

멀티미디어 부호화를 이용한 내용기반 검색에 관한 연구

박재필^{**}, 강진석^{*}, 고석만^{**}, 최연성^{***}, 김장형^{*}

제주대학교 정보공학과^{*},

제주산업정보대학 컴퓨터정보계열^{**},

군산대학교 정보통신공학과^{***}

A Study of base-contents retrieval for using Multimedia code

Jin-suk kang^{*}, Jae-pil Park^{*}, Suk-Man Koh^{**}, Yeon-sung Choi^{***}, Jang-Hyung Kim^{*}

^{*}Dept. of Information Engineering, Cheju National Univ.

^{**}Dept. of Computer Information, Cheju College of Technology.

^{***}Dept. of Telecommunication Engineering, Kunsan National Univ.

E-Mail : parkjaepil@hanmail.net

요 약

최근 들어, 컴퓨터 및 관련 기술의 발전은 통신망, 정보 처리 형태 및 통신 서비스의 발전을 동반하여 정보화 사회로의 진입을 가속화시키고 있다. 특히 인터넷의 WWW 등장으로 인하여 전자 신문, 전자 컨퍼런스, 멀티미디어 문서 검색 등 헤아릴 수 없는 많은 서비스들이 손쉽게 개발되었고, 이는 일반인들의 정보 통신망의 수요를 더 한층 증가시키고 있으며 실제적인 정보화 시대를 앞당기고 있다. 최근 연구되고 있는 주요 기반 기술들로는 정보의 형태가 멀티미디어 화에 따른 내용 기반 정보 검색 기술 연구가 활발히 진행되고 있고, 멀티미디어 정보가 방대함에 따라 기존의 하드디스크 대신에 3차 저장 장치를 하부 저장 시스템의 일부로 확장시키는 연구가 한창이다. 또한 정보 검색 분야 외에도 정보 구축 기술로, 정보들을 체계적으로 관리할 수 있도록 모델링 기법에 대해 기술한다.

ABSTRACT

Recent progress on computer and related technology, especially including communication network, multimedia information processing and communication service technology, accelerates the entrance of information society. Especially, WWW brought the information crisis in its demand and size on the communication network, by making it easy to develop information service, like electronic (internet) news, electronic conference, multimedia information retrieval, and so on, on the internet. Due to this change, DBMS should provide efficient ways to store and manage various types of multimedia data and to model complex information structures. In order to satisfy these requirements, there have been many researches on architecture of multimedia DBMS, content-based retrieval for multimedia information, tertiary storage system for huge multimedia data, multimedia information modeling,

1. 서 론

시대를 살며서 컴퓨터를 전문적으로 사용하는 사용자뿐만 아니라, 일반인에게까지 친숙하게 들려온 멀티미디어(Multimedia)란 용어를 많이 들어왔을 것이다. 여기서 멀티미디어는 '컴퓨터에서 문자, 음향, 동화상, 정지화상 등을 동시에 적용하는 것'이며, 어원상으로 볼 때, '멀티(Multi)'는 다중이라는 의미이며, '미디어(Media)'는 매체인 문자, 음향, 영상, 정지화상을 의미한다. 이러한 멀티미디어(Multimedia)에 있어서 우리가 연구되어

야 할 영상처리 중 정지화상 검출, 영상의 압축과 압축을 통한 저 전송률의 데이터 전송, 문자, 음향, 영상, 영지 화상 등이 통합된 MMDBMS의 구축의 최종 목표라 할 수 있다.

여기서 볼 때, 멀티미디어가 중요한 의미를 갖는 이유는, 인간에게는 시각, 청각, 촉각, 후각, 미각 등의 이른바 5감이라는 것이 있어서 이를 통합적으로 인지하여 사물이나 상황을 파악하는데 반해 기존의 컴퓨터에서는 단순히 시각적 문자만을 보여주는데 그쳤기 때문이다.

또한, 저장 장치, 통신, 데이터 압축 기법등의 발

달로 다양한 형태의 멀티미디어 응용분야가 가능해지고 있다. 여러 종류의 멀티미디어 응용 중 주문형 비디오(Video On Demand, VOD)서비스는 기존의 TV 유선 방송을 대체할 수 있는 획기적인 서비스 분야로 각광을 받고 있다. 본 논문에서는 멀티미디어의 핵심 기술에 있어서 비디오 데이터의 검색 요청에 대해 데이터를 검색하여 사용자에게 공급해주는 측면인 파일 서버를 이용한 디지털 비디오 데이터의 검색에 대해 기술한다.

II. 기반 기술

1. 멀티미디어 시스템 요소 기술

최초의 전자식 범용 디지털 컴퓨터인 ENIAC이 개발된 이후 오늘날에 이르기까지 컴퓨터 기술의 발달과 저장 능력 측면에서의 메모리 IC의 대표적인 DRAM 용량의 증가할 만큼의 급격한 발전과 반도체 기술인 메모리와 마이크로프로세서를 중심으로한 디지털 신호 처리용 반도체(Digital Signal Processor: DSP)의 기술, NTSC 방식 TV의 아날로그 신호를 디지털화로 변화의 발전과정인 부호화 기술과 통신 기술의 발전으로 멀티미디어 환경도 여러 개의 시스템이 통신망에 연결되어 정보의 수신, 수신 공유가 가능한 MPC 방식의 Stand-alone 시스템의 주종인 네트워크 기술의 발전으로 인한 자유로운 데이터의 교환으로 기반이 조성되고 있다.

2. 부호화 기술의 필요성 및 원리

숫자나 문자로 된 텍스트 데이터들로부터 정보를 얻어내는 기존의 방식 대신 오디오나 이미지 등 좀더 인간의 감각에 접근한 형태로 표현된 데이터들로부터 직관적으로 정보를 얻어내려는 욕구가 강해졌다. 이러한 요구는 관련된 기술들, 즉 센싱, 신호처리, 통신, 디스플레이에 관련된 기술과 이들의 기초가 되는 소재 및 반도체, 그리고 소프트웨어 기술의 결합에 힘입어 현실 응용화 되는 등 날로 충족되어 가고 있다. 따라서 이들을 저장, 처리, 전송하기 위해서는 이들은 부호화하는 기술이 필연적이다.

가. 중복성 제거의 원리

불필요한 정보를 삭제하고 반복되는 정보를 간략히 기술함으로써 데이터의 크기를 줄이는 원리이다. 전화 음성인 경우 말을 하는 때와 말을 하지 않은 때가 분리되는 점을 이용해서 말을 하지 않을 때는 정보를 보내지 않음과 동영상의 경우 1초에 비슷한 그림을 60장 정도 빠르게 바꾸어줌으로써 움직이는 그림을 구현하기 때문에 직전 그림과 현재 그림이 별 차이가 없다는 점을 이용하여 정보의 양을 줄인다.

나. 가시성·가청성의 원리

가시성·가청성의 원리는 사람이 감지할 정보만을 추려서 남기고, 감지 할 수 없는 정보를 삭제함으로써 데이터의 양을 줄인다.

3. 부호화 기법 및 요소 기술

부호화는 여러 가지 응용에 따라 다양한 기법이 존재하지만 크게 무손실 부호화 기법과 손실 부호화 기법으로 분류할 수 있다. 본 절에서는 이와 같은 부호화 기법을 영상 데이터 부호화를 중심으로 기술하였다.

가. 무손실 부호화 기법

무손실 부호화 기법이란 부호화한 데이터를 복원했을 때 복원한 데이터가 부호화 전의 데이터와 완전히 일치하는 것을 말한다. 이 기법은 부호화한 데이터가 부호화 할 데이터에 어떤 변경이나 수정은 가하지 않는다. 따라서 멀티미디어 정보에서 정확성을 요하는 의료 영상의 부호화에 사용할 수 있다. 무손실 부호화의 경우 복원시 원래의 데이터와 동일한 정보를 얻을 수 있지만, 손실 부호화 기법에 비해 부호화율이 낮다는 단점이 있다.

(1) 허프만 부호화(Huffman Coding)

허프만 부호화는 일반 응용에서 가장 많이 사용되며, 기본 개념은 발생 확률이 높은 심볼에 대하여는 짧은 길이의 부호어를 발생 확률이 낮은 심볼에 대하여는 상대적으로 긴 부호어를 할당하여 평균 부호어 길이를 짧게 하는 것이다.

(2) 산술 부호화(Arithmetic Coding)

허프만 부호의 부호화 효율을 높이기 위한 방법의 한 가지로 소스 열을 몇 개의 블록으로 나눈 후, 각 블록을 가변장 부호어로 할당하는 것이다. 물론 복호기 측에서는 각각의 부호어에 대응하도록 가변 블록으로 분리한 후, 원래의 부호로 변환한다. 즉 소스열 블록과 부호어 블록사이에는 일대일 대응 관계가 존재한다. 이에 비해 산술부호는 비블럭 부호(또는 트리(Tree)부호)로 길이 m 개의 심볼로 구성되는 전체 입력 데이터 열에 단 하나의 부호어를 할당하는 기법이다. 산술부호에서 부호어 길이는 대략 이상적 이산 부호어 길이 $-\log_2 p^{(s_m)}$ 에 근접하며, 여기서 $p^{(s_m)}$ 은 소스열(s_m)의 발생확률을 나타낸다. 산술부호의 기본개념을 이해하기 위하여 이산 무기역 소스로 발생되어진 2진 심볼열(s_m)의 부호화를 고려해 보면 우선 심볼열에서 '0'이 나올 확률을 p 라하면 '1'이 나올 확률은 $q=1-p$ 가 된다. 여기서 왼쪽값은 포함하고 오른쪽 값은 포함하지 않는 구간(0,1)을 l 라 정의하기로 한다. 한편 길이 m 의 소스열(s_m)에 있어서 발생할 모든 조합은 2^m 이 되며, 이의 발

생 확률 $p^{(m)}$ 의 총 합은 1이 된다. 따라서 서브 (Sub) 구간, $I_l, l=1,2,3,\dots, 2^m$ 은 각 소스열 s_m 에 일시적으로 할당되어 질 수 있으며, I_l 의 길이는 $p^{(m)}$ 과 같으며 서브구간은 중첩되지 않는 방법이다.

(3) 반복길이 부호화(Run-Length(RL) Coding) 기법
이 기법은 반복되는 픽셀들을 하나의 색상 값과 그 색상이 반복되는 횟수로 치환하는 것으로 동일한 값을 가지는 픽셀들이 인접하게 모여있는 특성을 부호화 방법에 이용한 것으로써 영상의 종류에 따라 부호화율이 크게 달라진다. RL 부호화 방법은 총 16횟수의 영상의 표현은 색상 값과 그 색상이 연속적으로 나타나는 횟수의 쌍으로 표현되고 있다. 횟수는 0번이라는 것이 없기 때문에 '0000'이 1번 발생을 의미하고 '1111'은 16번의 반복을 의미한다.

나. 손실 부호화 기법
손실 부호화 기법은 높은 효율의 부호화율을 얻기 위하여 부호화시 화질 열화를 허용하는 기법으로 기본 구성은 그림 1과 같다.

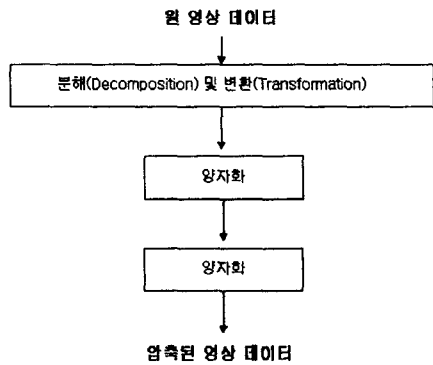


그림 1. 손실 부호화의 일반적인 구조

이와 같은 손실 부호화 기법으로는 색상표(Color Lookup table)를 이용하는 방식, 보정 (Interpolative)기법, 예측 부호화(Predictive Coding)기법, 변환 부호화(Transform Coding) 기법 등이 있다.

(1) 색상표를 이용하는 방식
각 픽셀의 값은 실제 표현될 색상을 나타내는 것이 아니라 색상표의 색인이다. 즉 픽셀의 색상은 픽셀의 값과 색상표에 의해서 결정된다. 이 방법은 색상표의 크기에 따라서 부호화율이 결정된다. 예를 들어 24bpp로 된 원천 영상을 256개의 색상을 가진 (즉 8bpp) 색상표를 이용하여 변환한다면 부호화율은 3:1이 된다. 이 방법은 영상 전체에서 사용할 수 있는 색상을 256개 혹은 더 적은 수의 색상으로 표현하는데 주로 이용된다.

(2) 보정(Interpolative) 기법
보정 부호화 기법은 영상을 구성하는 픽셀들의 부분 집합만을 전송하고 픽셀들 사이에 생략된 픽셀은 보정에 의해서 생성하는 기법을 말한다. 이 기법은 생략된 픽셀의 값을 그 주변의 값을 이용하여 추정하여도 사람이 인식하기 어려울 때 사용한다.

(3) 예측 부호화(Predictive Coding) 기법
일반적인 화상의 경우 인접 픽셀 값 사이에는 상관관계가 존재한다. 따라서 임의의 한 픽셀 값에 대한 많은 정보를 인접 픽셀 값으로부터 얻는 것이 가능하다. 예측 부호화 기법에서는 위와 같은 성질을 이용하여 부호화를 행하는데, 부호화의 기본 개념은 원하는 픽셀값의 예측은 주위의 픽셀 값을 이용하여 행하는 것이다. 예측을 기반으로 한 부호화 기법들로는 차이 PCM(Differential PCM: DPCM) 과 적응적 DPCM(Adaptive PCM: ADPCM)등이 있다.

III. 내용기반 정보 검색

본 절에서의 내용기반 검색은 색상을 이용한 특정 물체를 추출해 낼 수 있는 방법인 색상을 이용한 물체의 추출에 대해 기술한다. 일반적으로 색상은 약 400nm에서 700nm까지의 파장을 지닌 가시광선 영역의 전자기파를 의미하는 것이다. 모든 색상은 소위 주요 컬러(RGB)라는 서로 다른 파장의 길이를 가지는 세 가지 성분의 조합과 빛의 휘도 성분으로 표현된다. 인간의 망막으로 받아들이는 위와 같은 세 가지의 주요 컬러 성분은 서로 독립적인 컬러 매칭 함수와 적절한 스펙트럼 함수로서 표현된다.

$$\begin{aligned}
 R &= \int_{r_1}^{r_2} r(v)U(v)dv \\
 G &= \int_{g_1}^{g_2} g(v)U(v)dv \\
 B &= \int_{b_1}^{b_2} b(v)U(v)dv
 \end{aligned} \tag{1}$$

여기서 U 는 빛의 주파수이다. 대부분 사용되는 하드웨어의 종류에 따라 여러 가지 컬러 모델을 사용한다. 일반적으로 비디오 카메라는 RGB모델을 가장 많이 사용하고 있으며 이 모델은 다른 컬러 모델로 쉽게 변환이 가능하다. 비디오 카메라나 프레임 그레이는 아날로그 신호를 각각의 화소에 대해 세 개의 $[0,1,\dots,255]$ 의 값으로 변환한다. 즉, 한 화소는 각각 $[0,1,\dots,255]$ 값을 가지는 RGB의 조합 $Q = (R, G, B)$ 으로 표현된다. 하지만 이 모델은 하나의 색상뿐만 아니라 빛의 휘도까지 포함한 값이다. RGB공간상에서의 한 점 $Q = (R, G, B)$ 는 색상뿐만 아니라 휘도를 지니고 있다. 만약, 공간의 두 점 $Q_1 = [r_1, g_1, b_1]$ 와 $Q_2 = [r_2, g_2, b_2]$ 이 (2)와 같은 관계를 지니고 있다

면,

$$\frac{r_1}{r_2} = \frac{g_1}{g_2} = \frac{b_1}{b_2} \quad (2)$$

이 두 점은 같은 색상을 가지고 있으나 다른 휘도값을 가지고 있다는 것을 의미한다. 인간의 시각 체계는 여러 가지 다른 밝기와 환경에 의한 조명의 변화에 대해 색상을 구분할 수 있는 능력을 지니고 있으며 순수한 하나의 색상을 표현할 때 휘도는 적은 영향을 미친다. 그리고, 조명의 변화에 색상은 영향을 받지 않으나 휘도는 큰 변화를 가지게 된다. 따라서, RGB 공간상의 한 점에 대해 정확한 색상을 유지하면서 조명의 변화를 제거할 수 있는 정규화 과정을 거칠 필요가 있다.

$$\begin{aligned} r &= \frac{R}{R+G+B} \\ g &= \frac{G}{R+G+B} \\ b &= \frac{B}{R+G+B} \end{aligned} \quad (3)$$

RGB 모델의 3차원 공간상의 점들이 위와 같은 정규화과정을 거치게 되면 순수한 컬러 정보만을 지니게 된다.

$$r + g + b = 1 \quad (4)$$

그리고 (4)와 같은 관계를 가지게 되어 세 개의 요소 중 두 개의 요소로 컬러를 표현할 수 있다. 즉 $R^3 \rightarrow R^2$ 의 매핑 관계로 차원을 낮출 수 있어 계산상에서도 효율적이다. 정규화 과정을 거친 색상 모델에서는 화소에 대한 색상 정보를 $q = (r, g)$ 로 나타낼 수 있으며, 이는 각 색상 성분이 8비트로 나타내 진다면 기존 RGB 모델에서 표현 가능한 256^2 가지의 색상 정보중 휘도의 변화에 덜 민감한 $255^2/2$ 개의 색상정보를 정규화 색상 모델에서 표현할 수 있음을 의미한다. 또한 이 정규화 과정은 인간의 시각체계를 잘 표현한다는 HSI 모델과 밀접한 관련을 지고 있다.

$$\begin{aligned} H &= \cos^{-1} \left\{ \frac{\frac{1}{2} [(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\} \\ S &= 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \\ I &= \frac{1}{3} (R+G+B) \end{aligned} \quad (5)$$

즉, 정규화 과정을 거친 RGB 모델은 HSI 모델에서 밝기 성분을 제거하고 Hue와 Saturation 성분을 구 좌표계에서 직교 좌표계로 나타냈다고 할 수 있다. 위의 식에서도 알 수 있듯이 HSI 모델보다는 정규화를 거친 RGB 모델이 더 수식적으로 간단하고 계산상에서도 효율적이다.

IV. 연구 및 고찰

본 절에서는 특징물 추출을 위해 MPEG-2 Video 상에서 압축된 영역의 특징물을 추출하는데 있으며, MPEG과 같은 압축방법으로 압축된 비디오에서 키 프레임(즉, I, P, B 프레임)이라는 것과 다양한 패턴으로 나타난다는 것이다. 또한 본 연구의 실험은 카메라를 통하여 저장되어진 비디오 데이터를 기반으로 실험한 내용에 대하여 기술한다. 또한 비디오 데이터 상에서의 특정 내용 검색에 대해 실험한다. 또한, 본 실험에서는 프레임 형태에 독립적인 표현이 필요하게 된다. 이 표현에서 인덱싱과 검색 단계를 위해 추출한 모든 특징들은 프레임 형태와 같은 요소들과는 독립적으로 얻어지게 된다. 또한 본 실험에서는 인덱싱과 검색 단계에서 MB가 인트라 부호화된 것인지, 순방향/역방향으로 예측된 것인지, 혹은 쌍방향으로 예측된 것인지에 상관없이 각각의 프레임 형태에 따라 0, 1, 2개의 움직임벡터들을 가질 수 있도록 MB 기반 검출 알고리즘에 대해 연구한다.

1. MB 기반 컷 검출 알고리즘

매크로블록 기반 컷 검출 알고리즘은 각 프레임의 MB들의 형태를 이용하여 컷 검출 여부를 결정한다. MB들의 총 개수와 스킵된 MB들의 개수를 (6)과 같이 정의한다.

$$\begin{aligned} N &= N_f + N_b + N_{fb} + N_s \\ N_s &= N_{sf} + N_{sb} + N_{sfb} \end{aligned} \quad (6)$$

변수 정의	
N_f	한 프레임에서의 총 MB들의 수
N_b	순방향으로 예측된 MB들의 수
N_{fb}	역방향으로 예측된 MB들의 수
N_s	쌍방향으로 예측된 MB들의 수
N_{sf}	스킵된 MB들의 수
N_{sb}	순방향으로 예측되어 스킵된 MB들의 수
N_{sfb}	역방향으로 예측되어 스킵된 MB들의 수
N	쌍방향으로 예측되어 스킵된 MB들의 수

표 1. MB 변수의 정의

그리고 비디오 클립이 I와 P 프레임만을 갖는다면, 프레임 i 에서 시작하는 새로운 시퀀스가 결정되어진다. ($i \in P$, if $SV_i/N < t_1$) (7)참조

$$SV_i = \begin{cases} 0 & \text{if } i=0 \\ SV_{i-1} & \text{if } i \in I \\ N + N_s & \text{if } i \in I \\ N + N_{sb} + N_f + N_{sf} & \text{if } i \in B \end{cases} \quad (7)$$

- 조 건
- 조건 1 만약 $R_j \in P$ 이면, $SV_i/N < t_2$
 - 조건 2 $\forall l \in D$ 일때,
 $SV_i/N < t_3 (l \neq B)$
 - 조건 3 $\forall D$ 에서 k 이전의 B frames,
 $N_f - N_b > N_b$
 - 조건 4 $\forall D$ 에서 k 이후의 B frames,
 $N_b - N_f > N_b$

표 2. MB 기반의 컷 검출 조건

만약 조건 1은 해당 프레임의 검사여부 결정에 이용된다. 만약 이 조건이 거짓이면, 남은 조건들에 대한 테스트는 모두 생략된다.

2. 특징들을 이용한 비디오 인덱싱

본 실험의 과장된 압축된 영역에서의 특징추출 기법과 움직임벡터를 이용하여 사용자가 관심을 갖게되는 비디오 데이터를 구조화하고 검색할 수 있는 인덱싱 기법을 연구한다. 그림 5.1은 MPEG 비디오에서 컷 검출에 의해 분해된 샷을 인덱싱 하는 과정을 간략하게 보여준다.

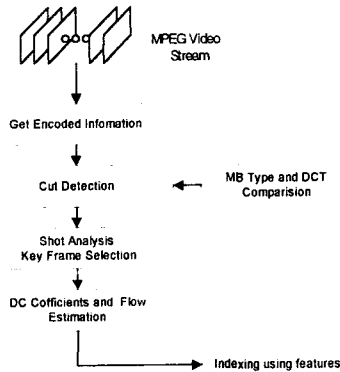


그림 2. Flow of video indexing in MPEG video

MPEG에서 일련의 비디오클립이 주어지면, 질의를 통해 비디오 클립에 대한 내용을 검색한다. 이러한 질의들이 가능하도록 비디오클립들을 구조화해야 한다. 비디오클립은 공간적 차원과 시간적 차원 모두를 가지고 있으므로, 본 논문에서는 각 장면들의 키 프레임은 공간적 차원을 표현하는 그 DC 계수들과 시간적 차원을 표현하는 그 흐름벡터들과 함께 그 장면에 대한 인덱스로서 받아들인다. 만약 질의가 단 한 프레임이라면, DB에 저장된 비디오들의 움직임정보와 대조하기 위해 끌어내야 하는 시간적 정보는 필요 없게 된다. 두 프레임들 사이의 공간적 유사성은 프레임들이 휘도, 색도, 질감, 형상과 같은 비슷한 공간

적 특성들을 갖는다는 것을 의미한다. 공간적 유사성을 찾는 데 가장 간단한 접근방법은 모든 클립의 매 프레임에 대한 모든 DCT 정보를 저장하고 비교 검색하는 것이다. 질의 프레임이 주어질 때, 가장 비슷한 한 프레임을 찾기 위해 가능한 한 모든 정보와 비교해야 할 것이다. 그러나, 이 접근방법은 시간과 공간모두에서 비효율적이다.

그리고, 휘도와 색도 성분의 DC 계수만을 이용하고 공간적 유사성을 분석하기 위해 그 계수들과 다른 프레임에서 부합하는 DC 계수들을 비교한다. 각각의 프레임을 DC 계수들의 벡터로서 다루었고 프레임들의 유사성을 결정하기 위해 벡터들사이의 거리함수를 취한 값을 이용한다.

3. 실험 결과 및 고찰

가. MPEG에서의 컷 검출 실험

MPEG 스트림에서 컷은 대부분 다음과 같은 세 가지 경우에 발생한다.

- ① 어떤 P 프레임이 인트라 부호화된 MB들이 대다수일 때
- ② B 프레임에서 대다수의 MB들이 이전 I나 P 프레임들로부터 순방향으로 예측되었을 때
- ③ B 프레임에서 대다수의 MB들이 역 방향으로 예측되었을 때.

아래의 표는 이러한 단계의 프레임에서 컷의 발생된 프레임에 따른 조건 값들을 보여준다. 여기서 B₂, I, B₁ Frame에서 B₂ Frame에서 컷이 검출된다. (표 3 참조)

	SV ₁ /N	SV ₂ /N	N _(k-1) - N _(k-1)	N _(k-1) - N _(k)	N _(k) - N _(k)	N _(k)
B ₂ Frame	Passed	Passed	27	317	16	307
I Frame	Passed	1.77 (1.4)	Passed	Passed		
B ₁ Frame	Passed	Passed	Passed		374	8
B ₂ Frame	Passed	Passed	374	8	385	1

표 3. Result of condition values

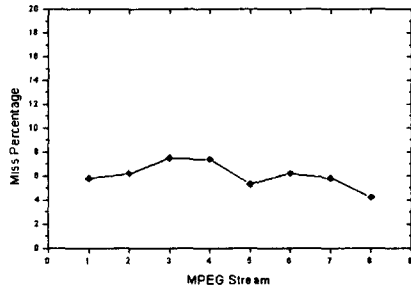
나. 비디오 질의 및 검색

DB에 저장된 키 프레임의 DC 계수들과 움직임 정보를 비교하고 질의의 키 프레임의 DC 계수들과 움직임 정보를 비교하여 비슷하게 매치(match)하는 시퀀스들을 돌려준다. 검색 시간을 줄이기 위해, 먼저 적응된 B*트리의 각 노드에 저장된 DC 계수들과 움직임벡터의 평균치를 비교한 후, 비슷하게 일치하면 더욱 정확한 질의결과를 산출하기 위해 그 샷을 검색한다. 그림 5.2은 하나의 질의 프레임이 주어졌을 때, DB에 인덱스된 정보와 추출된 질의 프레임의 특징들을 비교하여 비슷한 프레임들을 돌려주는 것을 보여준다.

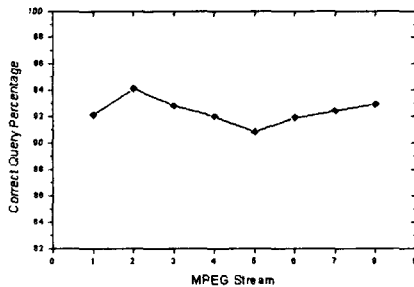


Query Frame Match 1(Frame 212)
그림 3. Retrieval results on MPEG Video Stream A (Continue)

이러한 기법들을 이용하여 픽셀영역과 압축된 영역에서 특징 추출 후에 인덱싱하는 과정까지의 소요된 시간을 보여주고 있다. 픽셀영역에서 인덱싱하는 것은 MPEG 비디오를 압축해제한 후에 그 실제적인 처리가 이루어지고 그 처리단위가 픽셀단위이므로 그 만큼 많은 시간이 소요되었다.



(a) Query Frame Percentage



(b) Match frame Percentage

그림 4. Percentages of retrieval results on MPEG stream

참고문헌

1. F. Arman, A. Hsu, and M.Y. Chiu, "Image Processing on Compressed Data for Large Video Databases", Proc. ACM Multimedia 93, Anaheim, CA, pp. 267-272, 1993.

2. D. Le Gall, "MPEG: A Video Compression Standard for Multimedia Applications", Communications of the ACM, 34(4), pp. 46-58, 1991.

3. R. M. Haralick and L. G. Shapiro, "Computer and Robot Vision", Vol. 2, Addison-Wisely, Reading, MA, 1993.

4. M. Mills, J. Cohen, and Y. Y. Wong, "A Magnifier Tool for Video Data", Proc. CHI '92, Monterey, CA, pp. 93-98, 1992.

5. A. Nagasaka and Y. Tanaka, "Automatic Video Indexing and Full-Video Search for Object Appearances", Visual Database Systems II, E. Knuth and L. M. Wegner, editors, North-Holland, pp. 119-133, 1991.

6. B. C. O'Connor, "Selecting Key Frames of Moving Image Documents: A Digital Environment for Analysis and Navigation", Microcomputers for Information Management 8(2), pp. 119-133, 1991.

7. R. Steinmetz, "Data Compression in Multimedia Computing-Standards Systems", Multimedia Systems 1(4), pp. 187-204, 1994.

8. D. Swanberg, C. F. Shu, and R. Jain, "Knowledge Guided Parsing in Video Databases", Proc. IS&T/SPIE Conf. on Storage and Retrieval for Image Video Databases, San Jose, CA, 1993.