

비디오 데이터를 위한 색상 히스토그램 기술

이종희
(주)리테일테크
multistar@freechal.com

Color Histogram Mechanism for Video Data

Lee Jong Hee
Technical Research Institute, Retail Tech Co., LTD

요 약

본 논문에서는 사용자의 키워드 학습과 비교 영역 학습을 이용하여 대용량의 비디오 데이터에 대한 사용자의 다양한 의미검색을 지원하는 에이전트 기반에서의 자동화된 비디오 검색 시스템을 제안한다. 사용자의 기본적인 질의와 질의에 의해 추출된 키 프레임의 이미지를 선택함으로써 에이전트는 추출된 키 프레임의 주석에 대한 의미를 더욱 구체화시킨다. 또한, 사용자에게 의해 선택된 키 프레임은 질의 이미지가 되어 색상 히스토그램 비교기법과 제안하는 비교 영역 학습 기법을 통해 가장 유사한 키 프레임을 검색한다.

1. 서론

현재 비디오 데이터에 대한 내용기반 검색에 대한 연구는 크게 두 가지로 분류할 수 있다. 첫째, 키 프레임에서의 특징을 추출하여 유사성을 이용한 특징기반 검색(feature-based retrieval)이 있으며, 둘째, 키 프레임에 대해 사용자의 주석을 입력하여 저장한 후 이러한 사용자의 주석을 비교 검색하는 주석기반 검색(annotation-based retrieval)으로 크게 두 가지로 분류할 수 있다. 하지만, 이 두 가지 비디오 데이터에 대한 내용기반 검색은 모두 단점을 가지고 있다[1,2].

본 논문에서는 사용자의 질의와 결과를 학습하여 비디오 데이터의 메타데이터를 지속적으로 자동 갱신하는 인텍싱 에이전트를 이용한 의미기반 비디오 검색 시스템을 제안한다.

2. 관련 연구

2.1. 특징기반 비디오 검색 기법

콜롬비아 대학에서 개발된 VisualSEEK[6]는 색상과 공간 질의를 지원하는 이미지 데이터베이스 시스템으로서 이미지의 구분은 색상과 히스토그램과 같은 특성에 의해 이루어지며, 이미지 비교를 위해 이미지의 영역과 색상, 크기 및 공간적 위치들을 추가적으로 사용하는 시스템이다.

2.2. 주석기반 비디오 검색기법

노르웨이 공대에서 개발한 VideoSTAR[9]는 관계형 데이터베이스 모델을 사용한 데이터베이스 시스템으로서 구조화된 비디오 데이터에 대해 인물, 위치, 사건 등의 속성으로 메타데이터를 구성하고, 이를 다시 기본, 1차, 2차 컨텍스트(context)로 구분하여 메타 데이터의 재사용과 공유를 용이하게 하며, 고정된 속성을 통해 사용자가 쉽게 질의를 구성할 수 있다.

2.3. 기존 연구의 문제점

특징기반 검색에 있어서 카메라 기법이나 장면의 변환이 거의 없는 경우 샷의 경계를 정확하게 검출하는 것은 매우 어려운 작업이며, 단일 스토리를 구성하는 장면의 경계 검출은 더욱 힘들다. 기존의 많은 연구에서 비디오 프레임에 대한 색상 히스토그램과 대표 색상을 기반으로 장면과 샷의 경계 검출이 진행되어 왔지만 사용자의 의미기반 검색을 완전히 지원하지는 못한다.

3. 의미기반 비디오 검색 시스템

3.1. 전체 시스템 구조

추출된 키 프레임은 주석자에 의해 가장 기본적인 하나의 단어로 주석이 입력되고 각 키 프레임과 기본적인 주석이 메타데이터로 저장된다. 사용자의 비

디오 의미기반 검색 질의가 전달되면 사용자 질의에서 필요한 키워드를 추출하여 주식 데이터베이스에 있는 프레임 정보와 비교한 후, 유사한 후보 키 프레임 리스트를 생성하여 사용자에게 1차적으로 디스플레이한다. 디스플레이된 후보 키 프레임 리스트 중 사용자가 선택한 특정 키 프레임은 멀티분할 히스토그램 비교법을 이용하여 특징기반 검색을 실시한 후 가장 유사한 키 프레임을 검출하여 사용자에게 2차적으로 디스플레이한다.

3.2. 키워드 학습에 의한 자동 주식 갱신 기법

키 프레임에서의 주식 키워드가 동일 키워드일 경우 새로운 의미 가중치는 다음과 같이 계산한다.

$$W_{Keyword_new} = W_{Keyword_old} + \frac{1}{N_{Kframe_SK}} \quad (1)$$

키 프레임에서의 주식 키워드가 비동일 키워드일 경우 새로운 의미 가중치는 다음과 같이 계산한다.

$$W_{Keyword_new} = W_{Keyword_old} - \frac{1}{N_{Kframe_SK}} \quad (2)$$

의미 가중치의 초기값은 0으로 하여 의미 가중치가 -1의 값을 갖게 되는 키워드는 의미 없는 키워드로 에이전트가 판단하여 해당 키 프레임에 대한 주식 정보에서 삭제한다. 따라서 사용자의 잘못된 입력한 키워드나 의미 부여가 잘못된 키워드에 대해 해당 키워드를 삭제시킬 수 있는 결과를 도출할 수 있다.

3.3. 키 프레임 특징처리 기법

3.3.1. 색상 히스토그램 비교

질의 이미지를 9개의 그룹으로 분할 한 후 질의 이미지 전체에 대한 R, G, B 평균값과 각 그룹의 R, G, B 평균값을 이용하여 멀티분할 색상 히스토그램 기법을 이용하여 질의 이미지와 1차적인 주식기반 검색에 의해 추출된 모든 키 프레임들의 색상 유사도를 계산한다.

3.3.2. 비교영역 학습

비교 영역의 색상 차이율을 계산하기 위한 색상 차이 가중치 계산은 다음과 같다.

- $W_{diff_color[n]}$: 비교 영역의 색상 차이 가중치
- $Rank$: 1부터 9까지의 정수
- $R_{[n]}$: 각 영역의 순위 값

- $A_{W_new[n]}$: 현재의 $W_{diff_color[n]}$ 의 누적의 평균
- $A_{W_old[n]}$: 이전의 $W_{diff_color[n]}$ 의 누적의 평균
- β : 비교 영역 추출 임계값 (0.5)
- N : 멀티 분할된 비교 영역 수

질의 이미지와 비교되는 키 프레임 이미지들의 9개의 비교 영역별 색상 히스토그램 값의 차를 각각 구한 후 값의 차가 적은 순으로 1부터 9까지의 순위 값을 부여한다. 다음 식에 의해 부여된 순위 값을 이용하여 비교 영역의 색상 차이 가중치를 구한다.

$$W_{diff_color[n]} = 1 - \frac{R_{[n]} - 1}{N} \quad (R_{[n]} | R_{[n]} \in Rank) \quad (3)$$

색상 차이 가중치를 사용자 검색이 이뤄질 때마다 지속적으로 누적 평균을 구한다. 따라서, 새로운 $W_{diff_color[n]}$ 의 누적 평균인 $A_{W_new[n]}$ 는 이전까지의 누적평균에 현재의 비교 영역 색상 차이 가중치를 합하여 다음과 같이 구한다.

$$A_{W_new[n]} = Avg[A_{W_old[n]} + W_{diff_color[n]}] \quad (4)$$

(단, $A_{W_new[n]} \geq \beta$)

0에서 1까지의 실수 값을 갖게 되는 색상 차이 가중치의 누적 평균값에서 임계값 $\beta(=0.5)$ 이상의 값을 가진 영역만을 선택하여 이를 최적 비교 영역으로 정한다.

4. 구현 및 실험 평가

4.1. 시스템 구현

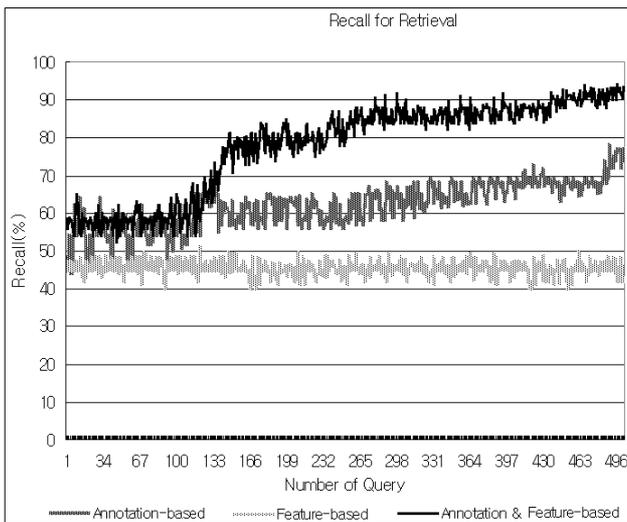
특징기반 검색을 수행하게 되면 질의 이미지에 대한 멀티 분할 색상 히스토그램 기법에 의해 유사도 계산을 하여 가장 유사한 키 프레임순으로 리스트가 다시 디스플레이된다. 그림 1은 특징기반 검색을 위한 사용자 인터페이스를 나타내고 있다.



[그림 1] 특징기반 검색

4.2. 실험 평가

제안하는 시스템의 실험 평가를 위한 비디오 데이터 도메인은 MPEG 포맷인 일반 국내외 영화 비디오 파일을 이용하였다. 실험 데이터 수는 국내영화 14편, 외국영화 16편으로 하여 총 30개의 영화를 대상으로 하였으며, 해당 비디오 클립 파일은 총 37개이다. 제안하는 시스템의 검색 적합성을 평가하기 위하여 재현율(Recall)과 정확율(Precision)에 의한 성능평가 척도를 이용하여 실험을 하였다. 그림 2는 시스템의 검색에 대한 재현율을 나타낸 것이다.



[그림 2] 제안 시스템의 검색 재현율

5. 결론

본 논문에서는 대용량의 비디오 데이터에 대한 주석기반 검색과 특징기반 검색을 이용하여 사용자의 다양한 의미검색을 지원하는 에이전트 기반에서의 통합된 비디오 내용 검색 시스템을 제안하고 설계하였

다. 또한, 실험을 위해 시스템을 구현하여 검색에 대한 재현율과 정확도를 측정하였다. 사용자의 1차적인 키워드 질의를 지속적으로 학습하고 질의에 의해 추출된 키 프레임의 이미지를 선택함으로써 에이전트는 추출된 키 프레임의 주석에 대한 의미를 더욱 구체화시킬 수 있다. 또한, 사용자에게 선택된 키 프레임은 색상 히스토그램 기법과 제안하는 영역비교 학습을 이용한 특징기반 검색기법을 통해 모든 비디오 데이터에서 가장 유사한 키 프레임을 비교하여 원하는 장면을 검색할 수 있다.

참고문헌

- [1] Nevenka Dimitrova, Avidesh Zakhor, and Thomas. Huang, "Applications of Video-content Analysis and Retrieval," IEEE Multimedia, Vol. 9 No. 3 pp. 42-55, 2002.
- [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] C. W. Ngo, T. C. Pong and H. J. Zhang, "Clustering and retrieval of video shots through temporal slices analysis," IEEE Trans. on Multimedia, Vol.04, No.04, pp.446-458, 2002.
- [4] M. S. Kankanhalli and T. S. Chua, "Video Modeling using Strata-based Annotation," IEEE Multimedia, Vol. 7, No. 1, pp. 68-74, 2000.
- [5] Myron Flickner and et. al, "Query by Image and Video Content : The QBIC system," IEEE Computer, Vol. 28, No. 9, pp. 23-32, 1995.
- [6] J. R. Smith and S. F. Chang, "VisualSEEK : A Fully Automated Content-based Image Query System," ACM Multimedia Conference, Boston, 1996.