 비디오 영상이 있는 폴더를 지정한다. 비디오 자료형을 설정한 후, 프레임 추출 버튼을 누른다. 그러면, 선택한 비디오 자료의 전체 프레임 수와 수행 시간, 추출된 대표키 프레임 수와 해당 대표키 프레임의 영상을 보여준다. 추출된 대표키 프레임 수가 화

면에 나타난 것보다 많을 경우에, 자세히 보기 버튼을 누르면 모든 대표키 프레임과 사용된 영상을 확인 할 수 있다.

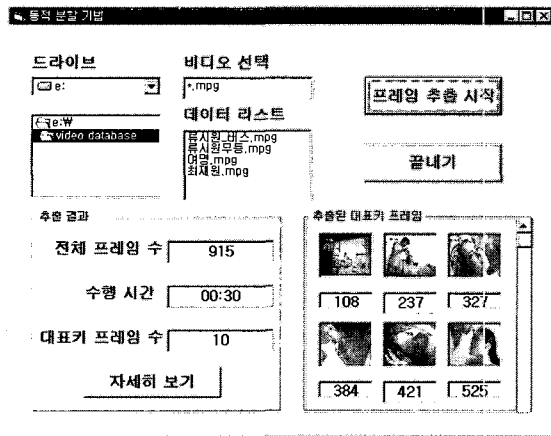


그림 3. 동적 분할 기법의 수행 화면
Fig. 3. Visualization of extraction of Representative Key-frame.

3. 제안 알고리즘의 효율성 및 결과 분석

실험에 사용된 분야별 동영상들을 표2에 제시하였다. 수집된 자료는 초당 프레임 수가 각기 달랐으며, 표2에서는 비교를 위하여 초당 프레임 수를 초당 30 프레임으로 변경하여 계산되었다. 표2에 의해 분야별, 내용별로 샷당 프레임 수에 차이가 있음을 알 수 있다.

정확성은 주어진 비디오 자료의 대표키 프레임을 추

출하도록 하여, 이를 사람이 판단할 수 있는 대표키 프레임 수와 비교하였다. 제안된 방법으로 추출된 대표키 프레임은 표3과 같이 실제 눈으로 확인된 거의 모든 대표키 프레임을 찾을 수 있다.

또한 제안한 비디오 자료의 특성을 적용한 동적 분할 방법은 추출할 대표키 프레임 수나 분야 등을 지정하지 않고, 어떤 동영상 자료에서도 대표키 프레임을 효율적이고 정확히 추출하였다.

표 3. 분야별 대표키 프레임 수 비교
Table 3. comparison with a number of Representative Key-frame for each fields.

	Commercials		Movie Clip		Music Video		News	
	C.D	F.D	C.D	F.D	C.D	F.D	C.D	F.D
균등 분할	6	4	5	4	4	3	4	2
클러스터링	4	6	4	5	3	4	3	3
특징 차이	8	2	6	3	6	1	5	1
동적 분할	9	1	7	2	6	1	6	0

(C.D=correct detection / F.D = false detection)

표 3의 결과는 실제 눈으로 확인 가능한 썸의 개수와 실험을 통하여 추출된 대표키 프레임과 비교하여 산출한 것으로, 제안한 방법이 다른 방법 보다 정확함을 보였다. 계승된 특징 차이를 이용한 방법에서는 사

표 2. 실험에 사용된 비디오 자료들
Table 2. Samples of video data.

Video Sequence	제 목	총길이 seconds	총 프레임수	총 샷 수	샷당 프레임수
Commercials	동산 C&G 섹시마일드(1)	30	915	11	83
	동산 C&G 섹시마일드(2)	30	915	16	57
	팸시 마운틴 듀	29	870	27	32
	아시아나 항공	19	579	11	52
Movie Clip	스텝맘	109	3326	90	36
	키스할까요	140	4275	69	61
	약속	155	4728	28	168
Music Video	타이타닉	201	6130	43	142
	카라	87	2653	34	78
	너를 사랑해	184	5612	31	181
News	MBC뉴스데스크	840	25628	210	122

용자가 원하는 대표키 프레임 수에 의한 균등 분할 방법을 사용하였기 때문에, 사용자가 원하는 대표키 프레임 수가 다를 경우 대표키가 아니거나 혹은 대표키임에도 불구하고 추출되지 않은 경우가 발생한다.

전체 프레임 수가 N 이고, 대표키 프레임의 수가 n , 분할 수가 p 인 비디오 자료를 가정하자. 제안한 알고리즘은 비디오 자료를 p 개로 분할하므로 특징 산출 시간이 p 만큼 필요하다. 다음의 각 단계부터는 추가적인 하나의 특징 산출 시간이 필요하고, 분할 기준을 위한 시간은 최대 $0.15N$ 이 필요하다. 그러므로 제안한 알고리즘은 n 개의 대표키 프레임이 추출 될 때까지 반복하게 되므로, 작업에 필요한 특징 산출 시간은 최대로 $O(0.15N+2p-n-1)$ 이다

반면에 일반적인 방법에서는 컷 검출을 통하여 키 프레임을 추출한 후, 비슷한 키 프레임을 집단화하여 대표키를 추출한다. 일반적으로 키 프레임을 찾기 위해서는 모든 프레임 N 에 대하여 조사해야 하므로 전체 특징 산출 시간은 $O(N)$ 가 필요하다. 여기서 $N \gg p$ 이고 대략 $p = (\frac{1}{30} \sim \frac{1}{180})N$ 이므로 기존 방식의 N 보다 제안된 방식은 $0.15N+2p-n-1$ 이므로 매우 효율적임이 증명된다.

V. 결 론

제출된 특징 차이를 이용한 방법이나 내용 기반 적 응 클러스터링 방법에서는 대표키를 추출하기 위해 균 등 분할 방법을 사용함으로써 비디오 자료를 필요 이상으로 분할하거나 적게 분할함으로써 대표키가 중복 추출되거나 추출 대상에서 제외되는 경우가 발생하는 문제가 있다.

이 연구에서는 이런 문제점을 해결하기 위하여 비디오 특성이 고려된 동적 분할 방법을 제안하였다. 제안된 동적 분할 기법은 비디오 자료가 분야별, 내용별로 샷의 길이가 다른 특징을 분석하기 위하여, 원시 비디오 자료의 각 부분의 특성을 파악하고, 분할 기준을 동적으로 적용하였다. 실험 결과, 제안한 방법은 기존의 방법에 비하여 대표키 프레임을 추출함에 있어 정확성을 높였을 뿐만 아니라 매우 효율적임을 증명하였다.

각종 동영상 자료가 저장된 데이터베이스에서 어떤 장면을 내용으로 검색하는 경우에 모든 프레임을 순차적으로 검색하는 데는 한계가 있다. 즉 씬과 같이 의미

를 전달하는 기본 단위를 대표하는 대표키 프레임에 의한 색인 구축이 꼭 필요하다. 본 논문은 내용 기반 검색을 위하여 색인을 구축하는데 사용자의 개입이 배제되면서도 효율적이고 정확성이 향상된 대표키 프레임 추출 방법을 제시하였다.

향후에는 추출된 대표키 프레임에, 내용에 적합한 구조로 주석을 입력 할 수 있는 연구가 계속되어야 할 것이다. 이 연구가 동영상 자료 중에서 대표키 프레임을 추출하여, 내용 기반 검색이 가능한 비디오 데이터 베이스를 구축하는데 많은 기여를 할 수 있으리라 기대한다.

참 고 문 헌

- [1] 김영희, 이순희, 박장춘, "대표키 프레임 추출을 위한 동적 분할 기법", 한국 정보 과학회 가을 학술 논문집 제 26권 2호 pp. 515-517, 1999
- [2] 이언배, 류근호, "시간지원 데이터베이스의 영상화를 위한 접속 시스템의 설계 및 구현", 한국 정보처리 학회 논문지 제1권 제 2호, pp. 435-447, 1994
- [3] 김기욱, 김형주, "비디오 주석 시스템의 설계 및 구현", 한국 정보 과학회 논문지(B) 제 24권 제 6 호, pp. 588-597, 1997
- [4] 정미영, 이원석, "효율적인 의미 검색을 위한 동영상 데이터 모델링", 한국 정보처리 학회 논문지 제 4권 제4호, pp. 908-922, 1997
- [5] 김기병, 김형주, "내용기반검색 및 주석 기반 검색을 통합하는 비디오 데이터 모델의 설계 및 구현", 한국 정보 과학회 논문지(C) 제3권 제2호 pp. 115-126 1997
- [6] H.J. Zhang, A. Kankanhalli, and S.W. Smoliar, "Automatic Partitioning of Full-Motion Video", ACM/Springer Multimedia Systems, Vol. 1, No. 1, pp. 10-28, 1993.
- [7] K. Otsuji, and Y. Tonomura, "Projection Detecting Filter for Video Cut Detection", Proc. of 1st ACM Intl. Conf. on Multimedia, Anaheim CA, pp. 251-257, 1993.
- [8] R. Kasturi and R. Jain, "Dynamic Vision, Computer Vision: Principles", Eds. R. Kasturi, R. Jain, IEEE Computer Society Press,

- Washington, pp. 469-480, 1991.
- [9] H. Ueda, T. Miyatake, and S. Yoshizawa, "IMPACT: An Interactive Natural-motion-picture Dedicated Multimedia Authoring System", Proc. CHI, New York, pp. 343-350, 1991.
- [10] Nagasaka A. and Tanaka Y, "Automatic Video Indexing and Full-Video Search for Object Appearances", Visual Database System, Proc. of 2nd working Conf. Visual Database System, pp. 113-127, 1991.
- [11] R. Zabih, J. Miller, and K. Mai, "A Feature-Based Algorithm for Detecting and Classifying Scene Breaks", Proc. of ACM Multimedia 95, San Francisco, CA, pp. 189-200, 1995.
- [12] F. Ahanger, A. Hsu, and M.Y. Chiu, "Image Processing on Compressed Data for Large Video Database", Proc. of 1st ACM intl. Conf. on Multimedia, Anaheim CA, pp. 267-272, 1993.
- [13] D. Swanberg, C.F. Shu and R. Jain, "Knowledge Guided Parsing in Video Database", Proc. of SPIE-Storage and Retrieval for Image and Video Database, San Jose, CA, Vol 1908, pp. 13-24, 1993.
- [14] 이미숙, "셋 경계 검출 및 카메라 동작 분석을 통한 비디오 자동 파싱 방법", 한국 정보 과학회 논문지(B), pp. 1249-1257 1999
- [15] 김우생, "계승된 특징 차이를 이용한 효율적인 대표키 프레임과 관련 키 프레임 추출 기법", 한국 정보 과학회 논문지(B) 제 26권 제 5호 pp. 567-664 1999
- [16] 루이스 자네티, "영화의 이해", pp. 257-278, 현암사, 1999
- [17] D. Zhang, and S.-F. Chang, "Clustering methods for video browsing and annotation", Proc. of IS&T/SLIP Storage and Retrieval for Still Image and Video Database IV, Vol.2670, pp. 239-246, 1996.
- [18] M. M. Yeung, B. Yeo, and B. Liu, "Segmentation of video by and graph analysis", Computer vision and Image Understanding, Vol. 71(1), pp. 94-109, 1998.
- [19] Smoliar. S. and Zhang, H., "Content-based Video Indexing and Retrieval", IEEE Multimedia, 1(2) pp. 62-72, 1994.

저 자 소 개



李淳熙(正會員)

1987년 한양대학교 전자과(공학사). 1993년 광운대학교 전산대학원 전산 전공(이학석사). 1995년~1997년 충북대학교 대학원 전자계산학과 박사 과정수료. 1987년~1993년 현대전자(주) 전산실. 1993년~1996년 현대정보기술(주) 기술개발팀. 1996년~현재 안산공과대학 컴퓨터정보과 조교수. 관심분야 : 멀티미디어 데이터베이스, 시공간 데이터베이스



柳根鎬(正會員)

1976년 숭실대학교 전산학과(이학사). 1980년 연세대학교 산업대학원 전산전공(공학석사). 1988년 연세대학교 대학원 전산전공(공학박사). 1976년~1986년 육군 군수 지원사 전산실(ROTC장교), 한국전자통신연구원(연구원), 한국방송통신대학교 전산과(조교수) 근무. 1989년~1991년 Univ. of Arizona. Research Staff (TempIS 연구원, Temporal DB). 1986년~현재 충북대학교 컴퓨터과학과 교수. 관심분야 : 시간 데이터베이스, 시공간 데이터베이스, 지식기반 정보검색, 객체 및 지식 데이터베이스 시스템



金 榮 熙(正會員)

1997년 성결대학교 컴퓨터공학과(공학사). 2000년 건국대학교 대학원 컴퓨터공학과 석사(공학석사). 2000년~현재 (주) 리드텍 코리아 근무 중. 2000년~현재 성균관대 대학원 전기전자컴퓨터공학부 박사과정. 관

심 분야 : 멀티미디어, 멀티미디어 데이터베이스, 웹 저작 원격 교육