

PVR 시스템에서 메타데이터 기반의 효과적인 브라우징 방법

김재목

삼성전자 정보통신연구소

jaemok.kim@samsung.com

A Metadata-based Effective Browsing Method for PVR System

Jaemok Kim

Samsung Electronics Telecommunication R&D Center

요 약

디지털 기술이 나날이 발전함에 따라 정지영상 압축 기술의 대표적이라 할 수 있는 JPEG이 탄생하면서 디지털 영상은 급격한 발전을 거듭해 왔다. 이에 동영상 압축 기술 MPEG이 탄생하였고, MPEG은 1, 2, 4를 비롯하여 7, 21로 발전을 거듭하고 있다.

PVR은 방송을 디지털로 저장하는 기기로서 시청자가 원하는 시간에 원하는 방송을 시청할 수 있게 해준다. PVR이 많은 콘텐츠를 저장할 수 있게 되자, 사용자는 콘텐츠에 대한 효과적인 검색 및 브라우징을 필요로 하고 있다. 본 논문에서는 콘텐츠를 기술하는 메타데이터를 이용한 의미기반 브라우징 방법을 제안한다. 제안하는 방법에서는 메타데이터에 있는 키워드와 인물정보, 장르, MPEG-7의 color descriptor를 이용한다.

1. 서 론

최근 아날로그 방송에서 디지털 방송과 케이블 방송, DMB 방송 등이 등장함에 따라 시청자는 넘쳐나는 방송 콘텐츠 속에 살고 있다. 디지털 영상기술의 발전으로 등장한 PVR(Personal Video Recorder)은 원하는 시간에 원하는 방송을 시청하는 것을 가능하게 해준다. PVR이 많은 방송 콘텐츠를 디지털 비디오 데이터로 저장할 수 있기 때문에, 사용자는 방대한 비디오 데이터를 보다 효과적으로 검색 및 브라우징을 하여 시청하는 것을 요구하고 있다. 효율적인 비디오 데이터를 브라우징하기 위해서는 방대한 양의 비디오 데이터에 대한 효율적인 관리가 필요하고, 이를 위해서는 대용량의 비디오 데이터를 체계적으로 분류하고 통합하는 기술이 필요하다[1].

비디오 데이터는 일반적인 텍스트 데이터와 달리 비디오 내에 데이터의 정보가 문자화 되어 있지 않아 비디오 데이터에 정보를 부여하는 것이 쉽지 않다. 따라서 디지털 방송은 이러한 비디오 데이터를 기술하는 메타데이터를 함께 제공한다. 비디오 데이터와 함께 제공되는 메타데이터는 제작자, 제작일자, 방송일자, 등장인물, 콘텐츠 제목과 시놉시스, 장르 등을 기본적으로 포함하고 있다. 이러한 정보를 이용하는 것은 원시 비디오 데이터를 이용하는 것 보다 검색속도에서 많은 향상을 기대할 수 있고, 메타데이터를 이용하기 때문에 사용자의 검색 목적에 부합하는 브라우징 결과를 기대할 수 있다[2].

기존 방법에는 브라우징을 위해 메타데이터보다 원시 데이터를 이용하는 방법이 제안되어 있으며, 최근 메타데이터를 일부 이용하는 방법이 제안되어 있다. 하지만, 이 방법은 원시 비디오 데이터의 의존율이 높기 때문에 많은 검색시간을 요구한다. 또한, 메타데이터를 이용한다고 해도 하나의 정보와 원시 데이터를 같이 사용하기 때문에 결국 많은 검색

색시간을 요구하고 브라우징의 결과도 좋지 못하다. 따라서 빠르고 정확한 브라우징을 위해서는 원시 비디오 데이터보다는 메타데이터에 중점을 둔 새로운 방법을 모색할 필요가 있다.

2. 관련 연구

기존의 연구에서는 비디오 데이터에 대한 내용기반 검색에 대한 연구는 크게 두 가지로 분류할 수 있다.

첫 번째로, 키 프레임에서 저차원의 특징을 추출하여 유사성을 이용하는 방법이 주를 이루고 있는 특징기반 검색(feature-based retrieval)이 있으며, 두 번째로, 키 프레임에 대해 사용자의 주석을 입력하여 저장한 후 이러한 사용자의 주석을 비교 검색하는 주석기반 검색(annotation-based retrieval)이 있다. 또한, 이 두 가지 방법을 병행해서 사용하는 방법이 있으나 단일 키워드의 사용으로 인해 의미있는 접근이 어렵고 이러한 까닭에 의미있는 브라우징에 대해서는 낮은 검색 효율을 보이고 있다.

2.1 특징기반 검색 (Feature-based Retrieval)

IBM Almaden 연구소에서 QBIC(Query by Image and Video Content)[3], 콜롬비아 대학에서 개발된 VisualSEEK[4], 대만 청화대학에서 개발한 Venus[5] 등은 질의 이미지를 이용하여 색상 및 질감 패턴, 공간 질의, 색상 히스토그램, 영상의 영역과 색상 크기 및 공간적 위치, 비디오의 각 프레임에 나타나는 객체들의 시간관계, 공간관계 등을 질의로 수행한다.

특징기반 검색은 비디오 데이터 자체로부터 그 데이터가 가지고 있는 Color, Texture, Region 정보, Spatial Color Distribution 등의 저차원 특징 정보들을 추출하여 검색하는 방법이다[6]. 따라서 특징기반 검색 방법은 비디오 자체의

시각적 특징을 추출하고, 이들에 대한 유사성을 계산 비교 검색하는 방식이므로 시각적인 특징을 추출하는 것이 매우 중요하다.

2.2 주석기반 검색 (Annotation-based Retrieval)

미국 메릴랜드 대학에서 개발한 AVIS(Advanced Video Information System)[7], 노르웨이 공대에서 개발한 VideoSTAR(Video Storage and Retrieval)[8], 일본 고베 대학에서 개발한 OVID(Object-oriented Video Information Database)[9] 등은 비디오 내용에 나타나는 객체, 사건 행위 유형, 인물, 위치, 사건과 비디오 객체라는 개념을 도입하여 각 비디오 객체마다 포함된 내용을 설명하는 속성과 속성 값을 부여하여 메타데이터를 구성하여 검색의 효율을 높이고자 하였다.

주석기반 검색은 각각의 비디오 데이터에 자동 인식이 어려운 의미 정보를 사용자가 직접 문자로 주석을 부여하여 저장한 후 검색시 미리 부여된 주석을 이용하는 방법이다 [10]. 이 방법은 사용자가 직접 비디오를 보면서 내용을 주석으로 처리할 수 있기 때문에 비디오의 내용을 정확하게 표현하고 검색할 수 있다는 장점이 있다. 그러나 각각의 비디오에 대해 문자로 사용자가 일일이 주석을 부여해야 하므로 많은 시간과 노력을 필요로 하며, 불필요하게 주석의 양이 증가할 수 있다는 단점이 있다. 또한, 주석을 입력하는 주석자 한 사람의 주관적인 의미부여로 인해 많은 사용자의 다양한 검색어로는 정확한 검색 결과를 얻을 수 없다.

3. 본 문

최근 이 두 가지 비디오 데이터의 검색 방법을 통합하여 비디오 데이터에 대한 검색 효율을 높이고자하는 연구가 진행 중에 있으나 의미부여시 키워드의 개수가 많고, 동일 의미에 대해 서로 다른 표현으로 인해 높은 검색 효율을 기대하기 힘들다. 또한, 원시 비디오 데이터에 대한 의존도가 높음에 따라 비디오 데이터가 많을수록 검색 속도는 현저히 떨어진다. 또한, 메타데이터를 고려한 방법에서도 아직까지는 장르에 대한 유사성을 고려하지 않았다. 이러한 문제를 해결하기 위하여, 본 논문에서는 유사성을 고려한 키워드와 인물정보의 개수가 한정된 콘텐츠의 메타데이터에 기반한 방법을 제안한다.

3.1 시스템 구조

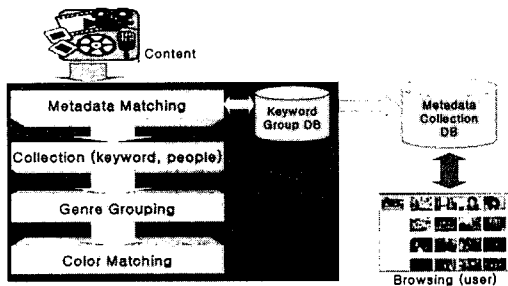


그림 1. 제안하는 시스템 구조

그림 1은 제안하는 시스템 구조로 입력되는 새로운 비디

오 데이터와 이를 기술하는 메타데이터를 이용하여 Metadata Matching, Collection, Genre Grouping, Color Matching의 4과정을 거쳐 구조화되며, 이 구조는 메타데이터 데이터베이스에 저장하여 효율적인 검색을 가능하게 한다.

3.2 메타데이터 매칭

비디오 데이터를 기술하는 메타데이터에는 비디오 데이터의 장르, 인물정보로 인물의 이름 그리고 키워드를 갖는다.

키워드는 단어로써 표현하기 때문에 의미는 같으나 표현이 달라질 수 있다. 이는 검색시 의미가 다른 키워드로 인식할 수 있다. 따라서 유사성을 고려한 키워드 인식을 위해 그림 2와 같이 데이터베이스에 키워드 그룹을 구성하고 이를 대표하는 대표 키워드를 통하여 검색에 사용함으로써 검색의 효율을 높인다.

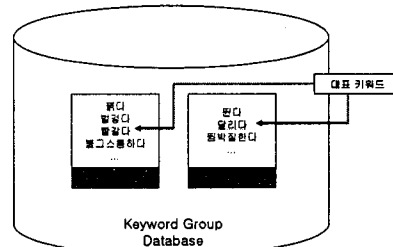


그림 2. 키워드 그룹과 대표 키워드

인물정보의 경우 인물의 고유한 이름을 사용하기 때문에 키워드 matching과 같이 유사성을 고려하기 위한 데이터베이스가 필요하지 않다. 동명이인이더라도 인물에 따른 ID를 부여한다.

3.3 Collection 구성

각 비디오 세그먼트들은 하나 이상의 키워드와 인물정보를 갖게 된다. 하나의 세그먼트가 갖는 키워드는 각각의 개별적인 동일 키워드 및 인물정보 세그먼트들의 collection을 구성한다. 따라서 하나의 비디오 세그먼트로 인해 하나 이상의 collection이 구성될 수 있으며, 기존에 존재하는 키워드 및 인물정보에 해당하는 세그먼트일 경우에는 collection 세그먼트를 추가한다.

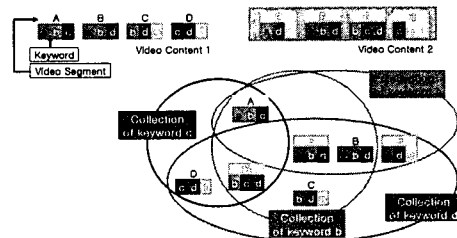


그림 3. collection 구성

그림 3과 같이 구성함으로써 하나의 세그먼트 A로 관계 있는 다른 세그먼트인 D와 G, C, F, B, E 세그먼트를 찾아 낼 수 있게 된다. 즉, 각각의 세그먼트는 키워드의 상관관계를 통해 유기적인 연결고리를 갖게 한다. 그러나 이러한 관

계를 형성하는데 있어서, 많은 키워드를 갖는 세그먼트들로 인해 유사성이 낮은 세그먼트들까지도 collection에 포함될 수 있으므로 TV-Anytime forum의 표준에 따라 3개로 제한한다[11].

3.4 장르 그룹화

기존의 연구에서는 장르별 유사성을 전혀 고려하지 않았다. 그러나 각 장르는 서로간의 유사성을 가지고 있다. 이는 다른 장르간의 비디오 컨텐츠라 할지라도 인물 또는 장면이 유사성을 가지고 있기 때문이다. 즉, 드라마와 영화 장르의 경우 각 세그먼트들 간의 유사성이 높다.

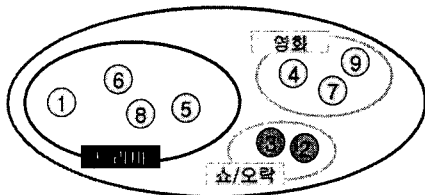


그림 4. 장르별 그룹화

그림 4와 같이 같은 키워드를 가진 collection을 구성하는 각각의 세그먼트는 기존의 연구에서와 같이 단일 키워드와 색상 유사도를 사용하면 드라마와 다큐멘터리 세그먼트가 유사성이 높은 것으로 검색될 수도 있다.

각 장르간의 유사성을 측정하기 위해서는 각 장르 그룹에 속해있는 세그먼트 개수로 결정한다. 이는 유사성을 높일수록 등장하는 빈도수가 높기 때문이다.

3.5 색상 매칭

색상 매칭 방법은 비디오 데이터의 검색 및 브라우징에 있어서 가장 널리 사용되고 있는 방법이다. 본 논문에서 사용하는 색상 매칭 방법은 MPEG-7 표준으로 채택된 Color Descriptor를 사용한다[12][13].

3.6 브라우징

사용자는 PVR에 저장되어 있는 비디오 컨텐츠를 보기 때문에 키워드와 현재 보고 있는 장면에 해당하는 비디오 세그먼트가 질의 비디오 세그먼트가 되고, 이는 데이터베이스에 저장된 해당 메타데이터 collection의 요소이다. 그렇기 때문에 질의 비디오 세그먼트에 해당하는 키워드와 인물정보에 따른 메타데이터 collection을 선정하고, 선정된 collection을 다시 하나의 collection으로 구성하여 브라우징한다. 이때, 색상 유사도를 이용하여 각 세그먼트들을 rank한다.

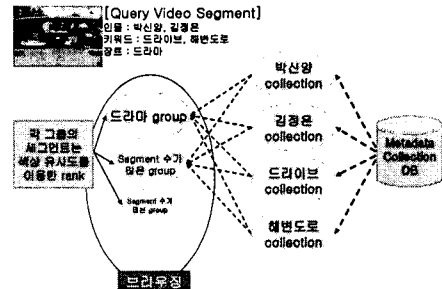


그림 5. 브라우징 방법

4. 결 론

PVR은 방송 컨텐츠를 저장할 수 있기 때문에, 메타데이터를 이용하여 검색을 수행 하므로 정확률이 높고, 검색시간이 빠르다. 그러나 비디오 세그먼트가 많아짐에 따라 collection의 개수도 많아지므로 유사도가 낮은 세그먼트도 검색될 수 있다. 따라서 color만이 아닌 다른 특징을 이용한 특징기반 검색방법이 추가되면 정확도를 더욱 높일 수 있을 것으로 기대된다.

참 고 문 헌

- [1] Sibel Adali, et. al., "The advanced Video Information System: data structure and query processing," *Multimedia System*, pp.172-186, 1996
- [2] N. Dimitrova, A. Zakhor and T. Huang, "Applications of video-content analysis and retrieval," *IEEE Multimedia*, Vol.9, No.3, pp.42-55, 2002.
- [3] Myron Flickner and et. al., "Query by Image and Video Content: The QBIC system," *IEEE Computer*, Vol.28, No.9, 1995.
- [4] J. R. Smith and S. F. Chang, "VisualSEEK: a fully automated content-based image query system," *ACM Multimedia*, Boston, 1996.
- [5] Tony C. T. Kuo and Arbee L. P. Chen, "A Content Based Query Language for Video Database," *IEEE M. M.* '96, pp.209-214, 1996
- [6] C. W. Ngo, T. C. Pong, H. J. Zhang, "Clustering and retrieval of video shots through temporal slices analysis," *IEEE Trans on Multimedia*, Vol.4, No.4, pp.446-458, 2002.
- [7] Sibel Adali, et. al., "the Advanced Video Information System: data structures and query processing," *Multimedia System*, pp.172-186, 1996.
- [8] R. Hjelmsvold, "VideoSTAR-A Database for Video Information System: data structures and query processing," *Multimedia System*, pp.172-186, 1996.
- [9] Eitetsu Oomoto, Katsumi Kanaka, "OVID: Design and Implementation of a Video Object Database System," *IEEE TKDE*, Vol.5, No.4, pp.629-643, 1993.
- [10] M. S. Kankanhalli and T. S. Chua, "Video modeling using strata-based annotation," *IEEE Multimedia*, Vol.7, No.1, pp.68-74, 2000.
- [11] TV-Anytime Forum: 2003, "Specification Series: S-3 on Metadata - Part A: Metadata Schemas," SP003v13 Part A, Version 1.3.
- [12] B. S. Manjunath, Philippe Salembier, Thomas Sikora, *Introduction to MPEG-7: Multimedia Content Description Interface*, John Wiley & Sons, 2002.
- [13] ISO/IEC 15938-8:2002, "Multimedia Content Description Interface - Part 8: Extraction and Use of MPEG-7 Descriptions," Version 1.