

写真共有サイトと観光Webサイトの  
マッシュアップに基づく観光情報提供システム

長尾光悦

北海道情報大学

Tourism Information System based on  
Mashup of Photo Sharing Site and Tourism Web Sites

Mitsuyoshi NAGAO

Hokkaido Information University

平成25年3月

北海道情報大学紀要 第24巻 第2号別刷

## 〈研究ノート〉

写真共有サイトと観光 Web サイトの  
マッシュアップに基づく観光情報提供システム

長尾光悦

Tourism Information System based on  
Mashup of Photo Sharing Site and Tourism Web Sites

Mitsuyoshi NAGAO

## 要旨

近年、映画やドラマのロケ地を巡るフィルムツーリズム、アニメ作品と関連のある土地を訪問する聖地巡礼観光、健康維持に主眼を置いたヘルスツーリズムなど、多種多様な旅行形態が生まれている。加えて、現在、団体旅行は減少の一途を辿り、家族や小グループ単位で旅行をする旅行の個人化が進んでいる。このような状況下において、我が国の観光を推進・発展させるためには、多様化する観光ニーズに適合した観光情報を提供可能とする必要がある。本稿では、写真共有サイトと観光 Web サイトのマッシュアップに基づく観光情報提供システムを提案する。ここでは、写真共有サイトに投稿されたジオタグと呼ばれる位置情報を持つ写真データをクラスタリングに基づき分析し、更に、観光 Web サイトから獲得される観光情報を組み合わせることによって、最新かつ多種多様なニーズに対応した観光情報を提供可能なシステムを実現する。また、提案システムのプロトタイプを構築し、実験に基づき提案システムの有効性を検証する。

## 欧文抄録

Recently, various types of tourism such as film tourism, pilgrimage to sacred places for anime fans and health tourism have appeared. In film tourism, tourists visit locations where shooting of film was performed, and tourists of anime fans visit locations related to Japanese animated works in pilgrimage to sacred places. Health tourism means tourism which focus on maintenance of good health. Moreover, the number of group tours which were main type of tourism in Japan keeps decreasing, and the personalization of tour that travels in units of family or small group is ahead. Under this situation, it is required to provide tourism information which respond to needs of tourists becoming diversified in order to promote and develop current tourism in Japan. In this paper, we propose a tourism information system based on mashup of photo sharing site and tourism web sites. Here, we perform an analysis based on clustering of photo data having position information called geotags in photo sharing site. In addition, we make possible to provide latest tourism information corresponding to various needs of tourists by combining the results of clustering and tourism information extracted from tourism websites. Finally, we construct a prototype system and then confirm the effectiveness of proposed system through an experiment.

## 1. はじめに

近年、映画やドラマのロケ地を訪問するフィルムツーリズムや、漫画やアニメの熱心なファンが作品と縁のある土地を訪問する聖地巡礼観光、更には、温泉療法や森林療法といった健康維持に主眼を置いたヘルスツーリズムなど多種多様な観光形態が生まれており、旅行者の観光に対するニーズは多種多様化してきている。加えて、我が国における主流の旅行形態であった団体旅行は減少の一途をたどっており、現在では、小グループや家族単位での個人旅行が主流となっている。このような個人旅行においては、団体旅行のような有名観光地のみを巡る観光ではなく、個人がテーマや目的を持ち、それに沿った観光地を訪問する。

このように、旅行形態の多様化と個人化が進む中、今後、我が国における観光を推進し、発展させるためには、個々のニーズに適合した多種多様な観光情報を提供していくことが必要とされることは明らかである。しかしながら、現在、観光雑誌等で提供されている観光情報は、多数の旅行者が望む、いわゆるステレオタイプの観光情報であり、多様化する観光ニーズに対応した観光情報を提供できているとは言い難い。

本稿では、写真共有サイトと観光 Web サイトのマッシュアップに基づく観光情報提供システムを提案する。写真共有サイトとは、ネットワークを通してユーザが撮影した写真を共有するためのサービスである。代表的なサイトとしては、フォト蔵、Flickr、Panoramio などが挙げられ、これらサイトにおいては、多くのユーザが撮影した写真を公開し、日々、膨大な量の写真データが蓄積されている。このような写真共有サイトにおいて、撮影した写真にジオタグと呼ばれる撮影場所に関するメタ情報を付加し、地図上で撮影した写真を共有することができるものが存在する。本研究では、写真共有サイトに投稿された位置情報を持つ写真データを利用し、更に、観光 Web サイトにおいて提供されている観光情報を組み合

わせることによって最新かつ多種多様なニーズに対応した観光情報を提供可能なシステムを実現する。具体的には、ジオタグを利用することが可能な写真共有サイトとして、Google Map や Google Earth においても採用されている Panoramio から写真データを取得し、写真データの持つ位置情報を使用したクラスタリングに基づく分析を行う。更に、トリップアドバイザー等の観光 Web サイトから観光情報の抽出を行い、抽出された観光情報とクラスタリング結果とのマッチング処理を行うことによって、効果的な観光情報の発見・提供を実現する。また、提案システムのプロトタイプを構築し、実験に基づき提案システムの有効性を検証する。

## 2. 写真共有サイト

現在、インターネットにおいて情報を公開・共有するための数多くのサービスが提供されており、それらサービスを利用することによって情報技術に関する特別な知識を有していなくても個人が簡単にインターネット上で情報の公開・共有を行うことができるようになっている。例えば、Ameba ブログ等のブログサービスを利用することによって日記のような情報を簡単に発信することができ、YouTube 等の動画投稿サイトを利用することによって動画や音声などのマルチメディア情報を発信することも可能である。このような情報発信の簡易化に伴い、インターネット上には、日々、多様かつ膨大な量の情報が蓄積されている。このインターネット上に存在する膨大な情報の中には有益な情報が多く含まれている可能性があり、このような情報を分析し、有益な情報を抽出するといった研究も多く実施されている [1][2][3]。

一方、これらインターネットにおける情報共有・発信サービスの中で、写真を扱うものは、写真共有サイトと呼ばれる。代表的な写真共有サイトとしては Panoramio、Flickr、Picasa、フォト

蔵などが挙げられる。写真共有サイトにおいては、ユーザが自身で撮影した写真データをインターネット上で公開し、他のユーザが閲覧、コメントすることが可能となっている。写真共有サイトの利用者は、増加傾向にあり、例えば、Panoramioであれば、2007年の時点で、ユーザ数は100万人以上、投稿写真データ数は500万枚以上となっている。

このような写真共有サイトの中で Panoramio と Flickr は、写真データにジオタグと呼ばれる緯度経度情報を付与することが可能となっている。また、この緯度経度情報によって写真データを地図上に配置し、共有することを可能としている。図1、及び、図2に Panoramio と Flickr における写真データの地図上での表示例を示す。図1に示されるように、Panoramio の場合には、Google社が提供している Google Maps と連携しており、この地図上で写真データを確認することが可能になっている。図2に示される Flickr においては、独自の地図を採用しており、地図上での写真データの表示方法も Panoramio とは異なっている。

近年、このようなジオタグを利用可能な写真共有サイトが観光の分野において注目され、写真データに基づき観光地を推定する研究等が行われている。例えば、Girardin らは、Flickr において、各ユーザが撮影した写真データの位置情報とタイムスタンプを利用し、旅行者の観光ルートの推定、及び、旅行者が集中している地点の分析を行っている[4][5]。また、Kurata は、Flickr をベースとして、横浜地区において写真が数多く撮影されている地点を潜在的な観光地とみなし、これをカーネル密度分析に基づき発見する研究を行っている [6]。王らは、Flickr 及び Panoramio からジオタグ付き画像を収集し、クラスタリングを行い、画像分類を行うことによって代表画像を決定し、その地点における代表画像を表示した観光マップの作成方法を提案している[7]。

しかしながら、これら研究においては、写真が



図1 Google Map 上に表示される Panoramio の写真データ

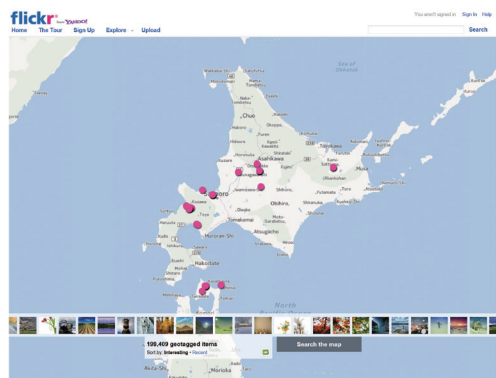


図2 地図上に表示される Flickr の写真データ

多く取られた場所を分析し、それを観光地として取り扱っており、その地点が、どのような観光地であるのか、すなわち、雑誌や Web サイトで紹介されるような著名な観光地であるのか、単に写真が多く取られているだけの観光地であるのか、更には、聖地巡礼などに代表される一般的な旅行者にとっては有益な観光情報とはならない趣味嗜好性が強い観光地であるのかといった情報は提供されず、ユーザがその場所の写真を確認することによって判断しなければならない。

そこで、本研究では、写真データにおける位置情報に基づく分析に加え、観光情報を提供している Web サイトから情報を抽出し、これらのマッシュアップを行うことによって、その地点がどの

ような観光地であるのかを分類し、情報の提供を行う。

また、ここでは、Panoramio をベースとした観光情報の提供を行う。Panoramio においては、指定範囲内の写真を検索することが可能であり、そのための API が提供されている。API を利用した写真検索を実行するためには、表 1 に示されるような検索範囲のための緯度経度情報や検索する写真のサイズといったパラメータを包含した HTTP リクエストを Panoramio に対して送信する。これにより、指定範囲内に含まれる写真に関する情報が返信される。レスポンスデータ内には、写真の撮影日時、撮影ユーザ名、写真 ID、写真が撮影された位置情報、写真のサイズ情報、写真のタイトル、撮影ユーザ URL、ユーザ ID、写真ファイル URL、写真 URL が含まれる。

Flickr も Panoramio と同様に写真データに対してジオタグを付与することが可能な写真共有サイトであり、REST により写真データを検索可能とする API も提供されている。しかしながら、Flickr においては、緯度経度情報のみによる写真データの検索が困難であり、関連するキーワードに基づき検索を行い、その中で指定された範囲の緯度経度情報を持つ写真データを獲得することが可能になっている。このように写真データの検索がキーワードをベースとするため、収集される写真データに偏りが出る可能性が考えられる。よって、本研究では、Panoramio を採用することとした。

### 3. 予備実験

本研究では、Panoramio において提供されている写真データをベースとして、旅行者の多様なニーズに適合する観光情報を提供することを目的としている。これを実現するための第一の前提として Panoramio における写真データとして、著名な観光地に関する写真データが提供されているのかを検証する必要がある。このための予備実

表 1 Panoramio における検索パラメータ

order	写真の表示順序 popularity (人気), upload_date (日付) から選択
set	人気のある写真(public),全ての写真(full),特定ユーザが投稿した写真(userID)から選択可能
from, to	検索写真数
minx, miny	検索範囲 (最小経度・緯度)
maxx, maxy	検索範囲 (最大経度・緯度)
size	original, mudium (default) , small, thumbnail, square, mini_squre から選択
mapfilter	品質や位置に基づく写真のフィルタリング

験を実施した。ここでは、観光雑誌、及び、観光情報を提供している Web サイトから観光地に関する情報を収集し、それらの観光地が撮影されている写真が存在するかを目視により検証した。検証地域としては札幌市を選定した。

予備実験において利用した観光雑誌は、ダイヤモンド社により発行されている札幌、旭川、小樽といった道内主要観光都市の観光情報が提供されている「地球の歩き方 MOOK 北海道の歩き方 2012」を利用した。この雑誌の中で札幌市の観光情報のみ抽出するものとした (図 3)。また、観光情報を取得するための Web サイトとしては、「ようこそ札幌」、「トリップアドバイザー」、「るるぶ.com」を利用した。ようこそ札幌は、札幌市が運営する公式観光情報提供サイトである。このサイト内において観光名所として提供されている情報を利用した (図 4)。トリップアドバイザーは、世界最大の口コミ情報を提供する旅行サイトであり、ホテルや観光地に関する情報を多く掲載している。この中で観光名所として提供されている情報を抽出することとした (図 5)。るるぶ.com



図3 予備実験における観光ガイドブック



図5 トリップアドバイザー



図4 ようこそ札幌



図6 るるぶ.com

は、JTB パブリッシングが運営する旅行雑誌として有名な「るるぶ」のための公式サイトであり、観光地に関する情報提供だけではなく、航空券や宿泊の予約なども可能な総合旅行サイトである。このサイトにおいて提供されている観光情報の中で、「見る」、「遊ぶ」、「買う」、「立ち寄り湯」、「定番」のカテゴリに分類されている情報を利用することとした(図6)。これらの雑誌、及び、Web サイトから抽出された情報としては、「大通公園」、「北海道庁旧本庁舎」、「時計台」といった有名観光地の他、「札幌市民防災センター」、「エ

リエールスクエア札幌渡辺淳一文学館」、「北欧館パン博物館」など、札幌市民にも認知度がそれ程高くないであろう観光地に関する情報も抽出された。

各媒体から抽出された情報と Panoramio において提供されている写真との照合結果を表1に示す。表1に示されるように、観光雑誌において提供されている観光情報は33箇所、ようこそ札幌では102箇所、トリップアドバイザーでは、77箇所、るるぶでは、70箇所が抽出された。雑誌の場合には、紙面の都合上、提供される情報がWebと比較して、2から3分の1程度であることが確

表 2 観光地の照合結果

	観光雑誌	ようこそ	トリップ	るるぶ
観光地数	33	102	77	70
一致数	32	83	85	54
一致率	97.0%	81.4%	84.4%	77.1%

※ようこそ：ようこそ札幌  
 ※トリップ：トリップアドバイザー  
 ※るるぶ：るるぶ.com

認められた。このことから示されるように、雑誌の場合には、掲載できる観光地が限られるため、厳選された主要な観光地情報のみが提供されていた。一方、Web では、提供可能な情報量に制限がないため認知度の低い観光情報も提供されていた。また、雑誌に掲載された情報の 84.8% の情報が Web サイトで提供されていることが確認された。

照合結果として、雑誌に掲載されるような著名な観光地の 97.0% が Panoramio 上の写真データとして提供されており、更に、Web サイトで提供されるような比較的認知度の低い観光地に対応する写真データでも、80%程度と高い割合で存在することが確認された。この予備実験の結果から、Panoramio をベースとして、主要な観光地に関する情報を提供できるシステムを実現可能であることが示された。

#### 4. プロトタイプシステムの構築

Panoramio と観光 Web サイトのマッシュアップに基づき観光情報を提供するためのプロトタイプシステムの構築を行った。図 7 にシステムの概要を示す。本システムでは、写真に付与された緯度経度情報を利用したクラスタリングに基づく分析、及び、観光 Web サイトから収集した観光地名と写真データに付与されたタイトル情報とのマッチングの二種類の操作により観光地に関する情報を、その地点がどのような種別の観光地であるかの情報を付与した形式で提供する。

このため、本システムでは、まず Panoramio

からの写真データの収集を行う。ここでは、観光情報を提供するための市町村を指定し、その市町村における役所を中心とした指定範囲内の写真データの収集を行う。続いて、収集された写真データの観光地の種別を分析するために、外部の観光 Web サイトから情報の収集を行う。収集情報としては、指定された地域に存在する観光地の名称情報が挙げられる。この名称に対して、形態素解析を適用し、それに含まれる名詞情報を抽出する。形態素解析においては、オープンソースの形態素解析エンジンである Mecab を利用した[8]。

Panoramio は世界的に利用されている写真共有サイトであり、写真の投稿者が外国人の場合には、写真データに付与されているタイトル情報は英文である可能性が高い。このため、観光 Web サイトから獲得された観光地名に対して英文翻訳を適用し、英文名への変換を行う。この操作のために、Microsoft 社が提供している翻訳 API である Microsoft Translator を利用する。この API では、自身のアプリケーションのための ID (AppID)、翻訳するテキスト (text)、翻訳前の言語 (from)、翻訳後の言語 (to) のパラメータを含む HTTP リクエストを送信することによって翻訳結果を受け取ることが可能である。

続いて、収集された写真データの位置情報を利用し、各写真データ間の距離を算出する。算出された距離情報に基づきクラスタリングを行う。各クラスタに含まれる写真データ数が一定数を越えた場合には、訪問者が多い潜在的観光地として判定される。また、クラスタ内に含まれる写真データ数が一定数を越えるが、同一のユーザによって撮影された写真が一定割合以上含まれる場合には、趣味嗜好性の強い観光地として判定される。各クラスタにおいては、クラスタに含まれる写真データの全タイトル情報に対して形態素解析が適用され、各名詞の出現数が算出される。クラスタに含まれる写真データのタイトルにおいて、出現頻度が高い名詞を最も多く含むものが、そのク

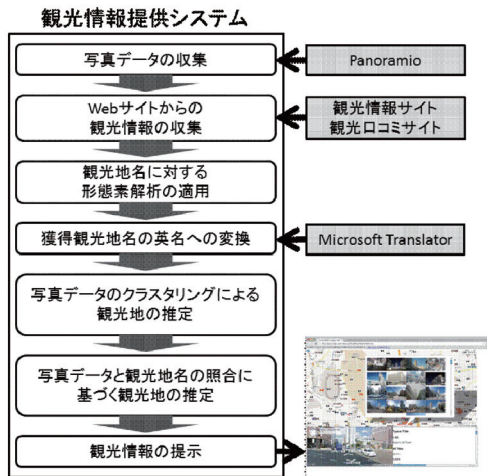


図7 プロトタイプシステムの概要

ラスタにおける代表タイトルとして採用される。

次に、形態素解析により得られた名詞情報及び英名称情報と写真データに付与されたタイトル情報とのマッチングを行う。マッチングにおいては、クラスタにおける代表タイトル、または、写真データに付与されたタイトル情報において形態素解析により得られた名詞情報、または、英単語情報が一定の割合以上含まれ、かつ、タイトル長が含まれた名詞または英単語長の合計長と比較し一定の割合以内であった場合に観光地名と一致するものと判定される。

最後に、分析された情報が写真データと共に Google Map をベースとしたインターフェース上に表示される。インターフェース上には、「観光 Web サイトに情報が掲載されている著名な観光地」、「写真が多く撮影されている潜在的な観光地」、「観光 Web サイトに情報が掲載され、かつ、写真が多く撮影されている観光地」、「趣味嗜好性の強い観光地」に分類され、対応するマーカーが配置される。また、マーカーをクリックすることによって、その地点に含まれる写真データ、周辺のストリートビュー画像、写真を撮影したユーザ数やタイトル情報等の付加情報が表示される。

## 5. 実験

本提案システムの有効性を検証するために実験を行った。実験では、予備実験と同様に札幌市を取り上げた。また、予備実験において、雑誌、及び、3つの Web サイト上に掲載されていた観光地のうち、Panoramio 上に写真データが存在することが確認された 107 箇所を利用し、本提案システムによって抽出された観光地情報との一致を検証した。また、この検証においては、本システムにより抽出された観光情報との一致は目視により判別することとした。更に、潜在的な観光地、及び、趣味嗜好性の強い観光地として、どのような情報が抽出されるかを検証した。

Panoramio における写真データの収集では、札幌市役所から東西南北へ 20km の範囲の写真データを収集した。これにより 8,312 枚の写真データが収集された。また、観光地名を収集するための外部 Web サイトとしては、予備実験で利用した 3 サイトから情報収集を行うこととした。更に、クラスタリングにおいては、100m 以内に位置する写真データをクラスタリングすることとした。この時、クラスタリングされた写真数が 10 枚以上の場合に、潜在的な観光地として判定するものとし、この中で、70% 以上の写真データが同一ユーザにより撮影されている場合には、趣味嗜好性の強い観光地として判定した。Web において提供されている観光情報との一致判定においては、観光地名から得られた名詞情報の 70% 以上がタイトルに含まれ、かつ、写真データのタイトル長が一致した名詞の合計長の 2 倍以内の場合に一致と判定した。これは、「○○○○○から見たテレビ塔」といったタイトルを有する写真データが観光地として誤判定されることを防ぐために設定したものである。これらのパラメータはヒューリスティックにより決定した。また、比較のため収集された観光地名に対して形態素解析を適用せず、観光地名が写真タイトルに完全に包含され、かつ、



タイトル長が観光地名の2倍を超えない場合に観光地として判定する実験を併せて実施した。

## 6. 実験結果

実験結果を、表3、図8、図9に示す。Panoramioにおいて収集された写真データは、本提案システムにおけるクラスタリング処理により、2,498箇所集約された。この中で、表3に示されるように観光地であると判別されたものは、240箇所であった。これらの観光地と雑誌や観光Webサイトで提供されている107箇所の観光地の一致を検証した結果、107箇所中88箇所、82.2%の観光地に関する情報が本システムによって提供可能であることが示された。本来であれば、全ての観光地に対応する写真データが存在するため、抽出率は100%とならなければならない。しかしながら、写真データには、タイトルが付与されていないもの、誤ったタイトルが付与されているもの、英語と日本語が混在したもの等が存在する。このため、観光Webサイトから抽出された名詞情報との一致率が低い、タイトルが長すぎ一致と判定されないといったことが発生したため、全ての観光地が抽出されなかった。しかしながら、現在のパラメータ設定において、観光Webサイトから得られた名詞・英単語情報との一致率が0%だったものは3カ所であった。この原因は、観光地が写真データの収集範囲に入っていなかった、著名な観光地の一部であるため、その観光地に包含された、異なるタイトルが付与されており名詞情報と全く一致しなかったというものである。したがって、この問題は、マッチングにおけるパラメータを変更することによって、理論的には、97.2%（107箇所中104箇所）まで上昇させることができる。しかしながら、名詞の一致率を下げるといったパラメータ調整を行うことによって、観光地情報との一致率は上昇するが、誤抽出も増加する。一致率と誤抽出はトレードオフとなるため、適切なパ

表3 観光地照合実験の結果

	抽出数	一致数	一致率
完全一致	213	79	73.8%
形態素解析	240	88	82.2%

ラメータ設定を発見する必要がある。

また、完全一致によって、観光Webサイトにおいて提供されている観光情報との一致を判定した場合、観光地として判定されるものは213箇所と減少し、一致率も73.8%と、形態素解析を利用した場合と比較して、低くなることが示された。これは、前述の写真データに付与されたタイトル表現の揺らぎが原因である。このため、Panoramioのような様々なユーザによって投稿される写真データから観光地に関する情報を抽出する場合には、本システムにおける形態素解析といった処理が有効に働いたことが明らかとなった。

図8に、写真が多く撮影されている潜在的観光地として判定された地点のいくつかを示す。図における上段は「幌平橋」、中段は「札幌護国神社」、下段は「北海道大学内銀杏並木」の写真がクラスタリングされたものである。幌平橋は、観光雑誌やWebサイトには掲載されづらいローカルな橋であるが、端にアーチ状の階段がかかっており、そこから札幌の有名河川である豊平川が見渡せる札幌市民にとっては馴染みの深い橋である。また、幌平橋は、映画「ガメラ2 レギオン襲来」において撮影にも利用されており、これにより多くの写真撮影が行われたために本システムによって抽出されたと考えられる。

また、札幌護国神社は、有名な観光地である中島公園の一角に存在する神社であり、秋には紅葉が美しく、穴場的な観光地として知られている。このような観光雑誌やWebサイトには掲載されづらい穴場的な観光情報も抽出されることが確認された。



図 8 潜在的な観光地

更に、北海道大学は、札幌市における有数の観光地であり、その敷地面積は約 178 万平方メートルと広大なキャンパスである。北海道大学の中には、クラーク像、ポプラ並木など著名な観光スポットが点在しているが、観光雑誌や観光 Web サイトでは、北海道大学として紹介される。本シ

テムでは、北海道大学の中において地元民には有名な銀杏並木というローカルな観光名所も個別に抽出可能であることが確認された。

一方、図 9 に、趣味嗜好性の強い観光地として抽出された地点のいくつかを示す。図における上段は、「北 24 条付近」、中段は「苗穂駅付近」、下段は「野幌森林公園」を撮影した写真が集約されたものである。札幌の歓楽街という北日本最大の歓楽街である「すすきの」が有名である。すすきのは観光雑誌や Web サイトなどでも必ず取り上げられる札幌における有数の観光地の一つである。これに対して、北 24 条は、全国的には有名ではない歓楽街であるが札幌市民にはなじみの深い歓楽街であり、そのため特定のユーザが多くの写真を撮影したためシステムによって抽出された。

また、苗穂駅付近は、鉄道の撮影ポイントであり、加えて、北海道鉄道技術館がある。この建物は、「さっぽろ・ふるさと文化百選」にも選ばれているものでもある。しかしながら、この苗穂駅付近は一般にはあまり周知宣伝されていない。一人のユーザが多くの鉄道写真を撮影していたため本システムにより趣味嗜好性の強い観光地として抽出された。

野幌森林公園は、札幌市、江別市、北広島市の 3 市にまたがる大規模な森林公園であり、その一部である野幌原生林は特別天然記念物にも指定されている。都市近隣に存在する平地林としては有名であり、600 種以上の野草やキノコ、140 種以上の鳥が生息する自然の宝庫となっている。しかしながら、札幌市中心部からは離れているため観光地としての知名度は高くはない。にもかかわらず、野鳥などの撮影を趣味とするユーザが多くの写真を投稿したため、趣味嗜好性の強い観光地として抽出された。これらから、提案システムを利用することによって観光地としての知名度は高くないものの一部の趣味嗜好性の強いユーザにとっては観光情報となりうる情報が抽出可能

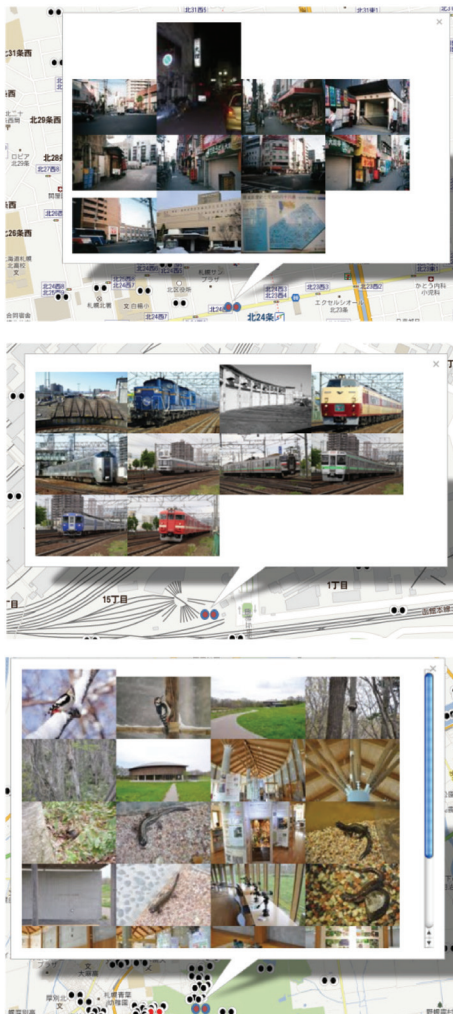


図9 趣味嗜好性の強い観光地

であることが確認された。

以上の実験結果から、提案システムは、知名度の高い観光地に関する情報から、趣味嗜好性の強い観光地に関する情報まで、多様化する旅行者ニーズに対応する観光情報を提供可能であることが確認された。

## 7. おわりに

本稿では、多様な観光ニーズに対応した観光情

報の提供を可能とするために、写真共有サイトと観光 Web サイトのマッシュアップに基づく観光情報の提供システムを提案した。提案システムでは、Panoramio においてジオタグを持つ写真データを収集し、位置情報に基づくクラスタリングを実施した。この結果に基づき、多くの写真が撮影されている潜在的観光地、または、趣味嗜好性の強い観光地の判別を行った。更に、観光 Web サイトから情報を収集し、観光地の名称データと写真データに付与されたタイトル情報とのマッチングを行うことによって知名度の高い観光地であるか否かの判別を行った。

実験においては、札幌市における観光地情報の抽出を行い、抽出された観光地情報を検証することで提案システムの有効性を確認した。実験結果から、提案システムを利用することによって、写真共有サイトをベースとして多種多様な観光ニーズを満たす観光情報を抽出・提供可能であることが確認された。

本提案システムで利用した Panoramio の写真データの中には、商品などの写真、ユーザ自身を撮影した写真など、観光地の情報としては不適切なものも含まれる。このため、今後の課題としては、写真データを分析し、観光情報を提供するために適切な写真データの分類の実現が挙げられる。更に、他の地域での提案システムの有効性の検証する予定である。

## 参考文献

- [1] 三田村保, 斎藤翔太, 大堀隆文: “Web マイニングによる北海道観光情報の調査分析”, 観光情報学会誌, Vol.3, No.1, pp.62-71 (2007)
- [2] 三田村保, 大町清隆, 大堀隆文: “Web 検索による北海道観光情報の分析”, 観光情報学会誌, Vol.1, No.1, pp.47-53 (2005)
- [3] 斎藤一: “Web における観光情報提供と分析”, 人工知能学会誌, Vol.26, No.3, pp.234-239 (2011)

- [4] F. Girardin, F. D. Fiore, J. Blat and C. Ratti: “Understanding of Tourist Dynamics from Explicitly Disclosed Location Information”, Proceedings of 4th International Symposium on LBS and Telecartography, Hong-Kong China (2007)
- [5] F. Girardin, J. Blat, F. Calabrese, F. D. Fiore and C. Ratti: “Digital Footprinting Uncovering Tourists with User-Generated Content”, IEEE in Pervasive Computing, IEEE, Vol.7, No.4, pp.36-43 (2008)
- [6] Y. Kurata: “Potential-of-Interest Maps for Mobile Tourist Information Services”, Proceedings of Information and Communication Technologies in Tourism 2012 (ENTER 2012), pp.239-248 (2012)
- [7] 王佳な, 野田雅文, 高橋友和, 出口大輔, 井出一郎, 村瀬洋: “Web 上の大量の写真に対する画像分類による観光マップの作成”, 情報処理学会論文誌, Vol.52, No.12, pp3588-3592 (2011)
- [8] 日本語形態素解析エンジン「Mecab」:  
<http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>