

지리정보를 이용한 자동사진분류

홍영진¹,김성운²,유명현³,이영범⁴,김상룡⁵
삼성종합기술원¹²³⁴⁵
{bhong¹, seongwoon.kim², mh.yoo³, leey⁴, srkim⁵}@samsung.com

Automatic Photograph Classification Using Geographical Information

Youngjin Hong¹,Seongwoon Kim²,Myunghyun Yoo³,Yongbeom Lee⁴,Sangryong Kim⁵
Samsung Advanced Institute of Technology¹²³⁴⁵

요약

점점 더 많은 디지털 카메라와 휴대폰이 고해상도 카메라가 장착되고 대용량의 저장공간이 제공되면서 사용자들의 사진촬영 빈도가 증대하고 있다. 조만간 휴대폰의 저장된 사진을 효과적으로 관리하고 브라우징할 수 있는 기술이 필요한 시기가 올 것이다. 본 논문은 휴대폰이나 디지털 카메라 혹은 카메라가 장착되어 사진을 찍을 수 있는 모든 형태의 휴대단말에서 촬영된 개인사진을 지리적 위치정보를 이용하여 자동으로 분류하는 시스템을 제시한다. 기존의 시간정보를 이용하여 촬영시간의 근접성을 이용해 순차적으로 자동 분류하는 시스템과는 달리 위치정보를 이용하여 촬영위치에 따라 비순차적으로 자동 분류한다. 촬영위치 근접성을 결정하기 위해 밀도기반 클러스터링 알고리즘을 사용하여 전체 사진을 대분류하고 기존의 자동사진 분류방식에서는 다루지 않았던 일상사진과 비일상사진을 분류하고, 대분류된 사진을 시간정보를 이용하여 소분류 함으로서 자동 사진분류 성능을 높이고자 한다.

Keyword : Photograph Clustering, Density-based Clustering, Time-based Clustering

1. 서론

최근 디지털 카메라뿐만 아니라 카메라가 장착된 휴대폰이나 PDA(Personal Digital Assistant) 등이 등장하고 그 성능이나 저장용량이 증가하면서 촬영한 사진을 효율적으로 관리하고 브라우징하는 방법에 대해 많은 관심이 생겼다. 현재까지는 휴대단말에서 촬영된 사진을 개인용 컴퓨터에 옮겨서 저장한 후 사용자가 직접 사진을 분류하거나 개인용 컴퓨터의 계산성능을 이용하여 자동으로 분류하는 소프트웨어를 사용하고 있다.

또한 휴대단말의 이동성을 최대한 활용하여 사진을 찍은 위치 등을 기록할 수 있는 방법도 등장하였다. 예를 들어 휴대폰 제조업체인 핀란드의 Nokia 사는 추후 대부분의 휴대기기에 GPS 센서가 장착되어 실외뿐만 아니라 실내에서도 어디서나 휴대단말이 있는 위치를 검출해 낼 수 있는 기술을 개발하고 있다. [1]에서 언급된 A-GPS 기술을

이용하면 모바일 통신 중계국의 위치정보를 이용하여 위성 GPS 신호가 잡히지 않는 경우에도 중계국과 단말기 간의 신호의 특성을 이용하여 단말기의 지리적 위치를 검출할 수 있다. 최근에는 GPS 센서가 없는 휴대 단말에서도 CI (Cell Identifier)나 SAI (Service Area Identifier)를 이용하여 휴대단말의 위치를 검출할 수 있는 기술도 등장하였다[1].

이렇게 휴대단말기를 이용하여 촬영된 사진의 지리적 위치정보를 검출할 수 있는 방법은 다양하게 제공되고 있다.

[2], [3] 그리고 [4]에 따르면 위치정보를 이용하여 사진을 분류 혹은 브라우징하는 것은 다음과 같은 장점을 가지고 있다.

- 사용자는 개인 사진을 분류할 때 사건(Event), 위치(Location), 주제(Subject), 시간(Time)을 기

준으로 사용하며, 이들 4 개의 기준 중 3 가지는 위치(Location) 정보와 관련된다. (예: 사진 = 시간 + 위치, 주제 = ‘누가, 무엇을, 언제, 어디서’의 조합)[2]

- 위치로 사진을 브라우징하는 것은 다른 브라우징 방법보다 사용자에게 직관적이며 이해하기 쉽다 [4].
- 위치정보는 특정 언어나 문화나 사용자와는 무관하게 만국 공통적 의미를 가지고 있다.
- 위치정보는 사진의 의미(Semantics)와 깊은 관련이 있다.

본 논문은 개인용 컴퓨터에 옮기거나 혹은 휴대 단말기에서 직접 사용자의 사진을 자동분류하는 방법에 대해 구체적으로 제시한다. 특히 이제까지 다루어지지 않았던 일상적인 사진들과 비일상적인 사진들(예: 여행사진, 출장사진 등)로 대분류하고 대분류된 사진을 다시 시간에 따라 소분류하는 방법을 다룰 것이다.

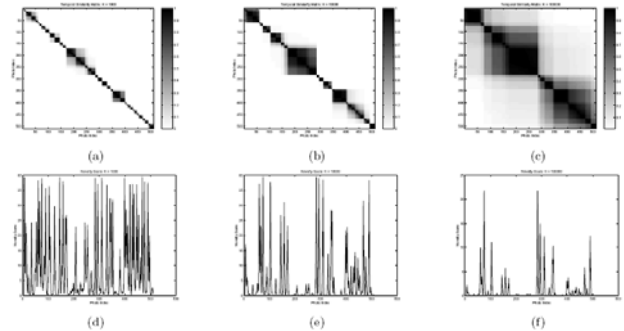
2. 시간기반 자동사진 분류 방법

일반적으로 촬영되는 사진의 헤더에는 촬영된 영상과 관련된 메타 데이터가 저장되어 있다[8]. 저장된 메타 데이터는 사진의 해상도, 저장형식, 셔터속도, 밝기, 초점거리, 촬영시간 및 기타 사진의 영상과 관련된 추가적인 정보가 포함되어 있다. 이러한 정보 중 촬영시간 정보를 이용하여 자동으로 사진을 분류하는 방법이 종래에 많이 연구되어 왔다. [5]에서는 자동분류를 위한 사진을 [그림 1]에서와 같은 유사도 매트릭스(Similarity Matrices)를 이용하여 자동사진 분류를 수행한다. 즉, N 개의 사진을 분류해야 하는 경우 사진의 메타정보 중 촬영시간을 기준으로 정렬한 후 $N \times N$ 매트릭스로 구성하여 유사도,

$$S_K(i, j) = \exp\left(-\frac{|t_i - t_j|}{K}\right)$$

S_K 를 각 항목별로 구한 다음 [그림 1]의 하단부의 Novelty Score 그래프를 도출하고 적당한 가우

시안 커널(Gaussian Kernel)을 적용하여 분류될 사진의 경계위치(Boundaries)를 결정한다. 이 때 최적의 경계위치를 자동으로 결정하기 위해 여러 K 값에 대해 분류를 시행하고 각 분류에 대해 Confidence Score를 구하여 K 를 결정하게 된다.



[그림 1] 여러 스케일에 따른 유사도 매트릭스

[6]에서는 [5]에서와 같이 시간정보를 이용하여 순차적으로 사진을 정렬하지만 최적 경계값(Threshold)을 결정하기 위한 방법의 제시 없이 실험에 의한 경험치를 이용하여 상수 $K = \log(17)$ 로 정하여 자동으로 사진을 순차적으로 분류하였다.

3. 위치기반 자동사진 분류 방법

본 논문에서는 사용자의 이동 패턴과 사진의 촬영 위치정보를 이용하여 일상적인 사진과 그렇지 않은 사진으로 대분류를 수행한다. 먼저 사진 분류를 위해 사용된 데이터의 예는 [그림 2]과 같다.

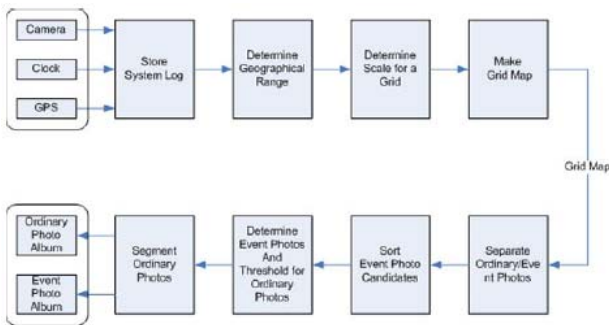
05/06/2005 14:55:24,207278,415792, ↓
05/06/2005 14:55:34,207278,415792, ↓
05/06/2005 14:55:44,207278,415792, ↓
05/06/2005 14:55:54,207278,415792, ↓
05/06/2005 14:56:04,207278,415792, ↓
05/06/2005 14:56:15,207278,415792,DSC04492.JPG ↓
05/06/2005 14:56:25,207278,415792, ↓
05/06/2005 14:56:35,207278,415792, ↓
05/06/2005 14:56:45,207278,415792, ↓
05/06/2005 14:56:55,207278,415792,DSC04493.JPG ↓
05/06/2005 14:57:05,207278,415792,DSC04494.JPG ↓
05/06/2005 14:57:15,207280,415792, ↓
05/06/2005 14:57:25,207282,415791, ↓

[그림 2] 휴대단말기의 시스템 로그 데이터 예제

사용자의 이동패턴의 분석을 위해 휴대 단말기는 사진을 촬영하지 않을 때도 휴대 단말기가

있는 위치를 일정한 주기로 반복하여 검출하게 된다. 검출되는 정보는 날짜, 시간, 경도, 위도, 사진 파일명이며 경도와 위도의 단위는 미터(m)이다. 논문에서 사용된 경도와 위도 데이터는 휴대 단말기에서 제공하는 TM 좌표계를 그대로 사용하였다. 검출주기는 임의로 결정할 수 있으나 여기서는 매 10초마다 한번씩 단말기의 위치를 저장하였다.

저장된 위치 데이터는 [그림 3]의 블록 다이어그램과 같이 처리가 이루어진다.



[그림 3] 전체 시스템의 블록 다이어그램

본 논문에서는 일상사진과 비일상사진을 대분류하기 위해 사용자의 이동패턴을 먼저 결정해야 한다. 이러한 대분류를 위한 자료구조로 본 논문에서는 그리드 맵 (Grid Map)를 이용한다.

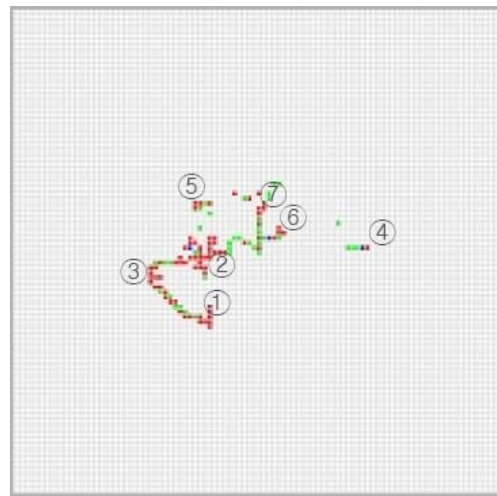
3-1 그리드 맵 생성

그리드 맵을 생성하기 위해 먼저 그리드 맵의 크기와 스케일을 결정해야 한다. 이를 위해 [그림 2]의 시스템 로그로부터 최대 및 최소 경도 위도 값을 구한다. 일반적으로 휴대폰의 저장공간은 다른 기기에 비해 상대적으로 한정적이기 때문에 경도, 위도의 범위에 따라 일정한 크기를 넘어가게 되면 하나의 그리드(Grid)가 포함하는 영역의 크기를 넓혀서 모든 범위가 그리드 맵에 들어가도록 한다. [그림 4]는 사용된 시스템 로그 데이터를 이용하여 생성한 그리드 맵의 예이다. 생성방법은 각각의 시스템 로그의 경도, 위도가 특정 그리드에 포함되는 경우 해당 그리드의 카운터 값을 증가시키는 방법으로 생성한다.

하나의 그리드가 가지고 있는 데이터의 내용은 다음과 같다.

(Grid_ID, numOfVisit, numOfPhoto)

Grid_ID 는 각 그리드를 구분하기 위한 인덱스 번호이고, numOfVisit 는 해당 그리드의 경도, 위도 범위에 시스템 로그 데이터가 있는 경우 1씩 증가하는 카운터값 그리고 numOfPhoto 는 해당 위치에서 촬영한 사진의 매수를 의미한다.



[그림 4] 그리드 맵의 예. (1 번은 회사, 2 번은 학교, 3 번은 집, 4 번은 병원, 5 번은 쇼핑몰)

생성된 그리드 맵은 실제 한 회사에서 아르바이트를 하고 있는 대학생의 휴대폰으로 검출한 데이터를 일주일간 저장한 후 그리드 맵에 표시한 것으로 [그림 4]의 경우 하나의 그리드가 $200 \times 200 m^2$ 영역을 포함하며, 빨간색으로 표시된 그리드는 사진이 촬영된 위치를 나타내고, 녹색으로 표시된 그리드는 사진을 촬영하지 않고 그냥 방문한 적이 있는 위치를 나타낸다. 1 번 영역은 회사가 있는 위치이며, 2 번 위치는 다니고 있는 학교, 3 번 위치는 살고 있는 집이 있는 위치이며, 4 번과 5 번은 각각 친구가 입원해 있는 병원과 대형 쇼핑몰의 위치이다.

3-2 밀도기반 클러스터링 알고리즘

그리드 맵을 생성한 후에는 사진을 일상사진과 그렇지 않은 사진으로 대분류하기 위해 밀도기반 클러스터링 알고리즘(Density-based Clustering)을

적용한다. 본 논문에서는 [7]에서 사용된 DBSCAN 알고리즘을 본 논문에서 생성한 그리드 맵을 처리할 수 있도록 변형하여 적용하였다. 즉, [7]에서는 잡음으로 처리되어 버려지는 영역이 본 논문에서는 비일상사진들로 분류하기 위해 필요한 중요한 정보로 사용된다. 오히려 해당 위치에서의 방문빈도는 적지만, 사진이 촬영된 빈도는 많은 위치를 검출함으로써 사진을 촬영한 사람에게는 중요한 여행지 사진 등으로 분류될 수 있도록 표시하게 된다. 변형된 밀도기반 클러스터링 알고리즘의 의사 코드(pseudo code)는 다음과 같다.

```

Do Density-based Clustering
{
  for(y = 0; y < NUM_OF_GRID; y++) {
    for(x = 0; x < NUM_OF_GRID; x++) {

      for(nx = (x-EPS); nx < (x+EPS); nx++)
        for(ny = (y-EPS); ny < (y+EPS); ny++)
          if( (nx-x)2 + (ny-y)2 <= EPS2 )
            if(grid[nx][ny].numOfVisit >= T_OF_VISIT)
              counter++;

      if( counter > 8 ) then
        add grid[x][y] into ordinary cluster;
      else if( grid[x][y].numOfPhoto > 1 ) then
        add grid[x][y] into candidates of event cluster;
      else
        drop grid[x][y];
    }
  }
}
Repeat Density-based Clustering for Candidates of Event Cluster;

```

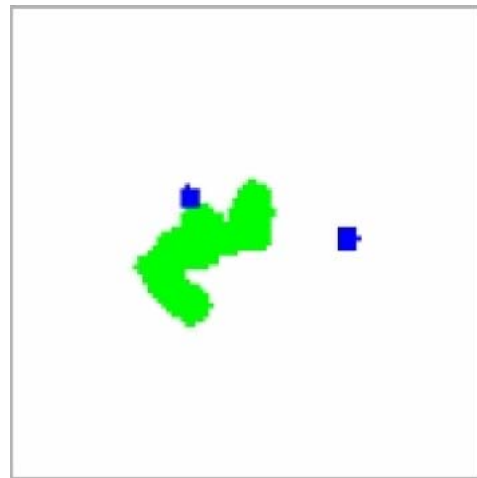
위의 코드에서 EPS 는 DBSCAN 알고리즘[7]의 입실론값(Epsilon Value)으로 현재 그리드와 이웃한 그리드 간의 밀도를 검출하기 위한 원형 영역의 반지름이고 T_OF_VISIT 은 사용자가 해당 그리드에서 머문 빈도수의 경계값(Threshold)으로 [7]에서와 동일한 값을 사용하였다.

변형된 밀도기반 클러스터링 알고리즘에서는 사용자가 자주 방문하는 그리드에 속한 사진들은 일상사진 클러스터 목록에 저장하고, 자주 방문하지 않았으나 사진촬영 빈도가 높은 그리드에 속한 사진은 비일상사진으로 저장한다.

만일 위의 조건 어느 것에도 해당하지 않는 경우에는 해당 사진은 대분류에 의해 분류되지 못하고 따로 저장해 둔다.

[그림 5]는 변형된 밀도기반 클러스터링 알고

리즘을 적용하여 그리드 맵에 대분류된 클러스터를 각각 녹색과 파란색으로 표시한 것이다.



[그림 5] 녹색 영역은 사용자가 자주 방문하는 위치 및 일상사진이 촬영된 위치이고, 파란색 영역은 비일상사진이 촬영된 위치를 나타냄

녹색으로 표시된 클러스터 영역은 사용자가 방문한 빈도가 높은 위치로 이 영역에서 촬영된 사진은 일상사진으로 자동 분류되고 추후에 촬영 시간의 유사도에 따라 재분류하게 된다. 파란색으로 표시된 클러스터 영역은 사용자의 방문 빈도에 비해 사진촬영 빈도가 높은 위치로 비일상적인 사진으로 자동 분류된다. 주로 여행지 사진 등이 여기에 속하게 되며 다른 자동 사진분류 방식과 달리 비순차적으로 분류되는 부분에 해당한다.

4. 계층적 자동사진 분류 방법

일상적으로 촬영되는 사진의 경우에는 여행지의 사진들과는 달리 비슷한 장소에서 시간의 차이를 두고 반복해서 촬영되는 사진들이 대부분이다. 예를 들면 성장하는 아이의 모습을 촬영한다거나 하는 사진들이 이런 일상사진에 해당한다. 따라서 제안된 대분류에 의해 자동으로 분류된 일상사진은 다시 촬영시간에 의해 재분류할 필요가 있다.

본 논문에서는 [6]에서 제안된 방식이 [5]에서 제안된 방식에 비해 계산량이 적고 상대적인 성능 평가에서도 큰 차이가 없어서 일상사진의 소분류를 위해 [6]에서 제안된 사진들간의 촬영시간 차(Time Gap)의 대수평균을 이용한 클러스터링 알고

리즘을 사용하였다.

4-1 경계값의 결정방법

PhotoTOC[6]에 적용된 대수평균차를 이용한 자동사진 분류 방법에서는 경계값(Threshold) K 를 얼마로 정하느냐에 따라 자동분류의 성능이 크게 차이가 난다. [6]에서는 사용자 입력 혹은 자동사진분류 시스템 개발자가 실험에 의해 최적값을 결정하는 것으로 되어 있다. 하지만, 대량의 사진이 아닌 개인사진에 이런 방식을 적용하는 경우 서로 다른 사용자에게 따라 그 성능차이가 크게 날 수 밖에 없다. 따라서 본 논문에서는 다음과 같은 가정으로 경계값을 자동으로 결정하여 사용하였다.

- 비일상사진 클러스터의 사진들은 단위시간에 촬영된 빈도는 높으나 사진과 사진간의 촬영 시간차에 의한 패턴은 일상사진의 촬영시간차의 패턴과 유사하다.
- 따라서 비 일상사진 클러스터의 사진들의 촬영시간차의 대수 평균차는 일상사진의 대수 평균차와 유사하다.

위의 가정에 따라 본 논문에서는 일상사진의 촬영시간차를 위한 경계값 K 를 결정하기 위해 [6]에서 제시된 대수 평균차 식으로부터 유도된 아래 식을 사용하였다.

$$K = \frac{\sum_{N=1}^C (\log(g_N) - \sum_{i=d}^d \log(g_{N+i}))}{C(2d+1)}$$

위 식에서 C 는 비일상사진 클러스터 중 하나의 크기, 즉 [그림 5]에서 파란색 클러스터 중 하나에 속한 전체 사진의 수이고 g 는 해당 클러스터에 속한 사진들의 촬영 시간차이다. d 는 [6]에서와 동일한 값($d=10$)을 사용하였다.

[그림 6]은 본 논문에서 제안된 방식으로 자동 분류된 사진의 Thumbnail 결과이다.



[그림 6] 자동분류 결과

5. 실험 및 평가

제안된 자동 사진분류 방법을 성능평가하기 위해 카메라가 장착된 휴대폰을 가진 대학생이 일주일간 촬영한 520 장의 사진을 사용하였다. 자동 사진분류 클러스터링의 성능을 비교하는 방법에는 여러 가지가 있지만 본 논문에서는 [5]와 [9]에서 제시된 성능비교 방법을 사용하였다.

먼저 사진을 촬영한 사용자에게 자신이 촬영한 사진을 분류하게 하고 이렇게 사용자가 직접 손으로 분류한 분류 목록 및 클러스터들을 Ground Truth 로 하고 F-score 값을 구하여 최종 성능을 [6]과 비교하였다.

F-score 를 구하기 위해 사용되는 식은 아래와 같다.

$$\text{precision} = \frac{\text{correctly detected boundaries}}{\text{total number of detected boundaries}}$$

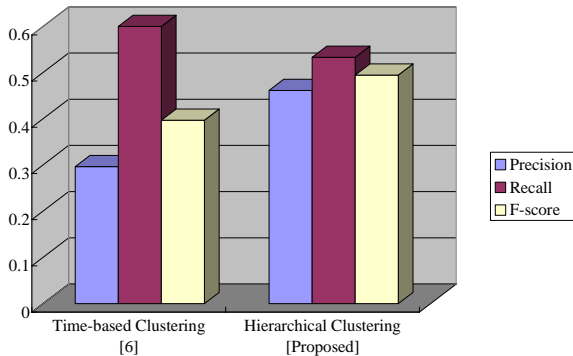
$$\text{recall} = \frac{\text{correctly detected boundaries}}{\text{total number of ground truth boundaries}}$$

$$F\text{-score} = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}}$$

위 식에서 Precision 은 전체 검출된 클러스터의 경계(Boundaries)에 대한 올바른 자동검출 경계의 비율이고, Recall 은 Ground Truth 에 대한 올바른 자동검출 경계의 비율 그리고 F-score 는 전체성능을 나타내기 위한 Precision 과 Recall 이 합성된 식의 결과이다.

공정한 비교를 위해 본 논문에서 제시된 경계값(Threshold) 결정방법에 의해 도출된 K 를 비교대상인 [6]에도 적용하여 비교하였다.

	Time-based Clustering [6]	Hierarchical Clustering [Proposed]
Correct number of clusters	27	24
Total number of detected clusters	91	52
Total number of ground truth clusters	45	45
Precision	0.2967	0.4615
Recall	0.6	0.5333
F-score	0.3971	0.4948



[그림 7] 검증실험결과

[그림 7]의 검증실험결과에서 보여지는 것처럼 기존의 Time-based Clustering 방식[6]이 논문에서 제안된 방식보다 더 많은 올바른 경계(boundaries)를 찾아 냈으나 (27 개), 경계검출을 위한 전체 클러스터의 수가 제안된 방식에 비해 현저하게 많아 (91 개) 전체성능을 나타내는 F-score 측면에서는 제안된 방식이 10% 정도 제안된 방식이 성능이 우수한 것을 알 수 있다.

6. 결론

기존의 시간기반의 순차적 자동사진분류방법과 달리 사용자와 사진의 위치정보를 이용하여 일

상사진과 그렇지 않은 사진을 대분류하고 시간적인 자동분류가 필요한 일상사진에 대해 다시 소분류를 적용하는 방식을 제안하여 실험에 의해 기존 시간기반 방식에 비해 개선된 성능을 얻을 수 있었다. 다만, 개선된 성능유지를 위해서 사용되는 사용자 위치검출에 필요한 비용이 소요된다는 단점은 추후 개선이 필요하다. 이러한 문제는 촬영된 사진만으로 사용자의 이동패턴도 추론해 내거나 좀더 저렴하게 사용자 위치를 검출하는 방법이 나온다면 해결할 수 있을 것이다.

7. 참고문헌

- [1] Location aware applications take off white paper. http://www.nokia.com/NOKIA_COM_1/Operators/Technologies/Network_Enhancements/Location_Technologies/location_wp.pdf
- [2] Lieberman, H., Liu, H. Adaptive Linking between Text and Photos Using Common Sense Reasoning. In Adaptive Hypermedia and Adaptive Web-Based Systems, P. deBra, P. Brusilovsky, R. Conejo (eds.), Springer-Verlag, Berlin, pp. 2-11, 2002.
- [3] K. Toyama, R. Lofan, and A. Roseway, Geographic Location Tags on Digital Images, ACM Multimedia 2003, Berkeley, CA, USA, Nov. 2003.
- [4] Morris, B. CARTO-NET: Graphic retrieval and management in an automated map library. Special Libraries Association, Geography and Map Division Bulletin, 152:19-35, 1988.
- [5] M. D. Cooper, J. Foote, A. Girgensohn, L. Wilcox, "Temporal Event Clustering for Digital Photo Collections", ACM Multimedia 2003, Berkeley, CA (2003)
- [6] J. Platt, M. Czerwinski, and B. Field. PhotoTOC:Automatic Clustering for Browsing Personal Photographs. Microsoft Research Technical Report MSR-TR-2002-17, 2002.
- [7] Martin Ester, Hans-Peter Kriegel, Jorg Sander, Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. Institute for Computer Science, University of

Munich (1996)

- [8] Digital Still Camera Image File Format Standard.
Japan Electronic Industry Development Association,
<http://www.pima.net/standards/it10/PIMA15740/exif.htm> (1998)
- [9] J. Boreczky, L. Rowe. Comparison of video shot
boundary detection techniques. Proc. SPIE Storage
and Retrieval for Image and Video Database (1996)