Detecting, Tracking and Recognizing License Plates

Michael Donoser, Clemens Arth, and Horst Bischof

Institute for Computer Graphics and Vision
Graz University of Technology
{donoser,arth,bischof}@icg.tu-graz.ac.at

ostract. This paper introduces a novel real-time framework which ables detection, tracking and recognition of license plates from video quences. An efficient algorithm based on analysis of Maximally Stable tremal Region (MSER) detection results allows localization of intersional license plates in single images without the need of any learn- scheme. After a one-time detection of a plate it is robustly tracked ough the sequence by applying a modified version of the MSER track- framework which provides accurate localization results and addition- y segmentations of the individual characters. Therefore, tracking and aracter segmentation is handled simultaneously. Finally, support vec- machines are used to recognize the characters on the plate. An experi- ntal evaluation shows the high accuracy and efficiency of the detection d tracking algorithm. Furthermore, promising results on a challenging a set are presented and the significant improvement of the recognition e due to the robust tracking scheme is proved.

roduction

a need for intelligent traffic management systems in order to cope with antly increasing traffic on todays roads. Video based traffic surveillance the key parts of such installations. Beside detection and tracking of identification by license plate recognition is important for a variety of ons like access-control, security or traffic monitoring.

eral, license plate recognition systems consist of two separate parts. tes are detected within a single frame of a traffic video sequence and haracter recognition is applied to identify the characters on the plate.

nt methods have been proposed for the detection of license plates. et al. [1] use a mixture of edge detection and vertical pixel projection detection module. In the work of Jia et al. [2] color images were seg- y the Mean Shift algorithm into candidate regions and subsequently as plate or not. The AdaBoost algorithm was used by Dlagnekov and [3] for license plate detection on rather low resolution video frames. d Zimmermann [4] proposed a different approach for the localization

everal approaches for recognizing the characters on the plate after suc-
etection were proposed. Shapiro et al. [1] use adaptive iterative thresh-
d analysis of connected components for segmentation. The classification
en performed with two sets of templates. Rahman et al. [5] used hori-
d vertical intensity projection for segmentation and template matching
ication. Dlagnekov and Belongie [3] use the normalized cross correlation
fication by analyzing the whole plate, hence skipping segmentation.
gh much scientific work has focused on recognizing license plates from
leo sequences, surprisingly little work has been done on integrating a
scheme to gather additional representations of the plate for improving
gnition rate. Furthermore, in the few systems that apply a tracking
nly simple and unstable approaches are used, as e. g. by Dlagnekov and
[3] who perform tracking by simply repeating detection for building
ndences.
ain contribution of this paper is a novel framework which unifies detec-
king and recognition of license plates in a robust and efficient way. The
g idea is to base detection, tracking and character segmentation on the
nciples which allows to provide segmentations of the individual charac-
ecognition in addition to accurate and robust license plate localization
subsequent frames. The framework is presented in detail in Section 2
xperimental evaluation is shown in Section 3.

ense Plate Recognition Framework

ion describes the entire framework for detection, tracking and recogni-
ense plates from traffic video sequences. The introduced system detects
pearing license plates from the sequence by a novel algorithm which is
the analysis of Maximally Stable Extremal Region (MSER) [6] detec-
ts. The concept, introduced in Section 2.1, does not require any learning
nd is capable of detecting different types of international plates. After
of a plate it is robustly tracked through the sequence by a modified
f the MSER tracking framework [7] as shown in Section 2.2. Therefore,
appearing car in the video sequence a set of license plate representa-
ollected which is used to improve the subsequent character recognition.
ection 2.3 describes how support vector machines are used to recognize
cters on the different representations of each plate and how the results
ed by a voting scheme to achieve the final recognition result.

ense Plate Detection

step of every license plate recognition system is the detection of the
ates within a single frame of the traffic video sequence. We propose

ne does not require any learning scheme and is able to detect different nternational license plates without adaption.

etection algorithm is based on analyzing the results of a Maximally xtremal Region (MSER) detection [8]. MSERs denote a set of distin-egions and have proven to be one of the best interest point detectors ter vision [9]. All of these regions are defined by an extremal prop-ne intensity function in the region and on its outer boundary. Special s form their superior performance as stable local detector. The set of s closed under continuous geometric transformations and is invariant ntensity changes. Furthermore, MSERs are detected in every scale. We antly exploit these properties for segmentation purposes.

eral, two variants of MSER detection can be distinguished denoted as nd MSER-. While MSER+ detects bright regions with darker bound-R- finds dark regions with brighter boundary. Figure 1(a) shows an m a traffic video sequence and Figure 1(b) and Figure 1(c) illustrate sponding MSER detection results as binary images. As can be seen, the ate itself is identified as MSER+, whereas the characters on the plate ted as MSER-.

a) Input frame (b) MSER+ (c) MSER-

SER detection results can be used for detecting license plates in video se-MSER+ finds the license plate, whereas MSER- identifies the individual char-it.

nderlying idea of our novel license plate detection scheme is to analyze R+ and MSER- detection results. We are looking for a larger MSER+ cense plate) that contains a set of smaller MSER- regions (characters). ombination is considered as license plate detection result. Furthermore, the detection by checking if the MSER- regions are approximately ed, if their center points approximately lie on a line and if the height of R+ is in the range of the average MSER- height. After verification, the s returned as license plate localization result and additionally segmen-f the characters are provided by the corresponding MSER- detections. gh the detection process is simple, it is effective and allows stable and

a) Proposed algorithm (b) AdaBoost result

cense plate detection results. Figure (a) shows the single detection of the algorithm indicated by the white border. 914 MSERs are detected, but only f regions fulfills the described criterion. Figure (b) shows the wrong and etections of an AdaBoost algorithm.

mparison, Figure 2(b) shows the result of an AdaBoost detector based ike features [10]. As can be seen the boosting framework returns wrong iple detections which need to be significantly post-processed, as e. g. by maximum suppression to remove multiple detections. Our result is also urate as the bounding box provided by the boosting variant. Further-ining of an AdaBoost based classifier is a rather complex procedure, is case was especially trained on Austrian license plates. Our approach require any learning scheme and is able to identify different types of in-al plates, because the simple criterion is fulfilled for almost all of them. shows detection results for different international types.

a) Input image (b) Characters (MSER-) (c) Plates (MSER+)

etection results on international license plates where (b) shows the segmented cted as MSER- and (c) the plate detected as MSER+. As can be seen the

acking of License Plates

ection of a newly appearing car in the traffic video sequence by its ate, a robust tracker is applied to increase the number of character ations for the subsequent recognition step. In general, any tracker can but we propose to apply a modified version of the MSER tracking k introduced by Donoser and Bischof [7] which has some advantages st to other tracking schemes. First, it is efficient and stable and can be o our specific requirements. Second, it provides an accurate segmenta- e license plate and third, in addition to the tracked plate it also returns tions of the individual characters on the plate (MSERs), thus tracking entation are handled simultaneously.

ISER tracking framework was designed to improve the stability and MSER detection results in video sequences. The tracker has to be initial- assing the region to be tracked R_t detected in image I_t of the sequence mework. The first step of the algorithm is to propagate the center point gion R_t to the next image I_{t+1} and to crop a region-of-interest (ROI) from the image. Then a data-structure named component tree [11] is this ROI. Every node of the component tree contains one candidate $_{+1}$ for the tracking and the algorithms looks for the node which is most the region R_t. The best fit is identified by comparing feature vectors built for each of the nodes of the component tree C^i_{t+1} and the input . The candidate C^i_{t+1} with the lowest Euclidean distance between its ector and the one of the region R_t is taken as tracked representation. ures calculated are mean gray value, region size, center of mass, width t of the bounding box and stability. All of these features are computed tally [12] during creation of the component tree. Thus, no additional sion time is required. After detection of the new representation the de- eps can be repeated for tracking the region through the entire sequence.

iginal MSER tracking framework was designed for tracking arbitrarily egion, but we adapt the method to our specific requirements of track- e plates. In our framework, the MSER tracking algorithm is initialized sult of the license plate detection algorithm presented in Section 2.1. we focus on license plates we reformulate the feature comparison ap- replacing the Euclidean distance computation by a simpler, but more equation based on comparison of two distinct features, the size and the arity of the region. Thus, in our framework the tracked representation oy calculating a distance value $\Delta(R_t, C^i_{t+1})$ for every candidate region

$$\Delta(R_t, C^i_{t+1}) = \frac{abs\left(|R_t| - \left|C^i_{t+1}\right|\right)}{|R_t|} + \left(1 - \vartheta\left(C^i_{t+1}\right)\right), \quad (1)$$

.| denotes the size of the region and $\vartheta(\dots)$ is the rectangularity. Then,

(a) Frame 1 (b) Frame 15

ustration of license plate tracking. The images show a traffic scene and the
l characters on the license plate are highlighted in white.

d through the sequence and the detected MSER− regions within the plate are provided as segmentations of the individual characters. Fur-, the tracking scheme is also used for discarding false positives of the step by removing non-moving or unstable plate tracks.

4 shows two frames of a traffic video sequence, where accurate license ections are provided in addition to the segmentation of the eight char- the plate.

aracter Recognition

step of our framework is to recognize the individual characters on ted plates based on support vector machines (SVMs). SVMs were first by Vapnik [13] and have proven to be an efficient tool for classification n optical character recognition (OCR) of handwritten digits or license tents [14]. For an introduction to SVMs and other kernel methods see ple [15].

eral, SVMs are designed for binary classification problems. Because recognition is a multi-class problem we apply a method based on a ion of several binary SVMs. The strategy is called *one-vs-one* where pair of output classes an individual SVM is trained, resulting in a total f $n \cdot (n-1)/2$ classifiers. Then all classifiers are evaluated, the votes ned up and the class with the maximum number is chosen.

he provided character segmentations are aligned, and then the *one-vs-* oach is used to classify each character independently of all the others. ented tracking approach provides several license plate representations car and therefore, we also have several classification results for every on the plate. A majority voting scheme is then used to determine the acter recognition result for every car.

5 shows a sequence of tracked license plates, the segmented characters corresponding classification results. As can be seen, the single image cognition provides wrong assignments, but the subsequent majority vot-

(a) Tracking Sequence (b) Classification Results

ecognition is improved by using several license plate representations provided acking scheme. The final result based on a majority voting for this plate s `G~19VAB` which matches with the real plate number.

ows to combine a sequence of single image based recognition results inal classification.

amework

ework is able to analyze traffic video sequences in real-time. It detects pearing cars by localization of their license plate. After a one-time , the plate is robustly tracked and several representations of the license collected. Until tracking fails the available repeated segmentations of racter on the plate are used to improve the recognition rate along the tracking sequence and to determine a final result by the majority voting The running times of the individual steps of the concept for analysis of equence of size 352×288 are shown in Table 1.

Running time per image of the individual steps of the framework for analyzing ream of size 352×288

	Detection	Tracking	Recognition
ime	$70ms$	$5ms$	$6ms$

perimental Evaluation

ated our framework on a challenging traffic video sequence in the type 5(a) which was filmed from a footbridge. The provided resolution was and therefore the characters on the plate only have an average size of

on the challenging data set and shows how the recognition rate is signif-
nproved by using several plate representations provided by the tracking

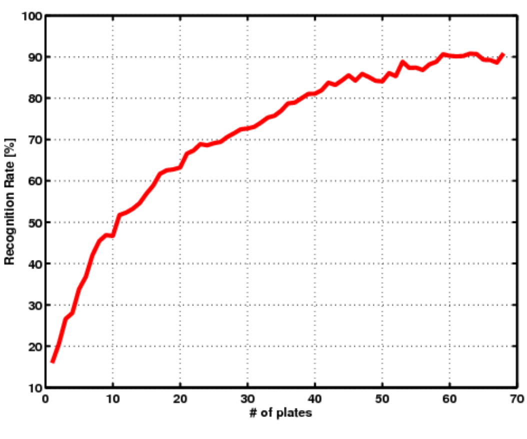
aracter Image Database

port vector machines are trained on approximately 2700 manually la-
ges of characters which were automatically extracted from high resolu-
es of parked cars using the license plate detection algorithm. According
solution of the test video sequence we resized all character images to
els.

cognition Results

ate the quality of the proposed framework we used it for recognizing
e plates of 109 cars passing the video sequence area. Due to the low
n and changing lighting conditions the average recognition rate for in-
t classification of every character in every detected license plate is only
As a consequence, in a single image based license plate recognition ap-
n average only 23.23% of the cars are totally correct classified. But
is significantly improved by analyzing additional plate representations
by the tracking scheme and combining the corresponding recognition
the presented majority voting scheme. Figure 6 analyzes the increase
ognition rate by using more representations. As the number of tracked
resentations gets close to 70, the percentage of totally correct classifi-
eaches more than 90%.

ng all available representations for recognizing the plate of every car
tes) our framework classifies 94.65% of all cars totally correct which is a



Recognition rates of totally correct classified plates (plate level) and correctly characters (character level) for a single image based approach in comparison cking based classification

	Single image approach	**Tracking approach**
level	80.74%	**97.16%**
l	23.23%	**94.65%**

g result for such a challenging data set. Please note, that no postprocess-checking the validity of license plates according to syntax restrictions, Table 2 analyzes the recognition rate on plate level, i. e. the number of rrect classified plates, and character level, i. e. the number of correctly characters, for the single image based approach in comparison to our variant. As can be seen analyzing several representations instead of a e significantly improves the recognition rate.

nclusion

er introduced a novel framework which allows detection, tracking and on of license plates. Detection is handled by analysis of Maximally Sta-mal Region (MSER) detection results and does not require any learning We introduced a robust tracking scheme, which provides accurate license alizations and segmented characters simultaneously. The experimental n showed that promising results are achieved on a challenging data set the robust tracking approach significantly improves the recognition thermore, due to the high efficiency of the individual components, the k can be used for real-time traffic video sequence analysis. To make the framework applicable in industrial scenarios, we also ported it to an ed embedded platform. Although the experiments were all performed on computer, the results also hold for a fully embedded implementation.

nces

ro, V., Gluhchev, G., Dimov, D.: Towards a multinational car license plate nition system 17(3), 173–183 (2006)

V., Zhang, H., He, X., Piccardi, M.: Mean shift for accurate license plate local-n. In: ITSC. Proceedings of the IEEE Conference on Intelligent Transporta-ystems, pp. 566–571. IEEE Computer Society Press, Los Alamitos (2005)

ekov, L., Belongie, S.: Recognizing cars. Technical Report CS2005-0833, University of California, San Diego (2005)

, J., Zimmermann, K.: Unconstrained licence plate and text localization and

an, C., Badawy, W., Radmanesh, A.: A real time vehicle's license plate recog- system. In: AVSS. Proceedings of the IEEE Conference on Advanced Video ignal Based Surveillance, pp. 163–166. IEEE Computer Society Press, Los tos (2003)
, J., Zimmermann, K.: Unconstrained licence plate detection. In: ITSC. Pro- gs of International Conference on Intelligent Transportation Systems, pp. 77 (2005)
er, M., Bischof, H.: Efficient maximally stable extremal region (MSER) track- : CVPR. Proceedings of Conference on Computer Vision and Pattern Recog- , pp. 553–560 (2006)
, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from nally stable extremal regions. In: BMVC. Proceedings of British Machine Conference, pp. 384–393 (2002)
ajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE actions on Pattern Analysis and Machine Intelligence (PAMI) 27(10), 1615– (2005)
, H., Jia, W., He, X., Wu, Q.: Learning-based license plate detection using and local features. In: ICPR. Proceedings of International Conference on rn Recognition, pp. 1102–1105 (2006)
an, L., Couprie, M.: Quasi-linear algorithm for the component tree. In: SPIE Geometry XII, vol. 5300, pp. 98–107 (2004)
, J., Zimmermann, K.: A new class of learnable detectors for categorisation. CIA. Proceedings of Scandinavian Conference of Image Analysis, pp. 541–550
k, V.N.: The nature of statistical learning theory. Springer-Verlag New York, New York, NY, USA (1995)
, L., He, X.: Number plate recognition based on support vector machines. VSS 2006. Proceedings of the IEEE International Conference on Video and Based Surveillance, Washington, DC, USA, p. 13. IEEE Computer Society Los Alamitos (2006)
kopf, B., Smola, A.J.: Learning with Kernels: Support Vector Machines, Reg- ation, Optimization, and Beyond. MIT Press, Cambridge, MA, USA (2001)