Principal Visual Word Discovery for Automatic License Plate Detection

Wengang Zhou, Houqiang Li, Yijuan Lu, *Member, IEEE*, and Qi Tian, *Senior Member, IEEE*

***Abstract*— License plates detection is widely considered a solved problem, with many systems already in operation. However, the existing algorithms or systems work well only under some controlled conditions. There are still many challenges for license plate detection in an open environment, such as various observation angles, background clutter, scale changes, multiple plates, uneven illumination, and so on. In this paper, we propose a novel scheme to automatically locate license plates by principal visual word (PVW), discovery and local feature matching. Observing that characters in different license plates are duplicates of each other, we bring in the idea of using the bag-of-words (BoW) model popularly applied in partial-duplicate image search. Unlike the classic BoW model, for each plate character, we automatically discover the PVW characterized with geometric context. Given a new image, the license plates are extracted by matching local features with PVW. Besides license plate detection, our approach can also be extended to the detection of logos and trademarks. Due to the invariance virtue of scale-invariant feature transform feature, our method can adaptively deal with various changes in the license plates, such as rotation, scaling, illumination, etc. Promising results of the proposed approach are demonstrated with an experimental study in license plate detection.**

***Index Terms*— Clustering, geometric context, object detection, principal visual word (PVW).**

I. Introduction

LICENSE plate detection plays a key role in intelligent transportation systems. It can be applied in vehicle management, such as security control, traffic monitoring, automatic vehicle ticketing, and so on. Recently, this topic has attracted more attention in privacy protection as a lot of images and videos containing private information such as license plates are shared on the Internet. For instance, in Google Street View, one important task is to blur license plates in order to protect privacy.

License plates detection is widely considered a solved problem, with many systems already in operation. Nevertheless, the existing algorithms or systems work well only under some controlled conditions. For instance, some systems require sophisticated video capture hardware, possibly combined with infrared strobe lights (PVW), or require that the images be taken with little distortion from view-point changes. Although many reported results are very good, with even perfect accuracy on their test datasets, it is still a challenging task to detect license plates in open environment.

Generally, a license plate detection system has to solve two problems: where a license plate is located and how big it is. Usually, the candidate position of characters in the license plate is first identified, and the bounding box of the license plate is determined later. There are many challenges in license plate detection in an open environment, such as various observation angles from cameras, background clutter, differently sized license plates, poor image quality from uneven lighting conditions, and multi-plate detection. Typical instances can be found in Figs. 9 and 10.

Recently, the bag-of-words (BoW) model [1] based on local invariant feature [2], [3] has attracted a lot of attention in the computer vision and multimedia community. Due to the merits of scale-invariant feature transform (SIFT), i.e, the invariance property in rotation, scale, and illumination, it has been widely applied in object recognition [2], video tracking [4], content-based image retrieval [1], [5], and especially partial-duplicate image search [6], [7].

Observing that a specific character in different license plates can be considered as duplicates of each other, we bring in the idea of BoW model based on local feature for license plate detection. Since visual words generated from unsupervised clustering [5] are sensitive to noisy features from image background, it is desirable to yield the principal (discriminative and descriptive) visual words that correspond to each unique character in the license plate. Besides, these principal visual words (PVW) are expected to contain geometric context, which can be used to deduce the size of the corresponding character.

Motivated by the above discussion, we formulate license plate detection as a visual matching problem. For each character, we collect SIFT features falling into the character region and generate PVW by unsupervised clustering. The amount of PVW for each plate character is determined automatically.

Manuscript received August 26, 2011; revised April 20, 2012; accepted April 30, 2012. Date of publication May 15, 2012; date of current version August 22, 2012. The work of H. Li was supported in part by the Fundamental Research Funds for the Central Universities of China under Grant WK2100230003. The work of Y. Lu was supported in part by the Research Enhancement Program, Start-up Funding from Texas State University, and the DoD HBCU/MI under Grant W911NF-12-1-0057. The work of Q. Tian was supported in part by the National Science Foundation under Grant IIS 1052851, a Faculty Research Award from Google, FXPAL, the NEC Laboratories of America, and the Army Research Office under Grant W911NF-12-1-0057. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Patrick Flynn.

W. Zhou and H. Li are with the Electrical Engineering and Information Science Department, University of Science and Technology of China, Hefei 230027, China (e-mail: zhwg@mail.ustc.edu.cn; lihq@ ustc.edu.cn).

Y. Lu is with the Computer Science Department, Texas State University, San Marcos, TX 78666 USA (e-mail: yl12@txstate.edu).

Q. Tian is with the Computer Science Department, University of Texas, San Antonio, TX 78249 USA (e-mail: qitian@cs.utsa.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIP.2012.2199506