Robust License Plate Extraction Using a Combination of Spatial and Frequency Domain Features

Erol Seke¹ Nihat Adar²
¹ Eskişehir Osmangazi University, Dept. of Electrical-Electronics Eng.
² Eskişehir Osmangazi University, Dept. of Computer Eng.
{eseke, nadar}@ogu.edu.tr

Abstract

A combination of familiar 'frequency spectrum' and 'variations about local mean' or so called 'signature' characteristics is used to locate and extract vehicle license plate from noisy digital images. Very promising results are obtained with relatively simple algorithm which is very suitable for real time license plate recognition systems.

1. Introduction

Automatic license plate recognition (LPR) is one of the main application fields for research on locating text in complex background. In the last decade, LPR systems became in the reach of many low cost applications like automated parking control/door along with being a part of complex traffic management and enforcement infrastructure. Although the problems in LPR systems are mostly identified and solved there is still room for further research on improvements like speed and accuracy especially when under difficult imaging conditions.

An LPR system can be separated into three nearly independent blocks;
1. One or more moving or fixed imaging system which consists of cameras, capturing/digitizing hardware and sometimes triggering systems. The output of this block is a single image or multiple images containing the license plate area of the vehicles.
2. License plate (LP) extractor finds and extracts the sub-image containing only the license plate.
3. Character recognition (CR) part determines the symbols contained in the LP and outputs what is detected as text.

The third block is obviously country dependent. As an additional complexity, license plates in many countries carry some symbols, stickers, province identifiers and such, other than the vehicle identifier characters. LPR systems can be developed and tuned for different country needs and diverse conditions.

Many researchers divide CR step into segmentation and recognition parts. However, recognition is possible with partly segmented or whole LP images, too. Binarization (threshold filtering) of segmented character sub-images is also a popular step before the recognition (an evidence of the fact that researchers ignore the information theory basics stating that any irreversible operation causes loss of information.)

LP extraction has been analyzed by several researchers. The algorithms for determination of LP region in images rely on the obvious (or sometimes not very obvious) characteristics of the LP. General features of LPs are:
1. There is more spectral power in somewhat higher spatial frequencies in the license plate region, due to the contrast and alternation between characters and the background. This feature is shown by not only LP text but also any proportionally sized text that is captured in the image.
2. A scan-line on LP has more mean-crossings than other regions. (related to feature 1)
3. Foreground of the characters usually consists of shades of a single color. Same argument may apply for background color in many countries. (This is only useful in color pictures and under strong lighting)
4. Width/Height ratio is usually close to 5 for single line plates.
5. The number of characters in a license plate is 5 to 10 (country specific). There may be additional symbols other than Latin characters and Arabic numerals. (However, in some countries Latin characters and Arabic numerals are not used at all.)

There are other constraints, measures and assumptions that algorithms utilize. For example, horizontal text assumption is very dependable for fixed camera setups. Some assumptions, on the other hand, may require handling of special cases. In many countries, foreground and background colors are fixed, but may have special cases such as inversion of these colors. In any case, algorithms use one or a combination of the features and
assumptions listed above. In this paper, we develop and analyze a technique which uses a combination of features 1 and 2 in LP extraction, and experimentally show that the technique achieves superior results with very poor images.

One of the simplest techniques for locating LP in a complex background image takes the advantage of background color of the plate. Lee, Kim and Kim exploited this feature in [1] using neural networks and assumed fixed width/height plate ratio. Obviously, the success of this technique strongly depends on the colors in the image, which makes it unreliable when used alone. An updated version of the technique ([2]) which uses colors, employed a support vector machine followed by an adaptive mean shift algorithm, making the approach inherently slow and not suitable for real time LPR. Earlier successful attempts for LP extraction use the technique developed for general text locating algorithms. Zhong, Karu and Jain, in [3], relied on the spatial variance over the text. Since their aim was to locate general text, not the LP, they did not use other features of LP. Barroso and colleagues suggested a direct search for character-like features in the image and expected at least three candidate patches horizontally positioned. In their 1997 paper [4], they also used the characteristic horizontal and vertical brightness changes (called as signature of LP) to improve detection accuracy. Vector quantization for compression generates some parameters related to the texture in the block of interest. These parameters are utilized by Zunino and Rovetta in [5] for detecting LP in the image. However, the complexity of this novel method is higher than many simpler techniques due to the inherent complexity of vector quantization.

A method directly related to the one in [3] has been proposed in [6] where local gradient, edge density and density variance maps obtained. The technique is somewhat slower than compared algorithms and open for further research. Another hybrid approach which employs edge detection, edge statistics, hierarchical-decisions and morphology was proposed in [7] claiming a success rate of 99.6%. However, their algorithm was tuned for stopped vehicles and took the advantage of it. They imply considerable improvement over the previous system used in a highway toll-collection boot, but do not mention the technique used in the old system.

Morphology and vertical edge detection is combined in [8] for 342 images with LPs of 6 countries, claiming 100% detection accuracy in algorithm runs tuned for single countries and 95% for entire library.

Broumandnia and Fathy, in [9], scanned one out of every N row in image horizontally and counted the edges stating that the number of vertical edges is an indication of presence of license plate. However, Anagnostopoulos and colleagues, in [10], dismissed this idea on the basis that the technique is too simple to have any use in LPR. They proposed an adaptive image segmentation technique which basically creates a region of interest map according to the statistical measures in an adaptively sized scanning window.

A different gradient-like approach is presented by Shi, Xu and Fu in [11]. Their technique finds the corners in the image using Harris' corner detection algorithm first and estimate the location by applying adaptive thresholds. Corners are assumed to be denser in LP region. However, the method fails in the presence of other text and text-like formations as expected.

The focus on the use of neural networks in LPR has been increasing recently, in parallel with the developments in neural network area. Several success stories have been announced both in LP extraction and in character segmentation ([12]). Success really depends on the selection of NN, its parameters and training images for each implementation. It is a tedious process which should be carried out very carefully.

Research aimed at the spatial frequency content of LP is rare. One implementation reported in [13] searches four quadrant corners for candidate LP locations and evaluates/eliminates them according to their frequency content. Hsieh, Huang and Jung, in [14], used wavelet transform to roughly identify the license plate location. They claim 92.4% accuracy with the possibility of detection of multiple LPs. However, reasoning and the possible advantages of wavelet transform is not very clear since no comparison is given with other methods.

Researchers usually argue that the complexity of the transforms required for a spatial frequency analysis is an obstacle in front of such approaches. However, a carefully designed method can eliminate unnecessary additional complexity introduced by transforms. This paper describes a method which employs short time discrete Fourier transform (STFT) and a measure of local variations.

2. The Setup and Assumptions

Test images are obtained from a 374x288 camera placed about 5 meters above a highway recording front views of the vehicles. Camera is triggered by the weight loops detecting the vehicles passing over. The resolution is low, relative size of LPs in the images is small and picture quality is not considered good depending on the lighting and weather conditions. But the assumption of horizontally aligned LPs is strongly valid as in many fixed camera setups.
3. Spatial Frequency Characteristics

One dimensional Discrete Fourier Transform (1D-DFT) is defined as

\[ F(k) = \sum_{n=0}^{N-1} f(n)g(n,k) \]

where \( f(n) \), \( g(n,k) \) and \( F(n) \) are input data set, transform kernel and output coefficients, also

\[ g(n,k) = (1/\sqrt{N}) \exp(-j2\pi nk / N) \] \( n,k = 0,...,N-1 \). Output coefficients of DFT are the weights in the finite sum

\[ f(n) = \sum_{k=0}^{N-1} F(k)h(k,n) \]

in order to reconstruct finite input data, where, just like \( g(n,k) \), \( h(k,n) \) represents sinusoidal waveforms of different frequency and phase. Eq-2 is called inverse DFT (IDFT). The magnitudes of the complex output from DFT represent the amplitude of the sinusoid at the corresponding frequency.

In Fig-1, a sample LP is shown. The white horizontal line shows where sample data is taken and in Fig-1b normalized data is plotted. In this plot, some sort of alternations ("signature"[4]) between bright and dark pixel values is seen, but the alterations are not at a specific frequency. The frequency spectrum plot in Fig-1c shows the relative magnitudes of the frequencies involved. In this plot zero-frequency component is suppressed since it is too big compared to other components. Spectrum in Fig-1c illustrates that relatively large portion of the spectral power scattered within lower half of the spectrum (depending on the size of the LP compared to number of samples).

In our analysis with different rows of several LP examples, similar results have been found. Depending on the resolution (horizontal size of LP in pixels), bandwidth of the major frequency components is increased or decreased as expected, but in no example the location of the main block was changed, nor any significant component was seen at the higher frequencies. This was not true for other subimages taken from vehicle pictures though. One of the most similar regions (other than other text in front of the vehicles) is shown in Fig-2.

What is shown in Fig-2 is a headlight protected with a cage illustrating vertical edges. Frequency spectrum of 8th row is also shown. Major difference between the spectrum in Fig-1c and Fig-2c is that the powerful components are shifted towards zero. Another example image containing a headlight at night illustrates a sinc-like formation in the frequency spectrum having significant components at higher frequencies. Such differences let us use

\[ f_m = \sqrt[3]{p_1p_2p_3} \]

where \( p_1, p_2 \) and \( p_3 \) are total powers in three almost equal-width sub-bands (\( \varepsilon-w/6 \), \( w/6-w/3 \) and \( w/3-w/2 \)) respectively where \( w \) is the width of the entire band. \( p_i \) s are calculated using

\[ p_i = \sum_k |F(k)|^2 \]

over the sub-band. Three lowest freq. components are not used (hence \( \varepsilon \)) in attempt to eliminate global brightness and slow gradual variations. Eq-3 is a geometric mean giving some measure of reasonable distribution among sub-bands and is not very reliable for locating LP alone. Other suggestions would include the use of neural networks trained by spectrums of LPs. However, our suggestion is simple and serves the purpose provided that additional constraints are also employed.

4. Count of Alternations

It is evident from Fig-1b that the "signature" of LP can be used for detecting LP location. This feature was used by many researchers previously ([4], [9]). One of the simplest ways to use this feature is to count the number of edges ([9]) or equivalently number of passes throw mean level. Counting is illustrated in Fig-3. Starting from the leftmost sample, data are scanned for mean passes. If a sample is above the gray area and, checking towards the right of this point, a sample below the gray area is found then we have a mean-pass at
hand. Inverse of this check is done starting from the newly found point. This way, the number of mean-crossings in the sample set is counted as a measure of the presence of LP and is expected that this count is higher in LP region.

The thickness of the gray region (threshold) is selected to be 5% of the difference of the max and min values in the entire image (not sample set). This hysteresis eliminates most of the false counts in noisy images.

![Figure 3. Alternation around mean of the data.](image)

**5. Two Features Combined**

The measures described in sections 3 and 4 are combined simply with multiplication

\[ f_P = f_m f_c \]  

where \( f_c \) is the count of mean-passes described in section 4. What is actually detected with these measures is not LP but the text in LP. Text other than LP is also detected. Therefore, employing other features listed in section 1 along Eq.5 would further reduce false detections. This, on the other hand, is left as future work.

Vehicle image is scanned and about \( N(M - L)/(N_k M_k) \) data sets of \( L \) pixels each are obtained, where \( N \) and \( M \) are image dimensions and \( N_k \) and \( M_k \) are vertical and horizontal downsampling rates respectively. That is, every \( N_k \)th row is processed and in the rows processed, \( (M - L)/M_k \) overlapping data sets are extracted where starting pixels of horizontal data sets are \( M_k \) pixels apart. The extraction of data sets is illustrated in Fig-4. Obviously both \( N_k \) and \( M_k \) can be 1 in which case the complexity would increase unnecessarily. Because the characteristics described do not quickly change from row to row or column to column. It is observed that when \( L \), \( N_k \), and \( M_k \) are selected to be around expected LP width, half of expected LP height and expected character length respectively, detection performance is not affected but the complexity is reduced considerably allowing usage in real time applications.

A measurement image is created using Eq.5 and it is smoothed in order to simplify the maximum point search. Fig-5 shows a typical vehicle image. Fig-6 illustrates where maximum feature values occur (lower-right). It is seen in Fig-6 that feature values are high over the headlight region because of abrupt brightness changes. Our algorithm was designed to find one LP. However, it is always possible to generate a confidence value from the maximum value at the feature image for rejecting input images with no visible LP.

![Figure 4. Extraction of the data sets from image](image)

![Figure 5. A test image (converted to gray-level).](image)

![Figure 6. a) Feature intensity image calculated using Eq.4. b) Smoothed feature image.](image)

A small subimage, which is 30% larger than expected LP, is extracted around the coordinate where the maximum is found. This subimage is trimmed above and below using the same measures but on every row (\( N_k = 1 \)). For trimming from left and right, total power in vertical DFT components (excluding 0 freq.) is checked. Decision on the trimming horizontal points are made based on the
comparison of these powers with an experimentally determined threshold (0.25 of the average). Final subimage showing extracted LP is shown in Fig-7.

Figure 7. Extracted LP (magnified here).

6. Tests Results and Conclusion

Algorithm is failed with only 4 of the 400 images used in the tests. One of the failed images is shown in Fig-8. It is seen that even if the algorithm succeeded, it is nearly impossible to recognize the characters in the LP of this vehicle. Although Fourier Transforms are used, the complexity of the algorithm is quite low due to downsampling caused by $k_x$ and $k_y$. Another reason is that the transforms are 1D. We believe that the proposed algorithm can be further tuned and additional constraints can be employed to eliminate most of the failures, which shall be the direction in future of this work.

Figure 8. An image for which the algorithm failed.

7. Acknowledgment

This work has been supported by Savronik Elektronik San. Tic. A.Ş., Eskişehir Turkey.

8. References