고유성분 분석과 휘도성분 흐름 특성을 이용한 내용기반 비디오 검색

Content-Based Video Search Using Eigen Component Analysis and Intensity Component Flow

전대홍, 강대성
Dae-Hong Jeon and Dae-Seong Kang

동아대학교 전기전자컴퓨터공학부
School of Electrical, Electronic and Computer Eng., Dong-A Univ.
E-mail: dskang@daunet.donga.ac.kr

요약

본 논문은 동영상의 대표프레임에서 eigen value와 휘도 성분을 이용한 내용기반 검색 기법에 관한 연구이다. video를 shot단위로 분할하여 shot을 대표하는 대표프레임을 얻어내고, 그 대표프레임을 Eigen Component Analysis(ECA)를 이용하여 데이터베이스를 생성하여 shot들의 휘도값 분포를 구한다. 생성된 코드북과 각 대표 프레임에 대한 코드북 인덱스 값, 휘도값을 database화하여 질의 영상과 video database간의 저장된 비디오들의 코드북과 코드워드간의 유클리디안 거리를 이용하여 유사도 높은 비디오를 찾고, 검색되어진 video에 포함된 대표프레임들의 코드북 인덱스 값과 입력 영상의 코드북 인덱스 값을 비교하여 입력 영상의 특징과 가장 유사한 대표프레임을 얻어낸다. 실험결과 제안된 방법이 검색에 있어서 영상의 형태에 대한 전체적인 특징을 제공하는 대표프레임의 eigen value와 휘도 성분을 이용함으로서 보다 검색 결과가 우수하며, 영상의 통계적인 특성을 이용함으로서 시간과 메모리 공간을 줄일 수 있음을 확인하였다.

Abstract

In this paper, we proposed a content-based video search method using the eigen value of key frame and intensity component. We divided the video stream into shot units to extract key frame representing each shot, and get the intensity distribution of the shot from the database generated by using ECA(Eigen Component Analysis). The generated codebook, their index value for each key frame, and the intensity values were used for database. The query image is utilized to find video stream that has the most similar frame by using the euclidean distance measure among the codewords in the codebook.

The experimental results showed that the proposed algorithm is superior to any other methols in the search outcome since it makes use of eigen value and intensity elements, and reduces the processing time etc.

key words: ECA, key frames, intensity, search, video stream

I. 서론

최근 컴퓨터와 초고속 통신의 급격한 발달은 디지털화된 멀티미디어 데이터의 폭발적인 증가를 가져왔다. 방대해진 멀티미디어 데이터들을 효과적으로 저장, 브라우징(browsing), 인덱싱(indexing), 검색(search)하기 위한 연구들이 이루어지고 있다. 특히 VOD 서비스, 디지털 박물관, 의료영상 검색 시스템과 같은 많은 응용 분야의 핵심 기술로서 멀티미디어 데이터베이스의 인덱싱 및 검색에 대해 활발한 연구가 진행되고 있다. 특히 기존의 텍스트 주석에 기반한 검색 방법의 한계를 극복하기 위하여 내용기반 검색(content-based search)에 대한 연구가 활발히 진행되고 있다. 초기의 영상 데이터베이스 시스템에서는 텍스트 기반 검색으로, 각 영상에 대해 그 영상을 나타내는 주석 정보를 기술한 뒤, 이 정보를 이용하여 영상을 검

색하게 된다. 그러나 멀티미디어 정보는 그 특성상 텍스트 주석 정보로 표현하기에 주관적이고, 대용량 매체의 경우 사람이 각 영상에 색인을 첨부해야 하므로 요구되는 일의 양이 늘어날 뿐만 아니라, 영상을 충분히 구분 할만한 주석이 부족하므로 질의 영상과 유사한 영상을 찾을 수 없다는 단점이 있다. 이에 반해 내용 기반 검색 방법은 물체의 움직임(motion), 형상(shape), 색(color), 질감(texture) 등을 이용하여 검색할 수 있어, 자동적으로 영상의 특징을 추출하고 정보를 저장할 수 있는 장점이 있다 [1].

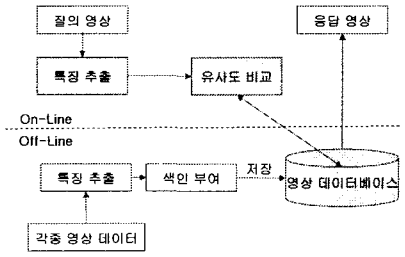


그림 1. 내용기반 검색 시스템
Fig. 1. Content-based search system

일반적으로 내용 기반의 검색은 그림 1에서와 같이 두 영상의 유사성에 대한 평가 척도로서, 두 영상 특징들 간의 차이값을 이용한다. 검색 대상 영상과 가장 유사한 최적의 영상을 찾기 위해, 영상 특징 정보의 차이값을 척도로 하여 데이터베이스 내의 모든 영상들과 검색 대상 영상을 비교한다. 이를 일반적으로 전역 탐색이라고 한다. 이 방법은 방대한 계산량을 요구하기 때문에, 최근에 여러 고속 탐색 기법이 개발되고 있다. 또한 내용 기반 검색 방법에는 여러 가지 방법들이 제시되었으며, 그 중 컬러 히스토그램을 사용하는 방법과 투영 기법을 이용하는 방법 등이 있다. 컬러 히스토그램을 사용한 방법에서는 영상의 전체 특성은 잘 나타내는 반면 영상의 위치정보의 손실로 인해 다른 영상을 같은 영상으로 오인식할 수 있다. 투영 기법을 사용한 방법은 영상의 크기에 따라 생성되는 데이터가 달라지기 때문에 다양한 형태의 데이터를 검색하는 데 문제가 발생하게 된다[2].

본 논문에서는 동영상의 전체특성을 이용하여 검색하기 위해 영상전체의 흐름을 나타낼 수 있는 intensity value 및 codeword의 발생빈도를 검색키로 사용하였다. 비디오 스트림을 shot 단위로 분할한 다음 분할된 shot의 시작 프레임을 대표프레임으로 사용하고, 추출된 대표 프레임에 ECA을 이용하여 입력 벡터의 차원을 축소하며 또한 입력 벡터의 통계적인 특성을 데이터베이스의 색인 정보로 저장한 후 입력되는 질의 영상에 대해 원하는 유사한 영상을 비디오 전체 내용을 검색하지 않고 대표 프레임과 intensity value를 사용한 영상 검색 기법을 제안한다.

II. 내용기반 멀티미디어 검색 기법

1. 정지영상 데이터에 대한 내용기반 영상검색

내용기반 영상 검색 기법으로는 색상 정보를 이용한 영상 검색 방법과 질감 정보를 이용한 영상 검색, 형태 정보를 이용한 영상검색, 위치 정보를 사용한 검색 방법과 압축 영상에서의 검색 방법이 있다[3][4][5].

2. 동영상 데이터에 대한 내용 기반 영상 검색

동영상에 대한 내용 기반 영상 검색을 위해서는 동영상 데이터를 색인하기 위한 비디오 파싱기법과 사용자가 원하는 데이터를 쉽게 검색할 수 있는 사용자 인터페이스 뿐만 아니라 제한된 저장공간에 대용량 비디오를 효율적으로 저장하기 위한 비디오 데이터 압축 및 저장 방법 등의 기술들이 필요하다. 동영상에서 영상의 정보를 검색에 사용하는 다른 방법은 영상의 정보들, 예를 들면 움직임, 색, 모양 등의 특징들을 사용해 일정한 영역의 움직임을 찾아내는 방식이다. 사용자는 원하는 특정 영상을 찾기 위해 정적인 스케치를 하는 것이 아니라 원하는 동영상의 일부를 찾기 위해서 동영상내의 특정 개체의 움직임을 스케치할 수 있으며 입력된 비디오는 비디오 분할에 의해서 장면단위의 shot으로 구분된 후 사용자는 각 샷의 물리적 정보 및 논리적 정보를 추출하여 shot단위로 인덱싱 할 수 있다 [6].

비디오내의 특정 객체의 움직임을 찾아내는 연구로 VideoQ 시스템에서는 사용자로 하여금 비디오내의 특정 영역의 움직임을 표현해주는 움직이는 스케치(animated sketch) 질의를 하게 해준다. 이 시스템은 먼저 비디오를 분할하여 shot으로 구분한 후 shot단위로 특정 영역들의 움직임을 추출하여 하나의 객체로 만든다. 시스템은 각 객체로부터 객체단위의 색이나, 질감, 모양, 움직임 등의 특징들을 추출하여 데이터베이스에 저장한 후 검색시에 이용한다. 이 시스템은 웹기반 비디오 검색 시스템으로 영상내의 특정 영역이 배경과 구분이 잘되고 스케치로 표현하기 쉬운 경우, 예를 들면 축구 사건, 스키 사건들의 경우 좋은 검색 결과를 보인다[7][8][9].

III. ECA를 이용한 멀티미디어 검색 시스템

1. 대표프레임의 검출

비디오 영상에서 대표 프레임을 검출하는 방법은 여러 가지가 있으나 본 논문에서는 비디오 영상을 shot단위로 분할한 다음 분할된 shot의 시작 프레임을 대표 프레임으로 사용하였다. 비디오 영상에서 shot이란 단위는 하나의 카메라에 의해서 찍혀진 비디오 스트림을 칭하는 단위로서 shot내에서는 프레임들간의 변화가 적다. 그리고 하나의 shot 길이는 짧게는 2~3초에서 길게는 1분 이상 되므로 4~5분 정도 되는 비디오 영상의 경우는 10개 내외로 그 수도 적다. 따라서 shot의 시작 프레임을 전체 비디오 영상을 대표하는 프레임으로 사용하여도 원하는 비디오 영상을 검색할 수 있다. 그림 2는 비디오 영상에서 shot의

경계를 검출하기 위한 알고리즘이다.

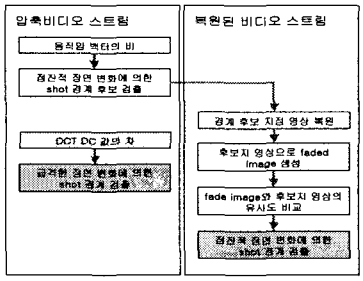


그림 2. shot 경계 검출 알고리즘

Fig. 2. Shot boundary detection algorithm

(가) 압축 영역에서의 작업

MPEG[10]으로 압축된 영상에서 DCT 변환된 값들 중 DC 값은 DCT 변환된 8 x 8 영상의 평균을 나타낸다. 따라서 DCT 변환된 프레임에서 DC 값들로 영상을 만들고, 만들어진 영상의 히스토그램을 구하면 복원 영상의 히스토그램과 유사한 특성을 갖는다. 이러한 DCT 변환의 성질을 이용하여 영상을 복원하지 않고 DC 영상의 히스토그램 값의 변화만으로도 shot이 변화하는 지점을 검출할 수 있다. DC 영상의 히스토그램을 이용할 경우 장면이 서서히 변화하는 shot 경계는 검출하지 못하므로 이를 보완하기 위해 움직임 예측 벡터의 비를 이용한다.

MPEG으로 압축된 영상의 B 프레임에는 전방향 움직임 예측 벡터(FMV)와 역방향 움직임 예측 벡터(BMV)가 있다. 이들 움직임 예측 벡터들은 프레임 간의 유사도를 이용하여 이전 혹은 이후의 프레임으로 현재 프레임을 표현하는 것들로 현재 프레임에서 장면 변화가 일어나는 경우 프레임 간 유사도가 적어져 유사도가 적은 방향의 움직임 벡터가 급격히 줄어들게 된다. 이러한 움직임 벡터의 성질을 이용하여 shot 경계를 검출할 경우 화면의 색상이나 명암의 급격한 변화에는 강인하나 물체의 움직임이 크거나 한 경우 잘못된 shot 경계를 검출한다. 하지만 장면이 서서히 변화하는 shot 경계도 검출해 내는 장점이 있어 점진적 장면 변화에 의한 shot 경계 후보지를 선정하는데 사용하였다. 급격한 장면 변화에 의한 shot 경계 지점 검출과 점진적 장면 변화에 의한 shot 경계 후보지 검출 과정은 다음과 같다.

Step 1. N-프레임에서의 DCT DC 값의 평균(DC_N)을 구한다.

Step 2. DC_N과 DC_{N-1} 값의 차가 문턱치($Threshold_{DC}$)보다 크면 급격한 장면 변화에 의한 shot 경계로 한다.

Step 3. 전체 비디오 스트림에서 FMV와 BMV의 비를 구한다.

Step 4. FMV와 BMV의 비의 구간 평균값을 계산할 window를 설정한다.

Step 5. Window의 중앙에 위치하는 프레임의 FMV/BMV 비와 window 내의 FMV/BMV 비의 평균값을 비교한다.

Step 6. 중앙 프레임의 FMV/BMV 비가 크면 점진적 장면 변화에 의한 shot 경계 후보로 설정한다.

(나) 점진적 장면 변화에 의한 shot 경계 검출

움직임 벡터의 비를 이용하여 검출한 shot 경계 후보지의 경우 물체의 움직임이 많은 장면이나 카메라의 panning이 일어나는 장면, 점진적 장면 변화에 의해 shot이 변화하는 경우를 모두 포함하고 있다. 이들 중 점진적 장면 변화에 의해 shot이 변화하는 지점을 찾기 위해서 영상 편집 시 fade 효과를 생성하는 원리를 이용하였다. Fade 효과는 아래와 같은 수식으로 표현할 수 있다[10].

$$E(t) = O(t_S)\alpha(t) + I(t_E)[1-\alpha(t)] + C, \quad \forall t \in (t_S, t_E) \quad (1)$$

여기서, $E(t)$는 시간 t에서 화면에 나타나는 영상이며, $O(t_S)$는 fade 효과가 시작되는 시간 t_S의 비디오 스트림이며, $I(t_E)$는 fade 효과가 끝나는 시간 t_E의 비디오 스트림이다. 이때 $O(t_S)$가 0인 경우를 fade in, $I(t_E)$가 0인 경우를 fade out, $O(t_S)$와 $I(t_E)$ 모두 0이 아닌 경우를 dissolve라 한다. $\alpha(t)$는 시간 t_S에서 t_E사이의 구간에서 1에서 0의 값으로 단조 감소하는 일차원 함수이다. 식(1)과 같은 fade 효과의 특징을 이용하여 n번째 후보지의 영상을 out going 영상으로, n+2째 후보지의 영상을 in coming 영상으로 하여 임의의 fade image를 만들어낼 수 있다. 이렇게 만들어진 fade image와 n+1 번째 후보지의 영상의 유사도를 비교하여 두 영상의 유사도가 높은 경우 n+1 번째 후보지는 점진적 장면 변화에 의한 shot이 변한 곳으로 검출해 낸다. 다음은 후보지들에서 점진적 장면 변화에 의한 shot 경계를 검출 과정을 나타낸 것이다.

Step 1. 점진적 장면 변화에 의한 shot 경계 후보지를 복원한다.

Step 2. 수식 (1)을 이용하여 n번째 후보지와 n+2번째 후보지로부터 fade image를 만든다.

이때 $\alpha(t)$는 0.9에서 0.1까지 0.1 step으로 하여 총 9개의 fade image를 만든다.

Step 3. Step 2에서 만들어진 fade image와 n+1번째 후보지 영상의 유사도를 비교한다.

Step 4. 유사도가 높으면 점진적 장면 변화로 검출한다.

Step 5. $\alpha(t)$는 0에서 1까지 0.1씩 증가시켜가면서 Step 2에서 Step 4까지의 과정을 반복한다.

Step 6. 전체 후보지들에 대해 Step 1에서 Step 5까지의 과정을 반복한다.

2. 주성분 해석 기법

본 논문에서는 앞절에서 기술한 대표 프레임 추출 알고리즘에 의해 추출된 대표 프레임에 대한 영상 색인 및 검색 기법으로 주성분 해석을 이용하는 방식에 대해 제안한다. 주성분 해석 기법은 입력 공간내의 변동량이 많은 방

향으로 주축을 발견하여 데이터의 중복성을 제거하는 기법으로 주성분 분석 기법이라고도 한다[11]. $n \times n$의 2차원 영상을 n개의 1차원 데이터 요소로 축소함으로서 상당한 양의 처리 시간과 메모리 공간을 줄일 수 있다. 또한 주성분 분석에 의하여 계산된 주축은 입력 벡터의 최대 고유치의 고유 벡터의 방향과 일치하며 입력 벡터의 변동량이 가장 큰 방향이다.

본 논문에서는 추출된 대표 프레임들을 2차원 입력 벡터로 나타내어 변동량이 가장 많은 주축을 구한다. 주축은 입력 벡터의 최대 고유치의 고유 벡터의 방향과 일치함으로 데이터의 중복성을 제거할 수 있다. 다음으로 영역 분할 알고리즘을 이용해서 주축과 수직을 이루면서 노드의 평균을 지나는 지점을 기준으로 영역을 2개로 분할한다. 각 영역의 평균 오차를 계산하여 평균 오차가 가장 큰 영역을 선별한 후 선택된 영역에 영역 분할 알고리즘을 재적용한다. 다음은 주축을 구하는 알고리즘과 영역 분할 알고리즘이다.

(가) 주축 및 영역분할 알고리즘

주축을 구하기 위해서는 우선 입력벡터에 대한 기하학적 평균(geometrical mean)과 입력 벡터에 대한 공분산 행렬(covariance matrix)과 가장 큰 고유치(the largest eigenvalue)를 구함으로써 주축의 각도를 계산 하게된다.

영역 분할을 구하기 위해서는 같은 영역에 할당된 입력 벡터의 평균 좌표를 계산하며 주축을 구하는 알고리즘을 이용하여 주축과 수직이면서 노드의 평균 좌표를 지나는 직선과의 각도(ϕ)를 계산하여 다음의 규칙에 의해서 해당 영역의 각 데이터는 두 영역, R1, R2로 분할된다.

if (cos(ϕ) ≤ 0) then R1 else R2

그림 3은 주축을 구하는 알고리즘과 영역 분할 알고리즘을 적용하여 전체 입력 벡터를 두 부분으로 나누는 과정을 나타낸 것이다[11].

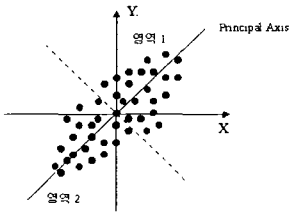


그림 3. 주축에 의한 영역 분할
Fig. 3. Domain division by principal axis

이상의 알고리즘에 의해 전체 입력 벡터는 여러 개의 영역으로 분할되고 각 영역의 평균값을 코드워드로 사용한다. 그리고 추출된 대표 프레임은 각 코드워드에 할당된 값을 카운트하여 코드북을 만든다. 이 카운트 값은 MPEG 비디오 스트림의 색인을 위한 파라메터로 사용된다.

(나) ECA 알고리즘

Sanger에 의해 제안된 주성분 해석[11]을 위한 학습 알고리즘 GHA의 문제점은 학습율이 입력 상관 행렬(covariance matrix)의 고유치에 따라 큰 고유치에 대해서는 발산하고 작은 고유치에 대해서는 아주 천천히 수렴하는 단점을 갖는다. 이를 극복하기 위해 ECA는 모든 고유치에 대해 비슷한 학습 속도를 갖는 적응 학습 알고리즘을 사용한다. 확률 분포 $P(\boldsymbol{x})$를 갖는 n차원 제로-평균 입력 패턴 벡터 $\boldsymbol{x}$에 대해, ECA는 처음 m개의 주요 성분 벡터를 구한다[11].

(다) 대표프레임의 휘도성분 분석

디지털 영상의 휘도 성분 분포도는 한 영상의 형태에 대한 전체적인 특징을 제공하며 동영상의 대표 프레임은 그 영상의 특징을 가장 잘 나타내고 있다.

$$I_{diff} = \sum_{i=1}^{n-1} [K_i(x,y) - K_{i+1}(x,y)] \quad (2)$$

수식 (2)는 각 프레임간의 휘도성분의 차분을 나타내며 n은 대표프레임 입력 개수이며, K_i는 i번째 대표프레임의 휘도성분이다. I_{diff}는 대표프레임간의 휘도성분의 차분값이 된다.

$$I_{ave} = \frac{1}{n} \sum_{i=1}^{n} I_{diff_i} \quad (3)$$

수식 (3)은 휘도차분의 평균이며 I_{diff_i}는 i번째의 휘도 차분값이며, n은 I_{diff_i}의 전체 갯수이다.

$$I_{dis} = \frac{1}{n} \sum_{i=1}^{n} I_{diff_i}^2 - I_{ave}^2 \quad (4)$$

수식 (4)는 휘도 차분값에 대한 분산에 관한 식이다.

본 논문에서는 동영상의 대표프레임을 추출 후 대표프레임간의 휘도성분의 차분을 구함으로서, 영상 전체의 휘도성분 변화와 성질을 파악한다. 그리고 검색오류를 줄이기 위해 휘도성분의 차분값의 분산을 구하였다. 이는 차분값의 평균값만을 이용하였을 때, 평균값은 같지만 휘도성분의 분포곡선이 다른 video stream을 동일영상으로 검출해 내는 것을 막을수 있다.

3. 검색 시스템의 구현

본 논문에서 제안하는 영상 검색 시스템은 그림 4와 같으며, 비디오 스트림에서 추출한 대표프레임을 ECA을 사용하여 얻은 코드북의 코드워드와 코드워드의 발생 빈도를 이용하여 DB를 구축하였고 Intensity value도 같이 추출하여 구축하였다. 질의 영상에 대한 응답은 질의 영상을 영상 데이터베이스의 코드북에 맵핑하였을 때의 오차값과 질의영상의 코드워드 발생빈도와 데이터베이스에 저장된 비디오 스트림의 대표프레임들의 코드워드 발생빈도와 Intensity value를 비교하여 얻었다.

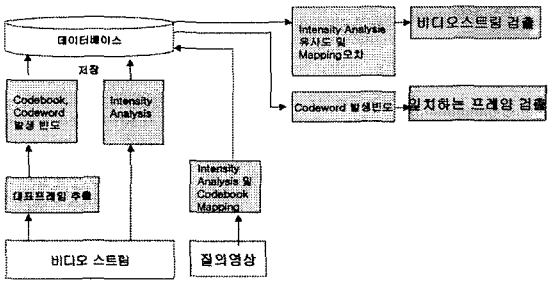


그림 4. 제안 알고리즘을 적용한 검색 시스템

Fig 4. The proposed search system

이는 식(5)와 (6)으로 표현가능 하며 intensity value의 비교부분은 실험에서 임의로 정한 유사도를 사용한다.

$$Code_{diff}(i) = \sum_{a=1}^{IMG-SIZE} \min(|Q(a) - C(i,1)|, |Q(a) - C(i,2)|, \cdots, |Q(a) - C(i,M)|) \quad (5)$$

수식 (5)에서 $Code_{diff}(i)$는 질의영상을 i번째 비디오 스트림의 코드북에 맵핑하였을 때 오차이며, IMG-SIZE는 질의영상의 크기, $Q(a)$는 질의영상의 a번째 픽셀의 값, M은 코드북의 코드워드 수, $C(i,1), C(i,2), \cdots, C(i,M)$는 i번째 비디오 스트림의 $1,2,\cdots,M$째 코드워드 값이다.

$$Frame_{diff}(i,j) = Code_{diff}(i) \sum_{k=1}^{M} |VF(i,j,k) - Q(i,k)| \quad (6)$$

수식 (6)에서 $Frame_{diff}(i,j)$는 각각 i 번째 비디오 스트림의 j 번째 대표프레임과 질의 영상과의 유사도이며, $VF(i,j,k)$는 i 번째 비디오 스트림의 j번째 대표프레임의 k 코드워드의 발생 빈도, $Q(i,k)$는 질의 영상을 i번째 비디오 스트림의 코드북에 맵핑하였을 때 k 코드워드의 발생 빈도이다. 먼저 수식 (5)를 이용하여 우선 질의 영상과 영상 데이터베이스에 저장된 코드북에 맵핑하여 $Code_{diff}(i)$를 얻어내고 이 값이 가장 작은 경우와 Inten-sity value가 질의영상과 유사도를 비교하여 비디오를 검출한다. 유사도는 실험데이타의 주관적 견해로 임의로 정한 값이며 유사도 범위는 $0 \sim 0.02$(분산값) 이다.

이렇게 비디오가 검색이 되어지면 검색되어진 비디오에 대해서 수식(6)을 이용하여 질의영상과 일치되는 대표프레임을 얻어낸다. 이렇게 하나의 질의영상에 대해 일치하는 비디오 스트림과 비디오 스트림 내의 대표프레임들에 대한 검색을 실시하여 원하는 shot까지 검색을 한다.

Ⅳ. 실험결과 및 고찰

표 1은 대표 프레임을 추출하기 위해서 사용한 shot 경계 검출 알고리즘의 성능을 검사하기 위해서 실험에 사용한 비디오 중에서 직접 shot 경계위치를 찾은 다음 제안한 알고리즘을 이용하여 검출한 결과와 비교한 것이다.

표 1. shot 경계를 검출한 결과

Table 1. The result of the shot boundaries searched

검출항목	전체 shot 경계	급격한 장면 변화 검출 후	후보지 검출 후	최종 결과
후보지	X	X	97	X
급격한 장면 변화	13	16	X	16
점진적 장면 변화	11	X	X	9
오검출	X	3	73	3
미검출	X	0	0	2

실험에 사용된 MPEG 스트림은 1700개의 프레임으로 구성이 되어 있고 13개의 급격한 장면 변화에 의한 shot 경계와 11개의 점진적 장면 변화에 의한 shot 경계를 가지고 있으며, 오검출된 것 중에는 fade in/fade out 시간이 짧은 dissolve를 급격한 장면 변화에 의한 shot 경계로 인식한 것과 카메라의 panning을 서로 다른 shot 경계로 인식한 경우가 있었으며, 장면 변화가 아주 긴 프레임에 걸쳐서 일어나는 경우 경계를 검출하지 못하는 경우가 있었다

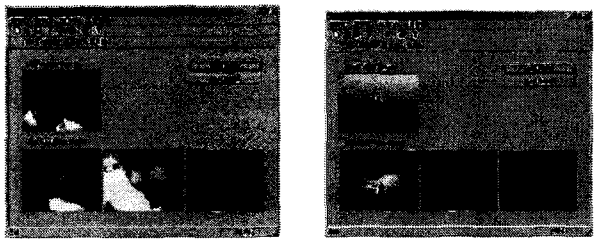


그림 5. Query video stream에 대한 검색결과

Fig. 5. The searched result by the query video stream

표 2와 그림 5는 20개의 동영상에서 얻은 100개의 query video stream 을 이용하여 DB에 질의하여 얻은 결과이다. 그림 5와 같이 query video stream을 비디오 데이터 베이스에 질의하였을 경우, search video stream 과 같은 video stream 이 검색되었다. 먼저 대표 프레임을 이용하여 질의하였을 경우에는 해당 프레임을 검색할 확률 및 질의 영상이 포함된 비디오를 검색할 확률을 나타낸 것이며, 임의의 프레임을 이용하여 질의를 하였을 경우는 질의 영상과 동일한 영상이 영상 DB 내엔 존재하지 않으므로 질의 영상을 추출한 비디오를 검색된 경우를 나타내었다. video stream의 유사성을 이용하여 database 내의 video stream을 검색한 결과 질의 영상의 intensity value 와 1순위, 2순위, 3순위로 검색된 비디오스트림의 intensity value 의 분포도는 그림 6과 같다.

표 2. 질의에 대한 검색 결과
Table 2. The searched result of the proposed algorithm

	대표프레임으로 질의		대표프레임이 아닌 경우
	Image Clip	Video Stream	Video Stream
1 순위 검색률	86%	85%	83%
2 순위 내 검색률	93%	90%	87%
3 순위 내 검색률	95%	96%	95%

표 3. ECA알고리즘을 사용한 알고리즘
Table 3. The searched result of the ECA algorithm

	대표프레임으로 질의		대표프레임이 아닌 경우
	Image Clip	Video Stream	Video Stream
1 순위 검색률	86%	82%	82%
2 순위 내 검색률	93%	88%	85%
3 순위 내 검색률	95%	92%	92%

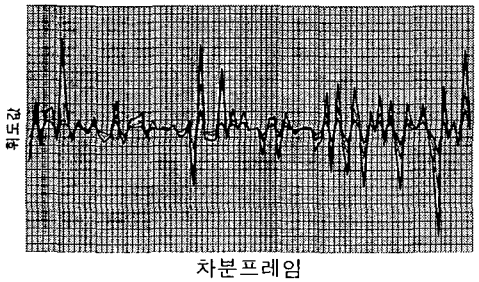


그림 6. 검색 되어진 Intensity Analysis 분포도
Fig. 6. The distribution of intensity analysis

대표프레임간 프레임의 휘도성분에 대한 차분값을 구하여 기존의 방식과 다르게 질의를 해보았다. 영상의 모든 특성을 표현하고 있는 휘도 성분을 검색키로 사용하였고, 제안 알고리즘을 검색에 적용한 결과, 질의 영상과 동일한 정지영상을 찾아내는 성능을 여타 정지영상 검색 알고리즘에 비해 떨어지지만 정지 영상을 포함하고 있는 비디오스트림을 찾는데는 그 성능이 우수함을 알 수 있다. 통계적인 특성의 ECA알고리즘만을 적용한 결과 intensity를 검색키로 적용한 결과에 비해서 정지 영상에서는 큰 차이는 없었으나, 동영상에서는 약 2% 정도의 검색율이 떨어짐을 확인해 볼수 있었다.

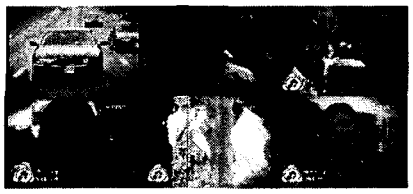


그림 7. 검색 결과가 좋은 비디오스트림들의 예
Fig. 7. The video streams with good search result etc.

그림 8. 검색 결과가 좋지 못했던 비디오 스트림들의 예
Fig. 8.The video streams with bad search result etc.

제안 알고리즘은 추후 사용자 정의 입력을 이용한 질의 기에서 생성된 이미지를 이용할 때에도 우수 검색효율을 나타낼 것으로 기대되어진다. 검색 엔진의 성능을 실험한 전체 결과는 위의 표 2와 같지만 실험 영상들의 특징에 따라 검색율의 차이를 보였다. 표 4는 질의와 출력 비디오 스트림의 intensity value 평균값이 동일하여 같은 비디오 스트림으로 오검출 할수 있었으나 비디오 스트림의 Intensity 분산특성을 검색키로 추가함으로써 오검출율을 줄일수 있었다.

표 4. Intensity value 분산 값을 이용한 검색 결과
Table 4. The search result used by variance of intensity value

	질의영상의 video stream	1순위내 검색 Video stream	2순위내 검색 Video stream	3순위내 검색 Video stream
Video stream의 차분평균값	0	0	0	0
Video Stream의 분산값	0.01	0.009	0.013	0.02

그림 5와 그림 7에서처럼 비디오 스트림 내에서 경계가 명확한 경우에는 대표프레임이 영상 전체를 대표할 수 있어 검색율이 높았으나 그림 8과 같이 비디오 스트림의 경계가 불명확할 경우는 검색율이 좋지 못하게 나왔다.

전체 비디오의 검색키로 intensity value외에 추가로 영상 내의 물체의 움직임이나 전체 비디오의 색상 변화 등 부가적인 검색키를 사용하여 경계가 불명확한 경우의 특성을 고려할 경우, 보다 높은 검색율을 나타낼 것이다.

V. 결론

본 논문에서는 비디오를 shot 단위로 분할하여 shot을 대표하는 대표 프레임을 얻어내고, 그 대표 프레임을 ECA를 사용하여 비디오에 대한 코드북과 각 프레임들에 대한 인덱스를 생성하고 intensity value을 분석한다. 여기서 얻어진 intensity value를 비디오 데이터베이스의 색인으로 사용하게된다. 질의 영상과 비디오 데이터베이스에 저장된 비디오들의 코드북의 코드워드 간의 유클리디안 거리와 intensity value를 이용하여 유사도가 높은 비디오를 찾고, 검색되어진 비디오에 포함된 대표프레임들의 코드북 인덱스 값과 입력 영상의 코드북 인덱스 값을 비교하여 입력영상의 특징과 가장 유사한 대표프레임을 얻었다. 본 눈문에서 제안한 알고리즘을 이용하여 검색한 결과 검색의 색인이 비디오나 입력영상 전체의 통계적 특성을 이용함으로서 원하는 영상을 효과적으로 찾을 수 있었다. 비디오 검색의 경우 질의 영상이 비디오 전체의 통계적 특성과 다른 영상을 입력하였을 경우 원하는 비디오를 정확히 찾아내지 못하는 경우가 있었으나, 원하는 대표프레임은 일정 순위 내에서 얻을 수 있었다. 자료 검색 시에도 영상의 특징 값 몇 개만을 비교함으로서 데이터 처리 시간과 메모리 공간을 줄일 수 있는 장점이 있다. 향후 제안한 방법은 색상과 같은 색인정보와 모션 데이터를 이용할 경우, 보다 나은 검색 결과를 제공할 것이다.

접수일자 : 2002. 6.28 수정완료 : 2002. 7.20

참고 문헌

[1] A.Pentlandetal., "Photobook: Tools for content-based manipulation of image databases." SPIE Proc. Storage and Retrieval for image and video databases, vol. 1908, pp 34-47, Feb. 1994.

[2] V.N. Gudivada and V.V.Raghvan, "Content-Based Image Retrieval Systems", IEEE Computer, Vol.28, No.9, pp.18-22, 1995.

[3] Myron Flickner et al, "Query by Image and Video Content: The QBIC System", IEEE Computer, Vol.28, No.9, pp.23-32, 1995.

[4] F.Idris and S.Panchanathan, "Review of Image and Video Indexing Techniques", Journal of Visual Communication and Image Representation, Vol.8, No.2, pp.146-166, 1997.

[5] Anil.K.jain, Aditya Vailaya, "Image Retrieval using Color and Shape," Pattern Recognition, Vol.29, No.8, pp.12-33, 1996.

[6] Markus Stricker and Alexander Dimai, "Color Indexing with Weak Spatial Constraint," Storage and Retrieval for Image and Video Database IV, SPIE proceedings, Vol. 2670, pp.29-40,1996.

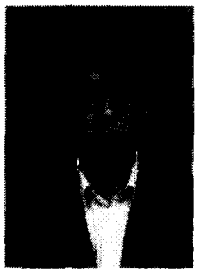
[7] I. K. Park, I. D. Yun. and S. U. Lee, "Color Image Retrieving using Hybrid Graph Representation," Image and Vision Computing, Vol.17, No.7, pp.465-474, 1999.

[8] D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos, "Distance Measures for Color Image Retrieval," International Conference on Image Processing, Vol.2, pp.770-774, 1998.

[9] Bo Tao & Bradley W. Dickinson, "Recognition and Retrieval of Textured Images using Gradient Indexing," International Conference on Image Processing, Vol.1 1998.

[10] 유영달, "Robust feautre를 이용한 key frame 추출과 MPEG video 검색," 동아대학교 대학원 공학석사학위논문, 1999.

[11] Oja, E., "A simplified neuron model as a principal component analyzer," Journal of Mathematical Biology 15, pp.267-273, 1982.



전대홍(Dae-Hong Jeon)
準會員
2001년 동아대학교 전자공학과 (공학사)
2001년-현재 동아대학교 전자공학과 석사과정
관심분야는 멀티미디어통신, 영상검색, 영상신호처리, Object Tracking

강대성(Dae-Seong Kang)
正會員
경북대학교 전자공학과 (공학사)
미국 Texas A&M대학교 (공학석사)
미국 Texas A&M대학교 (공학박사)
1984.3-1989.6 국방과학연구소 연구원
1994.7-1995.8 전자통신연구원 선임연
1995.9-현재 동아대학교 전기전자컴퓨터공학부 부교수
관심분야는 영상처리, 영상인식, MPEG