Scientific Journal Impact Factor (SJIF): 1.711
e-ISSN: 2349-9745
p-ISSN: 2393-8161
International Journal of Modern Trends in Engineering and Research
www.ijmter.com

CONTENT RECOVERY AND IMAGE RETRIVAL IN IMAGE DATABASE CONTENT RETRIVING IN TEXT IMAGES

K.Yuvasri[1], B.Arulkumar[2],S.Syed Farmhan[3]
[1]Computer Science and Engineering, Vivekanandha College Of Engineering For Women
[2]Computer Science and Engineering, Vivekanandha College Of Engineering For Women
[3]Computer Science and Engineering, Vivekanandha College Of Engineering For Women

Abstract - Digital Images are used in magazines, blogs, website, television and more. Digital image processing techniques are used for feature selection, pattern extraction classification and retrieval requirements. Color, texture and shape features are used in the image processing. Digital images processing also supports computer graphics and computer vision domains. Scene text recognition is performed with two schemes. They are character recognizer and binary character classifier models. A character recognizer is trained to predict the category of a character in an image patch. A binary character classifier is trained for each character class to predict the existence of this category in an image patch. Scene text recognition is performed on detected text regions. Pixel-based layout analysis method is adopted to extract text regions and segment text characters in images. Text character segmentation is carried out with color uniformity and horizontal alignment of text characters. Discriminative character descriptor is designed by combining several feature detectors and descriptors. Histogram of Oriented Gradients (HOG) is used to identify the character descriptors. Character structure is modeled at each character class by designing stroke configuration maps. The scene text extraction scheme is also supports for smart mobile devices. Text recognition methods are used with text understanding and text retrieval applications. The text recognition scheme is enhanced with content based image retrieval process. The system is integrated with additional representative and discriminative features for text structure modeling process. The system is enhanced to perform text and word level recognition using lexicon analysis. The training process is included with word database update task.

Keywords- Text recognition, character recognition, pixel based layout,character descriptor,lexicon analysis.

I. INTRODUCTION

Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. The term content in this context might refer colors, shapes, textures, or any other information that can be derived from the image itself. Without the ability to examine image content, searches must rely on metadata such as captions or keywords, which may be laborious or expensive to produce. The term CBIR seems to have originated in 1992, when it was used by T. Kato to describe experiments into automatic retrieval of images from a database, based on the colors and shapes present. Since then, the term has been used to describe the process of retrieving desired images from a large collection on the basis of syntactical image features. The techniques, tools and algorithms that are used originate from fields such as statistics, pattern recognition, signal processing and computer vision.

There is growing interest in CBIR because of the limitations inherent in metadata-based systems, as well as the large range of possible uses for efficient image retrieval. Textual information about images can be easily searched using existing technology, but requires humans to personally describe every image in the database. This is impractical for very large databases, or for images that are generated automatically, e.g. from surveillance cameras. It is also possible to miss images that use different synonyms in their descriptions. Systems based on categorizing images in semantic classes like cat as a subclass of animal avoid this problem but still face the same scaling issues.

CBIR systems can also make use of relevance feedback, where the user progressively refines the search results by marking images in the results as relevant, not relevant, or neutral to the search query, then repeating the search with the new information. Generally speaking, image content may include both visual and semantic content. Visual content can be very general or domain specific. General visual content include color, texture, shape, spatial relationship, etc. Domain specific visual content, like human faces, is application dependent and may involve domain knowledge. Semantic content is obtained either by textual annotation or by complex inference procedures based on visual content. This concentrates on general visual content descriptions.

II. LITERATURE SURVEY

2.1. Text Detection and Character Recognition in Scene Images with Unsupervised Feature Learning

Detection of text and identification of characters in scene images is a challenging visual recognition problem. As in much of computer vision, the challenges posed by the complexity of these images have been combated with handed signed features and models that incorporate various pieces of high-level prior knowledge. In this paper, we produce results from a system that attempts to learn the necessary features directly from the data as an alternative to using purpose-built, text-specific features or models. Among our results, we achieve performance among the best known on the ICDAR 2003 character recognition dataset.Feature learning algorithms have enjoyed a string of successes in other fields. To apply these algorithms to scene text applications, we will thus use a more scalable feature learning system.

Specifically, we use a variant of K-means clustering to train a bank of features. Armed with this tool, we will produce results showing the effect on recognition performance as we increase the number of learned features. Our results will show that it's possible to do quite well simply by learning many features from the data. Our approach contrasts with much prior work in scene text applications, as none of the features used here have been explicitly built for the application at hand. Indeed, the system follows closely the one proposed.

2.2. Top-down and Bottom-up Cues for Scene Text Recognition

The problem of understanding scenes semantically has been one of the challenging goals in computer vision for many decades. It has gained considerable attention over the past few years, in particular, in the context of street scenes. This problem has manifested itself in various forms, namely, object detection, object recognition and segmentation Although these approaches interpret most of the scene successfully, regions containing text tend to be ignored. One of the first things we notice in this scene is the sign board and the text it contains. Popular recognition methods ignore the text and identify other objects such as car, person, tree, regions such as road, sky. The importance of text in images is also highlighted in the experimental study conducted by Judd et al. They found that viewers fixate on text when shown images containing text and other objects. This is further evidence that text recognition forms a useful component of the scene understanding problem.Given the rapid growth of camera-based applications readily available on mobile phones, understanding scene text is more important than ever. Although character recognition forms an essential component of text understanding, extending this framework to recognize words is not trivial. It consists of “roughly front parallel” pictures of signs are quite similar to those found in a traditional OCR setting. In contrast, we show results on a more challenging street view dataset, where the words vary in appearance significantly.

2.3. Real-Time Scene Text Localization and Recognition

Text localization and recognition in real-world scene images is an open problem which has been receiving significant attention since it is a critical component in a number of computer vision applications like searching images by their textual content, reading labels on businesses in map applications or assisting visually impaired. Methods based on a sliding window limit the search to a subset of image rectangles. Methods in the second group find individual characters by grouping pixels into regions using connected component analysis assuming that pixels belonging to the same character have similar properties. Connected component methods differ in the properties used. The advantage of the connected component methods is that their complexity typically does not depend on the properties of the text and that they also provide a segmentation which can be exploited in the OCR step.. The real-time performance is achieved by posing the character detection problem as an efficient sequential selection from the set of Extremal Regions (ERs). In the first stage of the classification, the probability of each ER being a character is estimated using novel features calculated with O(1) complexity and only ERs with locally maximal probability are selected for the second stage, where the classification is improved using more computationally expensive features.

A highly efficient exhaustive search with feedback loops is then applied to group ERs into words and select the most probable character segmentation. It is further demonstrated that by inclusion of the gradient projection 94.8% of characters are detected by the ER detector.

2.4 Detecting Texts of Arbitrary Orientations in Natural Images

The great successes of smart phones and large demands in content-based image search/ understanding have made text detection a crucial task in human computer interaction. Although text detection has been studied extensively in the past, the problem remains unsolved. The difficulties mainly come from two aspects: (1) the diversity of the texts and (2) the complexity of the backgrounds.

On one hand, text is a high level concept but better defined than the generic objects; on the other hand, repeated patterns and random clutters may be similar to texts and thus lead to potential false positives.Hence, combining the strengths of specially designed features and discriminatively trained classifiers, our system is able to effectively detect texts of arbitrary orientations but produce fewer false positives.To evaluate the effectiveness of our system, we have conducted extensive experiments on both conventional and new image datasets. Compared with the state-of-the-art text detection algorithms, our system performs competitively in the conventional setting of horizontal texts.To evaluate the effectiveness of our system, we have conducted extensive experiments on both conventional and new image datasets. Compared with the state-of-the-art text detection algorithms, our system performs competitively in the conventional setting of horizontal texts. We have also tested our system on a very challenging large dataset of 500 natural images containing texts of various orientations in complex backgrounds. On this dataset, our system works significantly better than any of the existing systems, with an F-measure about 0.6, more than twice that of the closest competitor

2.5.Scene Text Recognition using Part-based Tree-structured Character Detection

With the rapid growth of camera-based applications readily available on smart phones and portable devices, understanding the pictures taken by these devices semantically has gained increasing attention from the computer vision community in recent years. Most of the previous work on scene text recognition could be roughly classified into two categories: traditional Optical Character Recognition (OCR) based and object recognition based. For traditional OCR based methods, various binarization methods have been proposed to get the binary image which is directly fed into the off-the-shelf OCR engine. Moreover, the loss of information during the binarization process is almost unrecoverable, which means if the binarization result is poor, the chance of correctly recognizing the text is quite small.

For scene character recognition, these methods directly extract features from original image and use various classifiers to recognize the character. While for scene text recognition, since there are no binarization and segmentation stages, most existing methods adopt multi-scale sliding window strategy to get the candidate character detection results. To recognize the scene text, we build the CRF model on the potential character locations. Character detection scores, spatial constraints and linguistic knowledge are used to define the unary and pairwise cost function. The final word recognition result is acquired by minimizing the cost function.

III. PROBLEM DESCRIPTION

Camera-Based text information serves as effective tags or clues for many mobile applications associated with media analysis, content retrieval, scene understanding and assistant navigation. In natural scene images and videos, text characters and strings usually appear in nearby sign boards and hand-held objects and provide significant knowledge of surrounding environment and objects. Text-based tags are much more applicable than barcode or quick response code because the latter techniques contain limited information and require pre-installed marks.

To extract text information by mobile devices from natural scene, automatic and efficient scene text detection and recognition algorithms are essential. Extracting scene text is a challenging task due to two main factors: 1) cluttered backgrounds with noise and non-text outliers and 2) diverse text patterns such as character types, fonts and sizes. The frequency of occurrence of text in natural scene is very low and a limited number of text characters are embedded into complex non-text background outliers. Background textures, such as grid, window and brick, even resemble text characters and strings. Although these challenging factors exist in face and car, many state-of-the-art algorithms have demonstrated effectiveness on those applications, because face and car, have relatively stable features. For example, a frontal face normally contains a mouth, a nose, two eyes and two brows as prior knowledge. It is difficult to model the structure of text characters in scene images due to the lack of discriminative pixel-level appearance and structure features from non-text background outliers. Further, text consists of different words where each word may contain different characters in various fonts, styles and sizes, resulting in large intra-variations of text patterns. To solve these challenging problems, scene text extraction is divided into two processes: text detection and text recognition. Text detection is to localize image regions containing text characters and strings. It aims to remove most non-text background outliers. Text recognition is to transform pixel-based text into readable code. It aims to accurately distinguish different text characters and properly compose text words. This paper will focus on text recognition method. It involves 62 identity categories of text characters, including 10 digits [0-9] and 26 English letters in upper case [A-Z] and lower case [a-z].

We propose effective algorithms of text recognition from detected text regions in scene image. In scene text detection process, we apply the methods presented in our previous work. Pixel-based layout analysis is adopted to extract text regions and segment text characters in images, based on color uniformity and horizontal alignment of text characters. In text recognition process, we design two schemes of scene text recognition. The first one is training a character recognizer to predict the category of a character in an image patch. The second one is training a binary character classifier for each character class to predict the existence of this category in an image patch. The two schemes are compatible with two promising applications related to scene text, which are text understanding and text retrieval. Text understanding is to acquire text information from natural scene to understand surrounding environment and objects.

Text retrieval is to verify whether a piece of text information exists in natural scene. These two applications can be widely used in smart mobile device.The main contributions of this paper are associated with the proposed two recognition schemes. Firstly, a character descriptor is proposed to extract representative and discriminative features from character patches. It combines several feature detectors and Histogram of Oriented Gradients (HOG) descriptors. Secondly, to generate a binary classifier for each character class in text retrieval, we propose a novel stroke configuration from character boundary and skeleton to model character structure.

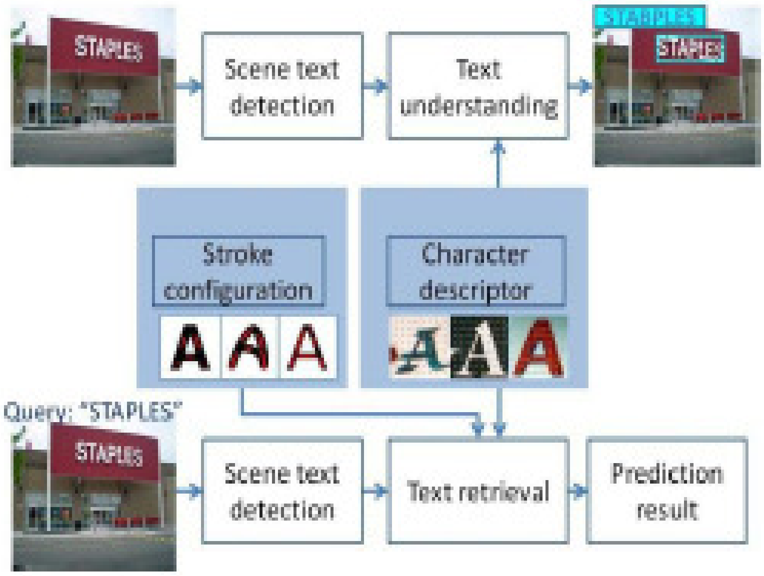

Fig. 1.1. The Flowchart Of Our Designed Scene Text Extraction Method

The proposed method combines scene text detection and scene text recognition algorithms. Similar to other methods, our proposed feature representation is based on the stateof- the-art low-level feature descriptors and coding/pooling schemes. Different from other methods, our method combines the low-level feature descriptors with stroke configuration to model text character structure. Also, we present the respective concepts of text understanding and text retrieval and evaluate our proposed character feature representation based on the two schemes in our experiments. Besides, previous work rarely presents the mobile implementation of scene text extraction and we transplant our method into an Android-based platform.

IV. PROPOSED SYSTEM

Scene text recognition process is performed to identify the text or string in a natural scene image. Text region selection, Character descriptor and character structure analysis methods are used for text recognition process. The system is enhanced to support text and word level recognition process. Content Based Image Retrieval (CBIR) scheme is integrated with the system.

V. CONCLUSION

Scene text recognition process is performed to identify the text or string in a natural scene image. Text region selection, Character descriptor and character structure analysis methods are used for text recognition process. The system is enhanced to support text and word level recognition process. Content Based Image Retrieval (CBIR) scheme is integrated with the system. The system improves the accuracy levels in the text recognition process. Content based image search is supported by the system. Text and word level recognition scheme is used for the scene understanding purpose. Text structure modeling is upgraded to improve the classification process.

REFERENCES

[1]A.Coates et al., "Text Detection And Character Recognition In Scene Images With Unsupervised Feature Learning," in Proc. ICDAR, Sep. 2011, pp. 440–445.
[2]A.Mishra, K. Alahari and C. V. Jawahar, "Top-Down And Bottom-Up Cues For Scene Text Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 1063–6919.
[3]A.Shahab, F. Shafait and A. Dengel, "ICDAR 2011 Robust Reading Competition Challenge 2: Reading Text In Scene Images," in Proc. Int. Conf. Document Anal. Recognit., Sep. 2011, pp. 1491–1496.
[4]C. Shi, C. Wang, B. Xiao and Z. Zhang, "Scene Text Recognition Using Part-Based Tree-Structured Character Detection," in Proc. CVPR, Jun. 2013.

[5]C. Yi and Y. Tian, "Localizing Text In Scene Images By Boundary Clustering, Stroke Segmentation, And String Fragment Classification," IEEE Trans. Image Process., vol. 21, no. 9, pp. 4256–4268, Sep. 2012.
[6]C. Yi, X. Yang and Y. Tian, "Feature Representations For Scene Text Character Recognition: A Comparative Study," in Proc. 12th ICDAR, 2013.
[7]C.Yao, X. Bai, W. Liu, Y. Ma and Z. Tu, "Detecting Texts Of Arbitrary Orientations In Natural Images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 1083–1090.
[8]Chucai Yi and Yingli Tian, "Scene Text Recognition in Mobile Applications by Character Descriptor and Structure Configuration" IEEE Transactions On Image Processing, Vol. 23, No. 7, July 2014
[9]D. L. Smith, J. Feild and E. Learned-Miller, "Enforcing Similarity Constraints With Integer Programming For Better Scene Text Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 73–80.
[10]K. Wang, B. Bbenko and S. Belongie, "End-to-End Scene Text Recognition," in Proc. Int. Conf. Comput. Vis., Nov. 2011, pp. 1457–1464.
[11]L. Neumann and J. Matas, "Real-Time Scene Text Localization And Recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2012, pp. 3538–3545.

