

휴대폰 상에서 개인용 사진 컬렉션에 대한 자동 이벤트 군집화 방법

유정수^o 낭종호

서강대학교 컴퓨터공학과

yjs@mlneptune.sogang.ac.kr, jhnang@sogang.ac.kr

Automatic Event Clustering Method for Personal Photo Collection on Mobile Phone

JeongSoo Yu^o JongHo Nang

Department of Computer Science and Engineering, The Sogang University of Korea

1. 서 론

최근 스마트 폰의 사용이 급격히 증가함에 따라 사용자는 휴대폰에 대량의 사진을 보관하는 추세이며, 휴대폰에 저장된 사진의 관리 방법의 중요성도 증가하고 있다. 일반적으로 사용자는 휴대폰에서 촬영한 개인용 사진 컬렉션에 대하여 이벤트 기준으로 관리 및 접근하는 것을 선호한다[1]. 따라서 많은 양의 사진에 대하여 이벤트 단위로 자동으로 그룹핑하는 기능은 휴대폰용 사진 관리에 유용할 것이며, [2-6]에서 여러 이벤트 그룹핑 방법들을 제안하였다. 직관적으로, 이벤트는 촬영 시간과 가장 밀접한 관계가 있다고 볼 수 있다[5,6]. 예를 들어 사용자는 졸업식과 같은 특정 시점에서 여러 장의 사진을 촬영한다. 또한 최근의 휴대기기용 카메라들은 EXIF와 같은 메타데이터를 지원하므로, 시간 정보는 이벤트 그룹핑에 매우 유용한 정보라고 볼 수 있다. 그러나, 사용자 중심 관점에서 보았을 때, 시간 단위의 그룹핑은 몇 가지 문제점들의 해결을 요구한다. 먼저, 이벤트는 기본적으로 계층적인 구조를 갖는다. 예를 들어 여행 이벤트는 식사, 파티, 등산 등의 세부 이벤트로 구성될 수 있다. 또한 이벤트의 종류에 따라서 시간 축 상에서의 스케일이 다르다. 즉, 같은 이벤트로 포함 시켜야 할 시간 간격이 모호해진다. 따라서 컬렉션에 포함된 사진들에 대하여 비교사-클러스터링(unsupervised clustering)을 통한 이벤트 그룹핑이 효과적인 방법 중 하나라고 볼 수 있다. 그러나 k-means, mean-shift 등과 같은 대부분의 비교사-클러스터링은 기대 값 최적화에 기반하기 때문에 off-line 연산에 적합하고, 증감(incremental)적인 클러스터링이 어려우며, 수렴 시간을 예측하기 힘들고, 반복(iteration)으로 인하여 높은 연산 양을 요구한다는 제한을 갖는다. 휴대폰상의 이벤트 그룹핑은 촬영 후 즉시 그룹핑이 결정되어야 하며, 이를 위해서는 증감적이어서 하고 적은 연산량을 사용해야 한다는 점이 요구된다. 따라서 기존의 비교사-클러스터링 방법은 휴대폰 상에서의 이벤트 클러스터링에 적합하지 않다고 볼 수 있다. 비교사-클러스터링과 다른 접근 방법으로 순차 분할 방법을 들 수 있다. 순차 분할 방법은 x 와 군집 C 의 거리가 임계 값보다 적을 경우 x 를 C 에 포함시키는 단순한 연산으로써, 비교사-클러스터링의 제한들을 모두 극복하는 것이 가능하지만, 임계 값을 필요로 한다. 앞서 설명한 바와 같이 이벤트는 종류에 따라 시간 범위가 다르기 때문에 임계값은 단순히(예를 들어 1 hour) 결정되기 어렵다. 그러나 실제 세상에서 발생하는 이벤트는 분명히 통계적인 특성을 갖는다는 점을 주목할 필요가 있다. 따라서 본 논문에서는 여러 제한점을 갖는 비교사-클러스터링 방법의 대안으로서 순차 분할 방법을 사용하며, 임계값 결정을 위하여 실제 샘플에 대한 통계적 분석을 수행함으로써 효율적인 이벤트 그룹핑 방법을 제안하였다. 제안한 방법의 기본 아이디어를 간단히 설명하자면 다음과 같다. 먼저 수집된 샘플 데이터를 통하여 이벤트내의 사진들에 대한 시간 간격의 확률 분포를 추정한다. 만일 촬영된 사진의 시간 경과치가 높은 확률 구간에 속한다면 최근 이벤트에 포함시키며, 모호한 구간에 속할 경우에는 시간 정보의 신뢰도가 상대적으로 감소한다고 볼 수 있으므로 촬영 위치, 시각 특징 등을 부가적으로 고려하여 이벤트 분할 여부를 결정한다. 본 논문은 우선 이벤트 클러스터의 계층성의 고려는 향후 연구로 미루고, 먼저 단위-이벤트 레벨에서의 클러스터링 방법을 연구하였다.

2. 실험 및 분석

본 논문에서 제안한 순차 분할 기반의 이벤트 클러스터링 방법을 ARM9 800MHz 스마트폰에서 구현 및 실험하였다. 먼저 클러스터링의 정확도를 측정하기 위하여 [7]에서 제안한 PEC(Partitioning Error

Cost)를 사용하였다. PEC는 각 결과 클러스터의 에러를 과분할과 미분할로 나누어서 측정한 후 모든 클러스터에 대하여 평균값으로 클러스터링 정확도를 측정하는 방식이다. PEC측정을 위하여 40개의 검증용 컬렉션들에 대하여 k-means, SOM^{self-organized-map}과 비교 실험 하였으며, GPS의 경우 기증자들의 폰에 GPS수신 장치가 없는 경우에는 Google Picasa를 사용하여 기증자들이 직접 수동 태깅을 하였다. <그림 1>은 제안한 방법이 모든 컬렉션에 대하여 우수하였음을 나타내고 있다. <그림 2>는 각각 시간 정보와 위치 정보 및 시각 정보를 사용하는 경우에 대한 측정이며, 시간 정보에 위치 정보가 추가 될 때 향상 치가 매우 큰 것을 나타낸다.

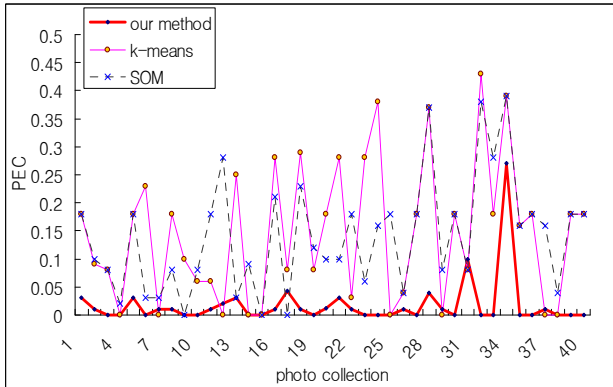


그림 1 각 클러스터링 알고리즘의 PEC 비교

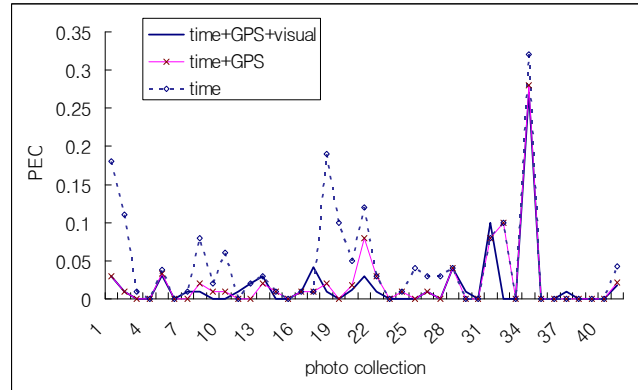


그림 2 시간, GPS, 시각 정보의 효과 비교

3. 결론

본 논문에서는 실제 휴대폰 사용자로부터 수집한 샘플 데이터에 대하여 통계적 분석을 통하여, 매우 적은 연산량을 필요로 하며 기존의 클러스터링 방법보다 높은 정확도를 유지하는 자동 이벤트 경계 검출 방법을 제안하였다. 제안한 방법은 ARM9 800MHz 스마트폰에서 사진 한 장당 평균 약0.2초의 속도로 자동 이벤트 클러스터링을 수행하였다. 본 논문은 Atomic event 클러스터링만을 대상으로 하였으며, 향후 이벤트의 중요한 특성인 계층성을 고려하여 연구를 진행할 계획이다. 또한 개인용 컬렉션을 대상으로 한다는 점에 초점을 맞추어 사용자의 컨텍스트 정보를 적극 이용하여 효율성을 높이는 방법도 고려할 계획이다.

참고 문헌

- [1] Lim J.H., Tian Q., Mulhem P., "Home photo content modeling for personalized event-based retrieval", *IEEE MultiMedia*, Vol. 10 No.4, pp.28-37, 2003.
- [2] PLATT, J., CZERWINSKI, M., AND FIELD, B, " Simplifying the management of large photo collections," In *Fourth IEEE Pacific Rim Conference on Multimedia*. IEEE Press, Los Alamitos, CA, pp. 6- 10, 2003.
- [3] Graham A., Molina H.G., Paepcke A., Winograd T., " Time as essence for photo browsing through personal digital libraries," In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital libraries*, Portland, USA, pp. 326- 335, 2002.
- [4] Naaman M., Song Y.J., Paepcke A., Molina H.G. , " Automatic organization for digital photographs with geographic coordinates," In *Proceedings of the ACM/IEEE-CS Joint Conference on Digital Libraries*, Tucson, USA, pp. 53- 62, 2004.
- [5] Cooper M., Foote J., Girgensohn A., Wilcox L. , " Temporal event clustering for digital photo collections," In *Proceedings of the ACM International Conference*, vol. 1, no. 3, pp. 269- 288, 2005.
- [6] A. Loui and A. Savakis, " Automated event clustering and quality screening of consumer pictures for digital albuming," *IEEE Transaction on Multimedia*, vol. 5, no. 3, pp. 390- 402, Jun. 2003.
- [7] Liangliang Cao, Jiebo Luo, Kautz H, Huang T.S., " Image Annotation Within the Context of Personal Photo Collections Using Hierarchical Event and Scene Models," *IEEE Transaction on Multimedia*, vol 11, no. 2, pp. 208-219, 2009.