

イベント構造の抽出に基づくコンテンツ管理

渡邊 裕子 (指導教員：小林 一郎)

1 研究概要

近年、デジタルカメラやカメラ付き携帯電話の急速な普及によって、個人が大量の画像ファイルを所有し、管理する機会が増加した。それにともない、Picasa [1]をはじめとする画像管理ツールも増加の一途をたどり、イベント毎に写真を整理し、各々の写真に対してコメントやタグなどの自由な付与が可能になった。

写真にタグを付ける方法は、人による手動タグ付けとコンピュータによる自動タグ付けの二通りが考えられる。人手によるタグ付与は、多大な時間と根気が必要となるが、付与されたタグはタグとして信頼性が高く検索に有効である。また、コンピュータによるタグ付与は、期待通りの結果が得られないことや、コンテンツにより効果に偏りが生じるため、様々な手法による事前学習が必要である。しかし、人手に比べタグ付与の手間を大幅に軽減できる。

このことから、本研究では人とコンピュータによるタグ付けを併用し、付与されるタグの質を保ちつつ、タグ付与に費やす手間を軽減する手法を提案する。具体的には、イベントに即した画像管理を行うために、イベントの半自動抽出を行う。コンピュータによって自動的に構造化されたイベントを用いれば、ユーザはそのイベント構造に即したタグ付けを行うことができる。また、構造化されたイベントは、ユーザにイベントの概要を思い出させる効果も期待できる。さらには、イベント構造を用いた画像共有にも拡張することができる。今まで個人がそれぞれ管理していたイベントのアルバムを統一し、イベント参加者が共有したイベントに通じた画像共有を可能にする。

2 イベント構造の抽出に基づく画像管理

本節では、提案手法について説明を行う。

2.1 画像管理に用いる情報

時間情報、GPS 情報と関連づけて画像ファイルを管理するため、撮影された画像ファイル自体が所有する以下の情報を定義する。

Exif 情報： ファイル名、撮影日時、GPS 情報など

補足情報： 撮影者、被写体、コメント

今回は、画像情報の中で特に Exif 情報に含まれる、撮影日時と GPS 情報をイベント構造化に用いる。

2.2 イベントの構造化

システム処理手順を以下に示し、図 1 に提案するシステムの概要図を示す。

step1. 画像ファイルから Exif 情報を取得する。

step2. 時間情報と GPS 情報に基づいてイベントのクラスタリングを行い、イベントを抽出する。

step3. ユーザがイベント構造を修正する。

人は、旅行や結婚式などイベントの発生に合わせて、写真を撮影する。さらには、イベントはサブイベントの集合とみなすことができる。例えば、結婚式披露宴中の衣装替えなどのサブイベントが発生すると、人は

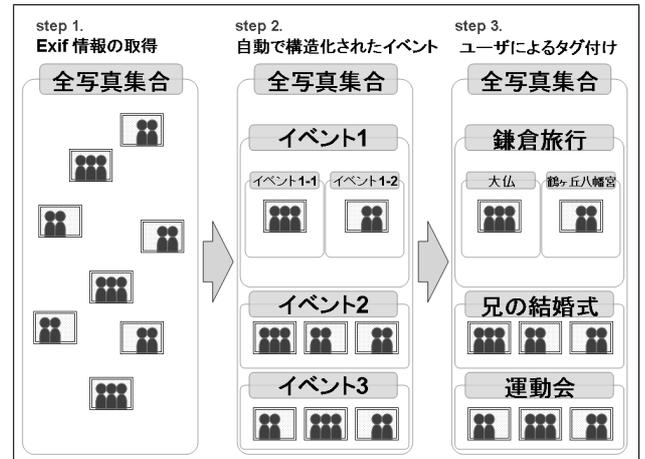


図 1: システム概要

写真を撮影する。つまり、人が写真を撮影する動機は、サブイベントの発生に依っていると考えられる [4]。したがってイベントに沿って写真を整理することは極自然であり、イベントの構造を用いればより自然な画像管理を行うことができる。このように撮影された写真とそれが含まれるイベントは、写真のもつ情報 (撮影日時や GPS 情報) によって結びつけられていると考えることができる。

イベントの構造を抽出するために、写真のもつ情報を用いてイベントを抽出するクラスタリングを行う。そのクラスタリングの手法として、以下の二通りの手法が考えられる。

- トップダウンな手法による構造化

1つのイベントを細分化する手法であり、事前にイベントに関するスケジュールが明記された詳細なプログラムがある場合に用いることができる。この場合、イベントのスケジュールと写真の撮影日時を合わせることで、プログラムに即した画像管理を行う。

- ボトムアップな手法による構造化

同一のイベントと思われるクラスタを統合していく手法であり、事前に詳細なプログラムがない場合でも用いることができる。

本研究では、あるイベントの画像データのみが与えられ、そのデータからイベントの構造を抽出をするため、ボトムアップな手法による構造化を行う。したがって、これに適した凝集型階層的クラスタリングを用いる。階層的クラスタリングにより、各イベントがクラスタに相当するため、構造化されたイベントは木構造をもつ。木構造のうち葉クラスタに相当するイベントをサブイベントと呼ぶ。

次に、クラスタ C_1 とクラスタ C_2 の距離関数を定義する。各クラスタ C_i は時刻情報 t_i と位置情報 (λ_i, ϕ_i) をもつ。ここで、 λ_i は経度、 ϕ_i は緯度を表す。本研究では群平均法を用いるので、各 t_i, λ_i, ϕ_i を次式で定義する。

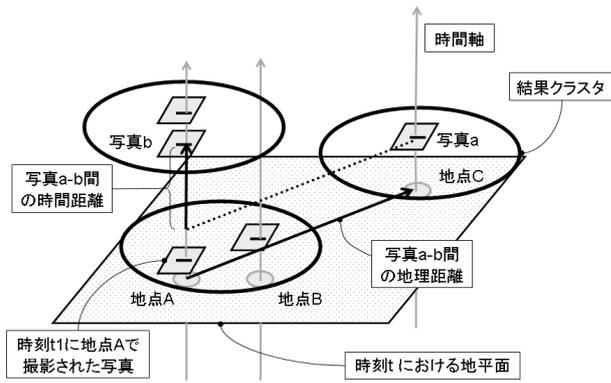


図 2: クラスタリングイメージ図

$$t_i = \frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}, \quad \lambda_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \lambda_{ij}, \quad \phi_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \phi_{ij}$$

n_i はクラスタ C_i に含まれる写真の総数であり, $t_{ij}, \lambda_{ij}, \phi_{ij}$ は, それぞれクラスタ C_i に含まれる各写真の時刻情報, 経度情報, 緯度情報である. 各クラスタ C_1, C_2 間の距離関数を, 以下のように定義する.

$$D(C_1, C_2) = w_t D_t(C_1, C_2) + w_p D_p(C_1, C_2)$$

D_t, D_p はそれぞれクラスタ間の時間距離, 位置距離を指し, w_t, w_p はそれぞれそれらに対する重みである. イベントクラスタの併合において, イベントが適切にサブイベントに細分化されるためには, クラスタを併合する閾値の設定が重要である. また, 時間距離, 位置距離のどちらをより優先するか重み付けによって, 形成されるクラスタは大きく変化する. クラスタリングのイメージを図 2 に示す.

3 提案手法に基づく実験

あるイベント E を撮影した写真サンプル 474 枚を用いて, 著者が定義したサブイベント (葉に相当するクラスタ) がイベントとして抽出されるように, 手動で閾値の調整を行った. ここで, イベント抽出の確かさを評価する指標として採用する, 適合率 (Precision) P と再現率 (Recall) R をそれぞれ定義する.

$$P = \max_{\hat{e} \in \hat{E}} \frac{|e \cap \hat{e}|}{|\hat{e}|}, \quad R = \max_{\hat{e} \in \hat{E}} \frac{|e \cap \hat{e}|}{|e|}$$

e は, 実際に起こったあるサブイベントを表し, $|e|$ は, サブイベント e に含まれる写真枚数を表す. \hat{E} は, 自動で分割されたサブイベント \hat{e} の全体集合を表す.

閾値を変化させて得られたサブイベントの適合率と再現率の平均値と最低値を表 1 に示す. データは, A 氏が撮影した写真 126 枚を用いた.

表 1: 閾値 (大小) の適合率と再現率

	閾値 (大)		閾値 (小)	
	適合率	再現率	適合率	再現率
平均値	0.43	1	0.71	0.93
最小値	0.04	1	0.04	0.39

表 1 に示すように, イベントの適合率と再現率は閾値によって変化する. 閾値を大きくとれば, 高い再現率を得られるが, 適合率は低くなる. 言い換えると,

複数のイベントを一つのイベントとみなすことが多くなる. 逆に閾値を小さくとると, 再現率を少し犠牲にするが, 高い適合率を得ることができる. つまり, 一つのイベントが複数のイベントとみなされることが増えてくる.

4 考察

今回の実験では, 閾値を手動で調整し, 各個人 (計 6 名) のデータに対して, 適合率と再現率のどちらも大きくなるようなクラスタリングを行った. 各個人毎に撮影したサブイベントに違いが生じ, 撮影間隔も異なるので, 適切なクラスタ間距離の閾値はそれぞれ異なる結果となった. つまり, 各個人の撮影状況に合わせた柔軟な閾値調整が必要となる.

また, 1 名分の場合と 6 名分の場合とを比較すると, 主要なサブイベント発生時にはほぼ全員が写真を撮影しているため, イベント撮影枚数が飛躍的に増加し, イベント構造がより顕著に現れている. つまり, 複数人でイベント写真の共有を行うことにより, 1 名分の画像から抽出したイベント構造に比べ, より正確なイベント構造の抽出を期待することができる.

5 おわりに

イベントで撮影した写真のもつ時間情報と位置情報を用いて, イベントを自動で構造化し, それに基づいて写真を管理する手法を提示した. それにより, 人がイベントを思い出すのにより近い形で, 写真の管理を行うことができ, また, 4 節の考察では, イベント構造を用いた画像共有への拡張の可能性も示すことができた. しかし, イベント構造化のために手動で閾値調整を行うことは困難を極め, 各個人に合わせた適切なイベント構造の自動抽出には未だ至っていない. 今後, 各個人の撮影枚数や撮影間隔に合わせた確からしいクラスタリングを行えるように, クラスタ間距離の閾値自動調整を可能にするつもりである. また, クラスタリングに用いる画像情報は, 撮影日時と GPS 情報のみにとどめているため, それから作られるイベント構造には限界がある. したがって, イベント構造をより適切に抽出するため, 撮影者, 被写体などを用いた手法 [5] などに倣い, 他の属性によるクラスタリングについて検討する.

参考文献

- [1] Picasa: <http://picasa.google.co.jp/>.
- [2] AJohn C. Platt, Mary Czerwinski, Brent A. Field: PhotoTOC: Automatic Clustering for Browsing Personal Photographs, *Fourth IEEE Pacific Rim Conference on Multimedia* (2003)
- [3] Mor Naaman, Yee Jiun Song, Andreas Paepcke, Hector GarciaMolina: Automatic Organization for Digital Photographs with Geographic Coordinates, *the 4th ACM/IEEE-CS joint conf. on Digital libraries*, pp. 53-62 (2004)
- [4] FKerry Rodden, Kenneth R. Wood: How Do People Manage Their Digital Photographs?, *the SIGCHI conf. on Human factors in computing systems*, pp. 409-416 (2003)
- [5] Fergal Monaghan, David O 'Sullivan: Automating Photo Annotation using Services and Ontologies, *the 7th International Conf. on Mobile Data Management (MDM'06)* (2006)