Unifying Approach for Fast License Plate Localization and Super-Resolution

Chu Duc Nguyen, Mohsen Ardabilian, Liming Chen
Universite de Lyon, CNRS
Ecole Centrale de Lyon, LIRIS, UMR5205, F-69134, France
{chu-duc.nguyen, mohsen.ardabilian, liming.chen}@ec-lyon.fr

Abstract—This paper addresses the localization and super-resolution of license plate in a unifying approach.

Higher quality license plate can be obtained using super-resolution on successive lower resolution plate images. All existing methods assume that plate zones are correctly extracted from every frame. However, the accurate localization needs a sufficient quality of the image, which is not always true in real video. Super-resolution on all pixels is a possible but much time consuming alternative.

We propose a framework which interlaces successfully these two modules. First, coarse candidates are found by an weak but fast license plate detection based on edge map sub-sampling. Then, an improved fast MAP-based super-resolution, using local phase accurate registration and edge preserving prior, is applied on these regions of interest. Finally, our robust ICHT-based localizer rejects false-alarms and localizes the high resolution license plate more accurately.

Experiments which were conducted on synthetic and real data, proved the robustness of our approach with real-time possibility.

Keywords-license plate; MAP super-resolution; interlaced localization and super-resolution;

I. INTRODUCTION

License Plate Recognition (LPR) is an active topic in automatic video-surveillance, such as systems designed for toll gates, stolen car detection, parking lots. Classically, a LPR system consists of three main modules: license plate localization, character segmentation, and character recognition. Among them, localization is considered as the most crucial stage, since a high accurate and fast recognition can be performed only if license plates had been correctly localized, and commercial systems for recognition of printed text are available [1]. Recently, super-resolution (SR) technology towards computer vision [2] is employed to enhance recognition rate in LPR [3]. SR reconstruction aims to estimate a high-resolution (HR) image from multiple low-resolution (LR) images. Given a set of shifted observations of one license plate, higher quality image of this plate could be generated using SR which facilitates the recognition procedure [4].

First approach which applied SR in LPR to identify moving vehicles was proposed by Suresh [3]. The author presented a robust Maximum a posteriori (MAP) based method with discontinuity adaptive Markov random field prior for enhancing edges in reconstruction process. Later, a fast MAP-based SR algorithm was proposed by Yuan [4]. A new cost function reduces the computation complexity that allows SR reconstruction can be treated as a deconvolution by using Wiener filter. But the fact that priors are removed makes the method unstable, due to registration errors. Most recently, Gambotto [5] proposed a SR based on non-uniform interpolation of registered LR images which was fast but even more sensible to motion estimates.

All existing SR approaches for LPR need the plate correctly cropped from successive frames, which is not a trivial task due to uncontrolled imaging conditions. License plate localization approaches diversify from rule based heuristic methods to training based classifiers which differ about used features [6] [7]. One main drawback of all published methods is that they generally make strong assumptions about license plate’s norm and context variability such as plate size, position, inclination and illumination condition. Moreover, the fact that only low features and their statistic have been exploited leads to high false-alarm rate. In [8], we proposed a robust and fast method which localizes license plate accurately with low false-alarm rate.

According to state-of-the-art on license plate localization and SR for LPR, an “egg and chicken” problem may occur if these techniques are used in real applications. SR assumes that license plates are localized and extracted correctly from every frame. On another hand, the exact localization requires a sufficient quality of the image, which is not always true due to motion blur, noise or compression artifacts. SR on all pixels is a possible but much time consuming alternative which makes real-time processing unlikely. While localization and SR are still treated separately, the problem persists.

In this paper we propose a unifying approach that efficiently interlaces both localization and SR in a complete framework (Fig. 1). Firstly, given a video frame, all vertical edge zones in this frame, e.g. license plate regions, are fast segmented from a sub-sampled edge map. Then these regions of interest (ROI) are tracked in sequences of the last and next frames. Secondly, these tracked ROIs are registered accurately under an affine model using local phase information. Thirdly, registered ROIs are used as the input of a fast MAP-based SR algorithm with edges preserving prior. Finally, we use our ICHT-based localizer [8] to reject false-alarms and localize the license plate in high quality ROIs more accurately.
The remainder of the paper is organized as follows. In Section II, we introduce the description of our approach and explain analytically how it is robust and fast. In Section III, we present experiments and results which prove the effectiveness of the proposed framework. Finally, Section IV gives our conclusions.

II. PROPOSED APPROACH

A. Fast License Plate ROIs Localization

The idea of a coarse but fast candidates detection comes from observations on text in license plate: characters have distinctive intensities from background; and characters always respect a base-line. Therefore license plate text forms a dense region of quasi-vertical strokes, regardless of viewpoint change effect.

Vertical edge map is generated by applying one vertical Sobel operator on a given video frame. The more resolution of this map decreases, the more vertical edges become shorter and come close. So we reduce the map resolution, a dense set of quasi-vertical strokes could be observed as one homogeneous quasi-horizontal region. All regions having that proprieties are detected quickly using our improved Hough connective transform, which proved robust to noise and quantification artifact [8]. Projection of these regions on the original frame gives some coarse rectangular ROIs. Therefore, the number of pixels passed to next steps is much reduced; while keeping true license plate among ROIs. Then, a fast 2D tracking of these ROIs in sequences of the last and next frames is yielded [9].

B. Accurate Local Phase Registration

SR algorithms use the information available from shifted LR observations to reconstruct an HR image. Accurate registration of these observations is important to the success of these algorithms [3]. License plate is a rigid body moving in a stable manner, only global transition, rotation and scale caused by perspective projection needs to be considered [4]. It means for LPR the affine model is sufficient under which two views of an object are related by:

\[
\begin{align*}
  x' &= xp_1 + yp_2 + p_3 \\
  y' &= xp_4 + yp_5 + p_6
\end{align*}
\]  

(1)

Given the model, sub-pixels accurate registration can be achieved with the exact knowledge of degradation parameters such as blur and non-uniform illumination. However, in practice, this information is rarely available. We overcome this problem using local phase to estimate registration parameters. Local phase is robust towards noise and smoothly varying illumination [10].

1) Local Phase: can be computed using Gabor filters, which are popular band pass filters as they achieve the theoretical minimum product of spatial width and bandwidth, desirable for better localization and accurate phase computation, respectively:

\[
g(x, y) = \frac{1}{2\pi \sigma_x \sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2} + \frac{y^2}{2\sigma_y^2}\right)} e^{j(\omega_x x + \omega_y y)},
\]

(2)

where \((\omega_x, \omega_y)\) is the angular frequency, \(\sigma_x\) and \(\sigma_y\) control the spatial width of the filter. At each frequency \((\omega_x, \omega_y)\), the image is convolved with \(g(x, y)\). The argument of the complex output is local phase.

2) Phase Difference: is computed by taking the difference of phase values at each location of the image pair at the given angular frequency. 2D translation components can be computed from the basis of Fourier Shift theorem, according to which, a shift of \((\Delta x, \Delta y)\) in the spatial domain would produce a phase difference of \(2\pi (\Delta x \omega_x + \Delta y \omega_y)\) at \((\omega_x, \omega_y)\). Locations \((x, y)\) and \((x', y')\) in (1) are translated by these shift values. Errors in phase difference computation may occur due to noise and the absence of the local frequencies with which the images are convolved. The degree of match in the amplitude values is used as a measure of confidence [10]. The confidence is high if the amplitudes of the Gabor filter response at \((x, y)\) in both the images are close.

C. Improved MAP-based SR Reconstruction

The imaging process in SR reconstruction can be formulated under matrix form as:

\[
y_r = DH_r W_r x + n_r, \quad 1 \leq r \leq m,
\]

(3)

where \(x\) is the original HR image \((N_1N_2 \times 1)\), \(y_r\) is the \(r\)th LR observation \((M_1M_2 \times 1)\), \(D\) is the downsampling matrix \((M_1M_2 \times N_1N_2)\), \(H_r\) is the blur matrix...
for $r$th frame ($N_1 N_2 \times N_1 N_2$) which describes motion blur and camera PSF, $W_r$ is the geometric warp for $r$th frame ($N_1 N_2 \times N_1 N_2$), $n_r$ is the noise in $r$th frame, and $m$ is the number of LR observations.

$W_r$ is estimated in the registration step and $D, H_r$ are generally assumed to be known. Solving for $x$ in (3) given the observations $y_r$ is an ill-posed inverse problem. Hence, it is important to use a priori information about $x$ that will reduce the space of solutions which conforms to the observed data. The Bayesian MAP formulation allows the incorporation of prior knowledge about $x$ to improve robustness during the reconstruction process. The MAP estimate of $x$ can be described as:

$$x = \text{arg max} \{ log[P(y_1, ..., y_m | x)] + log[P(x)] \}. \quad (4)$$

We propose a discontinuity adaptive Markov Random Field (DAMRF) prior in which the degree of interaction between pixels across edges is adjusted adaptively in order to preserve discontinuities [3]. Assuming that the noise fields are independent of $x$ and each other with variance $\sigma^2$, the MAP estimate is equivalent to:

$$\hat{x} = \text{arg max} \{ \frac{1}{2} \sum_{r=1}^{m} \frac{|y_r - DH_r W_r x|^2}{2\sigma^2} + \sum_{c \in C} g(d_c(x)) \}, \quad (5)$$

where $C$ is the set of cliques in the neighborhood system of a pixel, $d(x)$ is a spatial activity measure within the data and $g(n) = \gamma - \gamma e^{-u^2/\gamma}$ is the adaptive interaction non-convex function. A deterministic annealing GNC algorithm is used for optimization [3].

In order to improve the robustness of SR algorithm, we propose a process that verifies the quality of registration before including an image in the reconstruction. Let $\{X_r\}$ be the set of projections of $\{y_r\}$ onto the high-resolution grid using estimate $W_r$, $D$ and $H_r$. $\{X_r\}$ is the projection of the target observation. We define a registration error distance:

$$\text{dist}_r = \frac{1}{S} \sum_{p} |X_r - X_{r, \text{ref}}|, \quad 1 \leq r \leq m, \quad (6)$$

where $P$ is all image patches and each patch is of size $S$. An image of index $r$ is included in the MAP reconstruction if the corresponding distance satisfies:

$$\text{dist}_r \leq k \cdot \text{min}(\text{dist}_r), \quad 1 \leq r \leq m, r \neq \text{ref}, \quad (7)$$

where $k$ is a constant which is empirically set. Since mis-registered outliers can be removed, the SR algorithm converges faster and yields higher quality output.

D. ICHT-based Localization and Verification

The detail of our license plate localizer can be found in [8]. It should be noted that the weak localization in the section II.A is a part of our localizer. Then we use high level line segment features which describe directionality, regularity, similarity, alignment and connectivity, to discriminate plate / no-plate patterns. An improved connective Hough transform (ICHT) has been proposed for fast features extraction, and also for fine delimitation of plate pixels on image. High quality ROIs issued of SR step are used as input of the proposed localizer. False-alarm candidates are rejected and the boundary of the plate is more accurately localized.

III. EXPERIMENTS AND RESULTS

In this section, we demonstrate the robustness of the proposed approach in terms of registration, SR reconstruction and framework performance. Experiments were conducted on synthetic and real data with a Centrino PC (2.2 GHz) using C++.

First of all, three LR observations of size $60 \times 15$ pixels were degraded from a good quality license plate image by down-sampling, adding Gaussian blur, affine warping and Gaussian white noise of variance $\sigma^2 = 5$. Illumination variation was generated in the third LR observation (Fig. 2c). We compare our local phase registration with an invert computational gradient descent (ICGD) which uses point features correspondence for registration [11]. The performances were compared using absolute mean shift error $(\text{mse})$ [10]. Table I shows that the proposed algorithm performs much better than ICGD while illumination change exists which is more likely to appear in real video-surveillance, thanks to the illumination invariance of phase information.

Next, since benchmark data for evaluation of SR for LPR do not exist, we gather video sequences from UCSD project [12] as real data test. The input sequence for test consists of manually cropped license plates from 10 successive frames. We compare the proposed MAP-based SR method with a non-linear interpolation (NLI) [5] and the original DAMRF [3] and the results are shown in Fig. 3. The SR image using the NLI is quite blurred and low contrasted due to the high-frequency information lost during the process of sampling are not recovered. The DAMRF and the proposed algorithm visually yield the best result with more sharpened edges. Nevertheless, our algorithm converges faster after 18 iterations comparing to 30 for DAMRF. This comes from

![Figure 2. Synthetic degraded images: (a) blurred and noisy down-sampled, (b) affine warped of (a), (c) affine warped of (a) with illumination change.](image)

<table>
<thead>
<tr>
<th>Method</th>
<th>Const. Illum.</th>
<th>Illum. change</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICGD [11]</td>
<td>0.254</td>
<td>2.813</td>
</tr>
<tr>
<td>Proposed</td>
<td>0.247</td>
<td>0.587</td>
</tr>
</tbody>
</table>

Table I

REGISTRATION PERFORMANCE IN mse.
the automatic rejection of mis-registered images before the reconstruction stage.

Finally, we use a real sequence of 20 successive frames where there is so low quality license plate images that most existing localization methods [6] [7] [8] cannot localize (Fig. 3a). The proposed SR and localization [8] are tested in three different ways: SR on all pixels, then localize the license plate on the HR image; localize the plate in every frame, then do SR on cropped candidates; and the proposed framework (Fig. 1). The performance is evaluated by visual quality, localization precision and computation time. Fig. 4 shows that the output quality of Loc.→SR is inferior to those yielded by SR→Loc. and the proposed framework, due to errors in localization. SR→Loc. and the proposed framework localizes the license plate more accurately since the localization is performed in high quality image. However, the proposed framework is much faster than SR→Loc. (Tab. II) for the reason that only ROIs pixels are required to be estimated in SR reconstruction.

![Figure 3. Results of SR methods (up factor 2).](image)

![Figure 4. HR license plate outputs.](image)

<table>
<thead>
<tr>
<th>Framework</th>
<th>Computation time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SR → Loc.</td>
<td>0.21</td>
</tr>
<tr>
<td>Loc. → SR</td>
<td>2.05</td>
</tr>
<tr>
<td>Proposed</td>
<td>1.32</td>
</tr>
</tbody>
</table>

**IV. CONCLUSIONS**

In this paper, we introduce a unifying approach of license plate localization and SR for LPR in video-surveillance. Experiment results prove that our framework overcomes the instability of existing SR methods for LPR which require license plates are properly cropped from every frame; and vice versa localization methods expect image of sufficient quality. We also propose an accurate registration algorithm based on local phase which performs well even if illumination change happens. Another contribution is a robust MAP-based SR reconstruction excluding mis-registered images. In addition, our framework has real-time response (Tab. II) while tested with real video. In our further work, we propose to investigate the strategy of localization and SR interlacing in the MAP cost function.

**ACKNOWLEDGMENT**

This work has been partially supported by the French ANR Omnia project under the grant ANR-07-MDCO-009-02.

**REFERENCES**