# 모바일기기 사용자의 컨텍스트와 이미지 주석을 이용한 베이지안 네트워크기반 사진 자동요약

민준기° 조성배
연세대학교 컴퓨터과학과
loomlike@sclab.yonsei.ac.kr sbcho@cs.yonsei.ac.kr

# Bayesian Network based Automatic Summarization of Photos using User's Context on Mobile Device and Image Annotation

Jun-Ki Min° Sung-Bae Cho
Department of Computer Science, Yonsei University

**요 약**

모바일기기에 탑재되어있는 디지털 카메라의 성능이 향상됨에 따라 이를 이용한 사진의 촬영 및 수집이 용이해졌으며, 따라서 사용자 로그정보를 이용하여 방대한 양의 사진을 분석하거나 브라우징해주는 방법들이 연구되고 있다. 본 논문에서는 모바일기기의 불확실한 로그정보와 사진 주석정보를 베이지안 네트워크로 모델링하여 사용자가 겪은 이벤트들을 추론하고 사용자의 일과를 요약해주는 방법을 제안한다. 우선 사진들을 시간과 위치정보에 따라 분할하여 사진그룹목록을 생성하고, 이를 모바일기기에 입력되어있는 사용자의 일정목록과 합하여 임시이벤트목록을 생성한다. 그 뒤 베이지안 네트워크를 이용하여 각 이벤트를 인식하고 이를 가장 잘 나타내는 사진을 선택한다. 제안하는 방법은 선택된 사진들을 나열하여 사진다이어리형식으로 사용자의 일과를 요약하여주며, 이때 특정 이벤트와 매치되는 사진이 없을 경우 미리 정의되어있는 만화 컷을 대신 사용하여 내용이 매끄럽게 이어지도록 하였다.

## 1. 서 론

최근 모바일기기에 탑재되어 있는 디지털 카메라의 성능이 향상됨에 따라 손쉽게 사진을 촬영하거나 보관할 수 있게 되었으며, 따라서 방대한 양의 사진을 효과적으로 검색하거나 브라우징하기위한 방법들이 활발히 연구되고 있다. Lim[1] 등은 사진영상으로부터 비쥬얼 키워드를 추출하고 이들 간의 위치관계를 그래프로 표현하여 사진의 내용을 분석한 뒤 이벤트를 인식하는 방법을 제안하였으며, Kuo[2] 등은 사진의 검색, 처리 및 이해를 돕기 위해 사진의 의미정보인 사람, 물체, 장소 등과 촬영환경정보인 위치, 각도 등 총 12가지 속성을 MPEG7 형식의 의 메타데이터로 정의하였다. 이밖에도 사진인식이나 주석 태깅 기술을 적용하여 사진을 요약해주는 어플리케이션에 대한 연구도 활발히 진행 중이다[3, 4].

모바일기기는 사용자가 항상 휴대하기 때문에 사용자의 다양한 로그가 사진과 함께 저장되며, 이는 사진의 내용을 분석하거나 요약하는데 유용한 정보가 된다. 본 논문에서는 모바일기기로부터 수집한 사용자의 로그정보와 PIMS(Personal Information Management System)의 일정정보, 사진에 태깅되어있는 주석정보를 이용하여 사용자가 경험한 일들 중 식사, 영화관람 등과 같은 주요이벤트들을 인식하고 해당 사진들을 선택하여 사진다이어리 형식으로 사용자의 일과를 요약해주는 방법을 제안한다. 이때 모바일 로그의 불확실성을 해결하기 위해 제안하는 방법은 확률추론모델인 베이지안 네트워크를 이용하여 사용자 컨텍스트 정보로부터 이벤트를 인식한다.

## 2. 배경연구

### 2.1 베이지안 네트워크 기반 모바일 로그분석

모바일기기로부터 수집한 사용자 로그는 사용자가 경험한 이벤트나 사용자의 상황을 판단하는 근거로 활용되지만, 센싱값의 결측치나 불확실성 때문에 정확한 컨텍스트를 추출하는데 어려움이 있다. 따라서 이를 해결하기 위해 여러 요인들 간의 인과관계를 조건부확률로 모델링하는 베이지안 네트워크기반 방법들이 제안되어왔다. Krause 등은 베이지안 네트워크를 사용하여 모바일기기 컨텍스트로부터 사용자 상황을 예측하고 적합한 서비스를 선택하는 방법을 제안하였다[5]. Cho 등은 모바일기기에 수집된 사용자 컨텍스트를 만화형식의 다이어리로 구성해주는 방법을 제안하였는데, 컨텍스트 정보로부터 특이성(Landmark)을 추출하기위해 베이지안 네트워크를 사용하였다[6].

### 2.2 사진 요약

방대한 양의 디지털 사진첩은 중요하지 않거나 중복되는 사진들을 많이 포함하기 때문에 중요사진을 선택하여 요약하는 방법이 필요하다. Lim 등은 촬영한 시간과 칼라 히스토그램정보를 기반으로 사진을 그룹화한 후 얼굴탐색알고리즘을 이용하여 얼굴이 가장 잘 나타난 사진들을 중요사진으로 선택하였다. 이때 사진에 얼굴이 나타

나지 않은 경우에는 각 그룹의 첫번째 사진을 중요사진으로 선택하였다[3]. Smeaton은 마이크로소프트에서 개발한 센스캠(SenseCam)에 내장되어있는 빛, 적외선, 3축 가속도 센싱값의 변화량을 이용하여 센스캠에 의해 자동으로 촬영된 일련의 사진들로부터 중요사진을 선택하는 방법을 제안하였다[4]. 하지만 이와 같은 기존의 사진요약 연구는 영상특징이나 센서값 등 저수준 특징정보만을 이용하여 중요사진을 추출하기 때문에 사용자가 겪은 이벤트 등의 고수준 정보를 사진요약 결과에 반영할 수 없다.

## 3. 베이지안 네트워크 기반 사진 자동요약

사진을 자동으로 요약하기 위해서 본 논문에서 제안하는 방법은 다음과 같다. 우선 모바일로그로부터 위치정보와 일정목록을 사용자 컨텍스트로 추출하고, 사진을 촬영한 시간과 위치에 따라 분할하여 사진그룹목록을 생성한다. 그 후 일정목록과 사진그룹목록을 합하여 임시이벤트목록을 생성한다. 마지막으로 베이지안 네트워크를 이용하여 실제 이벤트를 인식하고 각 이벤트를 가장 잘 나타내는 사진을 선택하여 사진다이어리를 생성한다. 이때 만족하는 사진이 없을 경우 해당 이벤트를 나타내는 만화 컷으로 대체하여 사진다이어리를 구성한다. 그림 1은 제안하는 방법의 전체 흐름도를 나타낸다.

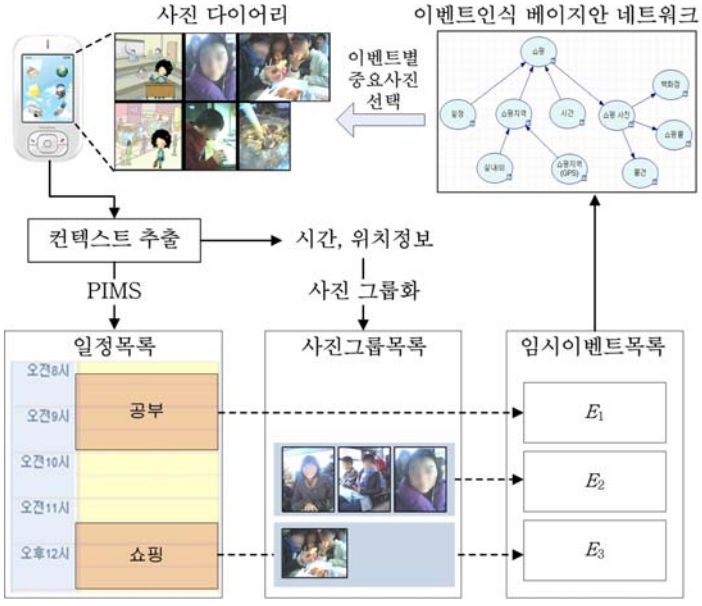


그림 1. 제안하는 방법의 전체 흐름도

### 3.1 모바일 컨텍스트 추출과 사진 그룹화

사용자의 위치정보(방문한 장소)는 GPS의 도, 분, 초 정보를 $x$, $y$좌표로 환산하고 이를 지도상에 다각형 영역으로 표시되어있는 지역정보와 매칭하여 추출한다. 이때 지도상의 지역이 넓을 경우 GPS좌표가 다각형의 내부에 있으면 사용자가 해당 장소를 방문한 것으로 판단하며, 지역이 좁을 경우 다각형의 중심(식 1)과 GPS좌표와의 거리가 건물반경 $R$(식 2)보다 작을 경우 방문한 것으로 판단한다[7]. $R_e$는 오차허용범위로, 본 논문에서는 0.3초로 설정하였다.

$$x_m = \sum_{i=1}^{n} \frac{x_i}{n}, \quad y_m = \sum_{i=1}^{n} \frac{y_i}{n} \tag{1}$$

$$R = \sqrt{(x_m - x_i)^2 + (y_m + y_i)^2} + R_e \tag{2}$$

일정목록은 모바일기기의 PIMS에 사용자가 입력한 일정정보로부터 얻을 수 있으며, 표 1은 본 논문에서 사용한 일정정보의 속성값을 나타낸다.

표 1. 일정정보의 속성과 값의 예

| 종류 | 값 |
|---|---|
| 날짜 | 2008년 3월 10일 |
| 시간 | 19:00~20:00 |
| 장소 | 연세대학교 백주년기념관 |
| 내용 | 동아리친구와 연주회 관람 |

사진그룹목록은 그림 2와 같이 $n$장의 사진 $P_1$, ... $P_n$을 촬영시간과 장소정보를 기반으로 이벤트 단위로 분할하여 생성한다. 첫 번째 단계에서는 [6]과 같이 시간 임계값을 설정하고 인접한 두 사진의 촬영시간 차가 임계값보다 클 경우 서로 다른 이벤트로 분할한다. 두 번째 단계에서는 동일 이벤트에 속하는 사진들 간의 촬영장소 정보가 바뀔 경우 세부 이벤트로 분할한다.

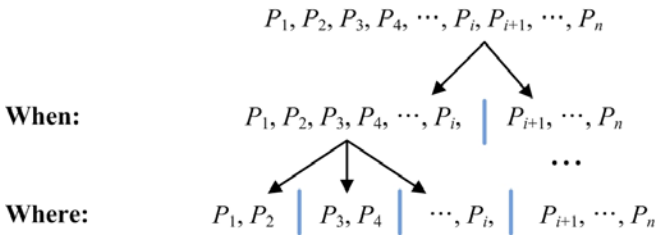


그림 2. 시공간 정보를 이용한 사진 그룹화

### 3.2 이벤트 인식과 사진다이어리 구성

일정목록의 이벤트들은 실제로 발생하였는지 여부가 불확실하며 사진그룹목록의 이벤트들은 종류를 알 수 없다. 따라서 앞 절에서 생성한 사진그룹목록과 일정목록의 이벤트들을 발생시간을 기준으로 합하여 임시이벤트목록을 생성한 후, 각 임시이벤트에 해당하는 모바일 로그와 사용자 컨텍스트, 사진 주석정보를 베이지안 네트워크의 증거값으로 사용하여 인식한다. 제안하는 방법에서는 인식 정확도를 높이고 추론 복잡도를 줄이기 위해 영화감상, 공부, 쇼핑 등 이벤트 종류별 베이지안 네트워크를 모델링 하고, 입력한 증거값에 대해서 가장 높은 확률을 보이는 네트워크로 이벤트를 인식하였다. 본 논문에서는 베이지안 네트워크의 구조 및 조건부확률테이블을 전문가가 직접 설계하였다. 그림 3은 쇼핑 이벤트를 인식하기 위한 베이지안 네트워크의 예를 보여준다. 실 내/외 노드의 증거값은 GPS의 수신이 가능한지 여부로 알 수 있으며, 쇼핑사진 노드의 하위인 사진 주석정보 노드들은 임시이벤트에 속하는 사진들의 주석정보 유무에 따라 증거값이 할당 된다. 예를 들어 임시이벤트에

"백화점에서 친구와 함께", "새로 산 물건", "친구와 식사"의 3장 사진이 포함되어있는 경우, 그림 3의 사진 주석정보 노드에서 백화점노드와 물건노드는 Yes로 증거값이 할당되고 쇼핑몰 노드는 No로 증거값이 할당된다.

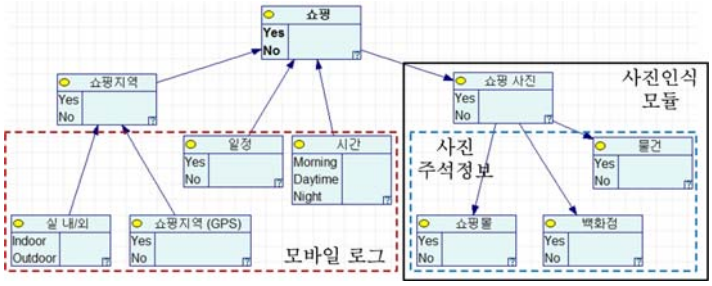


그림 3. 이벤트인식을 위한 베이지안 네트워크의 예

임시이벤트를 실제 이벤트로 인식한 뒤, 임시이벤트에 포함된 여러장의 사진들로부터 이벤트를 잘 나타내는 중요사진을 선택한다. 이를 위해 앞에서 사용한 베이지안 네트워크 중 사진인식모듈을 이용한다. 이벤트를 인식할 때에는 임시이벤트에 속하는 모든 사진의 의미정보를 사용하였지만, 중요사진 선택 시에는 각 사진별 의미정보만을 증거값으로 입력한다. 최종적으로 특정 임계값 이상되는 사진을 선택하고 발생시간 순으로 배치하여 사진 다이어리를 구성한다. 이때, 동일한 의미정보를 가지는 사진들이 있을 경우 사용자가 만족스러운 사진을 얻을 때까지 유사한 장면을 촬영한 것으로 간주하여 가장 마지막 사진을 선택함으로써 내용의 중복을 피하고, 이벤트 그룹 내에 임계값을 만족시키는 사진이 없는 경우에는 그림 4와 같이 해당 이벤트를 나타내는 만화 컷으로 대체하여 보여준다.

그림 4. 이벤트를 나타내는 만화 컷의 예 (1. 운동, 2. 이동, 3. 공부)

## 4. 실험 및 결과

### 4.1 실험환경

제안하는 방법의 적용 예를 보이기 위해 피험자를 선정하여 스마트폰에 GPS수신모듈을 장착하여 로그데이터를 수집하도록 하였다. 피험자는 여자대학생으로, 5일 동안 실제 스마트폰을 사용하며 총 74장의 사진을 수집하였다. 표 2와 그림 5는 각각 피험자로부터 받은 일과 보고서와 수집한 사진의 일부를 나타내며, 표 3은 5일 동안의 사용자 일정정보를 보여준다.

표 2. 첫째 날에 해당하는 사용자 보고서

| 시간 | 장소 | 행동 | 관련 사람 |
|---|---|---|---|
| 09:30 | 집 | 취침 | |
| 10:00 | 버스 | 집->학교 | |
| 11:00 | 신촌 할리스 | 공부 | |
| 11:30 | 신촌 니코니코 | 점심식사 | 애인 |
| 13:30 | 중도 | 토익 공부 | 애인 |
| 15:00 | 버스 | 코엑스 가는 길 | |
| 16:00 | 코엑스 | 코엑스 구경 | 애인 |
| 18:00 | 코엑스 메가박스 | 영화 봄 | 애인 |
| 20:00 | 코엑스 M-Zone | 30만번째 방문객 | 애인 |
| 20:30 | 코엑스 푸드코트 | 저녁으로 알밥먹음 | 애인 |
| 23:00 | 집 | 씻고 웹써핑 | |

그림 5. 첫째 날에 수집한 사진

표 3. 사용자 일정정보(셋째날에는 입력된 일정 없음)

| | 시간 | 장소 | 내용 |
|---|---|---|---|
| 첫째날 | 10:00~15:00 | 연세대학교 도서관 | 토익공부 |
| | 18:00~20:00 | 코엑스 메가박스 | 애인과 영화 관람 |
| 둘째날 | 09:00~12:00 | 미아동 영훈고등학교 | 토익시험 |
| | 14:00~16:00 | 명동 롯데백화점 | 언니와 쇼핑 |
| | 18:00~20:00 | 쌍문동 이모집 | 사촌 동생 과외 |
| 넷째날 | 13:00~15:00 | 성남 학생중앙군사학교 | 애인 임관식 |
| 다섯째날 | 12:00~13:00 | 연세대학교 외솔관 | 교양수업 |
| | 16:00~17:00 | 연세대학교 제1공학관 | 전공수업 |
| | 19:00~20:00 | 연세대학교 백주년기념관 | 친구와 유포니아 연주회 관람 |
| | 20:00~21:00 | 성신여대 앞 | 스터디모임 |

### 4.2 사진다이어리 생성결과

첫째 날 촬영한 사진들을 시간과 위치정보로 분할한 결과 4개의 그룹이 생성되었으며, 이중 2개는 표 3의 일정정보와 시간대가 일치하였다. 첫째 날에 해당하는 총 4개의 임시이벤트를 베이지안 네트워크로 인식한 결과 " 공부, 식사, 영화관람, 집에서 휴식" 으로 분류되었다. 이를 표 2의 사용자 보고서와 비교하였을 때 사용자의 일과를 적절한 수준에서 요약하였음을 확인할 수 있다. 첫째 날의 경우 피험자가 장시간 실내에 있어서 GPS신호의 대부분이 결측되었으나 사진의 주석정보를 이용하여 영화관람 이벤트 등을 인식할 수 있었다.

그림 5는 사진다이어리의 생성결과를 나타내는 것으로, 실제 수집한 사진들 중 중요사진과 만화 컷들이 사용자 일과를 적절히 요약함을 확인하였다. 하지만 둘째 날과 셋째 날의 경우 이벤트를 함께 수행한 대상보다는 자신의 사진이 집중적으로 선택되어 누구와 함께 있었는지를 알기 힘들었다. 또한 넷째날의 경우 수집된 사진의 수가 적고 로그정보가 정확하지 않아서 이벤트의 흐름이 다소 개연성 없게 표현되었다.

그림 6. 사진다이어리 생성결과 (각 그룹이 하루 일과를 나타냄)

## 5. 결 론

본 논문에서는 모바일기기로부터 수집한 사용자의 로그정보와 사진 주석정보를 분석하여 사용자의 일과를 요약하고 사진다이어리형식으로 표현해주는 방법을 제안하였다. 우선 사진들을 시간과 위치정보에 따라 분할하여 사진그룹목록을 생성하고, 이를 모바일기기상의 일정목록과 합하여 임시이벤트목록을 생성한다. 그 뒤 베이지안 네트워크를 이용하여 각 이벤트를 인식하고 중요사진을 선택한다. 이때 사진데이터가 없는 이벤트는 만화 컷으로 대체하여 사진다이어리를 구성하였다. 제안하는 방법은 사용자가 대량의 사진을 브라우징하거나 과거의 이벤트를 기억하는데 흥미와 편의를 제공할 수 있다. 하지만 사진의 주석정보가 태깅되어있어야 한다는 제약조건과 수집된 로그의 수가 적을 경우 이벤트를 인식하지 못한다는 한계점이 있다. 이를 해결하기 위해 로그정보의 결측값 개선방법과 사진영상 분석을 통한 사진주석의 자동생성방법 연구를 진행할 계획이다.

**감사의 글.** 이 논문은 지식경제부를 통해 IITA에서 지원받았음 (IITA-2008-(C1090-0801-0011))

### 참고문헌

1. J.H. Lim and Q. Tian, "Home photo content modeling for personalized event-based retrieval," *IEEE Multimedia*, vol. 10, no. 4, pp. 28-37, 2003.
2. P.-J. Kuo, T. Aoki, and H. Yasuda, "Building personal digital photograph libraries: An approach with ontology-based MPEG7 dozen dimensional digital content architecture," *IEEE Proc. Computer Graphics International*, pp. 482-489, 2004.
3. A.F. Smeaton, "Content vs. context for multimedia semantics: The case of SenseCam image structuring," *1st Int. Conf. on Semantics and Digital Media Technologies*, pp. 1-10, 2006.
4. J. Lim, J.H. Lim, and Q. Tian, "Automatic summarization for personal digital photos," *Proc. 2003 Joint Conf. of the 4th Int. Conf. on Information, Communications and Signal Processing, and the 4th Pacific Rim Conf. on Multimedia*, vol.3, pp.1536-1540, 2003.
5. A. Krause, A. Smailagic, and D.P. Siewiorek, "Context-aware mobile computing: Learning context-dependent personal preferences from a wearable sensor array," *IEEE Trans. on Mobile Computing*, vol. 5, no. 2, pp. 113-127, 2006.
6. S.-B. Cho, K.-J. Kim, K.-S. Hwang, and I.-J. Song, "AniDiary: Daily cartoon-style diary exploits bayesian networks," *IEEE Pervasive Computing*, vol. 6, no. 3, pp. 66-75, 2007.
7. 황금성, 조성배, 이종호, "모바일 디바이스 상에서의 특이성 탐지를 위한 베이지안 추론 모델," *정보과학회 논문지: 컴퓨팅의 실제*, vol. 13, no. 1, pp. 35-45, 2007.