

레터논문-10-15-1-13

소셜 북마크의 시간 정보 클러스터링을 이용한 비디오 클립 생성 자동화

한성희^{a)‡}, 이재호^{a)}, 강대갑^{a)}

Automated Video Clip Creation Using Time-based Social Bookmark Clustering

Sunghee Han^{a)‡}, Jaeho Lee^{a)}, and Daekap Kang^{a)}

요 약

최근 콘텐츠의 소비 방식 변화는 비디오 클립이라는 형식과 소셜 비디오 공유 플랫폼의 활성화를 야기했다. 이러한 비디오 클립 공급의 자동화를 위하여 여러 가지 방법이 시도되고 있다. 본 논문에서는 콘텐츠 자체의 특성에 기인한 방법이 아닌 집단 지성에 의한 북마크 데이터의 클러스터링을 통하여 효과적인 클립을 획득하는 방법을 제안한다. 사용자에게 의한 북마크 데이터를 2차원 평면의 점으로 표현한 뒤, 1차원의 북마크 누적 횟수 그래프를 이용하여 분할 클러스터링을 하는 방법은 콘텐츠 특성에 대한 이해 없이도 효과적인 하이라이트 추출을 가능하게 한다. 제시하는 실험 결과는 이 방법의 유용함을 보여준다.

Abstract

Recently the change of content consumption trend activated the social video sharing platform and the video clip itself. There have been intensive interests and efforts to automatically abstract compact and meaningful video clips. In this paper, we propose a method which use the clustering of the bookmark data created by collective intelligence instead of using the video content analysis. The partitioned clustering of points in 2-dimensional space derived from the bookmark data make it possible to abstract highlights effectively. The method is enhanced by the 1-dimensional accumulated bookmark count graph. Experiments on the real data from KBS internet service show the effectiveness of the proposed method.

Keyword : Social video bookmark, Video clip, Clustering

1. 서 론

디지털 방송 서비스와 비디오 정보의 폭발적 증가는 콘

텐츠 소비 형식의 변화를 불러 왔다. 방대한 콘텐츠들 속에서 한 콘텐츠의 전체가 아닌 핵심 부분(highlights) 또는 몇 개의 중요 구간의 조합들(video summary)만의 소비는 비디오 클립 서비스를 활성화시켰고, 비디오 클립들 중에서 유사한 것들을 묶어주고, 필요한 클립만 검색해 주는 기술은 중요한 이슈가 되었다.

이와 관련된 방법론들은 주로 한 콘텐츠의 프레임, 샷, 클립 단위의 데이터의 유사도 분석을 키로 하여 영상 처리,

a) 한국방송 방송기술연구소

Broadcast Technical Research Institute, KBS

‡ 교신저자 : 한성희 (shhan9@kbs.co.kr)

※본 연구는 지식경제부 및 방송통신위원회 공동의 정보통신 미디어 산업기술개발사업의 일환으로 수행한 연구로부터 도출되었음. [2009-S-016-01, DTV 기반의 크로스미디어 서비스 송출플랫폼 기술개발]
· 접수일(2009년11월19일), 수정일(2010년1월29일), 게재확정일(2010년1월29일)

통계, 최적화의 관점에서 전개되어 왔다^{[1][2]}.

반면 콘텐츠의 본질과는 관계 없이 생성된 데이터를 자동화된 비디오 클립 추출이라는 동일 목적으로 적용을 위해 이용할 수 있다. 웹에서의 집단 지성(collective intelligence)에 의한 데이터는 이용 방법에 따라 검색이나 추천 시스템과 같은 사례에서 매우 효율적이고 지능적인 결과를 보일 수 있다.

본 논문에서는 웹 비디오 서비스 분야에서의 소셜 비디오 북마크 (social video bookmark) 데이터를 비디오 클립 생성에 적용하여 간결하고 의미 있는 구간을 추출하는 것을 구현한다. 소셜 비디오 북마크는 서비스 이용자들이 시청한 비디오의 시작점과 끝점을 지정하여 다른 이용자와 공유함으로써 비디오 소비의 방법과 기회의 다양화를 제공하는 서비스이다.

II. 제안하는 방법

1. 소셜 북마크 데이터의 시간 정보 표현

비디오 북마크 데이터는 프로그램 아이디, 작성자, 구간 시작 시점, 구간 종료 시점, 길이, 제목, 태그, 열람 횟수 등 여러 가지로 이루어질 수 있는데, 간결하지만 의미 있는 클러스터링을 위하여 [구간 시작 시점, 구간 종료 시점]의 2차원의 데이터만을 대상으로 한다. 더 많은 항목을 사용하는 것은 적용하는 방식에 따라 더 정확한 결과를 얻는 데 도움이 될 수 있으나 차원 축소(dimension reduction)의 문제를 야기할 수도 있다^[3].

북마크 시간 정보 BM 과 해당 시간에 누적된 북마크의 횟수인 Acc 는 아래와 같이 표현된다. 북마크 누적 횟수는 북마크 시간 정보를 처리할 때 이용하기 위해 생성한다.

$$BM(i) = (x_i, y_i) = (bookmark[i][t_{start}], bookmark[i][t_{end}]) \quad (1)$$

$$Acc(n) = \sum_{i=1}^{\max(x_i)} \left(\sum_{n=0}^{\max(y_i)} 1, \text{ if } (x_i \leq n \leq y_i) \right) \quad (2)$$

이 때, i 는 한 콘텐츠 내에서의 북마크 인덱스 번호를 나

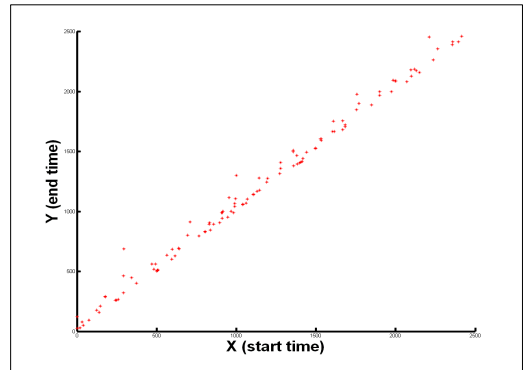


그림 1. 북마크 시간 정보
Fig. 1. Bookmark dataset

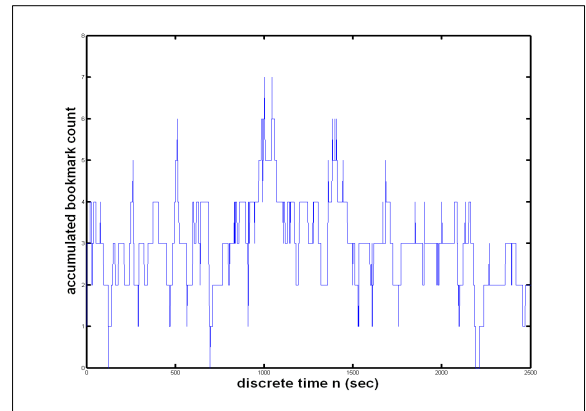


그림 2. 북마크 누적 횟수
Fig. 2. Accumulated bookmark count

타내고, x_i 와 y_i 는 각각 i 번째 북마크의 시작 시간과 종료 시간으로 동일한 초 단위의 값이 2차원 평면에 적용된 값이다. 2차원의 이 데이터는 가로와 세로가 동일한 단위로, 유클리디안 거리(euclidean distance)를 근집회를 위한 유사도 비교 지표(similarity score)로 사용하기에 적합하다. n 은 시간에 대한 북마크 누적 횟수 정보의 인덱스로 초 단위의 이산(discrete) 시간 값을 나타낸다.

2. 시간 데이터셋의 클러스터링

북마크 누적 횟수 $Acc(n)$ 는 시간 정보 $BM(i)$ 보다 이용자들의 선호도를 가시적으로 보여준다. 사용자의 북마크가

존재하고 있는 지점이 그렇지 않은 부분보다 선호도가 높다는 가정은 집단 지성에 대한 신뢰를 바탕으로 한다.

그렇지만 누적 횟수는 대략적인 피크 지점에 대한 정보를 제공하는 반면, 시간 구간의 경계나 크기에 대해서는 정확한 정보를 제공해 줄 수 없다. 따라서 주어진 시간 정보를 종합적으로 이해하기 위해서는 시간 데이터셋의 분석이 필요하다. 북마크 데이터셋의 점들을 클러스터링의 대상 항목으로 보고 북마크 누적 횟수를 클러스터링에 필요한 정보를 추출하는 데 이용한다.

클러스터링 기법^[4]은 주어진 데이터셋의 성격에 가장 잘 부합하는 것을 선택해야 좋은 결과를 얻을 수 있다. 데이터를 뚜렷한 그룹으로 쪼갤 수 있고, 2차원의 점들을 유클리디안 거리 유사도를 이용하기 위해서는 분할(partitional) 클러스터링을 사용한다. K-means는 연산 속도는 빠르나 초기 알고리즘 시작점(seed)의 결정에 따라 연산 결과가 달라지며, 군집의 개수를 사전에 결정해 줘야 한다는 단점이 있다.

본 논문에서는 K-means를 이용하되, 시작점과 군집의 개수를 결정하는데 북마크 누적 횟수를 이용하여 언제나 동일하며 정확하게 비디오 클립을 추출할 수 있는 군집을 도출해 낸다.

3. 북마크 누적 횟수 그래프의 반복적 국소 최대점을 이용한 클러스터링

북마크 누적 횟수 그래프의 국소 최대점(local maximum)은 사용자의 북마크가 갑자기 상승한 주요 지점일 확률이 높다. 그리고 그 국소 최대점이 얼마나 존재하는지에 따라 전체 비디오를 몇 개의 군집으로 나눌 수 있을 지 결정지을 수 있다. 그러나 해당 그래프는 매우 불규칙하고 많은 리플(ripple)들로 구성되어 있어서 의미 있는 국소 최대점을 찾아야 한다

반복적 국소 최대점 추출은 의미 없는 리플을 제거하고 참고해야 할 중요 국소 최대점을 도출해 준다.

$$P_0(i) = Acc(i), \quad Q_0(i) = i, \quad 0 < i \leq \max(t_{end}) \quad (3)$$

$$PQ(0) = ((P_0(1), Q_0(1)), ((P_0(2), Q_0(2)), \dots, (P_0(i), Q_0(i))) \quad (4)$$

$$PQ(k+1) = f(PQ(k)), \quad f(PQ(k)) = (localmax(P_k), \in dex(P_k)) \quad (5)$$

여기서 P_k 와 Q_k 는 각각 k 번째 반복에서의 북마크 누적 그래프의 값들과 해당 값에 대응하는 시간축의 인덱스이다. 그리고 $localmax()$ 는 주어진 값들 중에서 국소 최대값을 찾는 함수이며 $\in dex()$ 는 도출된 국소 최대값이 원 그래프에서 가지는 인덱스 위치를 찾아주는 함수이다. 식 (5)는 반복적 국소 최대점의 도출 과정으로 바로 이전 단계의 국소 최대점들을 대상으로 새로운 국소 최대점과 그 시간 위치값을 찾아낸다.

위 과정을 국소 최대점의 개수가 임계점 이하로 떨어지면 반복을 중지한다. 반복을 중지할 임계점은 최소 클러스터의 수로 미리 지정한다. 생성될 비디오 클립의 최대 시간(t_c)을 적용하여 $n \leq \frac{\max(t_{end})}{t_c}$ 일 때까지 국소 최대점 추출을 반복한다.

Algorithm step

- 1: 북마크 데이터에서 시간 정보와 북마크 누적 횟수 생성
 - 2: 북마크 누적 횟수로부터 임계점 이하의 국소 최대점 추출
 - 3: 국소 최대점의 수를 클러스터의 개수 k 로, 국소 최대점의 위치값을 시작 시간으로 가지는 북마크를 연산 초기 중심점으로 할당
 - 4: 북마크 시간 정보의 각 점을 유클리디안 거리가 최소가 되는 중심점으로 재설정
 - 5: 한 중심점으로 수렴된 점들로부터 새로운 중심점 계산
 - 6: 4~5 과정을 더 이상 군집의 변화가 없을 때까지 반복
 - 7: 군집 중 가장 밀집 지역의 중심점 순서로 비디오 클립 생성
-

III. 실험 및 결과

실험은 KBS 인터넷 홈페이지에서 서비스 중인 짬(zzim)의 데이터를 이용한다. 짬은 이용자가 마음에 드는 구간을 스크랩(scrap)하여 게시하면 다른 이용자들이 그 부분을 열람할 수 있는 비디오 북마크 서비스이다.

본 실험에서는 많은 데이터를 보유하고 있는 드라마 ‘꽃보다 남자’의 데이터를 사용한다. 줌 서비스는 디폴트 2분, 최대 3분으로 설정되어 있으며 화살표로 길이 조절이 가능 한데 사용자들이 거의 조절을 하지 않아서 줌 시간 데이터를 2차원에 도식해 보면 1시간 정도 사이에 2~3만개의 점이 $x \leq y \leq x+120$ 의 영역 사이에 뾰뾰하게 존재하여 거의 직선처럼 보인다. (이것은 줌 데이터만의 특수한 경우이다.)

최소 의미를 파악할 수 있는 국소 최대점을 표시한 북마크 누적 횟수 그래프와 최종 중심점을 표현한 그래프는 그림 3과 같다. t_c 는 120초로 했을 때 최대 중심점 개수는 31.9 개이다.

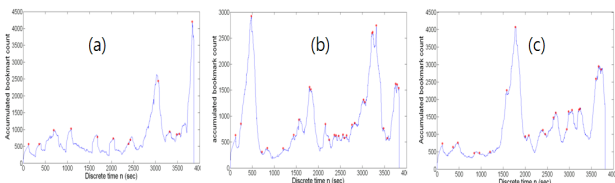


그림 3. 북마크 누적 횟수 그래프를 통한 클러스터링 초기점의 위치와 개수: (a) 꽃남 9회, 13개 초기점, (b)꽃남 12회, 30개 초기점, (c)꽃남 25회, 22개 초기점

Fig. 3. Clustering seed by accumulated bookmark count: (a) F4 episode9, 13 seeds, (b) F4 episode12, 30 seeds, (c) F4 episode 25, 22 seeds.

12회(그림 3(b))에 대하여 클러스터링 결과를 이용하여

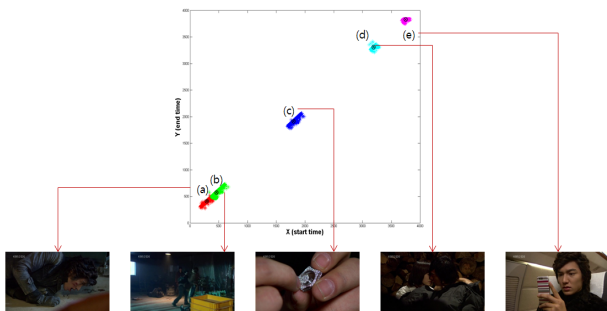


그림 4. 클립 데이터 상위 5개의 최고 조밀 클러스터, 중심⊗, (a) 폭행당하는 Joonpyo (b) F4의 복수 (c) 목걸이 선물 (d)금잔디와 Joonpyo의 키스인 (e)출국하는 Joonpyo

Fig. 4. The most populated Top 5 clusters with center⊗ (a) outraged Joonpyo (b) revenge of F4 (c) necklace present (d) kiss of Joonpyo and Jandi (e) leaving Korea

만들어지는 비디오 클립에 대한 도시는 그림 4와 같다. 상위 5개의 최고 조밀 영역은 그림 3에서 표시된 북마크 누적 횟수의 최고 영역을 반영하여 생성된 최종 중심점으로부터 만들어진 비디오 클립이다.

생성된 비디오 클립들은 KBS 홈페이지에서 제공하는 텍스트 줄거리^[7]와의 일치도가 높으며, 가장 많이 줌이 생성된 하이라이트 부분을 빠지지 않고 중복이 없이 잘 반영하고 있다.

IV. 결론

본 논문에서는 공급자형 클립 생성 자동화를 위한 방법을 제안하였다. 콘텐츠를 제공하는 플랫폼이 활성화되었을 때 공급자는 사용자들이 생성해 놓은 많은 양의 데이터를 얻게 된다. 이 데이터에 대한 신뢰를 바탕으로 좌표 단위가 동일한 2차원의 점들로 표현된 북마크의 특성을 활용한 클러스터링을 통하여 영상의 하이라이트 부분을 추출해 낼 수 있었다. 국소 최대점을 구하는 방식의 조절, 제작 시스템의 세그먼트(segment) 정보의 반영, 영상 썬체인지 디텍션 (scene change detection) 등과 결합되면 클립 구간 설정에 있어서 보다 정확한 결과를 얻을 수 있을 것이다.

참고 문헌

- [1] Li, Y., Zhang, T., Tretter, D. "An overview of video abstraction techniques", TR HPL-2001-191, July 2001, Imaging Systems Laboratory, HP Laboratories Palo Alto.
- [2] Ba Tu Truong, Svetha Venkatesh, "Video abstraction: A systematic review and classification", ACM Transactions on Multimedia Computing, Communications, and Applications (TOMCCAP), v.3 n.1, p.3-es, February 2007.
- [3] L.-Q. Xu and B. Luo, "Appearance-based video clustering in 2D locality preserving projection subspace," Proc. of ACM CIVR 2007, July 2007.
- [4] Pavel Berkhin, "A Survey of Clustering Data Mining Techniques. Technical Report, Accrue Software, 2002.
- [5] Yue Gao, Qiong-Hai Dai, "Clip based Video Summarization and Ranking", Conference on Image and Video Retrieval, 2008.
- [6] http://www.kbs.co.kr/end_program/drama/f4/view/view/1569466_28534.html