BlueFinder: Estimate Where a Beach Photo Was Taken

Liangliang Cao, John Smith, Zhen Wen
IBM Watson Research Center
{lcao,jsmith,zhenwen}@us.ibm.com

Zhijun Yin, Xin Jin, Jiawei Han
University of Illinois at Urbana-Champaign
{zyin3,xinjin3,hanj}@illinois.edu

ABSTRACT

This paper describes a system to estimate geographical locations for beach photos. We develop an iterative method that not only trains visual classifiers but also discovers geographical clusters for beach regions. The results show that it is possible to recognize different beaches using visual information with reasonable accuracy, and our system works 27 times better than random guess for the geographical localization task.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

General Terms

Algorithm

Keywords

Geographical estimation, beach photos, geographical clustering

1. INTRODUCTION

Beaches are among the most attractive places people would like to go to during vacations and holidays. Beautiful beaches, especially those famous ones, are often of unique characteristics and bear a lot of photos taken by tourists. People are often very interested in recognizing the location of beach photos. They like to organize photos by locations and share with their friends. These phenomena motivate us to ask one question: Is it possible to design a photo management system which automatically estimates the geographical location of beach photos?

The general problem of geographical estimation based on visual features is very difficult for unconstrained photos. As observed by previous studies, estimation based on visual features only is no better than random guess [2]. Previous studies usually rely on user tags [1] [2] or travel histories, which limits the applications when no user information is given. In this paper, we argue that localizing the beach photos are easier than general photos. Beach scenes are less diversified, and the photos enjoy special characteristics like sands, unique water color, and coastlines that are related to their geographical locations. Figure 1 illustrates some examples of beach photos clustered by different locations.It is easy to observe the visual similarities of beach photo in the same cluster.

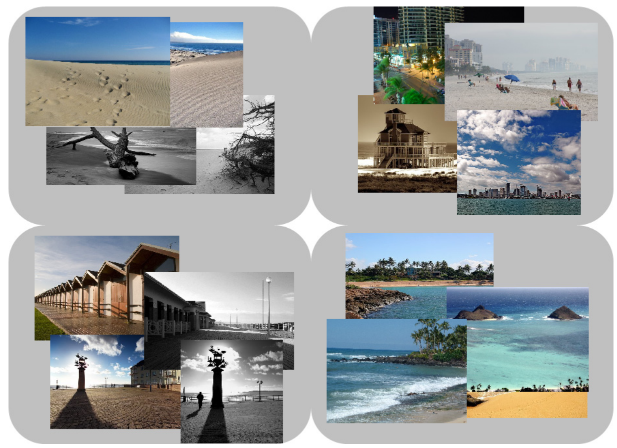

Figure 1: Examples of photos from different locations.

The problem of beach photo localization is also different from that of landmark recognition. One landmark usually corresponds one view or one subject with unique appearance, while a beach scene may contain a lot of clues including water, boats, people dresses, buildings and plants. Moreover, a landmark is usually limited to a point on the earth, while a beach usually covers a region. It is often inaccurate and also unnecessary to estimate the exact GPS coordinate for each query and we only need to estimate a coarse location for the beach photo.

To address the uniqueness of beach photo localization, an important component in beach photo localization is to find meaningful *geographical clusters* corresponding to different beaches. The use of geographical clusters benefits the problem of localization in two aspects: On the training stage, geographical clusters provide more training samples and hence lead to better recognition accuracy; on the testing stage, estimation of the most possible region for each query photo will be relatively easier than the estimation of GPS coordinates, while the information of geographical cluster will be good enough for trip planing and photo organization applications.

It is challenging to find meaningful geographical clusters. A naive way is to use fix grid to separate the earth surface into small regions. However, this method does not work for our system since the beach regions are of irregular shapes, and the grid may split a beach into several pieces. Another way is to use country borders to separate the geographical regions, which will become too coarse for large countries but too fine for small ones. This paper employs a density-based clustering method [1] to find the initialization of meaningful geographical clusters. The initialized cluster may not be optimal for beach localization since no visual features are considered in this stage. To develop a reliable beach recognition system, we need to take both visual similarity and geographical information into account. In this paper, we will introduce our geographical refining strategy which simultaneously learns visual characteristics and refines the geographical clusters. The details of our system will be explained in the following sections.

2. DATASET

A lot of beach photos are available in social media web sites associating with GPS tags. We collect 35K photos of Flickr photo with the tag of "beach" or "coast", of which each photo is labeled by a two-dimensional GPS coordinate vector. To provide a fair evaluation, we hand-select 1.1K testing photos with typical beach scenes, while removing photos with non-typical scenes such as personal portraits, and indoor pictures. The testing photos are evenly distributed in the world scale and there is no overlap between training and testing sets.

3. METHOD AND PERFORMANCE

Algorithm 1 illustrates the procedure of our system. We simultaneously train visual classifiers and refine geographical clusters in an iterative manner. At each iteration, we first refine geographical clusters using the GPS coordinates of beach photos. We select the efficient low dimensional mean-shift algorithm for the clustering task because it does not require predetermined cluster numbers and is scalable to millions or even billions of samples. Then we train a visual classification model for each cluster based on SIFT features. Given C geographical clusters, we decompose the learning problem into a series of $C(C-1)$ small problems, each of which corresponds to finding a decision boundary to separate photos from two clusters. In this case, each geographical cluster will at most win $C-1$ votes in the pairwise comparison. The estimated label is the class which won the largest number of votes. It is easy to see that an imperfect geographical cluster will harm the recognition accuracy. To refine geographical clusters, we evaluate the prediction scores for training photos, and compute the false alarm rate FA, together with the missed detection rate MD. For those with high FA scores, they usually cover a large region of land, we will split it into smaller regions. For those with low MD scores, they usually cover non-distinguishable regions and should be removed in training. Then the geographical clusters are refined and we will retrain the visual classification models for the next iteration. The iterative process will stop when geographical clusters are finalized with no splitting or removing operation. Figure 2 plot the distribution of beach photos and their corresponding geographical clusters in roughly different colors.

We compare the recognition results using our method, random

Algorithm 1 Simultaneously train visual classifiers and refine geographical clusters.

Input: Photos annotated with their GPS coordinates.
Initialize geographical clusters using efficient meanshift clustering.
repeat
- Train visual recognition model for each geographical cluster
- Based on the recognition model, associate each geographical cluster with photos according their visual recognition results
- Refine the clusters: split clusters with large FA and remove clusters with large MD

until The clusters are fixed.

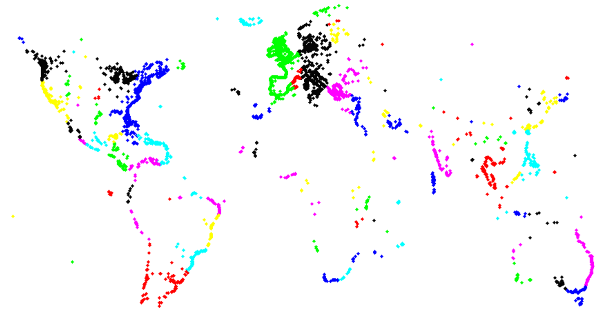

Figure 2: The geographical distribution of beach photos in the world. The horizontal and vertical axes denote longitude and latitude, respectively. The 144 geographical clusters are denoted by different colors.

Table 1: Recognition accuracy

Method	random guess	nearest-neighbor	our method
Accuracy	0.7%	10.3%	19.3%

Table 2: Percentage of localization within $\leq 5°$ neighborhood

Method	nearest-neighbor	our method
Percentage	6.7%	17.5%

guess and the nearest neighbor method [3]. The accuracy measures whether a testing photo is classified into the correct clusters. As shown in Table 1, our system using visual features works $19.3/0.7 = 27$ times better than random guess, and also better than the nearest neighbor method. Another intuitive way to evaluate our system is to check how far is the estimated location from its true GPS coordinates. We use the center of geographical cluster as the estimated location and compute the distance between estimation and GPS coordinates. As shown in Table 3, about 17.5% of testing photos are localized within a GPS coordinate neighborhood of 5 degrees.

4. DISCUSSIONS

This paper describes a system that can identify the geographical cluster and estimate the geographical locations for beach photos. In the future work, we will extend the current system for other concepts like buildings or landscapes, incorporate external knowledge of satellite photos, and design algorithms that are more reliable to noise effects.

5. REFERENCES

[1] L. Cao, J. Yu, J. Luo, and T. Huang. Enhancing semantic and geographic annotation of web images via logistic canonical correlation regression. In *ACM Multimedia*, 2009.

[2] D. J. Crandall, L. Backstrom, D. P. Huttenlocher, and J. M. Kleinberg. Mapping the world's photos. *WWW*, 2009.

[3] J. Hays and A. A. Efros. IM2PGS: estimating geographic information from a single image. In *CVPR*, 2008.