A Survey on Landmark Recognition Techniques

Prof. Rekhansh Rao[1] Prof. Toran Verma[2] Dipti Jaiswal[3]
[1]Assistant Professor [2]Associate Professor [3]M.Tech (CT) Scholar
[1,3]Department of Computer Science & Engineering [2]Department of Information & Technology
[1,2,3]Rungta College of Engineering and Technology, Bhilai, 490024, Chhattisgarh, India

Abstract— While it is probable to investigate the World Wide Web (WWW) for images, there is still no best way of searching the image. Hence recognizing and modeling landmarks at world-scale is a valuable and crucial task. First of all, the looks of any given landmark vary from one observation to successive. In addition to variations as a result of entirely diverse aspects, association in illuminated amendment, external litter, and dynamic geometry of the imaging devices are alternative factors affecting the variability of the ascertained landmarks. Finally, it is generally troublesome to create use of correct 3D information in landmark recognition applications. For those reasons, it is unacceptable to use several of the object recognition techniques. Hence it is vital to investigate the various techniques of landmark recognition so that we can find out the bottlenecks and how further improvement can be possible. In this paper, we perform a survey on various landmark recognition techniques and try to build a world-scale landmark identification engine.

Key words: WWW, GPS, PICASA

I. Introduction

As a traveller, how many times were you upset having to fail gathering interactive information about a couple of landmark or monument? The touristic landmarks are unit areas effortlessly recognizable and acknowledgeable sites and buildings, like a monument, church, etc.. as shown in Figure 1.1. They are the important parts of people's journeys, because of their illustrious physical appearance, cultural and historical options. The flare-up of non-public photography, along with web, has increased the exceptional growth of landmark picture sharing in several websites like Picasa Web Album.

With the massive amount of landmark pictures within the web, the time has arrived for computer conception to recognize landmarks universally, specifically to make a landmark identification engine, for the whole world. This engine shall not only have the capacity to identify the presence of destined landmarks in a picture, but also adds the information to a worldwide landmark information bank that arranges and catalogs landmarks, in terms of topographical places, admirations, traditional values and public meanings, etc. Such an earth-mapping landmark identification engine will be remarkably supportive for various hypermedia and vision uses. Primarily, by catching the pictorial features of landmarks, the engine will be able to deliver clean landmark pictures for constructing virtual sightseeing of a huge amount of landmarks. Additionally, by identifying landmarks, the engine will simplify both content understanding and geo-location discovery of pictures and videos. Lastly, by geologically establishing landmarks, the engine will help in an instinctive topographical investigation and navigation of landmarks in a local area, so as to provide tour guide suggestions and notions.

Fig. 1.1: Landmarks in India

To build such an earth-mapping landmark identification engine, the following issues, must be undertaken: (a) There is no properly prepared information bank of landmarks and landmark pictures in the world; (b) Even if there were such a list, it is still tough task to gather and organize correct landmark pictures; and (c) Efficiency is a nontrivial challenge for such a large-scale system.

II. Literature survey

While it is credible to mine the World Wide Web (WWW) for images, there is still no good solution for accessing the detail of the images from the web. Hence recognizing and modeling landmarks at world-scale is a valuable and crucial task.. Various works have been carried out in this field maintaining the integrity of the specifications.

Christos I. Colios et al. (2001) proposed a framework and presented a model for automated landmark identification to facilitate autonomous robotic navigation. Projective and point-permutation invariant vectors were employed to characterize landmark patterns. The use of these vectors relied on the existence of planar features in the workspace, which was valid in most indoor workspaces. The invariant vectors established direct point to point correspondences, which enabled the use of the convex hull and projectivity constraints. These constraints minimized the risk of mismatches across quintuples in different image frames. Moreover, the final selection of sub-landmarks as outlier patterns facilitated robust landmark recognition during navigation. Experiments with this approach had revealed its effectiveness for a realistic indoor environment.

M. Mata et al. (2001) defined a vision-based landmark identification approach for use with mobile robot routing tasks. A search procedure based on inherent techniques for pattern identification in digital images had been offered. The candidate structure allowed the topological localization of a mobile robot using natural and artificial landmarks. Text strings within landmarks picture could be captured and understood, if present. The resulting method had been verified on board a B21 mobile robot and proved useful.

Yan-Tao Zheng et al. (2009) explained about some efficiency issues of GPS based web enabled landmark recognition. In the processing pipeline, the geo-clustering of GPS tagged photos and landmark mining from tour guide corpus did not demand high efficiency requirement, due to the low dimensionality of GPS coordinates and relatively small tour guide corpus size. However, the large amount (~21.4 million) of raw input images and large magnitude of land mark models made efficiency essential in two aspects: (1) the landmark image mining and (2) landmark recognition of query images.To achieve efficiency, they exploited the following three measures.

- Parallel computing to mine true landmark images: The visual clustering process on each noisy image set did not interfere with each other. This enabled them to speed up the clustering process drastically by running parallel visual clustering on multiple machines.
- Efficiency in hierarchical clustering: By adopting single linkage, the shortest path between two clusters was equal to the shortest path of two regions in clusters, which had been computed in the phase of image matching. The clustering process was then equivalent to erasing graph edges above a certain distance threshold and collecting the remaining connected region sets as clusters.
- Indexing local feature for matching: To achieve fast image matching, they adopted the k-d tree to index local features of images. This allowed the local feature matching time to become sub-linear, thus enabling efficient recognition of query images. In their experiments, the time it took to recognize landmark in a query images is only ∼0.2 seconds in a P-IV computer.

Mohammed Elmogy et al. (2009) presented current effort toward building a robust and fast landmark recognition system. Author used a more natural approach in terms of computational efficiency to recognize the landmark online during robot navigation. The color histograms of the detected landmarks were used to provide the rough initial estimate of the landmark. Then they processed the resulted hypotheses to calculate an accurate estimation of the landmark. The route topological map was used to decrease the processing time by processing only the landmarks mentioned in the route description and ignoring other landmarks during navigation. In addition, authors used the robot’s stereo vision combined with the classified landmarks to find the nearest landmark to the robot and also to calculate the landmark’s position in the real world.

Pablo Sala et al. (2006) presented a novel graph theoretic formulation of the problem of automatically extracting an optimal set of landmarks from an environment for visual navigation. It’s intractable complexity motivated the need for approximation algorithms, and presented six algorithms. To systematically evaluate them, they first tested them on a simulator, where they could vary the shape of the world, the number and shape of obstacles, the distribution of the features, and the visibility of the features. The algorithm that achieved the best results on synthetic data was then demonstrated on real visibility data. The resulting decompositions found large regions in the world in which a small number of features could be tracked to support efficient online localization. Their formulation and solution of the problem were general, and could accommodate other classes of image features.

D. Meyer-Delius et al. (2011) presented a landmark placement approach that seeks to reduce the overall ambiguity in the environment to improve the localization performance of a mobile robot. To this extend they proposed a measure for the uniqueness of a robot pose based on the appearance of the environment as observed by the robot. Due to the combinatorial nature of the landmark placement problem, they introduced an approximate approach that incrementally selected landmark locations from a set of candidate locations and thereby maximized the average uniqueness in the environment. Furthermore, they described a concrete application in the context of localization with laser range scanners given a grid-based representation of the environment. They evaluated their approach for different environments in simulation and using real data. The results demonstrated that their approach yielded substantial improvements in the localization performance.

In Google Goggles, there had been interest in the problem of landmark identification in the computer vision community for about a decade. However, developments in landmark identification in a mobile application setting have been relatively recent. Advent of commercial large-scale applications like Google Goggles (see Fig.3.1) has spurred further interest in this area. However, most commercial applications rely heavily on server-side support and processing. The authors presented an approach that relies on incorporating priors from noisy user location data, in addition to using the image content itself for identification.
Monulens does not rely on any server-side support for processing, which helps in curbing the latency and power requirements associated with communication to and from the server. The entire processing cycle takes place on the Android device itself and the match is displayed. While client side implementations offer significant benefits in terms of power and latency. One of the major drawbacks of a server independent implementation is that data-intensive applications consume a lot of storage space on the portable device. In the context of applications such as landmark identification, storing data pertaining to tens of thousands of landmarks on a mobile device is practically infeasible. In order to overcome this drawback, they derived inspiration from “maps” applications. Many maps applications allow the user to download data pertaining only to the regions that they would be interested in. This data is downloaded as an array of latitude, longitude values, topology and contours. In the absence of an active connection, the user can still localize her on the map, since GPS does not rely on a network connection.

Fig. 3.1: Google Goggles

III. PROBLEM IDENTIFICATION

A landmark could be an object, created by man or naturally, with a high identification value. Sometimes a landmark is of exceptional size and is found on a solid and wild position of the earth. Samples of landmarks are buildings, monuments, statues, parks, mountains and different structures and places. Some problems regarding landmark identification are identified as follows:

- Key point detection/description and RANSAC [algorithm used to estimate the homography] are the computationally expensive steps.
- Due to Less efficient key point descriptors used in earlier approach, the performance of algorithm decreases.
- No cache-blocking mechanism therefore it takes time to compare with the reference image.
- Real time Computation of key points/descriptors.

IV. PROPOSED SOLUTION

A. Algorithm Optimization:

Here we try to develop such a landmark recognition engine that would overcome the problem used in the key point detection step. For this we will use Harris corner detection algorithm.

B. Cache-blocking/Prefetching:

Using Cache-blocking or prefetching, we can decrease the time taken to identify the landmarks that will improve the performance of the newly built engine.

C. Pre-Computing Key Points/Descriptors of Reference Images:

We can pre-compute the key points of the referenced images so that at the comparison time it does not require to repeat the same step.

Author Name	**Year**	**Title**	**Description**	**Outcome**
Pablo Sala, Robert Sim, Ali Shokoufandeh and Sven Dickinson	2006	Landmark Selection for Vision-Based Navigation	Decompositions from the real image data to measure the localization performance versus the non-decomposed map.	Outcome Efficiency depends on visual map ,Mean localization and localization time
D. Meyer-Delius, M. Beinhofer, Alexander Kleiner and W. Burgard	2011	Using artificial landmarks to reduce the ambiguity in the environment of a mobile robot	Presents a practical approach to compute a configuration of indistinguishable landmarks that decreases the overall ambiguity and thus increases the robustness of the localization process.	Depends on global localization and Uniform contour, space sampling
Andres Solis Montero, Hicham Sekkati	2012	Framework for Natural Landmark-based Robot Localization	Uses all matching points to improve pose estimation and an off-line target evaluation strategy to improve a priori map building.	80% True Detection depends on strongest keypoints
Aditya Srinivas Timmaraju and Anirban Chatterjee	2013	Real-time mobile-based landmark recognition application called Monulens. Contrary to most applications, Monulens does all processing on the mobile device alone, and does not require any server-side support.	Uses OpenCV functions though the Java Native Interface (JNI).ORB (Oriented and Rotated BRIEF) feature descriptor	Cross device validation done 90% True Detection

Table 2.1: Comparison

V. CONCLUSION

The phenomenal development of tourism related multimedia data in the Internet, such as the GPS-tagged photos and tour guide web pages, has optimistic computer vision investigators to think about landmarks globally. Here, we try to build a world-scale landmark identification engine, which organizes models and recognizes the landmarks on the scale of the entire planet Earth. Constructing such an engine is, in real meaning, a multi-source and multi-modal data mining task. Henceforth we need to analyze existing techniques. The objective of the mobile identification system is to analyze the effectiveness of the landmark recognition framework in a real world scenario, and to collect user feedback on the effectiveness and accuracy of the framework. In comparison table we have compared different literatures and find out some bottlenecks in existing techniques. In problem identification section we have mentioned some tailback of existing techniques such as

Monulens, Google Goggles if these tribulations can be solved then performance will increase.

References

[1] Magdalena Rischka and Stefan Conrad “Landmark Recognition: State-of-the-Art Methods in a Large-Scale Scenario",Proceedings of the LWA 2014Workshops: KDML, IR, FGWM, Aachen, Germany, 8-10 September 2014.
[2] Alexandre Alahi, Raphael Ortiz and Pierre Vandergheynst “FREAK:Fast retina keypoint” IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012.
[3] Andres Solis Montero, Hicham Sekkati Jochen Lang, Robert Laganiere and Jeremy James "Framework for Natural Landmark-based Robot Localization" ,Ninth IEEE Conference on Computer and Robot Vision,(2012), 131 – 138.
[4] Ethan Rublee et al. “ORB: an efficient alternative to SIFT or SURF”Proceedings of IEEE International Conference on Computer Vision (ICCV) 2011.
[5] David Chen et al. “City-Scale landmark identification on mobile devices”Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011.
[6] D. Meyer-Delius, M. Beinhofer, Alexander Kleiner and W. Burgard , ”Using artificial landmarks to reduce the ambiguity in the environment of a mobile robot”, IEEE International Conference on Robotics and Automation (ICRA) (2011), 5173-5178.
[7] Yan-Tao Zheng, Ming Zhao, Yang Song, Hartwig Adam Ulrich Busddemeier, Alessandro Bissacco, Fernando Brucher Tat-Seng Chua, and Hartmut Neven ” Tour the World: building a web-scale landmark recognition engine” Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE ConferenceComputer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference,1085 - 1092.
[8] Mohammed Elmogy and Jianwei Zhang” Robust Real-time Landmark Recognition for Humanoid Robot Navigation” International Conference on Robotics and Biomimetics Bangkok, Thailand, Proceedings of the 2008 IEEE February 21 - 26, 2009 .
[9] Pablo Sala, Robert Sim, Ali Shokoufandeh and Sven Dickinson’”Landmark Selection for Vision-Based Navigation” IEEE TRANSACTIONS ON ROBOTICS, VOL. 22, NO. 2, APRIL 2006.
[10] Aditya Srinivas Timmaraju and Anirban Chatterjee"Monulens: Real-time mobile-based Landmark Recognition",Stanford University 2013.
[11] David Lowe “Distinctive image features from scale-invariant keypoints”International Journal of Computer Vision 60.2 (2004): 91-110.
[12] Josef Sivic and Andrew Zisserman “Video Google: A text retrieval approach to object matching in videos” Proceedings of IEEE International Conference on Computer Vision (ICCV), 2003.
[13] Christos I. Colios , Panos E. Trahanias "A framework for visual landmark identification based on projective and point-permutation invariant vectors",C.I. Colios, P.E. Trahanias / Robotics and Autonomous Systems[Elsevier] 35 (2001) 37–51.
[14] M. Mata, J. M. Armingol, A. de la Escalera and M. A. Salichs “A Visual Landmark Recognition System for Topological Navigation of Mobile Robots” Robotics and Automation. Proceedings ICRA. IEEE International Conference on (Volume:2) (2001) 1124 - 1129 vol.2.
[15] Barbara Zitova , Jan FlusserBarbara Zitov, Jan Flusser "Landmark recognition using invariant features" Pattern Recognition Letters 20 (1999) 541-547.
[16] Martin A. Fischler and Robert C. Bolles “Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography” Communications of the ACM 24.6 (1981): 381-395.