IWSSIP 2010 - 17th International Conference on Systems, Signals and Image Processing

Doble Back-Projection License Plate Recognition Framework

Krzysztof Malczewski
Department of Electronics and Telecommunications
Poznan University of Technology
PL-60-965 Poznan, Polanka 3, Poland
kmal@et.put.poznan.pl

Ryszard Stasinski
Department of Electronics and Telecommunications
Poznan University of Technology
PL-60-965 Poznan, Polanka 3, Poland
rstasins@et.put.poznan.pl

Abstract— In this paper, a novel algorithm for a car license recognition system is presented. We employ a double back-projection super-resolution image enhancement technique. Modification of initial guess estimation idea has resulted in accuracy improvement and convergence speed up. Bi-lateral back-projection filtering scheme which has been employed as an advanced preprocessor can achieve edge-preserving image smoothing by incorporating the additional information from the feature domain. After computing such processed input video sequence with respect to affine motion model Irani and Peleg's back-projection algorithm may be applied. Moreover, we consider single mobile camera image acquisition case. Our solution is fully unsupervised and gives excellent optical recognition results.

Keywords-super-resolution, license plate recognition

I. INTRODUCTION

Video camera based traffic surveillance system is an important part of many practical applications. Information about present situation on the road may be immediately extracted by image processing algorithms. License plate recognition (LPR) applies image processing and optical character recognition technology to identify vehicles by automatically reading their license plates. The problem of automatic license plate (LP) recognition has been studied since early nineties. LP recognition applications require a lot of complex tasks, such as license plate detection, segmentation and finally character recognition. These tasks used to become more sophisticated when dealing with plate images taken in various angles or noisy plate images with various lighting conditions. The last advances in surveillance systems and sensor manufacturing technologies have implied the popularity of this kind of devices. Unfortunately, their resolution and image quality very often are poor, which makes their usage problematic. For instance, optical character recognition systems need to have sharp and possibly denoised high resolution inputs. Super-Resolution technique, which is presented here, seems to be very useful for this task. SR methods have found applications in a wide range image applications, such as surveillance imaging, medical imaging. Here, relative displacement between set of low resolution images are particularly important. Super-resolution algorithms are particularly attractive when extension of existing low resolution system is a requirement while its rebuilding is not possible. A novel solution, as proposed in this work, is the pre- and postprocessing image enhancement, which increases the optical character recognition accuracy. Most license plate recognition systems in operation today use video sequence and custom hardware to improve the quality. Simple interpolation methods tend to produce blurry results, while edge preserving algorithms may remove image details in regions without strong edges. The mobile camera motion trajectory is random, so that we have adopted very flexible affine motion model to its high approximate. Additionally, as our work is addressed to high performance optical character recognition systems, introductory unsharp mask filtering is applied. Processed in this way low resolution image sequence is then combined into one high resolution image frame. It is realized by modified Irani and Peleg's iterative back-projection scheme. What should be noted, its typical drawback has been eliminated here modified on bilateral back-projection SR initial guess estimation philosophy is presented. Although this method can improve the image quality significantly [6], it suffers from some unsatisfying artifacts, such as the ringing effect and the chessboard effect [1]. This motivates the usage of isotropic back-projection kernel. The term of the back-projection is used here two times. In this work, the use of bilateral back-projection method as the introductory step is proposed. As has been shown, the original back-projection algorithm can minimize the reconstruction error efficiently under certain conditions [1]. Then, bilateral filtering is employed to guide the back-projection process. The knowledge about image edges is integrated into the method to avoid a cross-edge projection, thus the ringing effect and chessboard effect can be removed. The preprocessing part has to improve the image and facilitate its analysis. Finally, just before the OCR part we also apply a quadratic unsharp masking filter which allows for highlight high frequencies and enhanced character edges. Additionally, to increase robustness against impulsive noise, the quadratic 2D Teager filter [2] is used instead of linear unsharp filters. Namely, quadratic non-linear filters have proven their value in enhancement of character edges, as discussed in [3]. Double Back Projection Technique

Being powerful, the iterative back-projection super-resolution algorithm [10], however accurate and powerful, suffers from a couple of drawbacks, i.e. high convergence dependence from the initial guess accuracy. The initial guess image is usually obtained by averaging low resolution images or up-interpolating the reference one. Unfortunately, those approaches may result in convergence and final result accuracy deterioration. This observation prompted us to consider higher complexity initial guess computing procedure, see fig. 1. In this paper we propose introductory, bilateral back-projection based initial guess estimation method. Image sequence is first processed and affine motion parameters with respect to affine motion estimation algorithm [12] are compensated [11]. Further parameters are used to compose appropriately positioned irregular set of samples. After the compensation step the bilateral back-projection algorithm is applied

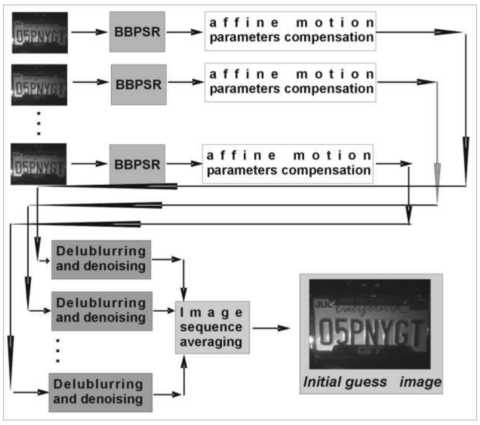

Figure 1. Structure of a back projection initial guess generation, BBPSR denotes Bilateral Back-Projection For Single Image Super Resolution

A. *Bilateral back-projection for initial guess reconstruction*

In general, the generation process of low resolution image can be modeled by a combination of the blur effect (due to atmospheric phenomena, the camera motion, and the sensor) and the down-sampling operation. Using a single filter g for the entire image, the generation process may be formulated as follows:

$$I^l = \left(I^h * g\right)\downarrow_s ,$$

where I^h and I^l are the high resolution (HR) and low resolution (LR) images respectively, $*$ is the convolution operator, and $\downarrow_s$ is the down-sampling operator with scaling factor s. The reconstruction error of an HR image I is formulated as the difference between the LR input image and the predicted LR image by I as follows

$$e_r(I) = I^l - \left(I * g\right)\downarrow_s \quad (2)$$

Back-projection

Back-projection [1] is an iterative algorithm of high resolution image reconstruction by minimizing the error function defined by (2). Starting with an initial guess [2] for the high resolution image, the imaging process is simulated to obtain the low resolution image corresponding to the observed input image. Given a LR input image sequence, the updating procedure can be summarized as doing the following two steps iteratively:

1) Compute the LR error $e_r\left(I_t^h\right)$ by (2).
2) Update the HR image by back-projecting the error as follows:

$$I_{t+1}^h = I_t^h + e_r\left(I_t^h\right)\uparrow_s * p ,$$

where I_t^h is the HR image at the t-th iteration, $\uparrow$ is the upsampling operator, p is a constant back-projection kernel.

where I_t^h is the HR image at the t-th iteration, $\uparrow$ is the upsampling operator, p is a constant back-projection kernel. In each step, the LR error $e_r\left(I^h\right)$ is back-projected to HR image by a kernel *p*.

Bilateral filtering

Bilateral filtering [1] is a method of non-linear filtering which can combine image information both from the space domain and the feature domain in the filtering process. In theory, it can be represented by the following equation:

$$h(x) = \frac{1}{k(x)} \sum_y I(y) c(x,y) s(I(x), I(y)),$$

where I and h are the input and output images respectively, x and y are pixel positions over the image grid, $c(x,y)$ and $s(I(x),I(y))$ measure the spatial and photometric affinity between pixel x and pixel y respectively, and $k(x) = \sum_y c(x,y) s(I(x),I(y))$ is the normalization factor at pixel *x*. The functions $c(\cdot)$ and $s(\cdot)$ are usually chosen as follows

$$c(x,y) = \exp\left(\frac{-\|x-y\|_2^2}{2\sigma_c^2}\right),$$

$$s(u,v) = \exp\left(\frac{-\|u-v\|_2^2}{2\sigma_s^2}\right),$$

The major idea of the bilateral filtering is to do the smoothing according to pixels not only close in the space domain, but close in the feature domain as well, thus the edge sharpness is preserved by avoiding the cross edge smoothing. Bilateral filtering is closely related to other edge preserving techniques widely applied in MRI image reconstruction, such as nonlinear diffusion and adaptive smoothing [1]

B. Back-projection multi-frame super-resolution image reconstruction

Starting with an initial guess I_0 [2] for the high resolution image, the imaging process is simulated to obtain a set of low resolution images $\{g_k^{(0)}\}$ corresponding to the observed input images $\{I_k^l\}$. In this paper we have extended the initial guess estimation issue. In [14] S. Borman, R. L. Stevenson argued compellingly of take an attention for initial guess image estimation procedure. These methods have the disadvantages of non-uniqueness of solution, dependence of the solution on the initial guess, slow convergence and high computational cost [14]. If I_0 were the correct high resolution image, then the simulated images $\{g_k\}$ should be identical to the observed images. The difference images $\{g_k - I_k^l\}$ are used to improve the initial guess by "back projecting" each value in the difference images onto its coresponding field in I_0, yielding an improved high resolution image I_1. This process is repeated iteratively to minimize the remaining error. This iterative update scheme can be expressed by:

$$I^{(n+1)} = I^{(n)} + \frac{1}{K}\sum_{k=1}^{K} T_k^{-1}(((g_k - I_k^{l(n)}) \uparrow s) * p)$$

where K is the number of low resolution images $\uparrow$ arrow an upsampling operator by a factor s and p is a back projection kernel determined by h and T_k. Taking the average of all discrepancies has the effect of reducing noise.

II. Character Segmentation of the License Plate

In contrast to the method of image binarization [5], the algorithm which is presented here uses the information of intensity and avoids the abruption and conglutination of characters that are the drawbacks of image binarization. In addition, because of using Hough transformation and the prior knowledge, the segmentation become more accurate and robust than the simple projection method [1]. The algorithm consists of three steps: preprocessing, horizontal segmentation and vertical segmentation.

A. Preprocessing

Preprocessing is very important for good performance of character segmentation. The preprocessing which is pre-sented in this paper includes size normalization (the license plate images are normalized), determination of plate kind and object enhancement. There are many kinds of license plates: black characters on a white background, white char-acters on a blue background, white characters on a black background and others.

B. Hough transformation based LP segmentation supported by adaboost algorithm

The segmentation part uses Hough transformation and the prior knowledge in horizontal, vertical segmentation and an information obtained from the AdaBoost algorithm.

Segmentation algorithm

The Hough transformation has been widely applied as a detector of lines in an image [1]. For each pixel in image space (x_0, y_0), using transformation. We get a curve $r = x_0 \cos\theta + y_0 \sin\theta$ in the parameter space (θ, r). Suppose that there are n points in the image space. After translating them to the parameter space, we obtain n curves in the parameter space. If these curves cross the same point (θ_0, r_0), then the n points in the image space are on a line. So we can find lines in the image space by searching the cross points in the parameter space. The Hough transformation is used on the midpoints of all subsection lines to eliminate the incorrect subsection lines and combine the correct subsection lines into a whole line. The vertical segmentation algorithm [1] is based on two approaches: projection analysis, constrained by the prior knowledge and AdaBoost algorithm. The vertical segmentation algorithm leads the estimation of the position of the left and right borders of the big interval, using the prior knowledge of character size. The knowledge about license plate characters is achieved from the AdaBoost character detector [8].

C. License plate detection based on the AdaBoost

Assuming that rough information according to the license plate position within the image plane is known a learning method based on the AdaBoost algorithm may be applied. Adaptive Boosting is a machine learning algorithm, formulated by Yoav Freund and Robert Schapire [7]. It is a meta-algorithm, and can be used in conjunction with many other learning algorithms to improve their performance. AdaBoost is adaptive in the sense that subsequent classifiers built are tweaked in favor of those instances misclassified by previous classifiers combines a set of 'important' weak classifier to form a strong classifier for object detection. Inspiring by [8] the integral image [9] to generate a bank of rectangle features for license plate detection is used. After feature, horizontal and vertical extraction, the AdaBoost framework is used to iteratively learn a strong classifier from a set of weak classifiers. At each step, a "good" weak classifier is selected and added in turn to form a strong classifier. Then, this approach and Viola and Jones method [9] are combined which results in a very fast license plate detector.

III. Post-Processing Non-Linear Filtering

To reconstruct image high frequencies, we need to enhance them in the LP image with appropriate filters. Significant high frequency part of image spectrum which is linked with character/background borders must be highlighted but impulsive perturbations interfere here. Appropriately chosen non-linear filtering approach satisfies these requirements. For instance, the 2D Teager filter, which is a class of quadratic Volterra filters [4] can be used to perform mean-weighted high pass filtering. Its impulse response is stronger in regions of high average intensity than in areas of low average intensity satisfying Weber's law [5]. Using *N* corresponding samples of images we perform Teager filtering to obtain $I_k^{h\eta}$, (*k=1,…, N*), being filtered images. This filter allows us to highlight character edges and suppress the noise level.

IV. The Proposed Algorithm

We propose a method in which preliminary image resolution enhancement is realized by bilateral back-projection for single frame image enhancement technique. This is the algorithm of image resolution enhancement. Typically, the bilateral filtering is a single frame super-resolution algorithm but in this paper it has been employed as the advanced initial guess multi-frame BP method. When the image resolution is enhanced license plate segmentation part may be applied. As its core we employ effective Hough transform supported by the AdaBoost learning system based technique. Furthermore, the Teager filter which value in OCR systems has been proved [3], is applied as pre-OCR processing part.

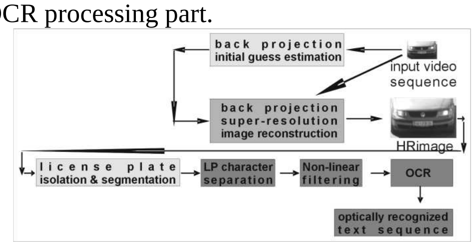

Figure 2. Structure of an all functional blocks in the proposed algorithm

Due to the use of the Teager filter the estimators proposed here are particularly well suited to enhancement of text sequence. In general, proposed improvements in OCR performance seem to result from increases in character separation rather than from the reduction of general noise. Finally, the ABBYY FineReader Optical Character Recognition software has been employed.

V. Experiment

In the experiment about 100 frames video sequence from built in mobile phone VGA camera have been used. The images were captured under different angles and positions. Results of direct use of Optical Character Recognition system has been shown in figure 3c. Robustness of the pro-posed approach is demonstrated by the fact that in approximately 99% of cases (when characters of the license plate were visible) the OCR system resulted in perfect recognition.

VI. Conclusion

Fully automated license plate recognition framework is presented. Multi frame based image resolution enhancement technique has been applied to license plate images refinement. Additionally, pre and post processing has been implemented, which has improved the OCR system performance.

"05PNYGT"

Figure 3. (a) One of the test video frames which have been used in the experiment, (b) results of the character segmentation technique, (c) the output of the OCR system.

It should be noted that this promising fully automated surveillance system may be used in many practical applications. The system may be useful when the information about car license plate is needed. The text form of information is easier to use then raster based one. This form is flexible and may be rapidly broadcasted and processed in many databases all over the world. Our approach may found applications in private transport systems such as travel−time measurements, parking lot traffic management, toll collection, speed−limit enforcement, and identification of stolen cars.

References

[1] S. Dai, M. Han, Y. Wu, Y. Gong, "Bilateral Back-Projection For Single Image Super Resolution," in Proc. IEEE Int'l Conf. on Multimedia and Expo (ICME'07), Bei-jing, China, July 2007.

[2] K. Malczewski and R. Stasinski, "Optical Character Recognition of Low Resolution Text Sequences from Hand-Held Device Supported by Super-Resolution," ELMAR, pp. 61-64, 48th International Symposium ELMAR-2006 focused on Multimedia Signal Processing and Communications, 2006.

[3] G.Ramponi, P.Fontanot, "Enhancing document images with a quadratic filter," in Signal Processing, 33:23-34,1993.

[4] S.Mitra, G.Sicuranza, "Nonlinear Image Processing" in Academic Press, 2000.

[5] A.K.Jain, " Fundamentals of Digital Image Processing," in Prentice Hall, 1989.

[6] A.K. Katsaggelos, R. Molina, and J. Mateos, "Super resolution of images and video ' in Synthesis Lectures on Image, Video, and Multimedia Processing. Morgan & Claypool, 2007.

[7] Y. Tai,W. Tong, and C. Tang,, "Perceptually-inspired and edge-directed color image super-resolution' in CVPR, 2006.

[8] Y. Freund, and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," In Proc. Second European Conference on Com-putational Learning Theory, pp. 23-37. 1995

[9] C.C. Chen and J. W. Hsieh, "License Plate Recognition from Low-Quality Videos," MVA2007 IAPR Conference on Machine Vision Applications, May 16-18, 2007, Tokyo, JAPAN

[10] P. Viola and M. J. Jones, "Robust real-time facedetec-tion," IJCV, vol. 57, no. 2, pp.137-154, May 2004.

[11] M. Irani and S. Peleg, Improving Resolution by Image Registration, CVGIP: Graphical Models and Image Processing Vol. 53, No. 3, May, pp. 231-239, 1991

[12] K. Malczewski, R. Stasinski, "Generalized Iterative Back-Projection Algorithm for Super-Resolution Moving and Static Object Extraction", Proc. IWSSIP'06 (CD), 2006

[13] G. Farnebäck, Fast and Accurate Motion Estimation using Orientation Tensors and Parametric Motion Models. In Proceedings of 15th International Conference on Pattern Recognition, volume 1, pages 135. 139, Barcelona, Spain, September 2000. IAPR

[14] S. Borman and RL Stevenson, "Super-resolution for image sequences a review", in Proc. IEEE Int. Symp. Cir-cuits and Systems , Jun. 1998, pp. 374-378.