

タグの時空間情報を考慮したジオタグ付き写真の可視化システムの開発

大森 雅己[†] 廣田 雅春^{††} 横山 昌平^{†††} 石川 博^{†††}

[†] 静岡大学情報学部 〒 432-8011 静岡県浜松市中区城北 3-5-1

^{††} 静岡大学創造科学技術大学院 〒 432-8011 静岡県浜松市中区城北 3-5-1

^{†††} 静岡大学情報学部 〒 432-8011 静岡県浜松市中区城北 3-5-1

E-mail: [†]cs09018@s.inf.shizuoka.ac.jp, ^{††}dgs11538@s.inf.shizuoka.ac.jp ,

^{†††}{yokoyama,ishikawa}@inf.shizuoka.ac.jp

あらまし 近年, GPS が搭載されたカメラやスマートフォンの普及により, ジオタグ付き写真が増加し, Flickr などの写真共有サイトに多数のジオタグ付き写真がアップロードされている. そこで, 本研究では, 写真に付与されたタグの時空間での分析の一環として, 写真へのタグ付けの時間的な変化, 地理的な偏りを可視化するシステムの構築を行う. 本システムでは, ユーザがタグや時間を指定し, 検索結果に含まれる写真の撮影位置を地図上に表示する. 例えば, 「sea」と検索した場合には, 多くの写真が海岸沿いで撮影されているため, 表示された写真だけで海岸線が把握できる. このような, タグの地理的・時間的な偏りが把握可能となるシステムの構築を行う.

キーワード ジオタグ, ソーシャルタギング, 可視化, GIS

Development of Visualization System of Geo-tagged Photographs Considering Space-time of Tags

Masaki OMORI[†], Masaharu HIROTA^{††}, Shohei YOKOYAMA^{†††}, and Hiroshi ISHIKAWA^{†††}

[†] Faculty of Informatics, Shizuoka University

3-5-1 Johoku, Naka-ku, Hamamatsu-Shi, Shizuoka, 432-8011 Japan

^{††} Graduate School of Science and Technology, Shizuoka University

3-5-1 Johoku, Naka-ku, Hamamatsu-Shi, Shizuoka, 432-8011 Japan

^{†††} Faculty of Informatics, Shizuoka University

3-5-1 Johoku, Naka-ku, Hamamatsu-Shi, Shizuoka, 432-8011 Japan

E-mail: [†]cs09018@s.inf.shizuoka.ac.jp, ^{††}dgs11538@s.inf.shizuoka.ac.jp ,

^{†††}{yokoyama,ishikawa}@inf.shizuoka.ac.jp

Abstract Many geo-tagged photographs are uploaded to social media site such as Flickr since camera and smart phone with GPS have become popular. Therefore, in this paper, we build a system visualizing the change based on time in tags of photographs and the variance in space, to research those photographs. Our system presents the map of photographs search results that are obtained from Flickr using the user-generated query as tags and/or time. For example, in this system, the almost photographs of a search result by keyword “sea” are displayed to near the shore, and we can see its coast. Therefore, using our system, we can see the change based on time in tags of photographs and the variance in space, to research those photographs.

Key words geo-tag, social tagging, visualization, GIS

1. はじめに

近年, GPS が搭載されたカメラやスマートフォンの普及により, ジオタグ付き写真が増加し, Flickr [1] や Panoramio [2] などの写真共有サイトにも多数のジオタグ付き写真がアップロー

ドされている. アップロードされた写真の多くにはタグが付与されており, ある特定のタグが付与された写真が地理的・時間的に偏在している. そこで, 本研究では, 写真共有サイトにアップロードされた大量の写真に付与されたタグの偏在から知識を抽出することを目指し, 写真に付与されたタグの地理的・

時間的偏在を可視化するシステムの構築を行う。

地理的な偏在の例として、「sea」と検索した場合、多くの写真が海岸沿いで撮影されており、表示された写真だけで海岸線を把握することが可能であると考えられる。時間的な偏在の例として、「snow」と検索した場合、一般的に雪が降るのは冬であるため、北半球が冬である12月から3月には北半球で撮影された写真が多く、南半球が冬である6月から9月には南半球で撮影された写真が多くなると考えられる。このような、タグの地理的・時間的な偏在が把握可能となるシステムの構築を目指す。

また、検索結果の写真に付与されたタグの内、多くの写真に付与されているタグを表示する。その中から、選択したタグが付与された写真のみを表示することで、検索したタグと、その検索結果の写真の多くに付与されているタグが共に使われる地域や時間を把握することが可能である。また、検索結果内でのタグの出現頻度に基づくクラスタリングを適用することで、類似するタグを含む写真の偏在を把握することが可能である。

本論文では、Flickrにアップロードされた写真に付与されたジオタグに基づいて、GoogleMaps [5] 上で写真へのタグ付けの時間的な変化、地理的な偏りを可視化するシステムを提案する。

本論文の構成は次のとおりである。2章では、本研究と関連研究の差分について述べる。3章では、提案システムの概要について述べる。4章では、実装したシステムの実行例を示し、それに対する考察を述べる。5章では、本研究で得られた成果のまとめについて述べる。

2. 関連研究

写真の位置情報と地図を用いた可視化システムの研究として、Zhang ら [3] は、雪、または緑に覆われている地域を写真に付与されているタグや、撮影時刻などのメタデータから推定し、地図上に表示するシステムを提案した。本研究では、あらゆるタグを対象にし、検索したいタグを入力することで、分析を支援可能なシステムの構築を行う。

Thomee ら [4] は、タグと位置情報から地図上に同じタグが頻出する領域を求め、タグごとに領域の可視化を行った。本研究では、タグで検索を行い、検索結果に対する地理的・時間的な偏り、検索したタグが付与された写真に頻繁に付与されているタグの地理的・時間的な分布などを可視化するシステムの構築を行う。

3. 提案システム

本研究では、ジオタグが付与された写真を用いて、写真の撮影位置を地図上に表示するシステムを提案する。システム構成図を図1に示す。本システムは、サーバサイドとクライアントサイドで構成されている。

3.1 サーバサイド

サーバサイドでは、写真の収集、検索、および検索結果の写真に対するクラスタリングを行う。クラスタリング手法は、 x -means 法 [8] を用いる。

前処理として、Flickr API を用いて、Flickr から写真を収集する。写真の収集は、ジオタグとタグが付与されたものを条件

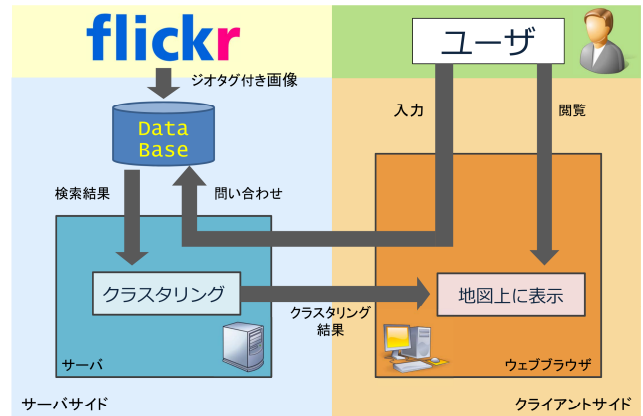


図1 システム構成図

とし、あらゆる時間を対象に行う。写真とともに、撮影時刻、ジオタグ、およびタグを収集する。検索処理においては、ユーザが入力したクエリをクライアントサイドから受け取り、写真の検索を行う。また、クラスタリングを行うよう指定された場合に、クラスタリングを行う。

検索結果には、写真に撮影されている物と、写真に付与されているタグの意味するところが直接的に結びつかない場合や、写真の意味するところを複数通りに解釈可能な場合が存在する。例えば、「apple」というタグで検索した場合、リンゴが撮影された写真や、iphoneなどのapple社の製品が撮影された写真が混在する可能性が考えられる。ユーザが必要なものだけを閲覧したい、また、各写真の枚数や、撮影された地域の違いによる写真の枚数を比較したい場合がある。本研究では、タグの出現頻度に基づいたクラスタリングを行い、撮影対象が類似している写真を1つのクラスタにまとめて表示する。これにより、視覚的に、同一クラスタに属する写真の撮影対象による地理的・時間的な分布を把握可能にする。

3.1.1 x -means 法

本システムでは、検索するクエリ次第で検索結果が大きく異なるため、最適なクラスタ数は一定ではない。そのため、検索結果に対して最適なクラスタ数が自動で決定される x -means 法を用いてクラスタリングを行う。

x -means 法は、 k -means 法を拡張したものである。 k -means 法では、あらかじめクラスタ数を決定しておく必要があるが、 x -means 法は、自動的に最適なクラスタ数 k を推定するという特徴がある。 x -means 法のアルゴリズムは、最初に、十分に小さいクラスタ数 k で k -means を行い、各クラスタに対し、さらに2分割することが適当である場合に再帰的に $k = 2$ として k -means 法によるクラスタリングを実行する。 k -means 法により分割を行うかどうかの判断は、クラスタの分割前と分割後のBIC(ベイズ情報量規準)を比較し、分割前の方が大きい場合は分割を行い、等しいまたは分割後のほうが大きい場合は分割を行わない。また、 k -means 法を行う際のセントロイドとデータ間の類似度は、ユークリッド距離を用いて算出する。 n 次元における、 p, q 間のユークリッド距離は以下の式で求められる。

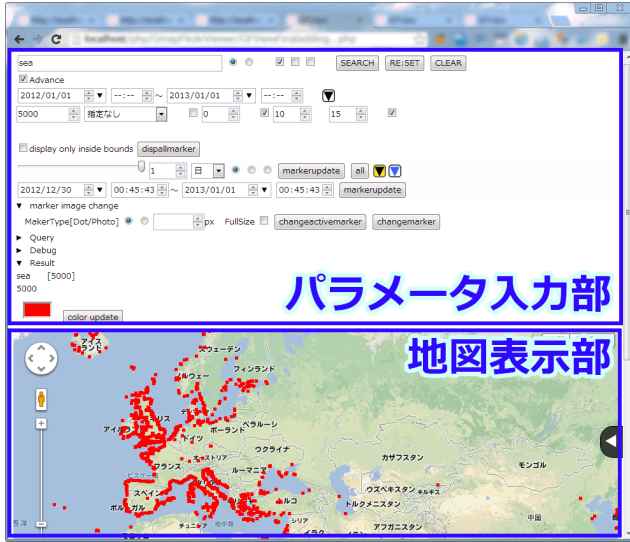


図 2 インターフェース

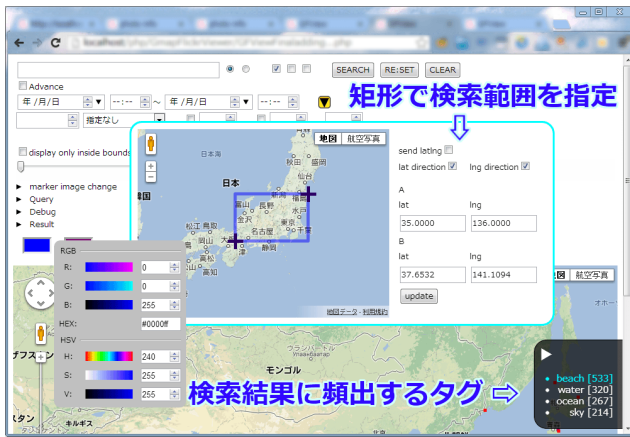


図 3 インターフェース (ポップアップ表示状態)

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (1)$$

3.1.2 特 徴 量

クラスタリングの特徴量には, $tf-idf$ を用いる. まず, 検索結果に含まれる各写真に付与された各タグに対して, tf_{pt} と idf_t を求め, $tf-idf_{pt}$ を求める. $tf-idf_{pt}$ は以下の式により求められる.

$$tf-idf_{pt} = tf_{pt} \cdot idf_t \quad (2)$$

$$idf = \log \left(\frac{D}{df} \right) \quad (3)$$

tf は一つの写真にあるタグが出現した回数, D は検索結果の写真の総数, df はあるタグが含まれる写真数である. 一つの写真に対して, 同じタグが複数付与されることはないため, tf は常に 1 または 0 となる. この $tf-idf$ を用いて, x -means クラスタリングを行う.

3.2 クライアントサイド

クライアントサイドでは, ユーザが検索したいタグや時間を

入力し, サーバサイドへ送信する. また, サーバサイドから検索結果を受け取り, 各写真の撮影位置を地図上に表示する. 表示には, Google Maps API を利用し, 写真の撮影位置にマーカーを設置する.

図 2 がシステムのインターフェースである. インターフェースは, 上部のパラメータの入力部分と下部の地図の表示部分で構成されている.

パラメータの入力部分では, クエリとして用いることが可能な, タグ, 日時, 緯度経度の範囲, 検索結果として取得する写真の枚数, 写真に付与されたタグの数の閾値, x -means クラスタリングを行うか, および x -means クラスタリングを行う際の初期クラスタ数を指定する. また, スライダーを使用した動的な表示マーカーの切り替え, 各写真の詳細なメタデータの表示, クラスタごとの表示の切り替え, 検索結果の写真に含まれるタグから頻出するタグを指定し, 表示する写真の絞り込みが可能である.

地図の表示部分では, これらの入力に応じて表示するマーカーの変更を行う.

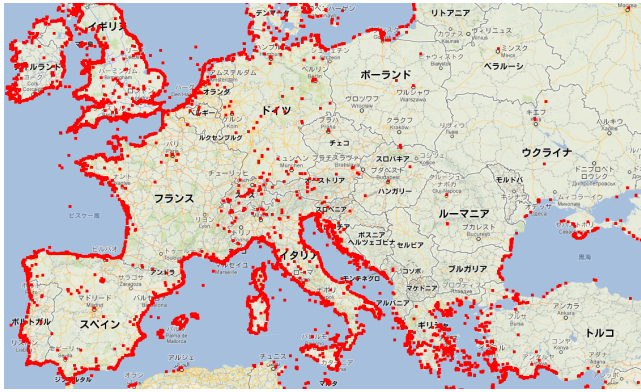
タグの入力では, 検索したいタグを入力する. その際, 複数のタグを用いた AND 検索, OR 検索, および NOT 検索を行うことが可能である. さらに, 括弧を使用して, 部分的な AND 検索と OR 検索を行うことも可能である. この機能を使用することで, 例えば, 日本とアメリカの海の写真が欲しい場合に「(日本 OR アメリカ) AND 海」というクエリで検索が可能になる.

日時の入力では, FROM となる日付, 時刻と, TO となる日付, 時刻を入力する. FROM を省略した場合, 1970 年 1 月 1 日, TO を省略した場合, 現在の時刻が検索クエリとして設定される. 日時の入力により, 指定した日時の範囲の写真のみを表示することが可能となる.

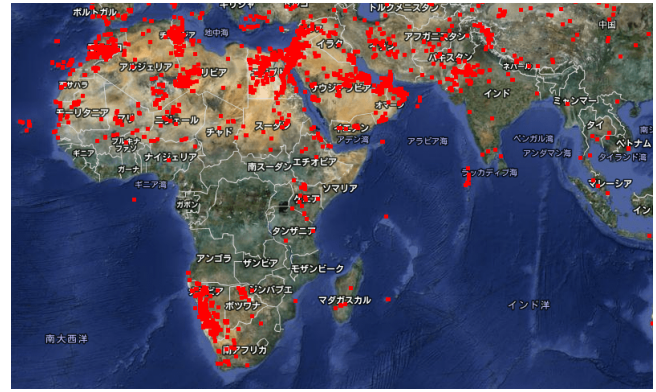
緯度経度による範囲の指定は, 図 3 のように, 範囲の指定を行う画面がポップアップで表示され, 2 つのマーカーを用いて, 矩形で検索範囲を入力する. タグの AND 検索機能を用い, 「日本 AND 海」のように地域を表す語を使用したクエリで検索した場合, 地域を表す語をタグとして持つ写真がその地域以外にも存在することや, その地域で撮影された写真であっても, 地域を表す語がタグとして付与されていないことがある. このような場合において, 緯度経度による検索範囲の指定が有効である.

検索結果の写真数の指定では, ユーザが検索結果として取得する写真の数を指定することが可能である. 取得する数を制限した場合は, 取得する順を写真の撮影時刻, Flickr へ写真をアップロードした時刻で指定することが可能である.

写真に付与されたタグの数の閾値の指定では, 閾値を指定することで, 大量のタグが付与された写真を検索結果から除外することが可能である. Flickr にアップロードされている写真には, 写真の内容と関連性のないタグが付与されているものが存在する. 中には, 1 つの写真に対し, 100 を超えるタグが付与されている写真も存在する. このような, 大量のタグが付与されている写真は, 写真と関連性のないタグが付与されている可



(a) 「sea」をクエリとした場合



(b) 「desert」をクエリとした場合

図 4 システム実行例 (地理的な分布)

能性が高い。本研究では、このような写真をノイズとして扱う。写真に付与されたタグの数に対して閾値を設定し、大量のタグが付与された写真を検索結果から除外することで、ノイズを減少させることが可能である。

x -means クラスタリングの初期クラスタ数の指定では、 x -means 法における、最初の k -means 法でクラスタリングする際のクラスタ数を指定することが可能である。 x -means 法はクラスタ数を自動推定するものであるが、最適と思われるクラスタリング結果を得られなかった場合、ユーザがクラスタ数を指定することで、クラスタ数を調整することが可能である。

スライダーを使用した、動的な表示マーカーの切り替えでは、検索結果の時間によるタグの利用分布の変化を視覚的に表示することが可能である。スライダーの左端が検索クエリの FROM として与えられた時刻、右端が検索クエリの TO として与えられた時刻となっており、スライダーの位置が特定の時刻を表す。現在のスライダーの時刻から、前後の期間に撮影された写真のマーカーのみを表示する。また、バーを使わず、詳細な時刻を指定することも可能である。例えば、ハロウィンのような特定の時期に発生するイベントに関するタグを検索し、イベント発生前後の時期をスライドさせると、イベントの時期に近づくに連れ、マーカーは増加し、イベントの発生時にマーカー数が最大となる。そして、イベントの終了後は、マーカーが減少する。このように、あるタグに関する写真の枚数や写真が撮影された場所、写真に付与されたタグの移り変わりを時間の流れに沿って閲覧可能とする。

各写真の詳細なメタデータの表示機能では、写真とその写真の撮影時刻や付与されたすべてのタグなど、写真についての詳細なメタデータを確認することが可能である。また、その場で簡易的に写真の内容を確認することも可能である。

クラスタごとの表示切り替え機能では、クラスタリングされた結果に対し、クラスタごとに表示と非表示を切り替えることが可能である。あるクラスタのみを閲覧したい場合、必要なクラスタのみを残し、他のクラスタを非表示にすることで目的のクラスタに属する写真の分布を閲覧することが可能である。また、全クラスタを表示して、クラスタごとに色を付けることにより、地域ごとに写真の分布を把握することが可能である。

検索結果の写真に付与されたタグから頻出するタグを指定し、表示する写真の絞り込みを行う機能では、検索結果の写真に付与されたタグの内、付与されている写真数の降順にタグを表示し、ユーザがその中からタグを選択する。タグが選択されると、選択したタグが付与された写真のマーカーのみを表示する。これにより、検索したタグが、選択したタグと共に用いられる地域や時間を把握することが可能である。

4. 実行結果

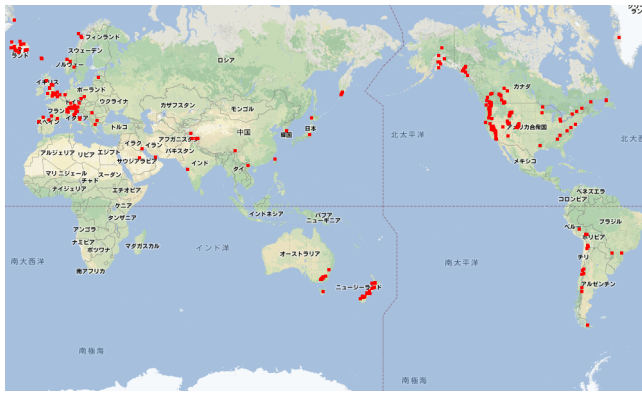
4.1 実行例

データセットには、Flickr から取得した写真、約 1,200 万枚を使用した。図 4、図 5 は、提案したシステムの実行例である。図 4(a) はタグ「sea」を、図 4(b) はタグ「desert」を、図 5(a) はタグ「snow」と日時 6 月～9 月を、図 5(b) はタグ「snow」と日時 12 月～3 月をクエリとし、得られた実行例である。

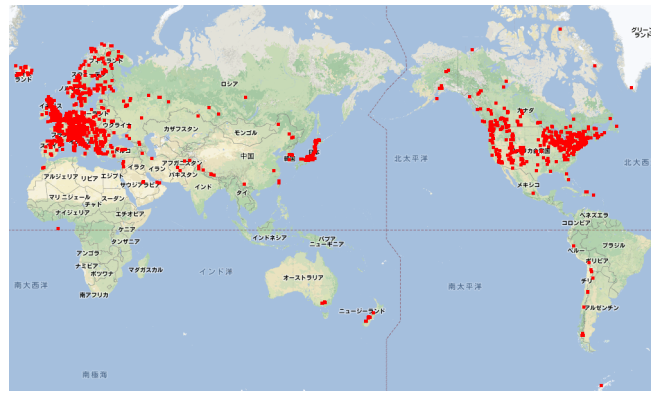
図 4(a) では、海沿いに写真が集まっており、マーカーだけで海岸線を把握可能である。また、図 5(a) と図 5(b) を比較すると、図 5(a) は、北半球が夏、南半球が冬の時期を指定しており、マーカーは、北半球(夏)に少なく、南半球(冬)に多い。それに対し、図 5(b) は、北半球が冬、南半球が夏の時期を指定しており、マーカーは、北半球(冬)に多く、南半球(夏)に少ない。これらから、「snow」というタグを持つ写真は、雪が降る冬に多いことが分かる。さらに、同じタグに関する写真でも、その撮影時期によって、撮影位置に偏りが生じることが分かる。

一方、図 4(b) では、多数のマーカーが衛星写真の茶色の砂漠部分と重なっているが、マーカーが配置されていない部分も存在する。また、図 5(a) では、豪雪地帯にもかかわらず、ロシアの北部にマーカーが存在しない。これは、砂漠の中心やロシアの北部など、人が少なく、旅行者も少ない地域では、写真を撮影する人が少ないことが原因であると考えられる。

図 6(a) は、「sea」をクエリとして検索した場合の頻出タグから「mar」を選択したものである。スペインやポルトガル、南米にマーカーが多く存在する。「mar」はスペイン語、ポルトガル語で海という意味があり、スペインやポルトガルなどスペイン語、ポルトガル語が利用されている地域が分かる。また、図 6(b) は、「sea」をクエリとして検索した場合の頻出タグから

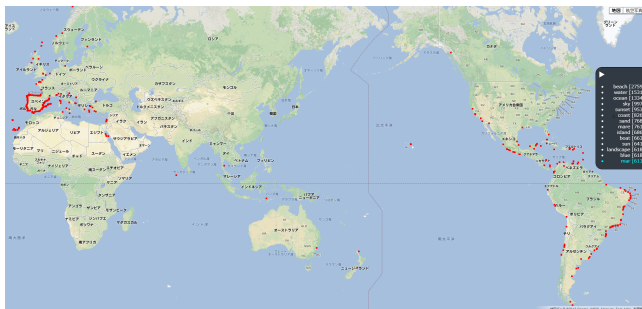


(a) 「snow」と6月～9月をクエリとした場合

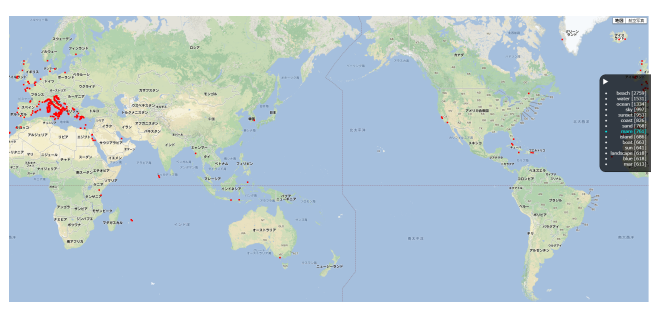


(b) 「snow」と12月～3月をクエリとした場合

図5 システム実行例(時間的な分布)

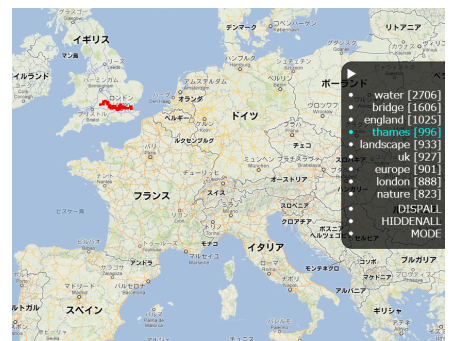
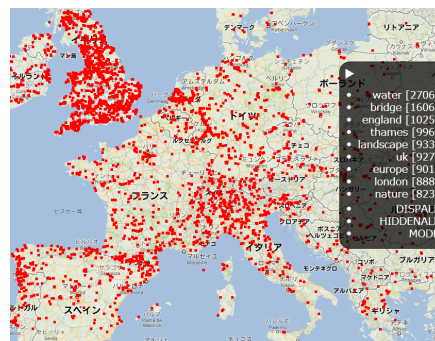
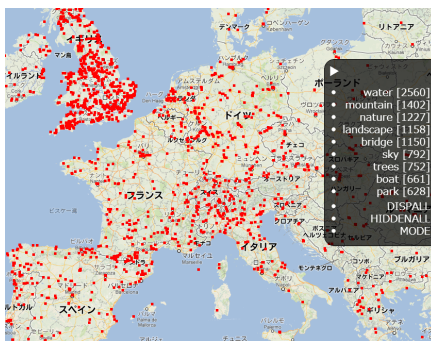


(a) 「sea」をクエリとし、「mar」を選択した場合



(b) 「sea」をクエリとし、「mare」を選択した場合

図6 システム実行例(頻出タグ表示「sea」)



(a) 「river」をクエリとし、検索範囲を指定していない場合

場合

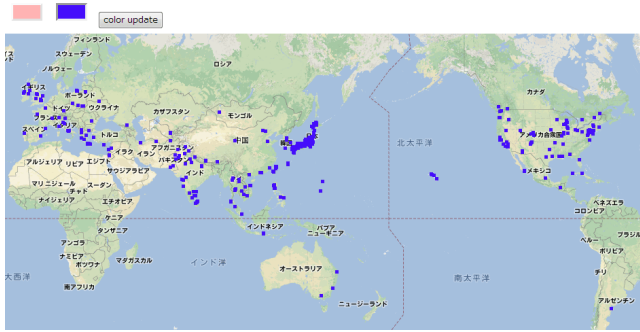
(c) 「river」をクエリとし、ヨーロッパを範囲とし、「thames」を選択した場合

図7 システム実行例(頻出タグ表示「river」)

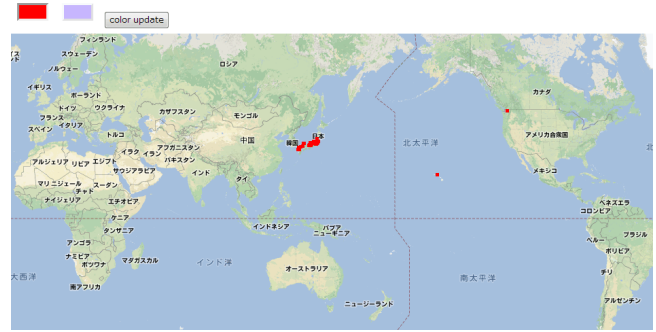
「mare」を選択したものである。イタリア付近にマーカーが多く存在する。「mare」はイタリア語、ラテン語で海という意味があり、イタリア語やラテン語が利用されている地域が分かる。

図7(a)は、「river」をクエリとした場合、図7(b)は、「river」をクエリとし、緯度経度でヨーロッパを指定した場合の実行結果である。地域を指定せず、世界中を対象とした場合の頻出タグは、「water」、「mountain」、「landscape」、「sky」などであり、ヨーロッパに関連するタグは存在しない。それに対し、ヨーロッパを指定した場合の頻出タグには、「uk」、「london」、「thames」、「europe」などのヨーロッパに関連するタグが存在する。図7(c)は、図7(b)の頻出タグから「thames」を選んだものである。マーカーにより、テムズ川の位置が把握可能である。

図8(a)、図8(b)は、「shrine」をクエリとし、 x -means法でクラスタリングを行った場合の実行結果である。図8(c)は、図8(a)に含まれる写真、図8(d)は、図8(b)に含まれる写真である。図8(a)には、日本の神社と関係ない写真が多く、図8(b)には、日本の神社に関する写真が多い。図8(d)の鳥居が写っている写真は、アメリカ、ヨーロッパで撮影されたものである。図8(a)に含まれる写真の中で、海外で撮影された写真を見ると、日本の神社に関係ないものが撮影された写真が多く、日本で撮影された写真を見ると、図8(c)の左下の写真のように、日本の神社に関係するものが撮影された写真が存在する。図8(b)では、日本の神社に関連する写真を抽出することができた。一方、図8(a)では、良い結果を得られなかった。 x -means法は、



(a) 「shrine」をクエリとし、クラスタリングを行った場合 (青クラスタのみを表示)



(b) 「shrine」をクエリとし、クラスタリングを行った場合 (赤クラスタのみを表示)



(c) 「shrine」をクエリとし、クラスタリングを行った場合の青クラスタに含まれる写真



(d) 「shrine」をクエリとし、クラスタリングを行った場合の赤クラスタに含まれる写真

図 8 システム実行例 (クラスタリング「shrine」良い例)

k -means 法を行うため、初期値がランダムである。そのため、同じクエリで検索した場合でも、良い結果が得られないことがある。図 9, 図 10 は、良い結果が得られなかったクラスタリング結果であり、図 9 は、同じ「shrine」を与えた場合の、異なるクラスタリング結果である。図 9(a) を見ると、日本の数ヶ所にしか写真が存在しない。また、図 10 は、「sea」をクエリとし、クラスタリングを行った場合の結果である。海が写った海岸沿いの写真と、海が写っていない内陸部の写真が別のクラスタに属することを期待したが、クラスタリング結果では、海岸沿いと内陸部の写真が各クラスタに混在していた。これは、類似する写真でも、タグを付与するユーザが異なれば、付与されるタグも異なるためであると考えられる。

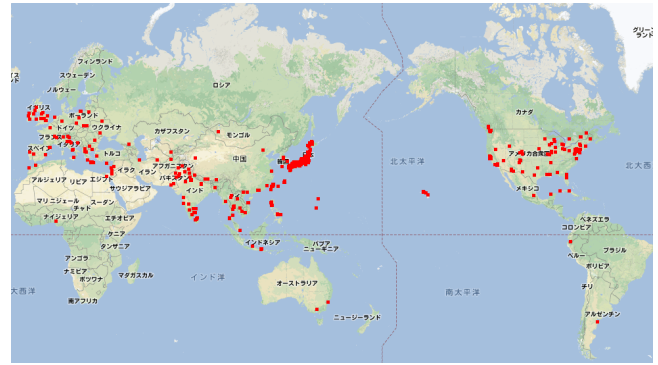
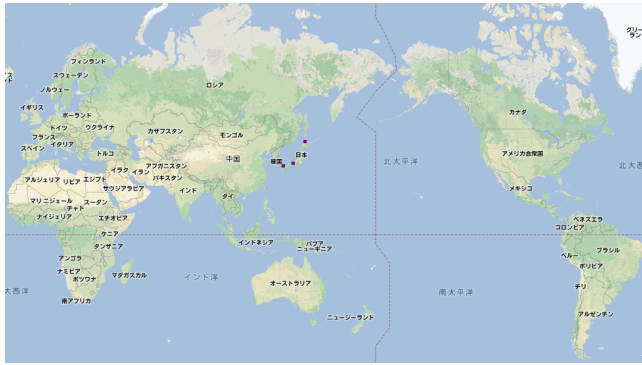
4.2 考 察

図 4(a) の内陸部や、図 5(a) の北半球など、クエリに対して適切でないと考えられる位置にも写真が存在する。これらの中には、海以外の「sea」という単語が含まれるものを撮影した写真や海ではないが海に関連するものが撮影されている写真、夏でも雪がある標高の高い山の上で撮影された写真など、タグにおける検索結果として正しい写真が存在する。しかし、写真の内容がタグと全く関係のない場合がある。これは、写真の内容と全く関連がないタグが付与されたことが原因であると考えられる。図 11 は、クエリを「sea」として検索した場合に含まれ

るノイズとなる写真である。

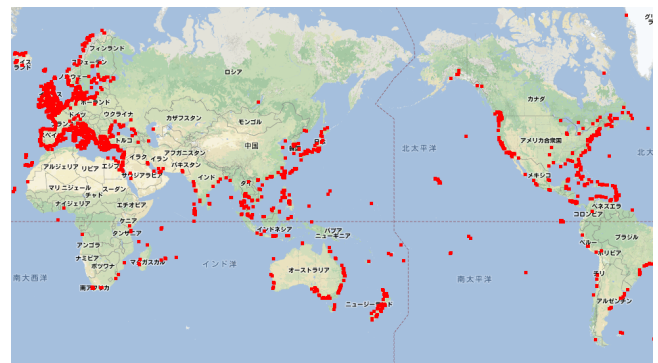
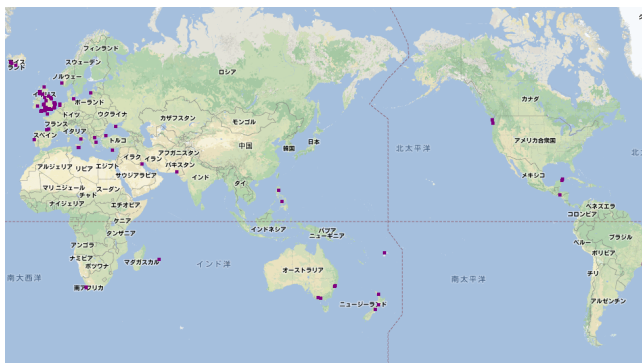
本システムでは、このような写真を除外するために、写真に付与されたタグの数が閾値を超えていた場合にノイズ写真として扱う。閾値は、検索の際にクエリとしてユーザが指定することが可能である。図 13(a) は、「sea」をクエリとし、タグ数によるノイズ除去を行っていない場合、図 13(b) は、「sea」をクエリとし、閾値を 15 としてタグ数によるノイズ除去を行った場合である。ノイズ除去を行った図 13(b) は、ノイズ除去を行っていない図 13(a) に比べ、ノイズと思われる内陸部のマーカーが減少していることが分かる。しかし、ノイズと考えられる写真がすべて除去されていない。この対策として、検索結果の写真の撮影位置の密度から飛び地を判断し、ノイズを除去することが考えられる。また、現在のシステムは、処理の速さの問題や、タグだけに着目した分布を可視化するため、画像処理を用いていない。画像処理を利用することで、タグと写真の内容が一致しないようなノイズについても対応が可能になると考えられる。

Flickr にアップロードされる写真の数は地域によって大きく異なる。現在の表示方法では、地域による写真の総数の違いを考慮していないため、異なる地域を比較する場合、マーカー数が多い地域が、必ずしもマーカー数の少ない地域に比べ、重要であるとは限らない。例えば、図 5(a) を見た場合、北半球と



(a) 「shrine」をクエリとし、クラスタリングを行った場合 (紫クラスタのみを表示) (b) 「shrine」をクエリとし、クラスタリングを行った場合 (赤クラスタのみを表示)

図 9 システム実行例 (クラスタリング「shrine」悪い例)



(a) 「sea」をクエリとし、クラスタリングを行った場合 (紫クラスタのみを表示) (b) 「sea」をクエリとし、クラスタリングを行った場合 (赤クラスタのみを表示)

図 10 システム実行例 (クラスタリング「sea」)

南半球でマーカー数を比較すると、夏である北半球のほうが「snow」とタグを付与されたマーカー数が多い。しかし、夏のほうが雪が多く降るというわけではない。これは、写真の総数が北半球のほうが多いためである。また、表示するマーカー数が増加すると、ノイズの影響が大きくなり、正確な分布の把握が困難となる。図 12 は、クエリを「sea」として取得件数の制限をしない場合の実行例である。イギリスやギリシャがマーカーで覆われていて、海岸線を把握することが困難である。これらの対策として、検索結果の写真数と同数のマーカーを設置するのではなく、検索結果がその地域の写真の総数に対して、どの程度の割合であるかを考慮し、表示するマーカー数を調節することで、写真総数の異なる地域の比較をより正確に行う事が可能になると考えられる。また、ノイズとなる写真は適切な写真と比較して少ないため、ノイズの除去にもつながると考えられる。

5. おわりに

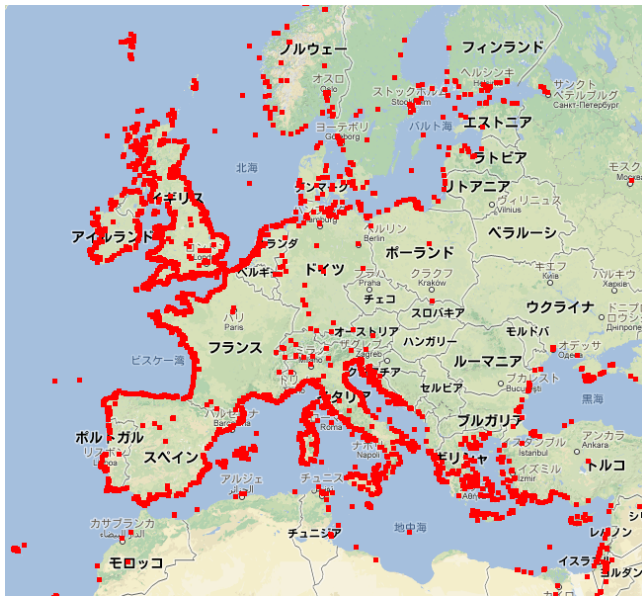
本研究では、写真に付与されているタグの地理的・時間的な偏在を可視化するシステムを構築した。これにより、写真に付与されたタグの偏在に意味があることを示した。

今後の課題としては、ノイズとなる写真の除去があげられる。検索結果の写真の撮影位置の密度から、飛び地であるノイズを除去することが考えられる。また、画像処理を用いることで、

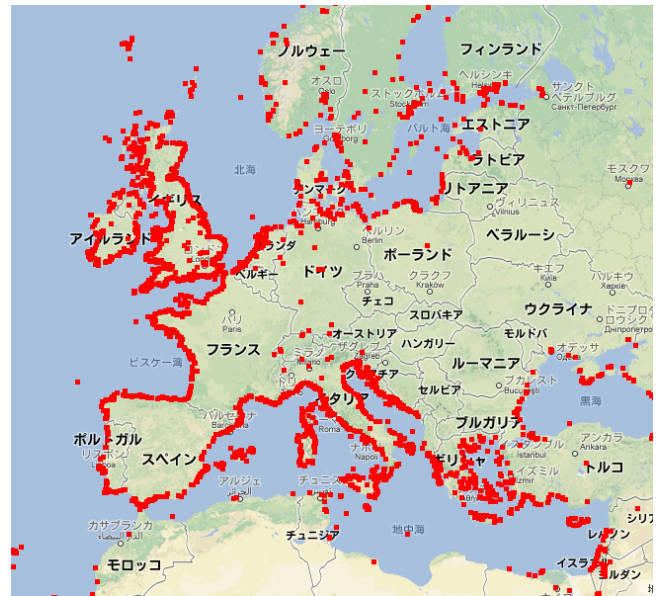


図 11 「sea」をクエリとした場合のノイズとなる写真の例

タグと写真の内容が一致しない場合にも対処できると考えられる。現在のシステムでは、検索結果の写真数と同数のマーカーを表示している。表示するマーカー数を地域ごとの写真の総数に応じて調整することで、より正確な分布が把握可能となると考えられる。現在、Flickr にはジオタグが付与された写真が増加しつつあるが、ジオタグが付与されていない写真も多く存在する。タグなどのメタデータを利用し、これらの写真の撮影位置を推定する研究が存在する [6], [7]。本システムで、これらの技術を利用することでジオタグが付与されていない写真を表示可能となる。また、閲覧性の向上のため、高解像度で表示する技術として、Tiled Display Wall [10] を利用することが考えられる。現在のシステムでは、検索結果内で頻出するタグのリ



(a) 「sea」をクエリとし、ノイズとするタグ数の閾値を指定しない場合



(b) 「sea」をクエリとし、ノイズとするタグ数の閾値を 15 とした場合

図 13 システム実行例 (ノイズ除去)

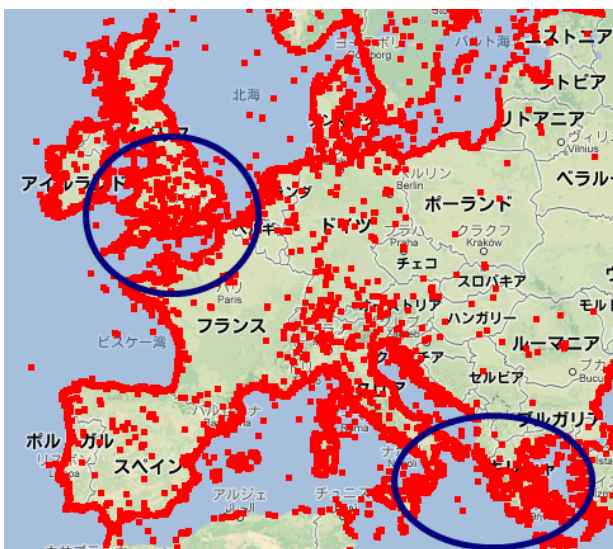


図 12 表示数が多い場合のノイズによる影響

ストを表示し、そこからタグを選択することで、検索したタグと選択したタグが共に使われる場合の分布を表示しているが、WordNet [9] を用いて、一般的に関連度の高いタグについても共に使われる場合の分布を表示することで、より多様な分布を把握可能になると考えられる。また、「sea」のようなタグに対して、実際の海岸線データと比較することで、タグの信憑性の評価を行うことがあげられる。

文 献

- [1] “Flickr”, <http://www.flickr.com/>
- [2] “Panoramio”, <http://www.panoramio.com/>
- [3] Haipeng Zhang, Mohammed Korayem and David J. Crandall, “Mining Photo-sharing Websites to Study Ecological Phenomena”, Proceedings of the 21th International Conference on World wide web, Pages 749-758, 2012
- [4] Bart Thomee and Adam Rae, “Exploring and Browsing Photos through Characteristic Geographic Tag Regions”, Proceedings of the 20th ACM International Multimedia Conference, 2012
- [5] “Google Maps”, <https://maps.google.co.jp/>
- [6] Pavel Serdyukov, Vanessa Murdock and Roelof van Zwol, “Placing Flickr Photos on a Map”, Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Pages 484-491, 2009
- [7] David Crandall, Lars Backstrom, Daniel Huttenlocher and Jon Kleinberg, “Mapping the World’s Photos”, Proceedings of the 18th International Conference on World wide web, Pages 761-770, 2009
- [8] 石岡恒憲, “クラスター数を自動決定する k -means アルゴリズムの拡張について”, 応用統計学, Vol.29, No.3, Pages 141-149, 2000
- [9] “WordNet”, <http://nlpwww.nict.go.jp/wn-ja/>
- [10] Shohei Yokoyama and Hiroshi Ishikawa, “Parallel Distributed Rendering of HTML5 Canvas Elements”, Proceedings of the 11th International Conference on Web engineering, Lecture Notes in Computer Science, Vol.6757, Pages 331-345, 2011