

ジオタグ付き写真を用いた風景カテゴリマップ作成手法の検討

Wang Jiani[†] 野田 雅文[†] 高橋 友和^{††} 出口 大輔[†] 井手 一郎[†]
村瀬 洋[†]

[†] 名古屋大学大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

^{††} 岐阜聖徳学園大学経済情報学部 〒500-8288 岐阜県岐阜市中鶉 1-38

E-mail: [†] {jwang,mnoda}@murase.m.is.nagoya-u.ac.jp, {ddeguchi,ide,murase}@is.nagoya-u.ac.jp

^{††} ttakahashi@gifu.shotoku.ac.jp

あらまし 近年、デジタルカメラの普及により、ユーザが撮影した大量のデジタル写真が、写真共有を行うソーシャルサイトにアップロードされている。また、これらの写真の多くには、GPS などによる位置情報がジオタグとして付随している。このようなジオタグを利用し、地図と関連付けて写真を閲覧することができる。しかしながら、例えば旅行を計画中のユーザがある地点から見える風景について直感的に理解したい場合、大量の写真を地図上に単純に配置しただけでは不十分であると考えられる。そのため、本研究では、ユーザにとって直感的に理解しやすい地図として、風景カテゴリマップを作成することを目的とする。風景カテゴリマップとは、地図中の注目する地点周辺で撮影された大量の写真を山や海などの風景カテゴリとして分類・整理し、写真の代わりに風景カテゴリのアイコンを用いて表現した地図である。このような表現方法によって、ユーザはある地点で多くの写真が注目する風景について直感的に理解できるようになることが期待される。

キーワード ジオタグ付き写真, 画像特徴, 風景カテゴリマップ

Study on the Creation of a Scenery Category Map using Geo-tagged Photographic Images

Jiani WANG[†], Masafumi NODA[†], Tomokazu TAKAHASHI^{††}, Daisuke DEGUCHI[†], Ichiro IDE[†],
and Hiroshi MURASE[†]

[†] Graduate School of Information Science, Nagoya University Furo-cho, Chikusa-ku, Nagoya-shi, Aichi,
464-8601 Japan

^{††} Faculty of Economics and Information, Gifu Shotoku Gakuen University 1-38, Nakauzura, Gifu-shi,
Gifu, 500-8288 Japan

E-mail: [†] {jwang,mnoda}@murase.m.is.nagoya-u.ac.jp, {ddeguchi,ide,murase}@is.nagoya-u.ac.jp

^{††} ttakahashi@gifu.shotoku.ac.jp

Abstract Recently, due to the spread of digital cameras, lots of digital photographs that are taken by users have been uploaded to social sites, where photographs are shared. Moreover, the location information from GPS etc. accompanies many of these photographs as a geo-tag. By using the geo-tag, the photographs and the map can be related. However, to intuitively understand the scenery of the spot where a user plans to travel, it is insufficient to simply arrange a large amount of photographs on the map. Therefore, this work aims to make the scenery category map as a map that can be intuitively understood by users. On the scenery category map, the photographs taken around the spot are classified and organized by scenery categories such as mountains and seas. And the photographs are replaced by icons of scenery categories on the map. According to the proposed method, it is expected that the user can intuitively understand the scenery to which lots of photographs pay attention from a place.

Key words photograph with geo-tag, image feature, scenery category map

1. はじめに

近年、デジタルカメラの普及により、ユーザが撮影した大量のデジタル写真が、写真共有を行うソーシャルサイトにアップロードされている。また、これらの写真の多くには、GPSなどによる位置情報がジオタグとして付随している。このようなジオタグを利用し、地図と関連付けて写真を閲覧することができる。しかしながら、例えば旅行を計画中のユーザがその地点から見える風景について直感的に理解したい場合、大量の写真を図上に単純に配置しただけでは不十分であると考えられる。そのため、本研究では、ユーザにとって直感的に理解しやすい地図を作成することを目的とする。

写真と地図を関連付けて閲覧できる従来のサービスとして、Flickr^(注1)(図1)やPanoramio^(注2)(図2)などが挙げられる。これらのサービスでは、写真に付与されたジオタグを用いて、対応する地図中の位置に写真を配置している。また、注目地点周辺で撮影された写真を、サムネイルにより一覧することもできる。写真のような画像情報は、テキスト情報だけでは伝えにくいような地域の雰囲気やユーザに対して視覚的に伝えることができるため、本研究でもこれを一つの利点と考えて、地図を作成する際にソーシャルサービスの写真を利用する。

これらのサービスは、ユーザが付与したテキストタグに基づいて写真を分類・整理することにより、指定されたテキストタグが付与された写真だけを地図上に表示することもできる。しかしながら、写真のテキストタグはユーザが明示的に記述した風景の内容であるため、その内容はユーザに依存する。そのため、適切なテキストタグが付与されない場合や表記ゆれ(city, townなど)の問題がある。その結果、テキストタグだけで分類・整理して作成した地図では、地域のイメージを明確に伝えるのは困難である。

これに対して、写真にはテキストタグには記述されない内容もそのまま含まれている。そのため、本研究では、画像特徴を用いて写真を分類することで、テキストタグによる方法の問題を回避する。

FlickrやPanoramioのようなサービスのもう一つの問題点として、ユーザが注目する地域に存在する写真が大量になるほど、何に注目して良いかを直感的に把握することが困難になることが挙げられる。そのため、本研究では、地域ごとに大量の写真进行分类・整理した結果を基に、多くの写真で注目されている中身を直感的に理解しやすい地図を作成することを目的とする。

具体的には、図3に示すような風景カテゴリマップを自動で作成する手法を検討している[1]。風景カテゴリマップとは、地図中で注目するある地点周辺で撮影された写真を山や海などの風景カテゴリとして分類・整理し、写真の代わりに風景カテゴリのアイコンを用いて表現した地図である。このような表現方法によって、ユーザは注目する地点からどのような風景が見え

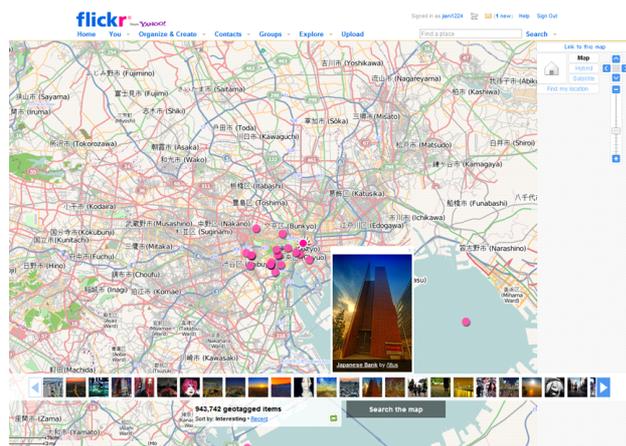


図1 Flickr^(注1)の利用例

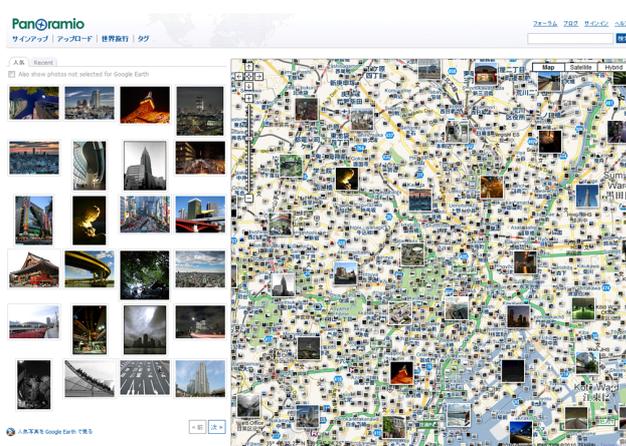


図2 Panoramio^(注2)の利用例

るかを直感的に理解できる。また、風景カテゴリマップを作成する際に、ユーザが撮影した大量の写真を用いているため、ランドマークのような観光資源だけでなく、ユーザ視点での地域のイメージを反映することができる。さらに、各アイコンが代表する実際の多数の写真に含まれた多様な風景内容の集まりを見ることによって、よりリアルティにあふれる地域の姿を浮かび上がらせる効果も期待できる。

以降、2.で風景カテゴリマップの作成手法について述べる。そして、3.で、実験方法を述べ、結果に対して考察する。最後に4.で本報告をまとめる。

2. 風景カテゴリマップの作成手法

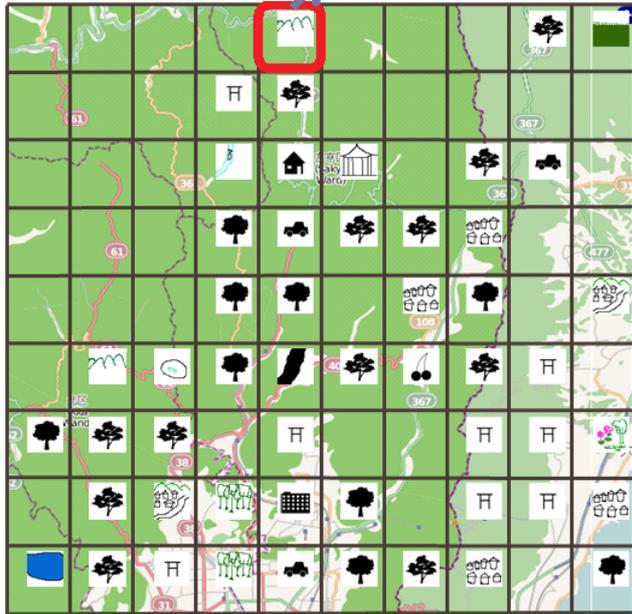
2.1 概要

写真共有サイト上にアップロードされた大量のデジタル写真を用いた風景カテゴリマップの作成手法を提案する。提案手法は、注目する地域範囲をブロックに分割し、各ブロックに含まれる写真のカテゴリ識別を行う。そして、その識別結果を用いてブロックの風景カテゴリを決定し、決定された風景カテゴリのアイコンを地図上に配置することで風景カテゴリマップを作成する。

以降、関連研究、および提案手法による風景カテゴリマップの作成について述べる。

(注1): <http://www.flickr.com/>

(注2): <http://www.panoramio.com/>



(c) OpenStreetMap contributors, CC-BY-SA

図 3 風景カテゴリマップの例

2.2 関連研究

各ブロックの風景カテゴリを決定するため、写真のカテゴリ識別を行う必要がある。このようなカテゴリ識別に関する研究として、ユーザが写真に対して付与したテキストタグと写真が持つ画像特徴を利用する研究が行われている。テキストタグを用いたカテゴリ識別は、キーワードによる分類処理である。現在、そのような処理は、画像検索エンジンの中で用いられている。例えば、Lycos Multimedia Search^(注3)、AltaVista Image Search^(注4)、Google Image Search^(注5)などは、その代表例である。しかしながら、テキストタグが付与されていない写真の存在とテキストタグの表記ゆれ (city, town など) などが問題となる。このため、すべての写真をカテゴリに分類することは難しく、分類検索精度は高くはない。

一方、画像特徴に基づく写真のカテゴリ識別手法として、局所特徴量のヒストグラム表現である BoF (Bag-of-Features) を用いた手法 [2] [3] が用いられる。また、色特徴を併用することで、より良いカテゴリ識別を行う手法も提案されている [4] [5]。さらに、位置情報を特徴の一つとして用いることにより、画像識別の精度を向上する手法も提案されている [6]。本研究では、これらの研究を踏まえ、BoF 特徴と色特徴を用いたカテゴリ識別を行う。これによって、テキストタグによらない識別ができる。

特に、本研究と関係する研究として、画像識別によってランドマークの写真と地図を関連付ける研究 [7] がある。この研究

(注3): <http://www.lycos.com/>

(注4): <http://www.altavista.com/image/>

(注5): <http://images.google.com/>

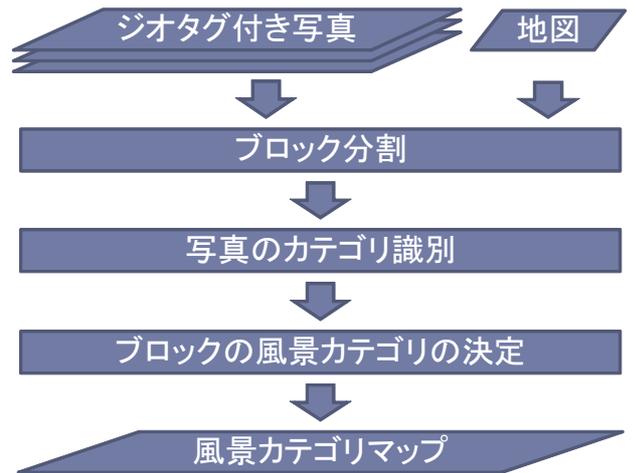


図 4 風景カテゴリマップ作成手法の処理の流れ

では、写真共有サイト上のランドマークの写真抽出・識別して、地図上に表現することが目的である。また、柳井 [8] はキーワードに基づく WWW 上の画像検索と、内容に基づく画像検索を組み合わせることを提案している。しかしながら、これらの研究は、特定のキーワードに合う写真や、似た見た目の写真の検索を対象としているため、風景の分布を表わす地図の生成を目的とする本研究とは異なる。

2.3 提案手法の流れ

図 4 に提案手法の処理の流れを示す。以降、風景カテゴリマップの作成手法の詳細を説明する。

2.4 ブロック分割

図 5 に示すように、注目する地域の地図を等間隔の複数のブロックに分割する。各ブロックは、一定の大きさの矩形とする。以降の処理はブロックごとに行う。

2.5 写真のカテゴリ識別

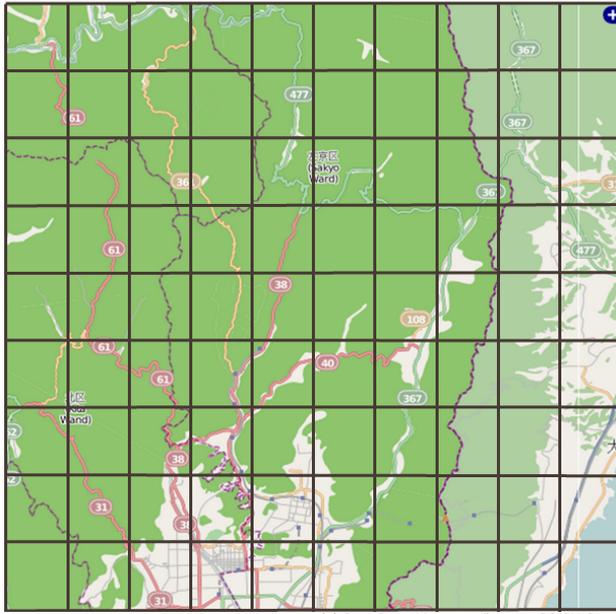
各ブロックの風景カテゴリを決定するため、注目するブロックに含まれる写真のカテゴリを識別する。提案手法では、各写真から画像特徴として BoF、および色特徴量を抽出し、統合したものをを用いる。そして、これらの特徴を学習した識別器により、識別を行う。識別器には、SVM (Support Vector Machine) を用いる。

2.5.1 BoF (Bag-of-Features) [2]

BoF は、画像を局所特徴量のヒストグラムで表現した特徴である。はじめに、SIFT (Scale-Invariant Feature Transform) [9] により特徴点を検出し、SIFT 特徴量を求める。全学習画像の全 SIFT 特徴量を k -means によってクラスタリングし、code book を作成する。そして、code book に基づいて各画像を SIFT 特徴量のヒストグラムとして表現する。Code book のピン数を N_B としたとき、BoF を $\mathbf{f}_B = [x_1, x_2, \dots, x_{N_B}]^T$ と表わす。なお、ノルムが 1 となるように正規化する。

2.5.2 色特徴量

提案手法では、色特徴量として、HSV 色空間を用いる。はじめに画像の各画素を RGB 色空間から HSV 色空間に変換する。そして、HSV 色空間の各軸を N_C 分割したものをピンとしたヒストグラムを作成する。ここで、ヒストグラムのピン数は



(c) OpenStreetMap contributors, CC-BY-SA

図5 実験の対象範囲(京都周辺)

N_C^3 となる。このとき、色特徴量を、 $f_C = [y_1, y_2, \dots, y_{N_C^3}]^T$ と表す。なお、ノルムが1となるように正規化する。

2.5.3 カテゴリ識別

以上の2つの特徴を統合した特徴量を $f = [f_B, f_C]^T$ とし、SVMの学習・識別に用いる。SVMは、カテゴリの正解ラベル付きの写真から f を抽出し、学習する。そして、そのSVMを用いてブロックに含まれる写真のカテゴリを識別する。

2.6 ブロックの風景カテゴリの決定

最後に、ブロックの風景カテゴリを決める。ここでは、同じ風景カテゴリとして識別された写真の枚数を数え、枚数が最も多い風景カテゴリをそのブロックの風景カテゴリとする。

以上の処理をブロックごとに行うことによって、風景カテゴリマップを作成する。

3. 実験と考察

3.1 概要

本実験では、図5に示す京都周辺を対象範囲とし、風景カテゴリマップの作成を行った。対象範囲は東経 $135.68699^\circ \sim 135.89298^\circ$ 、北緯 $35.03984^\circ \sim 35.20832^\circ$ であり、東西約 20 km、南北約 18 km の広さであった。ジオタグ付き写真は Panoramio から収集した。対象範囲に含まれる写真は、663 枚であった。

対象範囲を 10×9 のブロックに分割した。このとき、1つのブロックは、2 km 四方の矩形とした。各ブロックに含まれる写真の枚数は最小0枚、最大101枚であった。また、本実験では、カテゴリとして、図8に示す39種類を用いた。これらは、SUN Database [10] を参考にして設定した。カテゴリ識別に用いる識別器は SUN Database のラベル付きの写真を用いて学習した。学習に用いた写真は 16,689 枚であり、カテゴリごとの枚数の偏りはなかった。なお、学習画像を用いて、10-fold 交

表1 実験で使用した5種類の風景カテゴリ

風景カテゴリ	アイコン	SUN Database[9]のカテゴリ
街		alley, amusement park, bridge, building, gazebo, house, market, pagoda, plaza, railroad track, shop front, street, temple, tower, village, fountain
森		botanical garden, forest, forest path, park
水辺		bridge, canal, coast, creek, dam, hot spring, islet, lake, ocean, pond, river, sea cliff, waterfall
平地		badlands, desert, field, amphitheater
山		cliff, dam, mountain, sea cliff, valley

差検定をしたところ、識別率は 76.63% であった。

図8に示した39種類のカテゴリの中には、forest と forest path, creek と river のように画像特徴が類似し、人間の目でも判断が難しいカテゴリが含まれている。そこで、これらのカテゴリを画像特徴の類似性から5種類の風景カテゴリに再分類し、風景カテゴリマップを作成する際の風景カテゴリとして使用した。表1に実験で使用した5種類の風景カテゴリとそのアイコンを示す。

作成した風景カテゴリマップの評価は、正解の風景カテゴリマップと、識別器によるカテゴリ識別結果を用いて作成した風景カテゴリマップとの一致率により行った。ここでは、正解の風景カテゴリマップは、提案手法の写真のカテゴリ識別する際に、収集した各写真に対して識別器を使わず、人手によりカテゴリをラベル付けして、各ブロックの風景カテゴリを決定することで作成した。

3.2 実験結果

図6に提案手法により作成した風景カテゴリマップを示す。また、図7に正解の風景カテゴリマップを示す。一致率は 52.8% となった。なお、写真が存在しなかった、あるいはどのカテゴリにも属しないと人手により判断されたブロックはアイコンなしとし、一致率の計算から除外した。

3.3 考察

一致率が低かった要因の一つとして、識別器の性能の問題が挙げられる。実験で使用した663枚の写真の39種類でのカテゴリ識別率は9.4%であった。39種類のカテゴリを画像特徴の類似性から5種類での風景カテゴリに再分類した結果、風景カテゴリ識別率は55.4%まで向上した。しかしながら、例えば金閣寺を撮影した写真には、街 (temple) と水辺 (pond) などの複数の風景カテゴリが含まれているものが多く存在し、誤識別の要因となった。

また、本実験では画像特徴の類似性から風景カテゴリを定義したが、本来ならば、ユーザの利用目的に合わせて風景カテゴリを定義すべきである。以上のことから、正解カテゴリラベルの与え方も含めた風景カテゴリの定義に関する検討が必要であると考えられる。

さらに、各ブロックの風景カテゴリを決める際に、ブロック内の最低写真枚数を導入することで、識別結果の信頼性が向上できると考えられる。

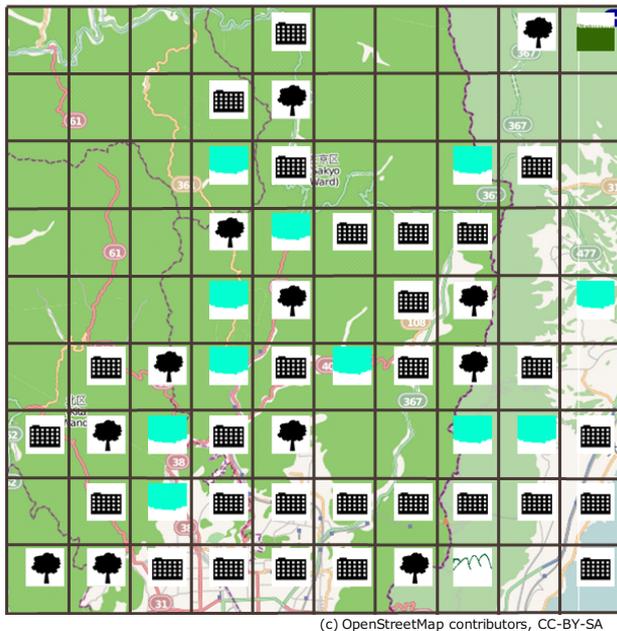


図 6 提案手法により作成された風景カテゴリマップ

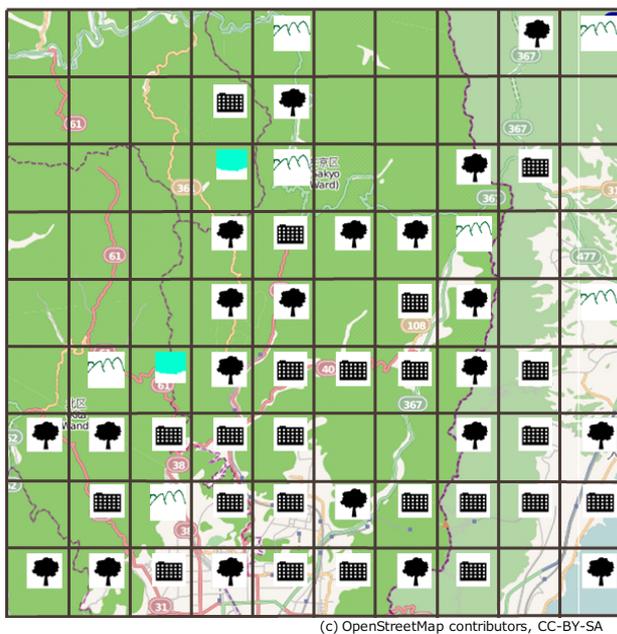


図 7 正解の風景カテゴリマップ

また、ブロック分割処理では、写真の位置情報だけに注目しているため、同じ場所の風景の写真が二つ以上のブロックに分割される可能性がある。これは、ユーザにとって直感的ではないと考えられる。このため、ブロックのサイズの自動調整などの処理も必要である。

4. む す び

本報告では、写真共有サイト上にアップロードされた大量のジオタグ付きデジタル写真に注目し、ユーザにとって直感的に理解しやすい風景カテゴリマップの作成手法を提案した。実験では、京都周辺を対象とした地域範囲に対して、提案手法を用いて風景カテゴリマップを作成し、人手により作成した風景

カテゴリマップと比較した結果、52.8%の一致率が得られた。今後の課題としては、風景カテゴリの定義に関する再検討が挙げられる。また、分割するブロックのサイズを可変にすることや隣接する同じ風景カテゴリのブロックを統合することなどによって、より柔軟に風景カテゴリマップを作成することを検討している。

文 献

- [1] Wang J., 野田雅文, 出口大輔, 井手一郎, 村瀬洋, “位置情報付き画像を用いた風景カテゴリマップ作成手法の検討”, 平成 22 年度電気系学会東海支部連合大会, N3-1, Aug. 2010.
- [2] Csurka G., Bray C., Dance C. and Fan L., Willamowski J., “Visual categorization with bags of keypoints”, Proc. ECCV International Workshop on Statistical Learning in Computer Vision, pp.1-22, Feb. 2004.
- [3] Li F. and Pietro P., “A Bayesian hierarchical model for learning natural scene categories”, Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp.524-531, May 2005.
- [4] Weijer J. and Schmid C., “Coloring local feature extraction”, Proc. European Conference on Computer Vision, pp. II.334-348, May 2006.
- [5] 串間和彦, 赤間浩樹, 紺谷精一, 山室 雅司, “色や形状等の表層の特徴量にもとづく画像内容検索技術”, 情報処理学会論文誌, Vol.40, No. SIG3(TOD1), pp.171-184, Feb. 1999.
- [6] 八重樫恵太, 柳井啓司, “撮影位置の情報をを用いた一般画像認識の可能性の検討”, 情報処理学会研究会, CVIM163-3, May 2008.
- [7] Zheng Y., Zhao M., Song Y., Adam H., Buddemeier U., Bissacco A., Brucher F., Chua T., and Neven H., “Tour the World: Building a webscale landmark recognition engine”, Proc. ACM MultiMedia, pp.961-962, Oct. 2009.
- [8] 柳井啓司, “キーワードと画像特徴を利用した WWW からの画像収集システム”, 情報処理学会論文誌, Vol.42, No. SIG10(TOD11), pp.79-91, Sep. 2001.
- [9] Lowe D., “Distinctive image features from scale-invariant keypoints”, International Journal of Computer Vision., Vol.60, No. 2, pp.91-110, Nov. 2004.
- [10] Xiao J., Hays J., Ehinger K., Oliva A., and Torralba A., “SUN Database: Large-scale scene recognition from abbey to zoo”, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, pp.3485-3492, June 2010.

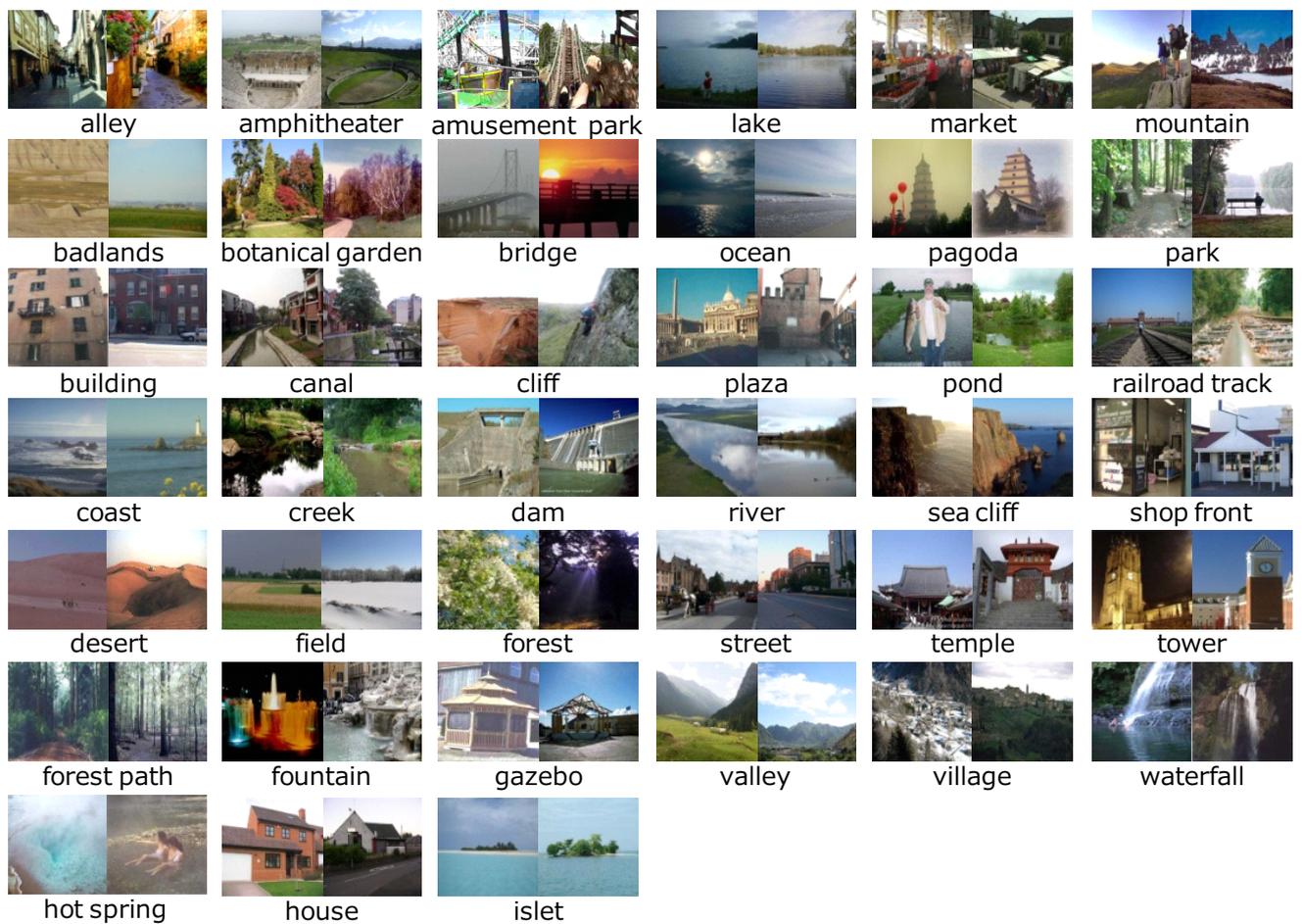


図 8 カテゴリと識別器の学習画像の例 (SUN Database [10])