

유사비디오 데이터의 공통 정보 마이닝을 위한 압축 기법

박동철^o, 장중혁, 이원석
연세대학교 컴퓨터과학과

Compression Methods for Knowledge Discovery of Similar Video Data

DongCheol Park, JoongHyuk Chang, WonSuk Lee
Dept. of Computer Science, Yonsei University
E-mail : {roser, jhchang, leewo}@amadeus.yonsei.ac.kr

요 약

산업정보사회가 발달함에 따라 다양한 형태의 비디오 데이터들이 여러 분야에서 대량으로 생성되고 있다. 이에 따라 이들의 가공을 통해 비디오에 나타난 의미 정보를 추출하려는 다양한 접근들이 시도되고 있으며, 최근 들어 데이터 마이닝을 이용한 기법에 대한 관심들이 증대되고 있다. 그러나 기존의 연구 대상에서 비디오 데이터를 기본으로 하여 지식 정보를 추출하기 위한 시도는 시공간적으로 방대한 비디오 데이터의 특징으로 인해 소극적으로 접근되어왔다. 본 논문에서는 유사한 비디오들로부터 효과적으로 비디오 데이터를 압축하고 특성을 추출하며 클러스터링을 통하여 형태론적인 비디오 정보로부터 직접 의미 있는 패턴을 추출하는 방법을 제안한다. 이를 위해 관심 영역 제한방법, 최소 반복도 제한방법 및 키 프레임 추출 방법 등이 포함된다. 최종적으로 실험용 비디오에 대한 마이닝 결과를 생성하고 최초의 트랜잭션과의 정확도를 비교하여 본 논문에서 제시한 기법들을 검증하였다.

1. 서론

산업정보사회가 발달함에 따라 다양한 비디오 데이터들이 증가되고 이들의 가공을 통해 비디오에서 표현되는 지식 정보를 추출하려는 필요성 과 연구 또한 증대되고 있다. 비디오 데이터에 대한 기존의 데이터 마이닝 연구는 비디오의 방대한 데이터 양과 그 특성에 기반한 효과적인 마이닝 방법의 부재로 인하여 통계적인 방법 수준의 한정된 연구에 집중되어 왔다. 특히 유사한 비디오를 대상으로 하는 경우 각 비디오에서 계속적으로 반복되는 특성정보

의 추출이 더욱 용이할 수 있으나 이 부분에 대한 연구는 활발히 이루어 지지 않았다.

본 논문에서는 유사한 비디오들로부터 영상 특징을 고려하여 데이터 마이닝을 위한 효과적인 데이터 압축기법을 제시하고 이를 활용한 데이터 마이닝 시스템을 설계한다. 비디오 특징 데이터를 처리하기 위해서 수정된 클러스터링 기법을 활용하였고, 최종적으로 유사 비디오들의 공통 지식에 대한 요약된 특성 정보를 생성하였다. 이러한 특성 정보와 각 원본 비디오 데이터와 비교하여 본 논문에서 제시한

압축 방법의 유효성을 검증한다.

본 논문의 구성은 다음과 같다. 2장에서는 기존의 비디오 데이터 베이스의 특성을 이용한 관련 연구에 대해서 기술하고, 3장에서는 효과적인 데이터 마이닝을 위한 데이터 압축 및 정제 기법들을 제안한다. 4장에서는 클러스터링 방법에 대해서 다루며 5장에서 성능 비교를 위한 모의 실험 테스트결과를 비교 분석하고 마지막으로 6장에서 최종 결론을 맺는다.

2. 관련연구

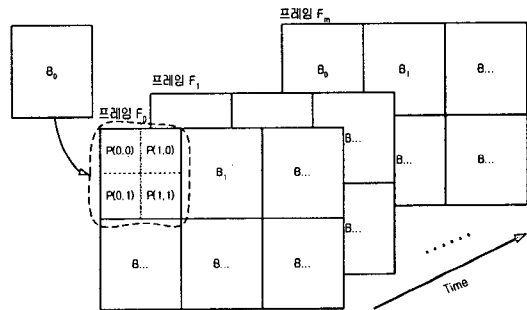
비디오에 관한 기존의 연구는 압축과 검색 등 다음의 몇 가지 영역으로 크게 분류될 수 있다. 비디오 데이터의 효과적인 저장이나 전송을 위해서 연구되어 온 비디오 압축 기법들은 비디오의 공간적 중복성을 제거하기 위한 JPEG과 화면간 부호화를 적용시킨 H.261 및 MPEG등의 표준이 발표되었다[1]. 그러나 이와 같은 압축 방법들은 저장, 전송 및 재생을 목적으로 하기 때문에 데이터 마이닝 기법을 추가적으로 사용하지 않았고 이에 효과적인 형태로 설계되지 않았다[2].

또한 비디오의 특징을 해석하는 관점에 따라 형태론적인 연구와 의미론적인 연구 측면이 있다. 형태론적인 관점으로는 비디오에 나타나는 개체들의 크기나 위치, 모양, 색상, 질감 등의 공간적인 특징과 비디오가 변하는 시간적인 특징 및 비디오 내부 개체들의 패턴 등의 변화에 대한 연구인 [3], [4], [5], [6] 등이 있다. 비디오내의 개체들의 동작과 의미 등을 분석 대상으로 하는 의미론적인 관점에서의 연구는 [7], [8] 등이 있으며, 배경 및 개체의 명확한 파악이 난해한 이유로 지금까지 소극적으로 다루어져 왔다.

공간에 분포되어 있는 데이터들의 군집성을 찾는 클러스터링에 관한 연구는 k개의 중심을 찾아 근접한 값들을 클러스터로 포함시키는 K-means 나 Forgy's 방법이 광범위하게 사용되고 있으며 대단위 데이터 용량과 신속한 처리에 적절한 BIRCH[9], DBSCAN[10] 등이 있다.

3. 데이터 압축에 관한 전처리

비디오는 일정한 개수의 픽셀로 구성되며 각 픽셀은 R, G, B등과 같은 특성정보 f 들로 구성된다. 따라서 픽셀은 하나의 특성정보 집합 $\{f_0, f_1, f_2, f_3, \dots, f_k\}$ 으로 정의될 수 있다. 사용자는 일정한 개수의 인접 픽셀들을 합병된 하나의 논리적인 블록으로 정의할 수 있고 이를 베이스 블록으로 정의한다. 정지 화면 또는 프레임은 [그림 1]과 같이 베이스 블록들의 집합이며 이러한 프레임들의 시간적인 연속 진행으로 비디오는 표현된다. 본 논문에서는 처리의 논리적 단위로써 트랜잭션을 사용하며 이것은 시작과 종료 시간 동안 단일 단위로 입력된 비디오를 의미한다. 따라서 트랜잭션 T는 [그림 1]과 같이 하나의 프레임 집합 $\{F_0, F_1, F_2, \dots, F_m\}$ 으로 정의된다.



[그림 1] 비디오 트랜잭션, 프레임, 베이스 블록 및 특성정보간의 관계

비디오는 공간적, 시간적인 중복성이 크므로 이러한 중복성을 제거하고 주요한 정보만 추출하는 경우 큰 의미적 손실이 없이 효과적인 마이닝 결과를 기대할 수 있다. 또한 주변 환경에 따른 잡음이 포함되는 경우가 많기 때문에 잡음을 제거하는 과정을 거치면 보다 정확한 특성정보를 추출할 수 있다. 본 논문에서는 상기 목적을 위해서 다음과 같은 압축 및 정제 방법을 제안한다.

3.1 관심영역 제한

트랜잭션 단위로 표현되는 비디오 데이터에서 선별적인 관심 영역만을 분석대상으로 정의하는 관심 영역 제한 방법을 고려할 수 있다. 이 방법은 데

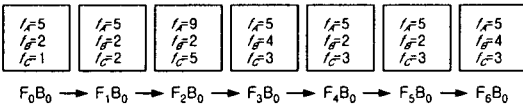
이터의 크기를 감소시킬 뿐만 아니라 비관심 영역으로부터 불필요한 특성정보가 추출되는 것을 방지하는 효과를 갖는다. 새로운 베이스블록 B'_i 는 B_i 가 관심영역에 속하는지 여부에 따라 다음과 같이 정해진다.

$$B'_i = \begin{cases} B_i, & \text{if } B_i \in \text{관심영역} \\ 0, & \text{otherwise} \end{cases}$$

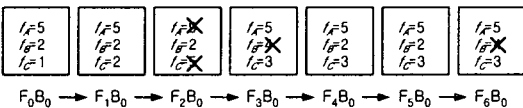
따라서 압축률은 (제거된 베이스 블록 수)*100 / (전체 베이스 블록 수)로 정의된다.

3.2 최소 반복도를 적용한 특성정보 추출

비디오 내에서 개체들이 출현한 후 비디오 내에 유지한 시간을 기준으로 중요도를 판별할 수 있다. 본 논문에서는 출현한 개체들의 특성정보들이 임계치 이내의 변화로 프레임 내에 연속적으로 유지되어야 할 최소 프레임 수를 정하고 이를 최소 반복도라고 한다. 이를 이용하여 트랜잭션에 나타나는 특정 특성정보가 최소 반복도를 만족하지 못하는 경우 그 특성정보가 제거된 로그를 생성하게 된다.



(a) 최소 반복도 적용 전



(b) 최소 반복도 적용 후

[그림 2] 최소 반복도를 적용하는 프레임 예

예를 들어 [그림 2]에서 (a)와 같이 하나의 비디오 트랜잭션이 프레임 F_0 부터 F_6 으로 구성될 때 프레임을 구성하는 베이스블록 F_0B_0 가 특성정보 f_A , f_B 및 f_C 를 갖는다고 가정하자. 이 경우 최소 반복도를 2, 임계치를 1로 두고 적용하면 F_2B_0 의 f_A 와 f_C 는 임계치 1을 초과하면서 유지되는 특성정보의 반복도는 1에 불과하므로 (b)와 같이 제거되며, F_3B_0 의 f_B 및 F_6B_0 의 f_B 역시 조건을 만족하지 못하므로 같이 제거된다.

이러한 과정을 통해 비디오 데이터의 특성 정보에 부정적인 영향을 미치는 순간적인 잡음 성분을 제거하고 최소 반복도를 만족하는 유효한 개체만 표현할 수 있으며 처리해야 할 데이터의 양을 감소시킬 수 있다.

3.3 키 프레임 추출

비디오는 전후 화면간 상관도가 높으므로 시간적으로 연결된 수많은 비디오로부터 주요한 정보를 지닌 키 프레임만을 선택하는 방법이 데이터의 양을 감소시킬 수 있다.

[정의 1] 프레임 유사도

프레임 F_i 와 F_j 간의 프레임 유사도 $\Delta(F_i \sim F_j)$ 는 다음과 같이 각 베이스 블록의 직전 프레임의 특성정보 $V_{(F-1)B_f}$ 에서 현재 프레임의 특성정보 V_{FB_f} 간 차이의 절대값을 누적한 값으로 정의된다.

$$\Delta(F_i \sim F_j) = \sum_{F=i+1}^j \sum_{B=1}^n \sum_{f=1}^k |V_{(F-1)B_f} - V_{FB_f}|$$

V_{FB_f} : 트랜잭션 내 프레임 F , 베이스 블록 B , 특성정보 f 의 값
 n : 베이스 블록의 수
 k : 특성정보의 개수

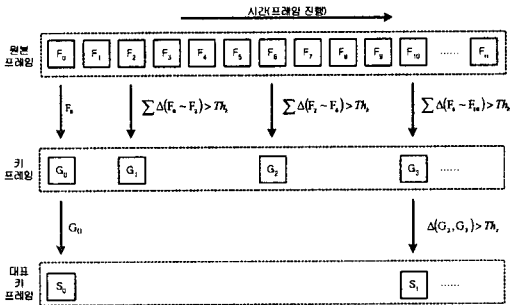
이러한 프레임 유사도는 직전 프레임으로부터 현재 프레임에 나타난 특성정보들의 변화량에 따라 비례하여 변화된다.

키 프레임을 선택하는 과정은 프레임 유사도를 기준으로 일련된 프레임의 특성정보 로그를 순차적으로 탐색하는 과정에서 프레임간의 특성정보 유사도가 임계치 Th_k 를 초과하는 순간의 프레임을 키 프레임으로 간주한다.

예를 들어 [그림 3]의 경우 키 프레임들은 $\{G_0, G_1, G_2, G_4\}$ 가 된다. 임계치 Th_k 뿐만 아니라 새로운 임계치인 Th_s 를 도입하여 대표 키 프레임을 추출할 수도 있으며, 이것은 각 키 프레임들 간의 누적된 유사도 대신에 직전 키 프레임과의 차이를 비교하여 $\Delta(F_i \sim F_{i-1}) > Th_s$ 인 경우 F_{i-1} 를 대표 키 프레임으로 간주한다.

키 프레임은 트랜잭션을 구성하는 모든 프레임

내에서 화면간의 중복성이 많은 프레임을 제거한 뒤 선택된 임계치 이상의 주요한 프레임만 선택하는 효과가 있다. 따라서 이러한 방법으로 재저장된 로그 파일은 원본 트랜잭션의 주요한 비디오 특성을 주요한 손실없이 유지한 상태로 압축된다.



[그림 3] 키 프레임 추출을 사용하는 예

4. 클러스터링

각 단계를 거친 특성정보 로그들은 클러스터링을 통하여 트랜잭션 비디오들에 대한 특성정보를 저장한 프로파일로 표현된다.

클러스터링 과정에서는 비디오 트랜잭션을 정형화한 형태의 특성정보 로그 파일들, 전체 트랜잭션 중에서 데이터가 나타난 트랜잭션 빈도를 의미하는 최소지지도 및 클러스터링 범위를 입력으로 한다.

클러스터링 후 결과 클러스터 C_i 는 클러스터 구별자 C_{id} , 베이스 블록 B_i , 특성 정보 f , 클러스터 최소값 C_{min} , 클러스터 최상값 C_{max} , 클러스터 평균값 C_{ave} 및 클러스터 지지도 C_{sup} 를 이용하여 아래와 같이 표현될 수 있다.

$$C_i = \{ C_{id}, B_i, f, C_{min}, C_{max}, C_{ave}, C_{sup} \mid 0 \leq x < W/\rho_w, 0 \leq y < H/\rho_h \}$$

최종적으로 프로파일 P는 클러스터 집합 $\{C_0, C_1, C_2, \dots, C_p\}$ 이 된다.

[정의 2] 프로파일 정확도

프로파일은 트랜잭션을 반영한 것이므로 생성된 프로파일과 클러스터링에 사용된 최초의 비디오 트랜잭션 간의 유사한 정도를 계산하여 오차를 비교할 수 있다. 프로파일 P의 클러스터들 C_{BF}^P 가 트랜잭션

T의 특성정보 V_{FB}^T 를 포함하는 정도에 따라 프로파일의 정확도 (T, P) 를 다음과 같이 정의할 수 있다.

$$\text{정확도}(T, P) = (\text{CNT}(P))^{-1} \sum_{F=1}^m \sum_{B=1}^n \sum_{f=1}^k \text{match}(V_{FB}^T, C_{BF}^P)$$

$\text{CNT}(P)$ = 프로파일 P의 클러스터 수

$$\text{match}(V_{FB}^T, C_{BF}^P) = \begin{cases} a_{sup}, & a_{min} \leq V_{FB}^T \leq a_{max} \\ 0, & \text{otherwise} \end{cases}$$

- V_{FB}^T : 트랜잭션 T, 프레임 F, 베이스 블록 B, 특성정보 f의 값
- C_{BF}^{cnt} : 프로파일내 베이스 블록 B, 특성정보 f의 클러스터 개수
- a: 클러스터
- m: 트랜잭션내 프레임 개수
- n: 프레임내 베이스 블록 개수
- k: 베이스블록내 특성정보의 개수

프로파일내의 클러스터가 트랜잭션의 특성정보를 많이 포함할수록 정확도는 증가한다.

5. 모의 실험 결과 및 분석

본 장에서는 지금까지 본 논문에서 제안한 압축 및 정제 기법을 실험을 통하여 검증한다.

5.1 모의 실험 환경

본 실험에서 사용한 비디오는 단일 색상을 가진 바닥 위에 설치된 폐쇄 레일 위를 등속으로 주행하는 모형 열차를 피사체로 하여 촬영한 것이다. 동일한 위치에서 촬영한 트랜잭션 10개를 동일한 카테고리에 포함하였으며 비디오 내부 개체들의 활성화도에 따라 [그림 4]와 같이 A-24, A-63, A-100의 3가지 카테고리로 분류하였다. 활성화도는 프레임간의 특성정보 변화량을 표현하는 프레임 유사도를 이용하여 계산하였으며 실험에 사용된 카테고리들은 A-24에서 A-63, A-100으로 갈수록 활성화도가 증가한다.

5.2 실험 결과 분석

각 카테고리별로 비디오 트랜잭션들을 클러스터링하고 결과 프로파일을 생성하였으며, 원본 트랜잭션과 각 프로파일 정확도를 비교하였다.

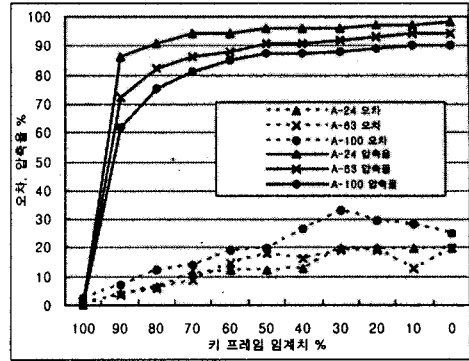
[그림 5]은 최소 반복도를 적용한 실험 결과로서 최소 반복도를 변화시키면서 각 카테고리에 적용

한 결과이다. 최소 반복도가 상향으로 조정될수록 최소 반복도를 만족하지 못하는 특성정보들의 빈도가 증가되고 동시에 베이스 블록 제거율 또한 증가되는 것을 보여준다.

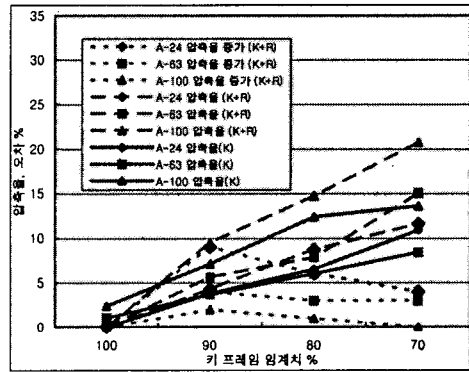
[그림 6]은 키 프레임 압축 기법을 사용한 실험 결과로서 활성화도 별로 구분된 카테고리별로 실험하였다. 활성화도가 낮을수록 프레임 압축율이 증가하는 것을 볼 수 있으며 이러한 키 프레임 압축은 큰 오차없이 효과적인 압축 성능을 보여준다.

[그림 7]는 최소 반복도 압축을 적용한 후 다음 단계인 키 프레임 압축을 연속으로 적용한 실험 결과이다. 이 실험에서 두 가지의 알고리즘을 동시에 적용하는 경우 최소 반복도에 의해서 잡음이 성공적으로 제거되고 압축율이 추가적으로 증가하지만 프로파일 오차는 거의 변화가 없음을 보여준다.

[그림 8]은 전처리 과정에서 각 활성화도 카테고리 100개의 트랜잭션을 키 프레임 압축한 경우 클러스터링 시간 단축 효과를 보여준다. 키 프레임 임계치를 감소시킬수록 제거되는 프레임수가 증가하고 로그 데이터가 감소되므로 클러스터링 시간이 급격히 감소됨을 알 수 있다.



[그림 6] 키 프레임 압축을 적용후 압축률 및 오차

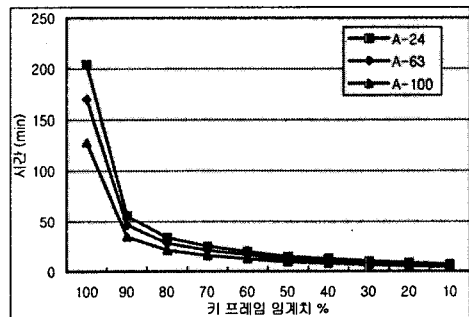


[그림 7] 최소 반복도와 키 프레임 압축을 동시에 적용한 경우 압축률 증가 및 오차

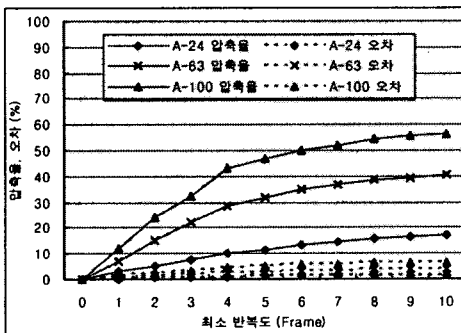


(a) A-24 (b) A-63 (c) A-100

[그림 4] 각 상대적 활성화도별 실험 비디오



[그림 8] 클러스터링 시간 단축



[그림 5] 최소 반복도를 적용 후 압축률 및 오차

6. 결론 및 향후 연구 방향

본 논문에서는 비디오 집합에서 추출된 방대한 비디오 로그로부터 데이터 마이닝을 통하여 효과적인

으로 비디오 특성을 추출하고 프로파일을 생성하는 방법에 대한 연구를 제안하였다. 이로써 형태론적인 비디오 데이터들로부터 지식을 추출하여 의미론적인 해석이 가능하도록 하였다. 또한 클러스터링 전에 관심 영역 선택, 최소 반복도 적용 및 키 프레임 추출 기법을 사용하여 비디오 트랜잭션으로부터 불필요한 정보를 제거하는 과정을 제안하였다. 추가적으로 비디오 트랜잭션들로부터 생성된 프로파일과 원본 트랜잭션과의 정확도를 측정하는 방법을 제시하였으며, 각 기법들이 프로파일의 정확도에 주요한 영향을 주지 않고서 클러스터링의 시간을 효과적으로 감소시키고 프로파일의 정확도를 유지하는 것을 실험을 통해 확인하였다.

향후 연구과제로는 클러스터링이 가지는 각 특성정보들이 독립적이기보다는 서로 결합된 하나의 의미를 표현하는 다차원이므로 동시에 각 클러스터들이 나타나는 것을 표현할 수 있는 다차원 클러스터링이 고려되어야 할 것이다. 추가로 비디오 시퀀셜 마이닝이나 비디오 데이터에 포함된 오디오 특성을 이용하는 방법에 대해서도 연구해 볼 수 있을 것이다.

[참고문헌]

[1] J. L. Mitchell, W. B. Pennebaker, C. E. Fogg, D. J. LeGall "MPEG Video Compression Standard". Digital Multimedia Standards Series, J. L. Mitchell and W. B. Pennebaker, Editors, Chapman and Hall. New York, 1997

[2] V. Kobla, D. S. Doermann, K. I. Lin "Indexing and Retrieval of MPEG Compressed Video". *Journal of Electronic Imaging*, Vol.7(2), 1998

[3] O. Kao, G. R. Joubert "Content based Internet search engine for analysis and archival of MPEG-1 compressed newsfeeds". *Multimedia and Expo, ICME, IEEE International Conference*, Vol.2, pages 1155-1158, Vol.2, 2000

[4] Chabane Djeraba "When image indexing meets knowledge discovery". In *Proc. of the*

Int. Workshop on Multimedia Data Mining (MDM/KDD'2000), in conjunction with ACM SIGKDD conference. Boston, USA, August 20, 2000

[5] Osmar T. Zaiane, Jiawei Han, Hua Zhu "Mining Recurrent Items in Multimedia with Progressive Resolution Refinement". In *Proc. 2000 Int. Conf. on Data Engineering, ICDE'00*, 2000

[6] O. R. Zaiane, J. Han, Z. N. Li, J. Y. Chiang, S. Chee "MultiMediaMiner: A system prototype for multimedia datamining". In *Proc. ACM-SIGMOD*, Seattle, 1998

[7] A. Natsev, R. Rastogi, K. Shim WALRUS: A similarity retrieval algorithm for image databases". In *Proc. ACM-SIGMOD*, Philadelphia, 1999

[8] A. Pentland, R. W. Picard, S. Scalroff "Photobook: Content-based Manipulation of Image databases". In *International Journal of Computer Vision*, Vol.18, No 3, pages 233-254, 1996

[9] Tian Zhang, Raghu Ramakrishnan, Miron Livny BIRCH: An efficient data clustering method for very large databases". In *Proc. of the ACM SIGMOD conference on management of data*, pages 103-114, Montreal, Canada, June 1996

[10] Martin Ester, Hans Peter Kriegel, Jrg Sander, Xiaowei Xu "A Density-Based algorithm for Discovering Clusters in Large Spatial Databases". *Proc. of Int'l. Conf. Knowledge Discovery and Data Mining*, pages 226-231, 1996