

## 움직임 벡터와 빛의 특징을 이용한 비디오 인덱스

이재현\*, 조진선\*

### Video Indexing using Motion vector and brightness features

Jae-Hyun Lee\*, Jin-Sun Cho\*

#### 요 약

본 논문에서는 움직임 벡터와 빛의 세기를 이용하여 비디오의 인덱싱과 검색 기법에 대하여 제안한다.

본 논문에서는 움직임 벡터의 특징과 빛의 세기를 계산하여 각 샷 당 하나의 대표프레임을 추출하였다. 각각의 대표프레임은 빛의 흐름을 계산하였다. 즉 움직임 벡터의 특징은 빛의 흐름으로부터 얻어냈고, BMA는 움직임 벡터를 찾기 위해 사용했다. 그리고 빛의 세기 값을 히스토그램으로 변환한 후 컷 검출에 사용하였다.

비디오 프레임의 움직임 벡터와 빛의 세기 특징을 기반으로 비디오 데이터를 구성하고 인덱싱 하였다. 비디오 데이터베이스는 비디오의 접근을 위해 내용기반을 제공하고, 인덱스 특징은 B+ 트리 검색을 사용했고, 내부적으로 구성되어 단 노드 방식으로 저장되어 컴퓨터 저장장치에 직접 접근할 수 있게 했다. 본 논문에서는 비디오 데이터 모델을 기반으로 한 비디오 인덱스의 문제를 정의하였다.

#### Abstract

In this paper we present a method for automatic motion vector and brightness based video indexing and retrieval. We extract a representational frame from each shot and compute some motion vector and brightness based features. For each R-frame we compute the optical flow field; motion vector features are then derived from this flow field, BMA(block matching algorithm) is used to find motion vectors and Brightness features are related to the cut detection of method brightness histogram.

A video database provided contents based access to video. This is achieved by organizing or indexing video data based on some set of features.

In this paper the index of features is based on a B+ search tree. It consists of internal and leaf nodes stores in a direct access a storage device.

This paper defines the problem of video indexing based on video data models.

---

\* 벽성대학 전산·사무자동화과 전임강사  
논문접수: 98.11.11. 심사완료: 98.12.15.

카메라 샷으로 나누는 작업이 필요하다. 그런 다음 개개의 카메라 샷으로부터 다양한 영상 성질을 추출하여 의미 기술(semantic description)을 설계한다.

## I 서론

### 1.1 비디오 정보검색의 필요성

현재 비디오 정보의 이용은 쉽고 효율적인 구성과 검색을 가능하게 하는 인덱싱과 검색시스템의 부족으로 단편적인 서비스만이 가능한 상태이다. 일반적으로 Logging툴을 사용하여 데이터 베이스에 입력된 구조화된 텍스트 방법으로 비디오 소스를 취급하고 있다. 하지만 이 방법은 이미지나 비디오의 내용을 찾는데 있어 제한된 기법만을 제공할 뿐이다. 특히 비디오는 각 클립당 정보를 범주화(categorizing)하기 어려워서 정지 영상보다 처리하고 이해하기가 더 힘들다. 그러므로 비디오를 그 속성에 따라 조직화하며, 용이하게 취급할 수 있는 패키지로 변환하는 것이 중요하다. 따라서 비디오 정보검색 시스템의 개발에서 대화형 부라우징과 동적인 질의를 지원함으로써 이러한 문제를 해결할 수 있다.[5]

본 논문에서는 비디오 부라우징 기술 개발에 관한 연구로 대형 저장장치에 저장된 비디오 정보를 사용자가 검색 재생할 수 있는 비디오 정보처리 기술로 다음과 같은 세가지 타스크를 제공하는 툴을 연구하였다.

- ① 인덱싱 하기 위하여 클립 단위로 비디오를 분할한다(parse).
- ② 클립을 식별하기 위하여 영상의 하위(Low-level)특성과 의미론적(semantic) 특성을 알아야 한다.
- ③ 질의로부터 결과를 비디오 자원으로 부터 선택하는 검색과 부라우징이 필요하다.

알고리즘과 시스템의 여러 종류가 첫 번째 타스크를 지원하기 위하여 계속적으로 개발되어 왔다. 일반화된 비디오 자동 파싱은 거의 불가능하며, 만일 비디오 구조가 명확히 식별되거나 비디오 모델을 설립할 수 있다면 가능하다.

파싱은 자동 비디오 인덱싱의 전단계로 비디오의 저장과 검색, 멀티미디어 DB구축에 없어서는 안되는 중요한 단계이다. 파싱을 위해서는 먼저 비디오 시퀀스를 일련의

### 1.2 관련연구

현재까지, 첫 번째 단계는 다양한 알고리즘이 개발되어 있다.[6][9]) 다음 단계는 하위 컴퓨터시각(low-level computer vision)을 응용하여 다각도로 시도되고 있으나, 아직 뚜렷한 결과는 없다. 비디오 영상은 매우 다양하고 추상화 단계가 낮아서 비디오 씬의 모델링에 있어서 일반적인 방법은 없다. 다만 비디오 구조를 명확히 식별할 수 있는 일부 특별한 형태의 비디오 영상 파싱은 가능하다.

M.Flickner[6]은 이미지와 비디오를 동시에 취급하는 내용기반 검색 시스템에서 비디오를 샷으로 나눈 다음 각 샷을 카메라 움직임과 객체의 움직임을 통해 특징화한 후 사용자의 질의는 객체의 움직임 또는 카메라의 움직임에 의해 검색할 수 있게 하였다.

La, cascia[8] 등은 칼라와 질감 정보만을 사용해서 자동 비디오 인덱싱 시스템을 개발했다. 비디오는 샷 안에서 나뉘지고 샷은 하나 또는 그 이상의 대표프레임을 추출하여 각 샷을 표현하게 했다. P.M Kelly[9]등은 내용기반 저장과 디지털 영상의 검색 시스템에서, 각각의 질감, 칼라, 모양정보를 통합시켜 특징을 저장한 후 컴퓨터가 자동적으로 특징을 추출하게 했다.

이상의 발표사례에서 보면 비디오 정보를 표현하기 위한 연구는 내용 기반검색을 위한 비구조화된 모델과 멀티미디어 데이터의 임시 구성 모델, 특정 응용 멀티미디어 데이터 모델로 구분지어 진다.

## II 비디오 분할

### 2.1 컷 검색

비디오에서 컷은 한가지 영상에서 다른 영상으로 바뀌는 스위치라고 정의되며, 컷 포인트는 비디오에서 불연속

이다. 컷 검출은 비디오를 샷들로 나누는 방법중의 하나로서, 개발된 많은 알고리즘들은 프레임간의 차이를 이용하여 주요 비디오 파라미터의 불연속된 곳을 탐색한다. 이러한 파라미터들은 빛의 세기, RGB, 색조컬러, 그리고 움직임 벡터를 포함한다.

### 2.2 밝기 정보를 이용한 비디오 분할

본 논문에서는 비디오 분할에 연속되는 프레임의 각각에 대한 밝기 정보를 이용하여 비디오를 분할한다. 또한 본 논문에서 밝기 정보는 평균 밝기 정보를 이용했으며, 다음의 식으로 얻어진다.

$$brightness_i = \frac{1}{N} \sum_{j=1}^N P_{ij} \dots (1)$$

$P_{ij}$ 는  $j$ 번째 영상 픽셀의  $i$ 번째 밝기 요소값이다.

컷을 구성하는 프레임들의 밝기 평균을 얻을 수 있으며 또한 컷에 구별 없이 프레임당 밝기 평균을 얻어 인덱싱 정보로 이용할 수 있다. 또한 밝기 정보는 프레임내의 객체의 특징을 추출하는 데는 문제가 있지만 컷 검출에 있어 모든 비디오 포맷에 일괄 적용되고 컷 검출에 필요한 연산속도 향상을 위한 알고리즘 적용이 용이하다. 식 (1)에 의한 밝기 정보의 추출은 비교적 용이하며 그 결과는 그림 1과 같다. 컷 검출의 인덱스로 사용하기 위해 정렬한다.

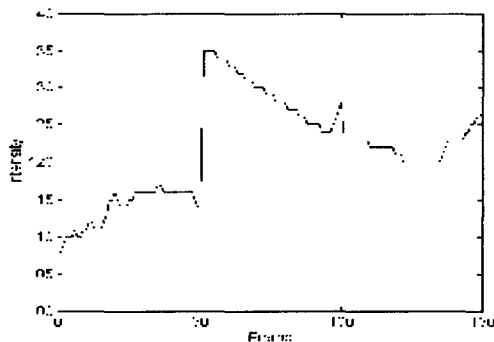


그림 1. 비디오 데이터 빛의 밝기에 의한 값  
Fig 1. Result from brightness of video data

## III 비디오 특징 추출

### 3.1 샷의 특징

샷에 들어있는 정보는 대개 카메라 동작조건과 장면내의 물체들이다. 특징분야를 위하여 장면 조건을 제한하지 않고는 컴퓨터로 각 물체를 인식하기는 곤란하다. 그러나 비디오를 취급하기 위한 인덱스로 중요한 비디오 특징은 추출할 수 있다.

### 3.2 움직임 벡터를 이용한 특징 추출

연속되는 프레임에 객체의 움직임을 검출하기 위해 제공되는 특별한 특징은 컴퓨터 비전에 근원을 둔 빛 흐름(optical flow)이다. 따라서 영상 움직임은 일정한 방향을 가지게 된다. 카메라 움직임에 기인한 움직임 벡터는 공간적으로 나타나며, 한 쌍의 프레임에 대한 각 부분특에 대한 좌표치가 계산된다. 그 계산된 공간적 패턴의 특징을 보면 움직이는 물체를 제외한 모든 물체의 방향이 어느 일정한 방향으로 향한다는 것을 볼 수 있다. 예를 들면 패닝이나 트래킹의 경우 빛 흐름의 패턴에서의 움직임 벡터는 왼쪽이나 오른쪽으로 향하며, 틸팅과 부밍은 위쪽이나 아래쪽으로 향한다. 또한 줌의 경우 움직임 벡터는 중심을 향하거나 중심 바깥쪽으로 향한다. 예를 들면, 패닝 및 트래킹하는 동안 벡터의 방향은 일정하며, 움직임 벡터의 대부분은 대표 벡터(modal vector)에 나란할 것이다.

다음과 같은 표현할 수 있다.[9]

$$\sum_k^N |\theta_k - \theta_m| \leq \theta_p \quad (2)$$

여기서 는 움직임 벡터의 방향  $k$  이고  $\theta_m$ 은 대표 벡터의 방향이다. 그리고  $N$ 은 프레임에서의 움직임 벡터의 총 개수이다.  $|\theta_k - \theta_m|$  는 두 벡터가 정확히 평행일 때 0이 된다. 그러므로 식 (2)을 이용하여 대

표 벡터 방향으로부터 모든 움직임 벡터의 변화량을 계산한다. 만약에 이러한 변화량이  $\theta_p$ 보다 작거나 같으면 카메라는 패닝 또는 트래킹이다.

줌ming은 인이든 아웃이든 초점 가지며, 만약 초점이 프레임의 중심에 배치해 있고 물체의 움직임이 없다면 모든 움직임 벡터의 평균은 0이 될 것이다. 그러나 프레임의 한가운데에 초점을 배치하는 것이 그리 쉬운 방법은 아니다. 다른 동작과 복합되지 않는 한 영상의 중앙을 기준으로 상하좌우의 움직임 벡터의 방향은 대칭이다. 즉, 모든 행에서 수직요소의 차의 크기는 언제나 개별 요소의 크기를 초과할 것이다. 좌단 또는 우단행에서 움직임 벡터의 수평 요소 또한 마찬가지로 해석할 수 있다. 대부분의 움직임 벡터가 식 (3)과 (4)의 조건을 만족할 때 좁이다.

$$|v_k^{top} - v_k^{bottom}| \geq \max(|v_k^{top}|, |v_k^{bottom}|) \dots\dots\dots (3)$$

$$|u_k^{top} - u_k^{bottom}| \geq \max(|u_k^{top}|, |u_k^{bottom}|) \dots\dots\dots (4)$$

카메라 움직임에 의한 특수 효과나 점진적인 변화는 카메라 움직임에 의해 발생하는 줌ming이나 패닝의 변화는 서로 구별되어야 한다. 카메라 움직임에 의한 변화는 점진적인 변화와 같은 순서로 연속적인 차이에 기인하는 경향이 있다. 또한 디졸브, 페이드-인 그리고 페이드-아웃과 같은 특수 효과에 의한 변화를 움직임 벡터 필드들이 검출되고 분석될 수 있다면, 카메라 움직임에 의한 변화는 특수 효과에 의한 변화로부터 구분될 수 있다. 빛 흐름 계산은 움직임 벡터를 만드는 것과 같이 두개의 연속된 프레임 사이의 차로 표현된다.

#### IV 특징 기반 데이터 모델 인덱스

비디오 데이터는 종래의 데이터 처리 기법과는 전혀 다른 처리 기법이 요구된다. 또한 비디오 데이터베이스는 특징을 기반으로 비디오를 접근할 수 있게 제공 되어

야 하며, 특징을 기반으로 데이터를 구성하거나 비디오를 인덱싱화하여 특정 정보를 검색하게 된다. 시간성을 갖는 멀티미디어의 데이터를 정확하게 표현하기 위해서는 멀티미디어 데이터의 시간성에 관한 정보를 유지, 관리하여야 한다. 따라서 멀티미디어의 동기화 정보를 표현하는 모델링 (modeling)기법이 필요하다. 본 논문에서는 비디오 데이터 모델 구조와 영상 데이터 시스템에 관한 정보를 [그림 2]의 헤더에 포함시켰다.

본 연구에서는 순차적인 비디오를 분할하여 임의의 위치의 비디오를 검색하기 위하여 임의의 위치 접근이 용이한 B-Tree 구조를 응용하여 그림 3과 같은 형태로 인덱스 구조를 설계하였다. 그리고 각 노드의 헤더에 본 논문에서 검출된 움직임 벡터의 시공간적 특성, 밝기정보를 다단계 참조 방식을 적용하여 각각의 특징들을 헤더에 삽입하여 순차적 저장방식의 특징을 유지하면서 빠른 검색이 되도록 하였다.

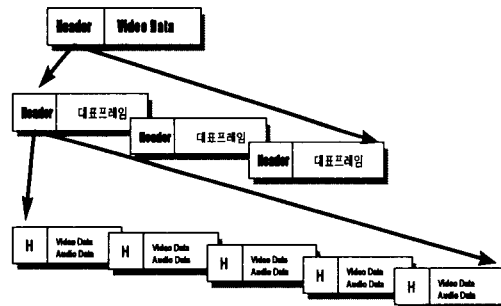


그림 2. 인덱스 데이터 저장 방식  
Fig 2. Storage method of index data

본 연구에서에서 제시할 Tree구조 노드는 3 단계의 레벨로 구성하며, 레벨 0은 자료 파일의 샷을 가리키는 포인터를 가지고 있다. 0이상의 레벨에서 각 키의 포인터는 그 키 값 이하의 키를 가지고 있는 노드를 가리킨다. 0이상의 레벨에 있는 각 노드에는 키는 없고 포인터만 있는 엔트리를 하나씩 가지고 있다. 이 포인터는 그 노드에 있는 가장 큰 키보다 큰 키들을 가진 노드를 가리킨다. 한 노드가 가지는 포인터의 개수는 키의 개수와 같거나 하나 더 있다.

각 단계별 헤더의 구조는 그림3과 같이 설계하였다.

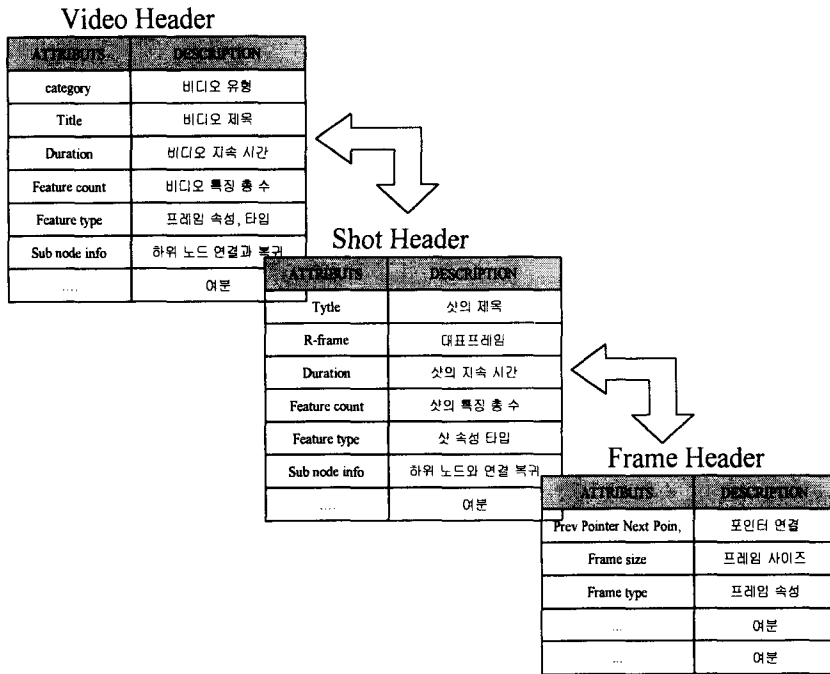


그림 3. 비디오 분할 계층적 헤더 구조  
Fig 3. A hierarchical Header structure of Video partition

### V. 정보검색 틀의 설계

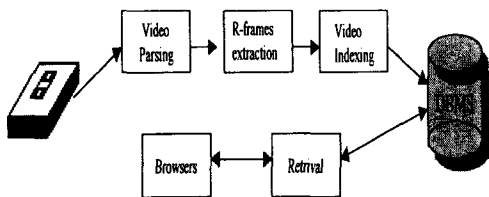


그림 4. 영상 정보 시스템  
Fig 4. Image Information System

본 논문의 시스템 구성도는 그림 4와 같으며 5단계로 이루어진다.

첫 번째 단계로 연속되는 비디오 열을 입력받아 본 논문에서 제시한 밝기 정보 추출 알고리즘을 용하여 컷 검출 파싱 하고, 두 번째 단계는 본 논문에서 제시한 움직임 벡터 연산 알고리즘을 이용하여 프레임 내에 물체의

이동 정보를 연산하여 대표 프레임을 검출하고, 세 번째 단계는 첫 번째 그리고 두 번째 단계에서 획득한 정보를 이용하여 4 장에서 제시한 인덱스를 구성하고 헤더 정보에 포함 시켜 데이터 베이스에 저장했다.

마지막 단계는 브라우저를 이용하여 비디오를 검색하는 과정으로 비디오의 저장 구조를 이용하여 사용자가 아이콘을 선택함으로써 검색 가능하게 사용자 인터페이스를 구성하였다.

### VI. 실험

본 논문에서 제시한 비디오 정보검색 시스템은 속도문제를 고려하여 TGstation(SDT-600) X-window환경에서 다중 윈도우를 구성한 그래픽 처리로 사용자 인터페이스를 구성하였으며, 사용된 언어는 OSF-Motif와 C언어를 사용하여 구현하였다. 실험 데이터는 인터넷에서 다운로드한 "주라기공원", 과 만화영화 "라이언 킹" 을 PC

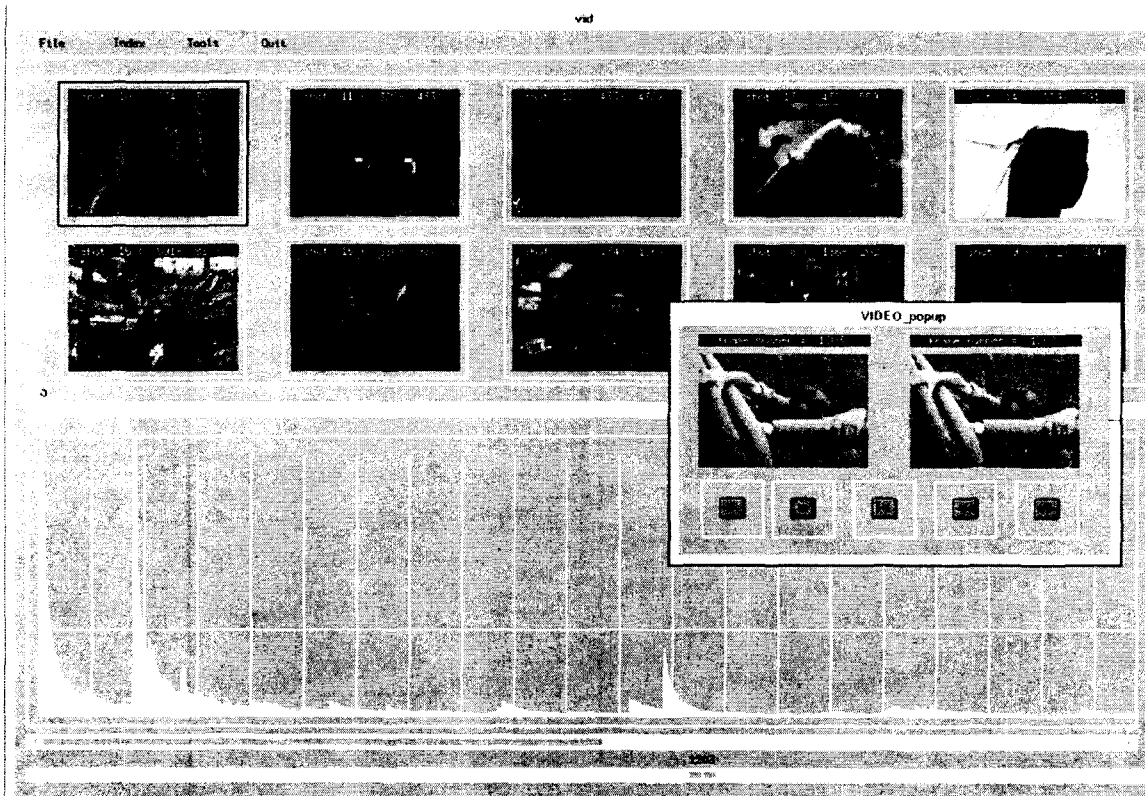


그림 5. 인덱싱 화면  
Fig 5. Screen of indexing

Video Blaster(RT300)에서 캡처(capture)하여 실험 고찰하였다. 구현 시스템에서는 음성은 고려하지 않고 비디오를 입력으로 받아 처리하였다. 그림 5는 구현된 시스템에서 비디오를 인덱싱하는 과정을 보여주고 있으며, 그림 6은 아이콘화된 분할된 샷을 마우스로 클릭하여 샷에 대한 카메라 동작을 보여주는 화면이다.

## VII 결론

멀티미디어 정보 시스템을 구축하기 위한 가장 핵심적인 기술인 멀티미디어 데이터베이스 기술은, 다양한 정보 검색을 위한 인덱싱을 필요로 하게 되었다. 또한 사용자의 입장에서 어떻게 정보를 접근해 나갈 수 있을까? 하는 문제를 생각하게 한다.

본 논문에서는 비디오 프레임의 밝기 정보와 움직임

벡터를 사용하여 비디오 인덱싱을 위한 특징으로 사용하였으며, 그 움직임 벡터는 블록매칭 알고리즘(BMA)으로 구하였다. 블록매칭 알고리즘으로 구해진 움직임 벡터와 밝기 정보를 이용하여 비디오 인덱싱 필드에 활용했으며 샷을 대표하는 프레임의 자동검출 기법을 제안하여 효율적인 비디오 정보검색에 활용하도록 했다. 인덱싱 검색을 위한 구조로는 B+ Tree구조를 이용하였으며, 데이터 모델의 키로는 움직임 벡터와 밝기 정보의 중복기를 사용하여 인덱싱을 제시하였다.

본 논문에서 검출된 움직임 정보를 활용하여 MPEG-7에 직접 활용하여 데이터량의 감소와 처리시간 단축에 집중적인 연구가 필요하다. 또한 대용량 데이터에 대한 회복 작업은 전체 시스템 성능을 크게 좌우하기 때문에 설계된 저장 구조에 적합하며 부가 비용을 최소로 줄이는 새로운 회복 기법이 필요하며, 또한 브라우징 틀에 대한 연구도 필요하다.

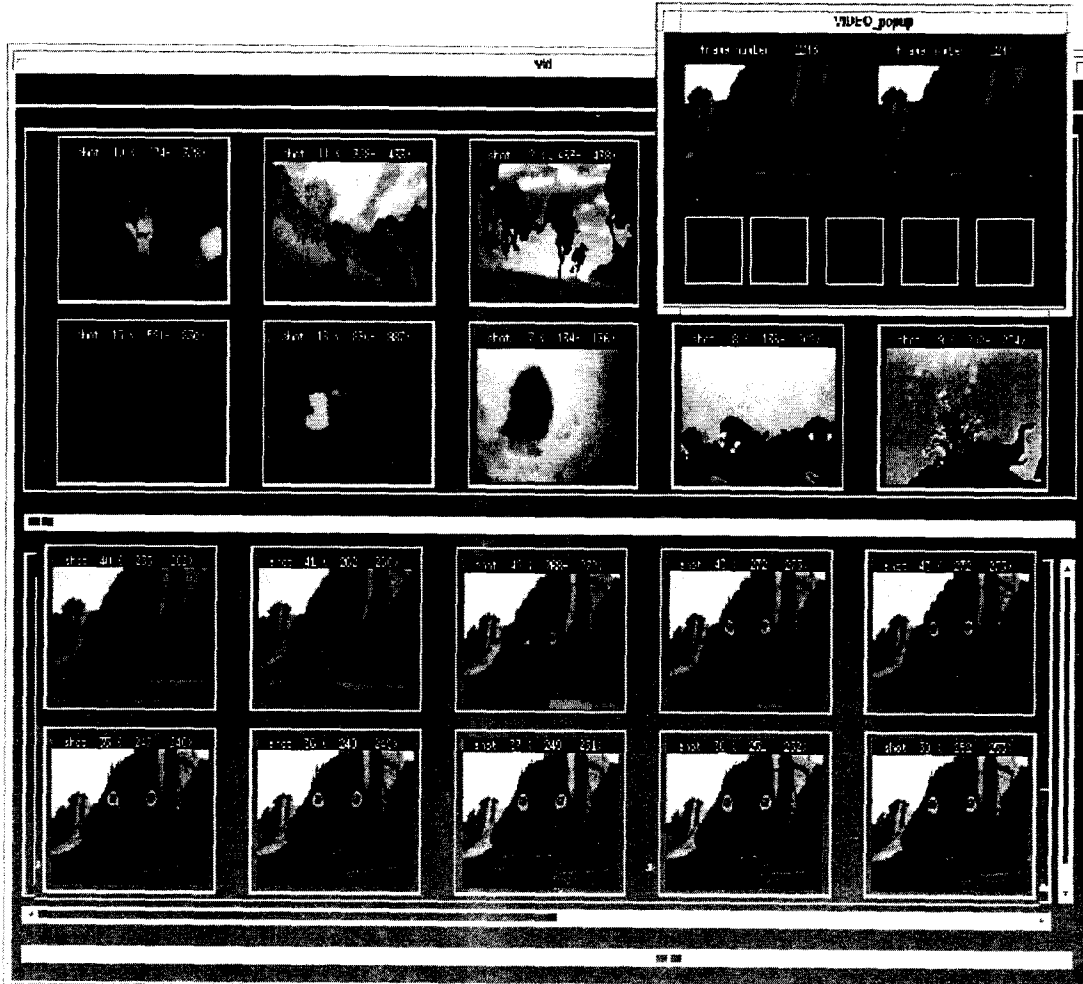


그림 6. shot 검색화면  
Fig 6. A Screen for shot retrieval

### 참고문헌

[1] A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances", Visual Database System, Vol. 11, Elsevier, Amsterdam, 1992, pp. 113-127.  
[2] H. Ueda et al., "Automatic Structure

Visualization for Video Editing", Proc. InterCHI '93 ACM Press, New York, 1993, pp. 137-141.  
[3] H. Zang, A. Kankanhalli, and S. Smoliar, "Automatic Partitioning of Full-Motion Video", Proc. ACM Multimedia System, ACM Press, New York, Vol. 1, 1993, pp. 10-28.  
[4] A. Akutsu et al., "Video Indexing Using Motion Vector" Proc. SPIE Visual Comm. and Image Proc., 1818, 1992, pp. 1522-1530.  
[5] S. W. Smoliar, and H. Zhang,

- "Content-Based Video Indexing and Retrieval", IEEE Multimedia 1994 summer, pp. 63-72.
- [6] A. Hampapur, R. Jain, T.E. Weymouth, "Indexing in Video Databases", SPIE Vol. 2420, 1995, pp. 292-306.
- [7] P.M. Killy, T.M. Cannon, and D.R. Hush: Query by image example: the CANDID approach. Proc. of the SPIE: storage and Retrieval for Image and Video Databases III, vol. 2420. pp. 238-248, 1995
- [8] M. La Cascia, E. Ardzone: JACOB: Just a content-based query system for video databases, Proc. of ICASSP '96, May 7-10, 1996 Atlanta.
- [9] Jae-Hyun Lee, Yeun-Sung Choi and Ok-Bae Jang "Gradual Cut Detection Low Level Vision for Digital Video" SPIE Vol. 2952. pp 683-688, 1996.
- [10] Yeo, B. L., and Liu, B. Rapid scene analysis on compressed videos. IEEE Trans, Circuits Syst. Video Technol. 5,6(Doc, 1995) 533-544.
- [11] Y. Tonomura, A. Akutsu, Y. Taniguchi, and G. Suzuki, "Structured Video Computing", IEEE 1994 fall, pp. 34-43.

저자 소개



이재현

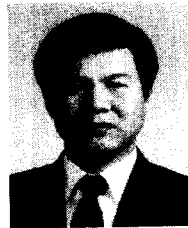
호원대학교 전자계산과 졸업(이학사)

조선대학교 대학원 컴퓨터공학과 졸업(공학석사)

전북대학교 대학원 전산통계학과 졸업(이학박사)

현재 벽성대학 전자계산과 전임강사

※관심분야 : 멀티미디어 정보검색, Visual Interface, 비디오 인덱싱



조진선

건국대학교 문리과대학 수학과 졸업(이학사)

건국대학교 대학원 수학과 졸업(이학석사)

현재 벽성대학 사무자동화과 전임강사

※관심분야 : FUZZY이론, 프랙탈 이론, 멀티미디어