

A Review on Advanced Mechanism in Learning and Recognition of Ops from Weakly Labeled Street View Images

Vaishali A. Mahale¹, Prashant M. Yawalkar²

Student, Computer Department, MET BKC College, Nashik, India¹

Professor, Computer Department, MET BKC College, Nashik, India²

Abstract: Mobile phones are powerful image and video processing device containing the other various features like high-resolution cameras, display color, and hardware-accelerated graphics. The various applications give rise to a key technique of daily life visual object recognition. On premise sign is a popular form of commercial advertising, widely used in our daily life. The OPSs containing visual diversity associate with complex environmental conditions. Observing that, today's in most of existing image data set OPSs characteristics are lacking. In this, first proposed an OPS-62 data set, in which totally 4649 OPS images of 62 different businesses are collected from Google's Street View. For addressing the problem of real-world OPS learning and recognition, was developed a probabilistic framework based on the distributional clustering, in order to exploit the distributional information of each visual feature. Learning OPS images for more accurate recognitions and less false alarms. Main approach is simple, linear, and can be executed in a parallel fashion, making it practical and scalable for large scale multimedia applications.

Keywords: Index Terms-Real-world objects, street view scenes, learning and recognition, object image data set.

I. INTRODUCTION

The presence of mobile devices in our daily life has changed the way we connect with the world. Users depend on mobile devices to maintain information and updates related to all general purpose applications. OPSs (On Premise Sign) are commonly erected on buildings in modern cities. These OPSs exhibit great visual diversity. Everyday objects or stores by using image based mobile interactions through their devices. For example, as users walk on the street, they might simply point the mobile camera to a store on the street to quickly access its related information, inquire special offers, and make reservations through their mobile devices without physically entering the store.

To build a recognizable image for a business to attract customers, each store has its own on-premise sign which is a visually consistent image for a brand and contains a mixture of text and graphics [5].

Therefore, In this paper proposes a probabilistic framework for learning and recognizing OPSs in real scene images has been discussed. Real-world characteristics of erected OPSs, such as arbitrary size, position, viewing angles, perspective distortion, occlusions, varying lighting conditions, foreground and background clutter, etc., make logos, trademarks, or texts in OPSs occupy a relatively smaller area or be occluded by other objects in real scene images. These characteristics make the existing solutions fail to identify logos, trademarks, or texts in OPSs [2]. The main approach exploits a probabilistic framework to extract discriminative visual words of each OPS category, and therefore is able to recognize and localize each OPS within images by using the learnt OPS model.

II. RELATED WORK

The task of learning and recognizing OPSs in real world images can be viewed as a problem of object recognition and localization. Firstly, To understand the value of OPS, the importance of signage in the business community D. Conroy survey and explain The role of the signage in businesses and organizations for attract new customers, Brand your site in the minds of consumers, ed.,the cost effectiveness in building, installing, and maintaining a sign, as well as provide analysis of how it can impact on profitability. Designing the right signage for business consider the factors such as visibility, readability, size, scale, location, cone of the vision and angle, brightness, letter style and also consider the types of signs. Also introduced about legal considerations for a sign codes and also describe about the legal rights for the sign owner [1].

2.1. Bag-Of-Features Image classification:

Bag-of-features representations contain a content based image classification owing to their simplicity and good performance. The basic idea is to treat images as loose collections of independent patches, sampling a representative set of patches, evaluating a visual descriptor vector for each patch independently, and using the resulting distribution of samples in descriptor space as a characterization of the image. In this paper, shows that through experimentation that for a representative selection of commonly used test databases and for moderate to large numbers of samples, random sampling gives equal or better classifiers than the sophisticated multi-scale interest operators that are in common use. Although interest operators work well for small numbers of samples, the single most important factor governing performance is the number of patches sampled from the test image and

ultimately interest operators cannot provide enough patches to compete [2].

2.2. Graph Based Visual Saliency:

A new visual saliency model is Graph-Based Visual Saliency (GBVS). It consists of two steps, first is forming activation maps on certain feature channels, and second is normalizing them in a way which highlights conspicuity and admits combination with other maps. This model is simple and biologically plausible insofar as it is naturally parallelized. Through the experiment, Observed that the GBVS model powerfully predicts human fixations on 749 variations of 108 natural images, achieving 98% of the ROC area of a human-based control, whereas the classical algorithms achieve only 84%. [3]

2.3. Spatial Pyramid Mining:

spatial pyramid mining, to discover association rules at multiple resolutions in order to identify frequent spatial configurations of local features that correspond to classes of logos appearing in real world scenes. By indexing representative examples by the mined rules they can efficiently detect a variety of different lettering or design marks associated with a brand. Features in an image are marked by matching rules to representative examples selected via a weighted cosine similarity measure. Logos are localized in an image via density-based clustering of matched features. Precision vs. recall curves are presented for experiments on a dataset of web images of nearly 1,000 images. An important limitation of this method is image resolution, as multiple local features are required to mine robust spatial configurations [4].

2.4 (pLSA) technique for text document analysis

The internet contains vast quantity of visual data in form of images. To effectively search the visual content of the Internet, to tackle this problem using recognize the object categories present in an image. This task requires learning models of the visual appearance of object categories which typically uses data sets of images which have been manually "labeled" to indicate the objects present. To analysis Introduced a simple approach to learning models of visual object categories from images gathered from Internet image search engines. The images for a given keyword are typically highly variable, with a large fraction being unrelated to the query term, and thus pose a challenging environment from which to learn. By training given models directly from Internet images, they remove the need to laboriously compile training data sets, required by most other recognition approaches- this opens up the possibility of learning object category models "on-the-fly." They describe two simple approaches, derived from the probabilistic latent semantic analysis (pLSA) technique for text document analysis that can be used to automatically learn object models from these data. They show two applications of the learned model: first, to re-rank the images returned by the search engine, thus improving the quality of the search engine; and second, to recognize objects in other image data sets [5].

2.5 Auxiliary Visual Words Discovery:

In this paper, observing the problems in large-scale image object retrieval by conventional Bag of Words model.

After observing the problem, they proposed a auxiliary visual words discovery through visual and textual clusters in an unsupervised and scalable fashion. Investigate the variant optimization methods for efficiency and accuracy in Auxiliary Visual Words discovery conducting experiments on consumer photos and showing great improvement of recall rate for image object retrieval. This paper elaborates the problems of current BoW model and the needs for semantic visual words to improve the recall rate for image object retrieval. They proposed to augment each database image with semantically related auxiliary visual words by propagating and selecting those informative and representative VWs in visual and textual clusters. Also they formulate the processes as unsupervised optimization problems. Experiment results show that they can greatly improve the retrieval accuracy compared to the BoW model (111% relatively) [6].

2.6 Cascaded Index:

To achieve a highly effective and scalable framework for recognizing logos in images proposed a cascaded index technique was proposed. At the core of their approach lays a method for encoding and indexing the relative spatial layout of local features detected in the logo images. Based on the analysis of the local features and the composition of basic spatial structures, such as edges and triangles, one can derive a quantized representation of the regions in the logos and minimize the false positive detections. Furthermore, they propose a cascaded index for scalable multi-class recognition of logos. For the evaluation of their given system, they had construct and released a logo recognition benchmark which consists of manually labeled logo images, complemented with non- logo images, all posted on Flickr. The dataset consists of training, validation, and test set with 32 logo-classes. Thoroughly evaluated their system with this benchmark and showed that proposed approach becomes effective in recognizing different logo classes with high precision [7].

2.7 Mobile Queue System:

To make effective and scalable use for the customers a system which consists of A mobile queue-card management system that offers more freedom to customers by enabling image-based queue-card retrieving and service-information querying actions using mobile phones. Mobile Queue interacts with cloud services allowing customers to query summary description and availability of services provided by stores. Mobile Queue also offers suggestions to waiting customers such as potentially interesting substitute activities and stores [8].

2.8 Detecting Logo Hypotheses and Burstiness Handling:

Detecting logos in photos is challenging task. A reason is that logos locally resemble patterns frequently seen in random images. J.Revaud proposed to learn a statistical model for the distribution of incorrect detections output by an image matching algorithm. It results in a novel scoring criterion in which the weight of correlated key point matches is reduced, penalizing irrelevant logo detections. Through the experiments on two very different logo retrieval benchmarks, it can be seen that the proposed

approach largely improves over the standard matching criterion as well as other state-of-the-art approaches [9].

3. DESIGN IMPLEMENTATION AND STRATEGY

• Overview of the proposed model

The motive of this research is to addressing the problem of real world OPS learning and recognition of weakly labeled street view images. Here, to tackle this problem by developing a probabilistic framework based on distributional clustering, in which Tsung-hung tsai proposed to exploit distributional information of each visual features. Through the experiments on OPS-62 dataset demonstrated the performance of given approach over a state of art probabilistic latent semantic analysis model for more accurate recognition and improvement in average recognition rate. To implement the proposed approach two main algorithms are used:

- 1] Visual saliency based codebook generation of OPS categories.
- 2] OPS modeling and recognition using distributional clustering.

System design:

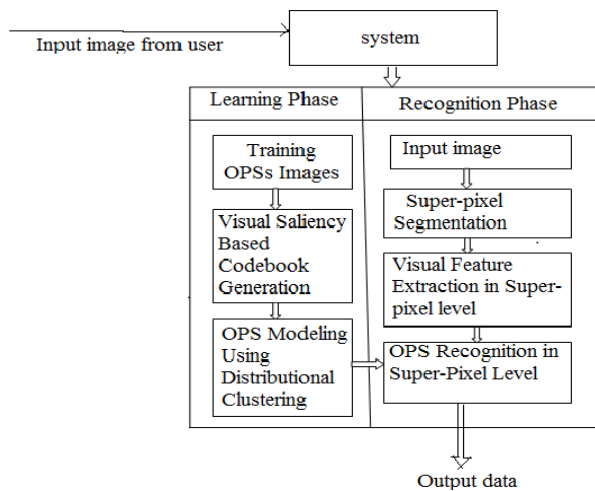


Fig 1: System Architecture

In this system, user gives input as image which is captured through mobile cameras. System gives input images and perform the actual proposed framework on given input image. In this framework two basic algorithms are used: first is visual saliency based codebook generation of OPS categories. In this algorithm first, filter out the background region for reducing the number of noisy visual word using visual saliency analysis. To filter out the background noise here Graph based visual saliency technique is implemented. After removing the background noise, visual feature are extracted using dense sampling strategy and Opponent SIFT descriptor for codebook Generation. After obtaining a codebook for each OPS category compute a discriminative subset and apply a distributional clustering to group all the code words in OPS categories into two disjoint clusters. Then conduct the concurrent OPS recognition and localization in super-pixel level using obtained OPS and background models.

4. CONCLUSION

Learning framework and text recognition have been solving the problem of real world street view challenges to identify business entities in street view images. Learning technique is exploited to benefit the selection of discriminative visual words and construct effective OPS model by using distribution clustering. The OPS -62 dataset which contain more real world images for visual object recognition. The visual saliency based codebook generation and OPS recognition using distributional clustering algorithms will able to improve average recognition rate as compared to the state-of-the-art pLSA models. However, in view of the low average recall values relatively, the OPS recognition in real-world scenes is still challenging problem. Further it may be possible to extend the technique for giving the offers relevant to retrieved image to the user through email or messages.

REFERENCES

- [1] D. Conroy, "What's Your Signage (How On-Premise Signs Help Small Businesses Tap into a Hidden Profit Center)." New York, NY, USA: State Small Bus. Develop. Center, 2004.
- [2] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in Proc. ECCV, 2006, pp. 490-503.
- [3] J. Harel, C. Koch, and P. Perona, "Graph-based visual saliency," in Proc. ANIPS, 2007, pp. 545-552.
- [4] J. Kleban, X. Xie, and W.-Y. Ma, "Spatial pyramid mining for logo detection in natural scenes," in Proc. IEEE ICME, Apr. 2008, pp. 1077-1080.
- [5] R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, "Learning object categories from internet image searches," Proc. IEEE, vol. 98, no. 8, pp. 1453-1466, Aug. 2010.
- [6] Y.-H. Kuo, H.-T. Lin, W.-H. Cheng, Y.-H. Yang, and W. H. Hsu, "Unsupervised auxiliary visual words discovery for large-scale image object retrieval," in Proc. IEEE Conf. CVPR, Jun. 2011, pp. 905-912.
- [7] S. Romberg, L. G. Pueyo, R. Lienhart, and R. van Zwol, "Scalable logo recognition in real-world images," in Proc. 1st ACM ICMR, 2011, pp. 1-25.
- [8] C.-W. You, W.-H. Cheng, A. W. Tsui, T.-H. Tsai, and A. Campbell, "MobileQueue: An image-based queue card retrieving system through augmented reality phones," in Proc. 14th ACM Int. Conf. Ubiquitous Comput, 2012, pp. 1-2.
- [9] J. Revaud, M. Douze, and C. Schmid, "Correlation-based burstiness for logo retrieval," in Proc. 20th ACM Int. Conf. Multimedia, 2012, pp. 965-968.
- [10] Tsung-Hung Tsai, Wen-Huang Cheng, Chuang-Wen You, Min-Chun Hu, Arvin Wen Tsui, and Heng-Yu Chi "Learning and Recognition of On-Premise Signs from Weakly Labeled Street View Images", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 23, NO. 3, MARCH 2014