

압축공간에서 활동도 기술자를 이용한 비디오 요약

윤진성*, 김계영*, 최형일*

*송실대학교 컴퓨터학부

e-mail:jsyoon@vision.soongsil.ac.kr

Video Summarization Using Activity Descriptor In Compressed Domain

Jin-Sun Yoon*, Gye-Young Kim*, Hyung-Il Choi*

*School of Computing, Soong-Sil University

요 약

본 논문에서는 MPEG-7의 활동도 기술자를 이용한 비디오 요약 기술을 제안한다. 제안한 방법은 압축상태의 비디오 자료에서 직접 움직임 벡터들을 추출, 각 프레임들의 활동도의 강도를 계산하고 프레임의 흐름에 따라 계산된 활동도의 변화량에 대해 푸리에 변환을 적용하여 얻어진 주파수 성분을 분석하여 활동도의 시간적 분포도를 계산한다. 계산된 강도 및 분포도는 MPEG-7의 표준에 따르기 위해 양자화하여 비디오 요약에 이용한다.

1. 서론

최근 디지털 영상, 비디오, 음악 등 멀티미디어 정보의 공유가 가속화되면서 이들을 효과적으로 검색할 수 있는 시스템의 필요성이 대두되고 있다. 검색 시스템들은 주로 텍스트 주석 등을 검색에 이용하였지만 이러한 방식은 멀티미디어 정보를 검색하는데 비효율적이며 텍스트 주석 작성에 많은 시간이 소요된다. 때문에 근래 들어 자신이 원하는 그림, 영화의 한 장면, 특정 음악의 일부나, 효과음 등으로 질의하는 내용기반 검색 시스템에 대해 많은 연구가 이루어지고 있다. 멀티미디어 데이터가 지니는 특징을 이용하여 원하는 데이터를 검색한다는 것이 이러한 시스템의 특징이다. 이러한 이유로 최근 "Multimedia Content Description Interface" 이하, MPEG-7 이라는 불리는 표준안이 제정되었다. 이는 멀티미디어 콘텐츠를 표현하는 방식을 표준화함으로써 멀티미디어 시스템의 기능을 향상시키고자 하는데 그 목적이 있다[1]. 본 논문에서 MPEG-7의 움직임 기술자의 한 종류인 활동도 기술자를 이용한 내용기반 비디오 요약 기술을 제안하고자 한다.

비디오에 대한 내용기반 검색을 위해서는 동영상 데이터를 색인하기 위한 비디오 파싱 기법과 사용자

가 원하는 데이터를 쉽게 검색할 수 있는 사용자 인터페이스, 제한된 저장 공간에 대용량 비디오를 효율적으로 저장하기 위한 비디오 자료 압축 및 저장 방법 등의 기술들이 필요하다[2][3].

비디오에 대한 내용기반 검색기법은 영상정보, 문자정보, 오디오 정보 중 어떤 정보들을 사용하는가에 따라 여러 가지로 분류할 수 있으며 비디오를 요약하기 위해서 가장 일반적으로 사용할 수 있는 정보는 영상정보이다. 영상정보는 주로 비디오를 장면 분할할 때 사용되며 이를 통하여 구조적인 비디오 브라우징이 가능하다. 특히, 비디오 압축기술이 발전하면서 압축상태의 자료를 이용한 비디오 검색, 요약, 브라우징 기술에 대해 많은 연구가 이루어지고 있다[4][5]. 본 논문에서는 압축상태의 비디오 자료에서 움직임 벡터를 추출하여 비디오 요약에 이용하고자 한다.

본 논문의 2장에서는 MPEG-7의 움직임 기술자와 압축상태의 비디오 자료에서 움직임 벡터 추출 방법을 소개한다. 3장에서는 본 논문에서 제안하는 활동도 기술자를 설명하고 4장에서는 활동도 기술자를 이용한 실험결과를 5장에서는 결론과 향후 연구 과제를 기술한다.

2. 연구 배경

MPEG-7에서는 동영상을 위한 여러 가지 기술자를 정의하고 있다 그 중 활동도 기술자는 어떤 장면에서 나타나는 액션의 강도 또는 속도를 표현한다. MPEG-7에서 정의한 활동도 기술자는 다음과 같이 4개의 파라미터들로 표현된다[6].

- 1) 활동도의 강도 : 3비트의 정수형 파라미터로 표현되며 활동도의 강도를 나타낸다.
- 2) 활동도의 방향 : 3비트의 정수형 파라미터로 표현되며 액션의 주 방향을 나타낸다.
- 3) 공간적 분포도 : 세 개의 16비트 정수들로 표현되며 활동도의 공간적 분포를 나타낸다.
- 4) 시간적 분포도 : 다섯 개의 6비트 정수들로 표현되며 활동도의 시간적 분포를 나타낸다.

MPEG-7에서는 기술자에 대해 표준화하였지만 기술자에 사용되는 특징을 어떻게 추출하는가에 대해서는 응용에 따라 적절한 알고리즘이나 기술을 사용할 것을 권고하고 있다.

비디오를 구성하는 최소단위는 필름 한 장에 해당하는 프레임이다. 이 프레임들간의 차이를 비교할 수 있어야만 비디오를 분석할 수 있다. 비디오에서 장면의 전환이 이루어지는 부분을 컷(cut)이라고 하고, 하나의 카메라 동작에 의해 촬영된 작은 비디오 단위를 샷(shot)라 한다. 비디오를 샷으로 구분하는 작업을 비디오 분할이라고 하며, 비디오 분할을 위해 장면의 전환점인 컷을 검출하는 작업을 컷 검출이라고 한다. 비디오는 연속된 프레임의 집합이므로 연속된 장면에서는 인접한 프레임 사이의 유사성이 강하고 장면의 전환이 이루어지는 부분에서는 프레임 사이의 유사성이 상대적으로 약하다. 따라서 컷을 검출하기 위해서는 비디오 요소의 프레임간의 차이를 이용하여 그 요소의 연속성을 계산하고 불연속 지점을 컷으로 간주한다. 지금까지 컷 검출을 위한 다양한 알고리즘이 연구되었다. 자동으로 컷을 검출하기 위한 방법으로는 히스토그램의 차이 비교, 화소간의 차이 비교, 에지 변화, 비교, 압축 상관 계수 비교, 유사율 측정법, 그리고 움직임 벡터 비교 등이 있다.

MPEG-1/2에서는 동영상의 각 프레임에 대해 움직임 벡터를 계산하여 시간적 압축에, DCT 계수를 계산하여 공간적 압축에 이용한다. 움직임 벡터는 프레임 사이의 상관관계를 이용해서 중복되는 화소들을 줄이기 위해서 16×16화소의 매크로 블록 단위의 움직임 예측 과정에 의해 계산된다. 예측은 프

레이미간 화소 데이터의 차이를 구하는 것으로 현재 처리하고자 하는 프레임상의 어떤 매크로 블록과 과거에 처리한 다른 프레임상의 데이터가 가장 잘 일치하는 매크로 블록을 탐색해서 일치된 매크로 블록이 어느 방향으로 얼마나 움직이는가를 움직임 벡터로 검출한다. 이러한 과정을 통해 프레임 내에서 시간적으로 중복되는 매크로 블록은 무시하고 움직임이 있는 매크로 블록의 움직임 벡터만을 코딩함으로써 시간적 압축 효과를 얻을 수 있다. 즉, 장면 내에서 움직임의 정도에 따라서 움직임 벡터의 크기와 수가 결정된다.

본 논문에서는 압축상태의 비디오 자료에서 직접 움직임 벡터를 추출하여 비디오를 이루는 각 프레임들과 비디오 세그먼트에 대해 활동도의 강도와 시간적 분포도를 계산하고자 한다. 압축상태의 비디오 자료에서 벡터를 추출하기 위해 실험에서 MPEG-1/2 라이브러리를 이용하여 압축 파일을 복호화 하면서 동시에 각 프레임들의 움직임 벡터를 추출하였다.

3. 활동도 기술자

1) 활동도의 강도

활동도의 강도는 어떤 장면에서 움직임이 얼마나 많이 일어나는가를 표현하며 0에서 5의 레벨을 지닌다. 예를 들어 자동차가 빠른 속도로 지나가는 장면은 레벨 5를, 아나운서가 인터뷰하고 있는 장면은 상대적으로 낮은 레벨 0에 가까운 활동도를 나타낼 것이다.

$$IntOfAct = \frac{TotalVec}{NumOfNotSkip} \quad \text{식(1)}$$

식(1)은 활동도의 강도를 구하는 식이다.

$TotalVec$ 은 프레임 내의 모든 움직임 벡터의 크기를 누적한 값이며 $NumOfNotSkip$ 은 프레임 내에서 코딩된 매크로 블록의 수를 나타낸다. MPEG-1/2의 움직임 예측과정에서 움직임이 없는 매크로 블록은 코딩되지 않으며 무시된다. 반면에 움직임이 있는 매크로 블록은 움직임 벡터를 계산, 코딩된다. 누적된 벡터의 값을 코딩된 매크로 블록의 수로 나눈 값을 활동도의 강도로 이용한다. MPEG-7에서는 활동도의 강도를 3비트의 정수형 파라미터로 정의하였다. 때문에 위해 본 논문에서는 P 프레임에서 계산된 활동도의 강도를 6가지의 값으로 양자화 하였다. I 프레임은 움직임 예측 과정을 거치

지 않기 때문에 사용하지 않았으며 B 프레임의 벡터는 프레임 사이의 보간 성격을 지니는 벡터이기 때문에 활동도의 강도 계산에 부적합, 사용하지 않았다.

2) 활동도의 시간적 분포도

비디오가 세그먼트로 분할 되어있고 각 세그먼트는 수 개의 프레임으로 이루어져 있다고 가정해보자 활동도의 시간적 분포도는 비디오 세그먼트에 대한 활동도의 지속성을 나타낸다. Sun은 프레임에 따른 시간적 분포도를 얻기 위해 프레임의 전체 매크로수 대 No_MC 매크로 블록수의 비율을 계산하였다 [7]. No_MC 매크로 블록은 보간 과정을 거치지 않은 매크로 블록을 의미한다. 즉, No_MC 매크로 블록은 움직임 예측 과정에서 코딩되지 않고 무시된 블록을 의미하므로 그 수는 프레임에서 일어난 움직임의 정도에 비례한다. 각 비디오 세그먼트에 대한 위 비율의 히스토그램을 작성하여 비디오 세그먼트의 활동도의 시간적 분포도를 분석하였다.

본 논문에서는 활동도의 시간적 분포도를 계산하기 위해 프레임의 흐름에 따른 활동도의 강도 변화량을 이용하고자 한다.

[그림 1]은 축구 경기 비디오 파일에서 P 프레임들에 대한 활동도의 강도를 그래프로 나타낸 것이다. 그래프에서 높은 수치는 나타내는 부분은 슈팅 장면이나 선수들이 격렬하게 움직이는 장면을 보여주는 프레임의 활동도의 강도다. 프레임이 흐르면서 각 프레임 별로 변화하는 활동도의 강도를 눈으로 확인할 수 있다. 보통 격렬한 액션을 보여주는 프레임들은 이 변화가 심하다. 본 논문에서는 이 변화량을 분석하여 활동도의 시간적 분포도를 계산하고자 한다.

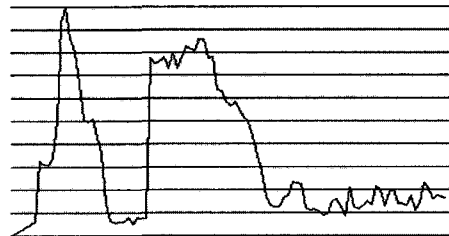
$$F(x) = \int_{-\infty}^{\infty} f(x)e^{-j2\pi mx} dx \quad \text{식(2)}$$

식(2)는 푸리에 변환 식이다. 여기서 입력 값 x는 P 프레임들의 활동도의 강도 값이다. 어떤 신호에 대해 푸리에 변환을 적용하면 그 신호에 어떤 주파수가 얼마나 나타나는지 알 수 있다. 변화가 심한 신호는 고 주파수 성분을 지녔다고 하고 반대로 변화가 적은 신호는 저 주파수 성분을 지녔다고 한다. 활동도의 시간적 분포도는 활동도가 세그먼트 내에서 얼마나 변화하는 가를 나타낸다. 즉 활동도의 변

화가 심하다면 그 변화량에 대한 주파수 성분을 분석하면 고 주파수 성분을 많이 포함하고있을 것이며 반대로 변화가 약하다면 저 주파수 성분을 많이 포함하고 있을 것이다. 때문에 [그림 1]처럼 프레임이 흐르면서 생기는 활동도의 강도에 대한 변화량에 푸리에 변환을 적용하면 각 비디오 세그먼트 내에서 활동도가 얼마나 변화하는지 알 수 있다.

$$TempDistOfAct = [Pvl, Pl, Pm, Ph, Pvh] \quad \text{식(3)}$$

식(3)은 식(2)를 이용하여 비디오 세그먼트가 지니는 강도의 변화량을 주파수 공간에서 분석하여 작성한 히스토그램을 표현한 식이다. 식에서 Pvl, Pl, Pm, Ph, Pvh은 주파수를 다섯 가지의 값으로 양자화 하여 누적한 값들이다. 만약 세그먼트 내에 활동도의 변화가 없거나 적을수록 위 다섯 가지의 값 중 저 주파수에 해당하는 Pvl,나 Pl에 많은 값이 누적될 것이며 반대로 변화가 많을수록 고 주파수에 해당하는 Ph, Pvh에 많은 값이 누적될 것이다.



[그림 1] 프레임의 흐름에 따른 활동도의 강도 변화

4. 실험

본 논문에서 제안한 방법을 펜티엄-4 PC 환경에서 VC++ 6.0 언어를 사용하여 구현하였다. 실험에 사용된 비디오는 축구 경기를 보여주며 MPEG-1의 기술로 압축되었다. MPEG-2 역시 거의 동일한 개념으로 압축을 시행하기 때문에 실험이 가능하다. 비디오는 Sun이 제안한 방법[8] 등을 참고, 장면의 전환이 이루어지는 프레임을 기준으로 분할하였다. 분할된 비디오 세그먼트는 여러 개의 프레임으로 이루어져있으며 그 중 첫 번째 P 프레임을 그 세그먼트의 대표 프레임으로 사용하였다.

[그림 2]는 질의로써 사용한 프레임으로써 슈팅 장면을 보여주고 있으며 높은 활동도의 강도를 나타내었다. 또한 선수들이 움직임이 격렬해서 질의 프레임을 포함한 비디오 세그먼트에 대한 활동도의 시간적

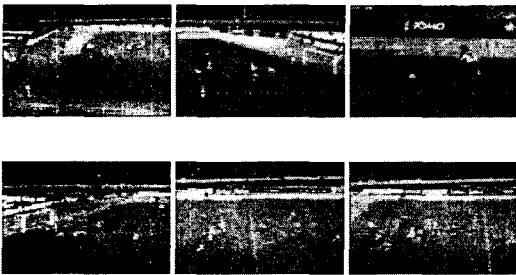
분포도를 분석하였을 때 주파수를 성분을 많이 지니고 있다는 것을 알 수 있었다.

[그림 3]은 시간적 분포도를 이용하여 검색한 결과를 순위대로 나열한 영상이다. 슷 장면이나 여러 선수가 격렬하게 움직이는 장면을 포함한 프레임이 검색되었다는 것을 알 수 있다.

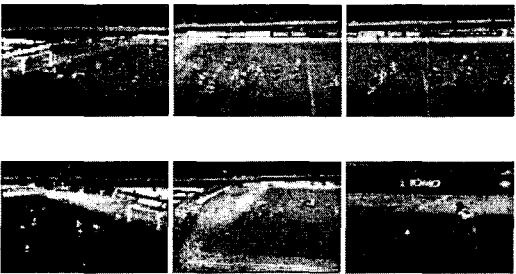
[그림 4]는 검색된 프레임들에 활동도의 강도를 이용하여 재 정렬시킨 결과를 순위대로 나열한 영상이다.



[그림 2] 질의프레임



[그림 3] 시간적 분포도를 이용한 검색결과



[그림 4] 활동도의 강도를 이용한 재 정렬 결과

5. 결론

본 논문에서는 비디오를 요약하기 위해 압축된 상태의 비디오 자료에서 추출한 벡터를 MPEG-7에서 정의하고 움직임 기술자의 한 종류인 활동도 기술자에 적용, 비디오 세그먼트의 브라우징에 이용하였다. 제안한 방법은 압축공간에서 브라우징이 이루어지기 때문에 시간이 절약되며 MPEG-1/2 비디오

의 특징을 MPEG-7의 기술자가 적절하게 이용할 수 있다는 예를 보여주었다. 향후 각 프레임이 나타나는 활동도의 방향을 계산할 수 있다면 더 좋은 결과를 얻을 수 있을 것으로 예상된다.

Acknowledgement

본 논문은 첨단정보기술연구센터(AITrc)를 통하여 과학재단의 일부를 지원 받았음

참고문헌

- [1] Requirement Group, Overview of the MPEG-7 Standard(version 6.0) ISO/IEC/JTC1/SC29/WG11/N4509 December 2001.
- [2] Wayne Niblack, "The QBIC Project: Querying Images by Content, Using Color, Texture, and Shape", In Proceedings of SPIE conference on Storage and Retrieval for Image and Video Databases, Vol 1908, pp.173-187, February, 1993.
- [3] Shih-Fu Chang, "A Fully automated content-based video search engine supporting spatiotemporal queries", IEEE Transactions On Circuits and Systems for Video Technology, Vol 8, No5 pp.602-615, September 1998.
- [4] V. Kobla, Compressed domain video indexing techniques using DCT and motion vector information in MPEG video, In Proceedings of SPIE conference on Storage and Retrieval for Image and Video Databases, Vol 3022, p p.200-211, February, 1997.
- [5] A. Divakaran, "Content-Based Browsing System for Personal Video Recorders", IEEE International Conference on Consumer Electronics (ICCE), pps 114-115, June 2002
- [6] S. Jeannin, "MPEG-7 Visual Motion Descriptors", IEEE Transactions On Circuits and Systems for Video Technology, Vol 11, No6 pp.720-724 June 2001.
- [7] X. Sun, "A Motion Activity Descriptor and Its Extraction in the Compressed Domain", IEEE Pacific-Rim Conference on Multimedia (PCM), LNCS 2195, pp. 450, October 2001
- [8] X. Sun, "Content-Based Representative Frame Extraction for Digital Video", International Conference on Multimedia Computing and System, pp. 190-194, 1998