

# イベント構造の抽出に基づく画像管理法

## An Image Management Method Based on Extraction of an Event Structure

渡邊 裕子<sup>1\*</sup> 小林 一郎<sup>1</sup> 和泉 憲明<sup>2</sup> 橋田 浩一<sup>2</sup>

Yuko WATANABE,<sup>1</sup> Ichiro KOBAYASHI,<sup>1</sup> Noriaki IZUMI,<sup>2</sup> Kôiti HASIDA<sup>2</sup>

<sup>1</sup> お茶の水女子大学大学院人間文化創成科学研究科

<sup>1</sup> Ochanomizu University, Graduate School of Humanities and Sciences

<sup>2</sup> 産業技術総合研究所

<sup>2</sup> National Institute of Advanced Industrial Science and Technology (AIST)

**Abstract:** 本稿では、日常生活のイベントの中で撮影した画像群の管理を目的として、イベント構造に基づいた画像管理法を提案する。画像群の時間情報、位置情報から自動的に抽出したイベント構造を用いることによって、日常生活のイベントに基づいた画像管理を可能にする。

### 1 はじめに

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、管理する機会が増えている。それにともない、Picasa [1], photology [2] 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、各々の写真にコメントや Geo タグなどを自由に付与することができるようになった。

写真にタグを付ける方法は、大きく分けて、

1. 人による手動タグ付け
2. コンピュータによる自動タグ付け

の2通りの手法を考えることができる。人手で付与されたタグは、情報の質が高くタグが付けられた後に検索をするのには有効だが、タグの付与には多大な時間と根気が必要とされる。また、コンピュータによる自動タグ付けは、タグ付与の手間を大幅に軽減できるが、その結果はユーザが意図したものではないことが多々あったり、自動タグ付けの効果はコンテンツにより偏りが生じるため、様々な手法による事前学習が必要となってしまう。

このことから、本研究では付与されるタグ情報の質を保ちつつ、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、上記した人とコンピュータによるタグ付けを併用したクラスタリング手法を提案する。すなわち、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、ユーザにどのようなイベントであったか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。イベント構造に即して画像共有を行うことにより、今まで各個人がそれぞれの形で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を通じて画像を共有することができる。

本稿は、以下のように構成される。第2節では、関連する研究を紹介する。第3節では、提案手法について説明を行う。第4節では、イベント構造化の実験結果を示す。第5節では考察を行い、最後に、第6節では提案手法をまとめ、今後の課題を述べる。

### 2 関連研究

本研究に深く関係する研究である PhotoTOC[3] と PhotoCompas[4] について記す。

PhotoTOC[3] は、2つの異なるクラスタリング手法を用いた画像管理を提案している。基本的には写真の撮影日時に即した時間間隔に基づいたクラスタリングを用い、写真の時刻情報が欠落している場合には写真の色情報に基づいたクラスタリングを用いる手法である。この手法はユーザインタフェースを重視する手法で、個々の抽出されたイベントの確からしさを評価しているとは言えない。

PhotoCompas[4] は、写真の撮影日時と GPS 情報を併用し、イベントを構造化しており、本研究に類似した研究であるように見えるが、PhotoCompas は最初に時刻情報によるクラスタリングを行い、次に GPS 情報

\*連絡先：お茶の水女子大学 大学院人間文化創成科学研究科 理学専攻情報科学コース 小林研究室  
〒112-0012 東京都文京区大塚 2-1-1 理学部 3号館 506号室  
E-mail: yuko\_w@koba.is.ocha.ac.jp

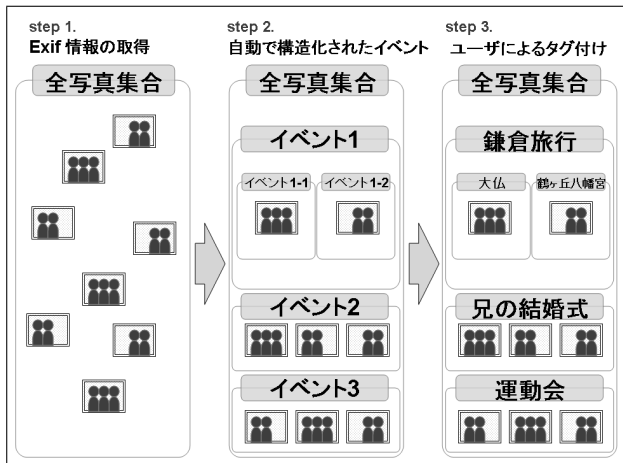


図 1: システム概要

によるクラスタリングの結果を用いて齟齬のないイベント階層構造を得る手法である。

本研究では、時刻情報と GPS から取得した位置情報の両方によって定義される距離を用いたクラスタリング手法を提案し、その性能を評価する。

### 3 イベント構造の抽出に基づく画像管理法

本節では、提案手法について説明を行う。

#### 3.1 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、メーカー独自情報など

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、撮影日時と GPS 情報を用いる。

#### 3.2 イベントの構造化

図 1 に、提案するシステムの概要図を示す。処理手順を以下に示す。

- step1. 画像ファイルの 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抽出を行う。
- step3. ユーザー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 [5]。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、画像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 2 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画像データのタイムスタンプを合わせることにより、プログラムに即して画像の管理を行う。イベントスケジュールがあらかじめ分かっている必要がある。

#### 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し、画像を管理することから、ボトムアップな手法による構造化を用いる。

ボトムアップな手法による構造化には、階層的クラスタリングが適しているため、本研究では、凝集型階層的クラスタリング [6] を用いる。以下に手順を示す。クラスタ間の距離の求め方は後述する。

- step 1.  $N$  個の写真からなるデータが与えられたとき、1 個の写真だけを含む  $N$  個のクラスタがある初期状態を作る。
- step 2. 新しく形成されたクラスタとその他のクラスタとの距離を求める。
- step 3. 全てのクラスタ間の距離のうち最も近い 2 つを結合して新しくクラスタを作る。
- step 4. 全てのクラスタが 1 つのクラスタに結合されるまで繰り返す。

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。本稿では、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  $(\lambda_i, \phi_i)$  をもつ。ここで、 $\lambda_i$  は経度、 $\phi_i$  は緯度を表す。本研究では群平均法を用いるので、各  $t_i, \lambda_i, \phi_i$  を次式で定義する。

$$t_i = \frac{1}{n_i} \sum_{j=1}^{n_i} t_{ij}$$

$$\lambda_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \lambda_{ij}$$

$$\phi_i = \frac{1}{n_i} \sum_{j=1}^{n_i} \phi_{ij}$$

ここで、 $n_i$  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  $C_1, C_2$  間の距離関数は、

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

で定義する。ここで、 $D_t, D_p$  はそれぞれクラスタ間の時間距離、位置距離であり、以下に定義する。また、 $w_t, w_p$  はそれぞれそれらに対する重みである。

各クラスタ  $C_1, C_2$  間の時間距離  $D_t(C_1, C_2)$  は、次式で定義する。

$$D_t(C_1, C_2) = |t_1 - t_2|$$

また、クラスタ  $C_1, C_2$  間の位置距離  $D_p(C_1, C_2)$  を次のように定義する。地球を完全な球体と仮定して求める。地球の半径を  $R = 6378137(\text{m})$  とする。このとき、経度の変位を  $\Delta\lambda = \lambda_2 - \lambda_1$ 、緯度の変位を  $\Delta\phi = \phi_2 - \phi_1$  と表す。東西に  $x$  軸、南北に  $y$  軸を取ったとき、 $x$  軸方向の変位を  $\Delta x = R\Delta\lambda \cos \phi_1$ 、 $y$  軸方向の変位を  $\Delta y = R\Delta\phi$  と表すことができ、 $C_1, C_2$  間の位置距離  $D_p(C_1, C_2)$  は、次式で定義することができる。

$$D_p(C_1, C_2) = \sqrt{\Delta x^2 + \Delta y^2}$$

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しきい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の重み付けによって、形成されるクラスタは大きく変化する。

例えば、観光地巡りや遊園地など大幅な場所移動を伴うイベントでは、位置距離に重きをおいたクラスタ

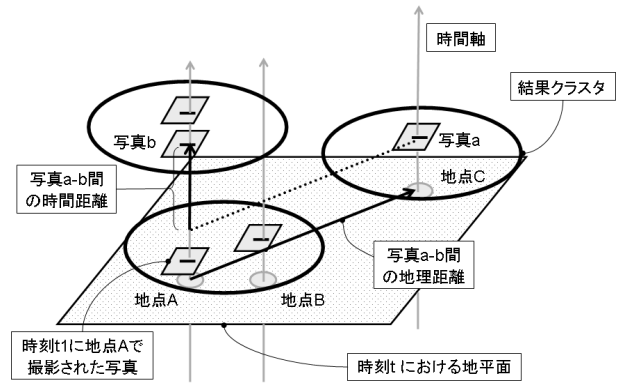


図 2: クラスタリングイメージ図

表 1: イベントの個人別撮影枚数

| カメラマン    | A  | K   | M  | T  | Y  | W   |
|----------|----|-----|----|----|----|-----|
| 撮影枚数 (枚) | 53 | 147 | 30 | 58 | 60 | 126 |

リングを行えば、時間距離に重きをおいたクラスタリングの結果に対して、地理的位置関係の近いものを優先したイベント構造を構成することができる。クラスタリングのイメージ図を図 2 に示す。

## 4 提案手法に基づく実験

実験に用いるサンプルは、2008年7月30日～31日にて、静岡県伊東市で行われた研究室合宿を撮影したものである。室外の写真には、SONYのGPSユニットキットGPS-CS1K、室内の写真には、Picasa[1]とGoogle Earth[7]を用いて、全ての写真にGPS情報を付与した。写真は、6名によって撮影された計474枚を使用する。イベントの個人別撮影枚数を表1に示す。また、旅行中に行われた全てのサブイベントを著者が定義したものを表2に示し、サブイベントの中で伊東駅周辺で発生したサブイベントの位置関係図を図3に示す。

表2に示したサブイベント(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、手動でしきい値の調整を行った。ここで、イベント抽出の確からし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)  $R$  をそれぞれ式(1)、式(2)に定義する。

$e$  は、あるイベントを表し、 $|e|$  は、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の全体集合を表す。

$$P = \max_{\hat{e} \in \hat{E}} \frac{|e \cap \hat{e}|}{|\hat{e}|} \quad (1)$$

$$R = \max_{\hat{e} \in \hat{E}} \frac{|e \cap \hat{e}|}{|e|} \quad (2)$$

表 2: 全サブイベント

| 2008年7月30日     | 2008年7月31日                       |
|----------------|----------------------------------|
| 0 池袋駅<br>-移動-  | 8 宿：朝食<br>-移動-                   |
| 1 伊東駅<br>-移動-  | 9 海賊船ゆーみん<br>10 伊東マリンタウン<br>-移動- |
| 2 昼食<br>-移動-   | 11 昼食<br>12 解散                   |
| 3 宿            | -                                |
| 4 宿：着手報告会      | -                                |
| 5 宿：夕食<br>-移動- | -                                |
| 6 花火大会<br>-移動- | -                                |
| 7 宿：コンパ        | -                                |

表 3: しきい値 (大小) の適合率と再現率

| イベント           | しきい値 (大)      |              | しきい値 (小)      |              |
|----------------|---------------|--------------|---------------|--------------|
|                | 適合率<br>(枚/枚)  | 再現率<br>(枚/枚) | 適合率<br>(枚/枚)  | 再現率<br>(枚/枚) |
| 0 池袋駅<br>-移動-  | 1/1<br>4/21   | 1/1<br>4/4   | 1/1<br>4/4    | 1/1<br>4/4   |
| 1 伊東駅<br>-移動-  | -<br>1/21     | -<br>1/1     | -<br>1/7      | -<br>1/1     |
| 2 昼食<br>-移動-   | 6/21<br>3/21  | 6/6<br>3/3   | 6/7<br>3/10   | 6/6<br>3/3   |
| 3 宿            | 7/21          | 7/7          | 7/10          | 7/7          |
| 4 着手報告         | 18/26         | 18/18        | 18/18         | 7/18         |
| 5 夕食<br>-移動-   | 8/26<br>-     | 8/8<br>-     | 8/8<br>-      | 8/8<br>-     |
| 6 花火大会<br>-移動- | 48/50<br>2/50 | 48/48<br>2/2 | 48/50<br>2/50 | 48/48<br>2/2 |
| 7 コンパ          | 8/8           | 8/8          | 8/8           | 5/8          |
| 8 朝食<br>-移動-   | -<br>-        | -<br>-       | -<br>-        | -<br>-       |
| 9 海賊船          | 14/20         | 14/14        | 14/14         | 14/14        |
| 10 MT<br>-移動-  | 4/20<br>-     | 4/4<br>-     | 4/4<br>-      | 4/4<br>-     |
| 11 昼食          | 2/20          | 2/2          | 2/2           | 2/2          |
| 12 解散          | -             | -            | -             | -            |

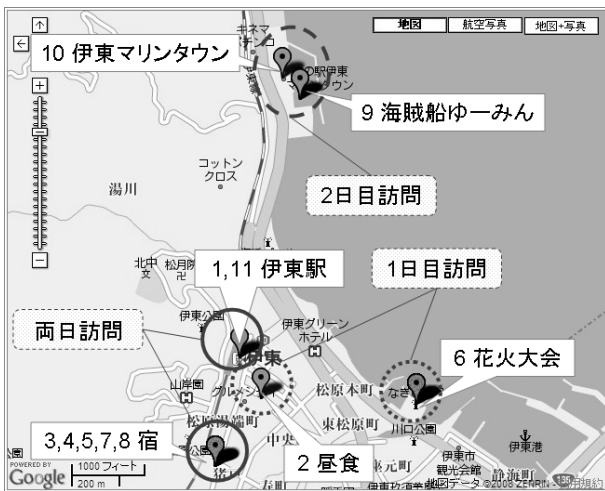


図 3: サブイベントの地理的位置

しきい値を変化させて得られたイベント毎の適合率と再現率を表 3 に示す。データは、W 氏の写真を用いた。

表 3 に示すように、イベントの適合率と再現率はしきい値によって変化する。しきい値を大きくとれば、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 1 つのイベントとみなすことが多くなる。逆にしきい値を小さくとると、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、1 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。

表 4 は、個人別に撮影したサブイベントをまとめた。

## 5 考察

### 実験結果の考察

今回の実験では、しきい値を手動で調整し、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十分に大きいクラスタリングを行った。

表 4 に示すように各個人毎に撮影したサブイベントに違いが生じた。撮影したサブイベントが異なれば撮影間隔も異なるので、適切なクラスタ間距離のしきい値はそれぞれ異なる結果となった。同じイベントを撮影していても、しきい値は一意に定まらないことが分かる。つまり、各個人の撮影状況に合わせたしきい値調整が必要となる。

また、複数のイベントを 1 つのイベントとみなす場合にはイベントを細分化し、1 つのイベントを複数のイベントとみなす場合にはイベントを統合するように、しきい値を動的に決定する柔軟な処理が求められる。例えば、表 3 を見ると、イベント 3, 4, 5 は時間的に連続し、場所移動もないため、しきい値を大きくとると、適合率が下がり、1 つのイベントとみなされてしまう。逆に、しきい値を小さくすれば、イベント 3, 4, 5 はそれぞれ 1 つのイベントとみなされるが、再現率が下がり、イベント 4 が複数のイベントに分断されてしまう。このような状況に対応するため、柔軟な構造化が必要と

表 4: 個人別サブイベント撮影有無

| イベント   | A  | K   | M  | T  | Y  | W   |
|--------|----|-----|----|----|----|-----|
| 0 池袋駅  | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 1 伊東駅  | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 2 昼食   | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 3 宿    | -  | -   | -  | -  | -  | -   |
| 4 着手報告 | -  | -   | -  | -  | -  | -   |
| 5 夕食   | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 6 花火大会 | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 7 コンパ  | -  | -   | -  | -  | -  | -   |
| 8 朝食   | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 9 海賊船  | -  | -   | -  | -  | -  | -   |
| 10 MT  | -  | -   | -  | -  | -  | -   |
| -移動-   | -  | -   | -  | -  | -  | -   |
| 11 昼食  | -  | -   | -  | -  | -  | -   |
| 12 解散  | -  | -   | -  | -  | -  | -   |
| 総撮影枚数  | 53 | 147 | 30 | 58 | 60 | 127 |

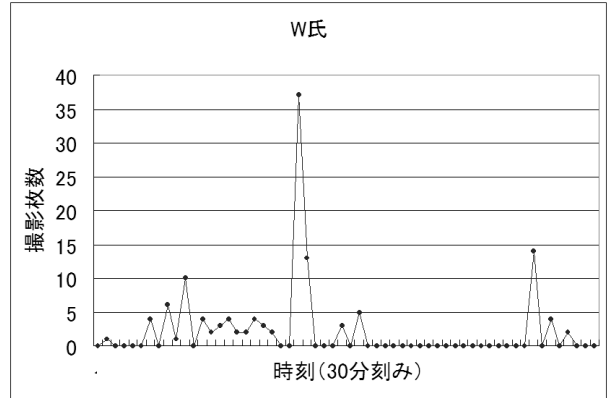


図 4: W 氏の撮影推移

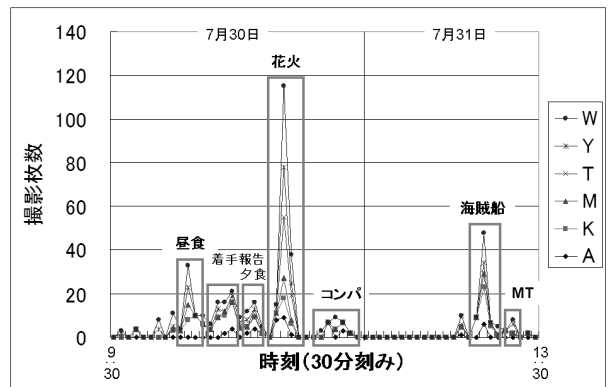


図 5: 6 名分の撮影推移

なる。

#### 画像共有への拡張に関する考察

図 4 は、W 氏のデータグラフ、図 5 は、6 名分のデータを積み上げたグラフである。グラフの X 軸が時間幅 30 分刻みの時間軸、Y 軸が撮影枚数 (枚) を表している。

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 (花火や海賊船など) 発生時には、撮影者 6 名中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、複数人でイベント画像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。

## 6 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、5 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でしきい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、クラスタ間距離のしきい値自動調整を可能にするつもりである。さらに、現在の手法では困難である、柔軟な構造化を実現するため、クラスタリング手法におけるしきい値やクラスタ間の距離関数の見直しが必要である。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 [8] 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] photology: <http://www.getphotology.com/>

- [3] AJohn C. Platt, Mary Czerwinski, Brent A. Field: PhotoTOC: Automatic Clustering for Browsing Personal Photographs, *Fourth IEEE Pacific Rim Conference on Multimedia* (2003)
- [4] 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)
- [5] 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)
- [6] 神鷹 敏弘, “データマイニング分野のクラスタリング手法 (1) - クラスタリングを使ってみよう! - ”, *人工知能学会誌*, vol. 18, no. 1, pp. 59–65 (2003)
- [7] Google Earth: <http://earth.google.co.jp/>
- [8] Fergal Monaghan, David O 'Sullivan: Automating Photo Annotation using Services and Ontologies, *the 7th International Conf. on Mobile Data Management (MDM'06)* (2006)