MAT JOURNALS

Journal of Image Processing and Artificial Intelligence
Volume 3 Issue 3

Automatic and Efficient Cleansing of Illustration Images in Web

***REVATHY* [1], *SUGANYADEVI*[2], *SUPRIYAA*[3]**
Department Of Computer Science and Engineering, SSN College of Engineering
***PRIYADHARSINI.R*[4], *MANISHA.S*[5]**
Assistant Professor, Department Of Computer Science and Engineering, SSN College of Engineering
priyadharsinir@ssn.edu.in

Abstract
The scope and nature of image data is crucial to understand and to determine the complexity of image search design. Interest in image retrieval has increased in large due to the rapid growth of the World Wide Web. There are huge number of high quality images for different image category available in web. When a search query is given, the information retrieval system gives us both relevant and irrelevant images to the users. In order to satisfy the requirement of the user and to give relevant details, there are many interactive and automatic methods that exists. The interactive methods are capable of building large collection of images with ground truth labels, but they depend heavily on human efforts. While Automatic methods leverage an object category model trained on text and visual features. The objective of this work is to review the works both interactive and automatic methods proposed for generating a large number of images for a specified object class.

Keywords: *Automatic Cleansing, Image Retrieval, Interactive Cleansing*.

INTRODUCTION

Image retrieval is a process of browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some methods of adding metadata such as captioning, includ ing key words or descriptions to the images, so that retrieval can be performed over the annotation words. Manual image annotation is time-consuming, laborious and expensive. To address this, there has been a large amount of research done on an automatic image annotation. The automatic image filtering is a focal problem in image processing and computer vision. Interactive cleansing requires human interaction and they don't address polysemy, which is the coexistence of many possible meanings for a word or phrase. Automatic cleansing or filtering of images nullifies human intervention and completely rectifies polysemy. We seek to build a large collection of images to be used for object detection and recognition research using automatic methods. Such data are useful for supervised learning and quantitative evaluation.

Noisy Web images are obtained by either text based or visual based web search querying. The web pages and the images are downloaded from the web. Irrelevant images are then removed to form a candidate web image collection through automatic cleansing. The prime advantage of automatic approach is that we can employ both text-based and visual-based image filtering to remove the illustration images, which have obvious differences with the images of the target object category in terms of text and visual features.

The paper is organized as follows. Section II involves the system overview. Section III portraits the comparison of interactive and automatic methods. Section IV deals with the detailed discussion of various methods. Section V encompasses the similarities and the differences in a tabular format. Conclusions are discussed in section VI.

OVERVIEW

In this paper, we compare and contrast the various methods of cleansing the web images. For interactive approach, LabelMe , ImageNet and Wordnet have been studied and discussed. The effectiveness of automatic approach has been witnessed through Optimol, probabilistic Latent Semantic Analysis and Ranking using Bayes Classifier methods .

INTERACTIVE AND AUTOMATIC METHODS

In interactive cleansing method, a collection of images are downloaded. This set includes illustration images which have obvious difference from the original images. As this method includes manual annotation, control points are connected and polygons(different shapes) are formed. Once the regions are marked, the objects can be labelled. Some labels may be generic and some may be specific. In order to avoid repetitions of same labels, the annotated images are regularly optimised. Thereby, properly labelled images are augmented to the existing image set. This way, cleansed image set without noisy images are obtained. The figure1. depicts the interactive fashion of cleansing images

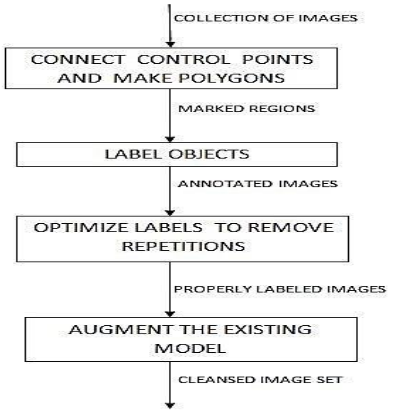

Fig1: *Flow chart representing the overview of Interactive Cleansing*

In automatic cleansing, images are gathered by either textual or visual based querying.These images contain noisy images including sketches, paintings, drawings, graphs, charts, maps etc. In order to remove these, initially few correct images are given as seed images. The image set is then classified as abstract and non-abstract images. Abstract set includes illustration images, while the non-abstract images refer to the realistic images.An object category model is learnt from the features of the images such as shape, texture etc. The model is then ranked based on textual and visual features. Based on the ranking, the images are categorized and augmented with the existing model. Thus cleansed image set is obtained without any manual Intervention.

The figure 2 emphasizes the steps involved in automatic cleansing of images.

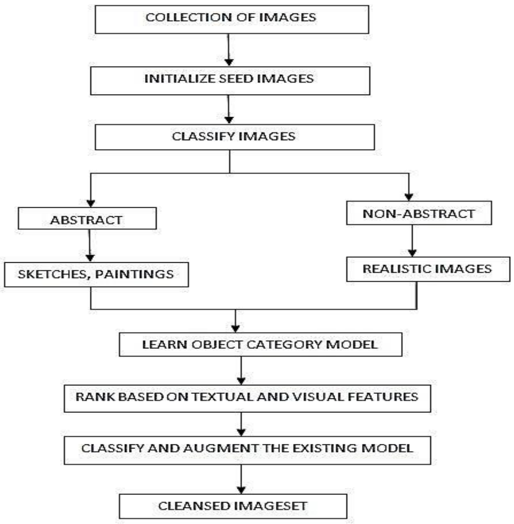

Fig2: *Flow chart representing the overview of Automatic Cleansing*

COMPARATIVE ANALYSIS

INTERACTIVE APPROACHES

LABELME

LabelMe is a web-based tool that allows easy image annotation and instant sharing of such annotations. This tool provides functionalities such as drawing polygons, querying images, and browsing the database. For each object present in an image, the labels should provide information about the object's identity, shape, location, and possibly other attributes such as pose. A Javascript drawing tool is designed. The user may label a new object by clicking control points along the object's boundary. The resulting labels are stored in the XML file format, which makes the annotations portable and easy to extend. The images and annotations are organized into folders. The folders are grouped into two main categories such as static pictures and sequences extracted from video. Numbers of control points determine the object that is being categorized on a large hand.The major advantages of LabelMeare :

- Designed for object class recognition as opposed to instance recognition and for learning about objects that are embedded in a scene.
- High quality labeling and many diverse object classes.

However, there are significant drawbacks such as quality control, complexity of the polygons provided by the users. The issue of unifying the terminologyis to properly index the dataset according to real object categories, uniform distribution of objects with respect to size and image location.

Fig3:*LabelMe tool*

Other major issues are: There can be a large variance of terms that describe the same object category. The level of description provided by the users may vary. Overlapping polygons has to be dealt seriously, since the problem of inferring depth ordering for overlapping regions is a simpler problem. Some objects are always on the bottom layer since they cannot occlude any objects. If two polygons overlap, the polygon that has more control points in the region of intersection is more likely to be on top. We can use histogram intersection to assign the region of intersection to the polygon with the closest colour histogram.

WORDNET

WordNetis an electronic dictionary, to extend the LabelMe descriptions.WordNet organizes semantic categories into a tree. The tree representation allows disambiguation of different senses of a word (polysemy) and relates different words with similar meanings (synonyms). For each possible description, different senses of meanings have been queried to WordNet. Among the returned senses the one that best matched the description was chosen. The major drawbacks are

i) Manually creating associations between the different
text descriptions and WordNet tree nodes.
ii)Necessary to frequently update these associations since the rate of new descriptions entered into LabelMe decreases over time

IMAGENET

Imagenet is a large-scale hierarchical Image Database. ImageNet aims to populate the majority of the 80,000 synsets of WordNet with an average of 500-1000 clean and full resolution images. The explosion of image data on the Internet has the potential to foster more sophisticated and robust models and algorithms to index, retrieve, organize and interact with imagesand multimedia data. But exactly how such data can be harnessed and organized remains a critical problem.ImageNet can be improvised in these ways : User labels can be evaluated to optimize the number of repetitions needed to accurately verify each image and also ImageNet can be made a central resource for broad range of vision related research. Through these improvisations, the construction process can be speeded up.

AUTOMATIC APPROACHES

OPTIMOL

OPTIMOL stands for Automatic Online Picture Collection via Incremental Model Learning. OPTIMOL is capable of automatically collecting much larger object category datasets for 22 randomly selected classes from the Caltech 101 dataset. A robust object category model and meaningful image annotation are provided by this algorithm.Current commercial image retrieval software is built upon text search techniques using the keywords embedded in the image link or tag. As we know, retrieved image is highly contaminated with visually irrelevant images. Extracting the useful information from this noisy pool of retrieved images is quite critical.Polysemy is common in the retrieved images, e.g. a “mouse” can be either a “computer mouse” or an “animal mouse”.

For every object category, dataset is being initialized with seed images. This can be done either manually or automatically. With this small dataset, the iterative process of model learning and dataset collection is begun. Learning is done via an incremental learning process. Given the current updated model of the object class, a binary classification on a subset of images downloaded from the web is performed.If an image is accepted based on statistical criteria , the existing dataset is augmented by appending this new image. The model is then updated with a subset of the newly accepted images. In this method, the already existing images in the dataset no longer participate in the iteration of learning. In the meantime, the background model will also be updated using a constant resource of background images. This process is repeated till a sufficient dataset is collected or all downloaded images are exhausted. Introducing better descriptive models can be the future work of Optimol.

PROBABILISTIC LATENT SEMANTIC ANALYSIS

Current approaches to object category recognition require datasets of training images to be manually prepared, with varying degrees of supervision. An approach that can learn an object category from just its name, by utilizing the raw output of image search engines available on the Internet was presented. A new model, TSI-pLSA, which extends pLSA (as applied to visual words) to include spatial information in a translation and scale invariant manner. It is evident that this approach can handle the high intra-class variability and large proportion of unrelated images returned by search engines.

A leading approach in this field is that of probabilistic Latent Semantic Analysis (pLSA) and its hierarchical Bayesian form, Latent Dirichlet Allocation (LDA).pLSA methods to incorporate spatial information in a translation and scale invariant manner has been adopted and extended. It is then applied to the more challenging problem of learning from search engine images. To enable comparison with existing object recognition approaches, the learnt models are tested on standard datasets. The training sets are extremely noisy yet, for the most part, the results are competitive (or close to) existing methods requiring hand gathered collections of images. This was achieved by improving state-of-the-art pLSA models with spatial information.

IMAGE HARVESTING ALGORITHM

MAT JOURNALS

Journal of Image Processing and Artificial Intelligence
Volume 3 Issue 3

Image Harvesting Algorithm deals with Candidate images, which are obtained by a textbasedwebsearchquerying on the object identifier. It aims to provide training databases so that a new object model can be learnt effortlessly. The method involves four steps.

- Images are downloaded through various searches such as Web Search- submits the query word to Google. Image search- each of the returned images is treated as a "seed" – further images are downloaded from the web page from where the seed image originated. Google Search-query can consist of a single word or more specific descriptions.Images divided into in class good, in class ok and no class. In class is further subdivided into abstract (not natural images) and non-abstract. Only WebSearch and GoogleImages are used, and their images are merged into one dataset per object class. However separating abstract images from all others automatically is very challenging for classifiers based on visual features.
- A filter is learnt based on three simple visual only features namely: a colorhistogram, a histogram of the L2-norm of the gradient, a histogram of the angles weighted by the L2-norm of the corresponding gradient. Histogram is being used since drawings&symbolic images are characterized by sharp edges in certain orientations and a distinctive color distribution.
- Seven features from the text and HTML-tags on the web page such as contextR, context10, filedir, filename, imagealt, imagetitle, website title are used to rank the images based on textual features.
- Ranking based on visual features is performed by resizing all the images to 300 pixels , followed by detection of regions and assignment of descriptor for each region . This method outperforms Google Image search engine , however high precision images were not obtained. Polysemy is rectified completely through this algorithm.

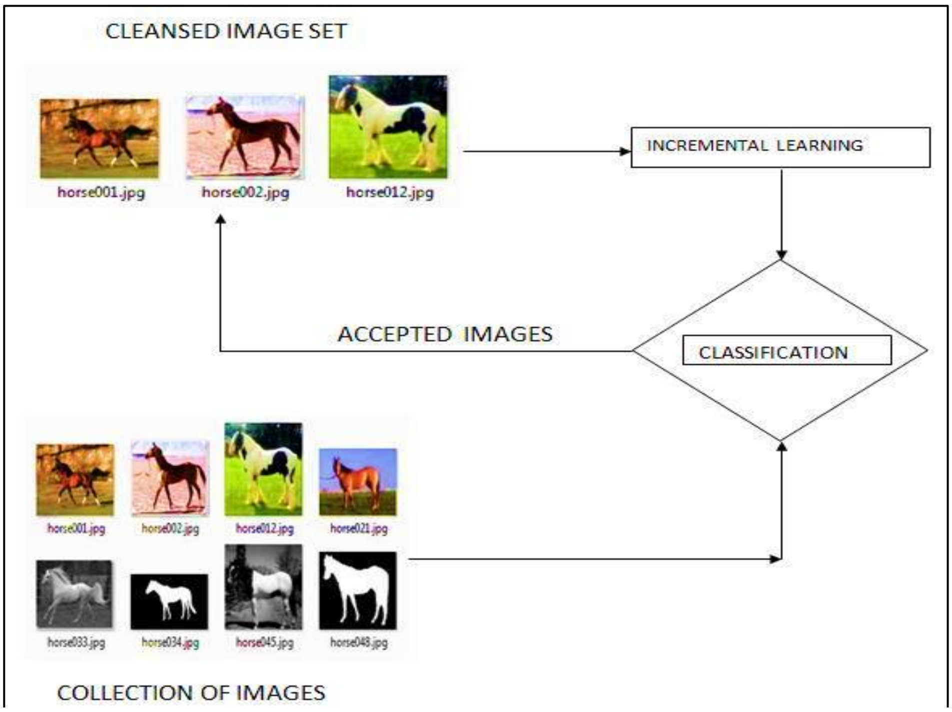

Fig4: *Automatic Cleansing*

COMPARISON WITH OTHER APPROACHES

TITLE	AUTHOR	ADVANTAGES	DRAWBACKS
ImageNet: A Large-Scale Hierarchical Image Database	Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li and Li Fei-Fei	1)Simple applications – object recognition, image classification and automatic object clustering ,object localization . 2) Hierarchical structure of WordNet.	Optimization of labels to avoid repetition
LabelMe: a database and web-based tool for image annotation	Bryan C.Russell, Antonio Torralba	1)Open and Dynamic 2) Many non-copyrighted images 3)High quality labelling. 4)Many diverse object classes.	1)Quality control. 2) Uniform distribution of objects 3)Complexity of polygons marked.
WordNet: An Electronic Lexical Database.	C.Fellbaum	1)Disambiguation of different sense of a word.	1)Necessary to frequently update associations

Fig 5: *Comparison of Interactive Approaches*

TITLE	AUTHOR	ADVANTAGES	DRAWBACKS
OPTIMOL: Automatic Online Picture Collection via Incremental Model Learning	Li-Jia Li · Li Fei-Fei	1) Capable of both learning highly effective object category models and collecting object category datasets significantly larger than that of Caltech 101 or LabelMe	1) More descriptive object models necessary.
Learning Object Categories from Google's Image Search	R. Fergus, L. Fei-Fei, P. Perona, A.Zisserman	1)the results are competitive (or close to) existing methods requiring hand gathered collections of images	1)better centroid proposals. 2)the use of fixed background densities to assist learning
Harvesting Image Databases from the Web	F. Schroff,A. Criminisi,A. Zisserman	1)Outperforms both Google Image Search and recent techniques which rely on manual intervention	1)High precision of top returned images

Fig 6: *Comparison of Automatic Approaches*

CONCLUSION

In this paper, we have explained both the interactive and automatic approaches used for image retrieval**.** These methods aim at relevant information for image retrieval and also the issues handled. Based on this survey, we prove that automatic method of cleansing the images is efficient and useful for supervised learning. In the various methods discussed above, we can witness that each algorithm has been brought up to improve some parameter which retained backward in existing methods.Polysemy, a problem with no automatic solution can be understood in different forms.

REFERENCES

1. Le Wang, Gang Hua, JianruXue, ZhanningGao and Nanning Zheng, Joint Segmentation and Recognition of Categorized Objects From Noisy Web Image Collection, IEEE Trans. Image Process., vol. 23, no. 9, pp.40702.86,Sep2014.
2. L. Mukherjee, V. Singh, J. Xu, and M. D. Collins,Analyzing the subspace structure of related images: Concurrent segmentation of image sets , in Proc. 12th ECCV, 2012, pp.128-142.
3. L.-J. Li and L. Fei-Fei,OPTIMOL: Automatic online picture collection via incremental model learning, Int. J. Comput. Vis., vol.88, no. 2,pp. 147168, 2010.
4. B. C. Russell, A. Torralba, K. P. Murphy, and W. T. Freeman,.,LabelMe: A database and web-based tool for image annotation, Int. J. Comput.Vis., vol. 77, nos. 13, pp.
5. 157173, 2008.
6. J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei,ImageNet: A

large-scale hierarchical image database, in Proc. IEEE CVPR,Jun. 2009, pp. 248255.
7. R. Fergus, L. Fei-Fei, P. Perona, and A. Zisserman, ,Learning object categories from Googles image search ,in Proc. 10th IEEE ICCV,Oct. 2005, pp. 18161823.
8. F. Schroff, A. Criminisi, and A. Zisserman,Harvesting image databases from the web ,in Proc. IEEE 11th ICCV, Oct. 2007, pp. 18.4
9. M. F. Porter,An algorithm for suffix stripping, Program, vol. 14, no. 3,pp. 130137, 1980.
10. C. Fellbaum. WordNet: An Electronic Lexical Database. Bradford Books, 1998.
11. N. Sawant, J. Z. Wang, and J. Li, “Enhancing training collections for image annotation: An instance-weighted mixture modeling approach,” IEEE Trans. Image Process., vol. 22, no. 9, pp. 3562–3577, Sep. 2013.