Fast Super-resolution for License Plate Image Reconstruction

YUAN Jie[1], DU Si-dan[2], ZHU Xiang[3]
Department of Electronic Science and Engineering, Nanjing University, Nanjing 210093, China
[1]yuanjie@nju.edu.cn, [2]dsd@ese.nju.edu.cn, [3]zhuxiang0703@gmail.com

Abstract

A fast super-resolution reconstruction algorithm designed for license plate recognition is proposed in this paper. It uses a new reduced cost function to produce images of higher resolution from low resolution frame sequences. Computational cost required in this algorithm is much lower compared with other methods. The effectiveness of the proposed algorithm is demonstrated through blind reconstruction experiments with real videos, whose result images are nearly equivalent to those yielded by classical MAP-based approaches. The presented algorithm can be applied in real-time recognition systems to improve their performances, and to reduce the requirement of imaging hardware.

1. Introduction

Super-resolution reconstruction (SR) aiming to form a high resolution image by combining multiple low-resolution images has received much attention these years. By now one main obstacle to the real application of this technology is the lack of accurate motion estimation of complex movements in general videos. However, under certain circumstances a simple global motion model could meet our needs. One example is the License Plate Recognition (LPR) field, in which the movement pattern of object is comparatively simple for estimating.

Most LPR algorithms and commercial LPR systems are designed for toll gates, parking lots, or other situations, where it is easy to take sharp and large plate photographs from vehicles that stay still or move in a very low speed. Unfortunately, if we employ such systems to identify faster-moving vehicles, plates might hardly be recognized because we need to consider motion blur problems then. For running vehicles, the shorter the distance between the camera and the object is, the more severe the motion blur becomes. On the other hand, if cameras were stationed far away or were zoomed out to avoid blur effect, the reduced resolution of plate images would make it even harder to recognize.

Another problem concerns additional information attached to the plate, such as province initials (as in China, these are Chinese characters), country flag (as in Europe), or state identification (as in USA). The recognition of such information would require even higher quality of images captured for LPR.

Some institutions, including leading companies in this field, developed high-speed systems for fast moving vehicles. However, these systems depend largely on sophisticated imaging hardware with much higher resolution and shutter speed to tackle the problem above. They are, of course, very expensive, and have hindered the extensive use of LPR system for automatic traffic monitoring.

Employing SR technology to facilitate recognition procedure is a promising method. The first article that clearly introduced the idea of combing SR and LPR to identify moving vehicles was written by Suresh and Kumar [2]. However, their Maximum a posteriori (MAP) based method is computationally demanding, which prevents it from real-time processing.

A fast MAP-based SR algorithm was proposed by Tanaka and Okutomi in [3]. It utilizes a new cost function, which greatly reduced the number of pixels required for reconstruction. Of course this method can be employed directly to enhance the quality of plate images, but we will see that the computational cost can be further reduced if the SR technology were applied specifically for LPR systems.

Actually for most optical character recognition (OCR) applications, the input image of characters would be converted into a binary one before feature vector extraction. Many details recovered by the SR process are useless and would be removed during later steps. Accordingly, we propose an optimized fast SR approach for moving LPR systems in this paper. This

978-1-4244-2175-6/08/$25.00 ©2008 IEEE

algorithm only costs a small amount of calculation, and can be applied in DSP devices for real-time processing. In Section 2, we briefly review the principle of Tanaka's fast MAP-based algorithm. In Section 3, we describe our method that can produce nearly equivalent binary results while cost even less computational time. Experiments and results are shown in Section 4. Finally, in Section 5, we offer our conclusions.

2. Fast MAP-based SR Algorithm

Let us denote by vector **f** the ideal high resolution (HR) image in lexicographical order, which generates the observed low resolution (LR) image sequence. Each LR image is denoted by vector $\mathbf{g}_k$, with k = 1, 2, …, K. We assume that the size of $\mathbf{g}_k$ is $(M\times N)\times 1$, and the HR image **f** is of $(PM\times PN)\times 1$, where P is referred to as the magnification factor. The imaging system model that produces each LR image can be described as:

$$\mathbf{g}_k = \mathbf{A}_k \mathbf{H}_k \mathbf{C}\left(\mathbf{d}_{k,r}\right)\mathbf{f}_r + \mathbf{e}_k \quad (1)$$

In this equation, the warp matrix $\mathbf{C}(\mathbf{d}_{k,r})$ of size $(PM\times PN)\times(PM\times PN)$ represents motion compensation operation that maps frame $\mathbf{f}_r$ to frame $\mathbf{f}_k$. This is matrix is determined by motion estimation parameter $\mathbf{d}_{k,r}$. $\mathbf{H}_k$ is a blur matrix of size $(PM\times PN)\times(PM\times PN)$, which can describe motion blur and sensor point spread function (PSF). The $(M\times N)\times(PM\times PN)$ matrix $\mathbf{A}_k$ denotes downsampling process, and the estimation noise is represented by $\mathbf{e}_k$.

We assume, in this section, that matrices $\mathbf{A}_k$, $\mathbf{H}_k$ and $\mathbf{C}(\mathbf{d}_{k,r})$ are all known before the reconstruction stage. For convenience, Eq. (1) can be rewritten in a simple form:

$$\mathbf{g}_k = \mathbf{B}_{k,r}\mathbf{f}_r + \mathbf{e}_k \quad (2)$$

where $\mathbf{B}_{k,r} = \mathbf{A}_k \mathbf{H}_k \mathbf{C}(\mathbf{d}_{k,r})$. Given these data and some prior information, the objective of SR is to estimate an accurate approximation of the ideal HR $\mathbf{f}_r$.

Classical MAP-based SR method, which is widely used to solve such problems, can be treated as an optimization process minimizing certain cost function like:

$$I_1 = \sum_{k=1}^{K}\sum_{l=1}^{M\square N}\left[\mathbf{b}_k(x_{k,l},y_{k,l})^T\cdot\mathbf{f} - g_{k,l}\right]^2 + \lambda\sum_{c=1}^{PM\square PN}\varphi_c(\mathbf{f}) \quad (3)$$

where $\mathbf{b}_k(x_{k,l}, y_{k,l})$ represents the PSF from HR image **f** to $g_{k,l}$, which denotes the lth LR pixel in image $\mathbf{g}_k$ for position $(x_{k,l}, y_{k,l})$. For different pixel positions with subpixel accuracy, their corresponding PSF varies from each other. $\varphi_c(\mathbf{f})$ here denotes smoothness constraint related to prior information [4]. The optimization of this term will penalize ringing noise in reconstructed images, and preserve edges and other high frequency content simultaneously.

Most optimization algorithms are iteration-based. During each iteration the number of pixels to be is at least $C = KMN$, which includes all the points in LR image sequence. Such huge number is the main reason that leads to the stupendous computational task of classical SR methods.

A reduced cost function was proposed in [3]. To deduce this function, each pixel position in LR images should first be mapped to the HR image coordinates. Then, these positions are discretized to the center of their nearest HR image pixels. We use R_j to denote the set of LR pixels corresponding to the jth HR pixel. In this set, the average value of LR pixel intensities can be obtained using:

$$h_j = \frac{1}{\omega_j}\sum_{m\in R_j} g_m \quad (4)$$

where ω_j is the number of elements of set R_j. The discretization makes PSF vectors for LR pixels corresponding to the same HR pixel become identical. So the cost function described in Eq. (3) can be reduced as:

$$I_2 = \sum_{j=1}^{C_h}\omega_j\left[\mathbf{b}(x_j,y_j)^T\cdot\mathbf{f} - h_j\right]^2 + \lambda\sum_{c=1}^{C_h}\varphi_c(\mathbf{f}) \quad (5)$$

In this function C_h = PM×PN. That means during one iteration only pixels in the HR image need to be estimated.

3. Proposed Fast SR for LPR

3.1. Further Reduced Function

The fast MAP-based SR approach described in Section 2 is applicable for various kinds of videos and for general motion. In the application for LPR, however, there exist some special characteristics which could be utilized for further cost reduction. For example, in most cases the license plate region of the image is composed of only two or three different colors: a background color, and one or two colors for characters. The background color and the colors for characters are usually with strong contrast to ensure that the message is clear. So when the image is converted into a grayscale bitmap before the SR or LPR process, there remain only two kinds of intensities in the plate region: the dark one and the light one, and they will finally be distinguished by certain threshold during the binary conversion step. We can see that smoothness constraint in Eq. (5) has a small effect on the binary image (see Figure 2). In other words, this constraint can be omitted.

In Eq. (5), parameter ω_j indicates that the variance of noise for h_j is proportional to the element number of set R_j. Although ω_j varies with each individual set, here we assume that the stroke shape features of result image would be influenced slightly if ω_j was approximated to be identical to every set. Experiments proved this assumption for most cases (see Section 4).

The further reduced cost function now become:

$$I_3 = \sum_{j=1}^{C_h} \left[\mathbf{b}(x_j, y_j)^T \cdot \mathbf{f} - h_j \right]^2 \qquad (6)$$

Since the PSF is space invariant, the optimization of Eq. (6) could be treated as a deconvolution problem of the following equation:

$$\mathbf{h} = \mathbf{b} * \mathbf{f} + \mathbf{e} \qquad (7)$$

Vector **h** is called the average image yielded by motion estimation, and by pre-averaging using Eq. (4). $*$ is the convolution operator, and vector **e** denotes estimation error. Many deconvolution methods, such as Wiener filter, can be used to obtain HR image according to the minimum mean square error (MMSE) standard.

3.2. Implementation

The overall SR algorithm for LPR is described as follows:

a. Get input sequence $\mathbf{g}_1$, $\mathbf{g}_2$… $\mathbf{g}_K$. Set image $\mathbf{g}_r$ as the reference image.

b. Locate and extract license plate region from each image. The plate extraction problem is not the main problem in this paper. Many algorithms can be used to solve it [5].

c. Estimate motion parameters for the image sequence using optical flow equation (OFE) based estimation method [6]. Since the license plate is a rigid body moving in a stable manner, only global transition, rotation, and some zooming caused by perspective projection needs to be considered. It means the motion can be characterized by a few parameters. We employ the affine model here:

$$\begin{aligned} x_2 &= a_1 x_1 + a_2 y_1 + a_3 \\ y_2 &= a_4 x_1 + a_5 y_1 + a_6 \end{aligned} \qquad (8)$$

d. Set the magnification factor P, map each LR image pixel to its corresponding HR image pixel according to the motion parameters that we estimated in step c, and calculate the average image **h** using Eq. (4).

e. Deconvolute image **f** from average image **h** through Wiener filter. The format of the space-invariant PSF is determined by the magnification factor P:

$$S_j(x_i, y_i) = \begin{cases} \dfrac{1}{Z} \exp\left(-\dfrac{\left((x_i - x_j)^2 + (y_i - y_j)^2\right)}{P} \right) & \begin{aligned} &\text{if } x_i - x_j \le Q \\ &\& \, y_i - y_j \le Q \end{aligned} \\ 0 & \text{other wise} \end{cases} \qquad (9)$$

where Z is the normalizing constant. Parameter Q is used to control the size of nonzero part of the PSF.

f. Further enlarge the image using bicubic interpolation, and convert it into binary image for the subsequent recognition procedure. In this step, the interpolation process can make strokes smoother, and easier to be identified.

4. Experiments

In this section, we report experimental results of the proposed fast SR algorithm. For comparison, we also show results obtained by other SR methods. All the algorithms have been run in MATLAB. The Wiener filter employed in our method was implemented by using MATLAB function *deconvwnr*. In order to estimate the performance of these algorithms in real application, we used some standard image sequences for super-resolution reconstruction, and some videos captured by a digital hand-held camera for blind reconstruction. Each sequence contains 20 up to 30 frames, with unknown motion parameters.

Comparisons of zoom-in reference image and reconstructed images using different methods are shown respectively in Figure 1. The input sequence consists of 30 frames. The magnification factor P is 3, and parameter Q is set 2. We can see that both the original LR plate image and the interpolated one can hardly be identified. The SR methods have improved the image quality, and make them possible to be recognized. The difference between results of MAP, maximum likelihood (ML, another SR algorithm similar to MAP with no prior term) and proposed Fast SR is slight. Although some ringing artifacts remain on image (d) and (e), they tend to be noticeable only in some flat regions, and have little influence on the reconstructed shape of characters. The binary images obtained from Figure 1 (c) – (e) are shown in Figure 2. They are nearly the same.

Figure 3 shows the results of another test. Due to the function reduction described in Section 3, there exist some grid noises in the result of fast SR. However, such noise influenced slightly to the shape characteristics.

All the tests were run in a PC with the Celeron CPU (2.66 GHz). The calculation time of the two sequences using three SR algorithms are summarized in Table 1. It is clear that although three SR methods produce similar results, the calculation time of our fast method is much less than the other two.

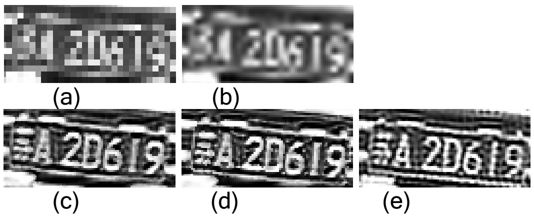

Figure 1. Test of a license plate sequence. (a) zoom-in original image. (b) bicubic interpolation. (c) ML. (d) MAP. (e) proposed fast SR.

Figure 2. Binary results yielded by (a) MAP and (b) proposed algorithm.

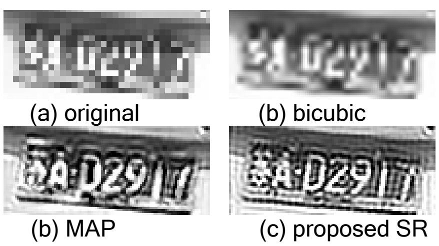

Figure 3. Results of license plate sequence taken by a hand-held camera.

Table 1. Comparison of calculation time

Resolution	Method	Calculation time (sec)
117×78 (A 2D619)	MAP	5.83
	ML	5.65
	Fast SR	0.86
114×81 (A D2917)	MAP	6.08
	ML	5.79
	Fast SR	0.88

5. Conclusions

In this paper, we introduced a specific fast SR algorithm for LPR systems. It was based on the optimization of a new cost function, which can drastically reduce the calculation task. Blind reconstruction experiments using real videos have proved that this algorithm can produce plate images nearly equivalent to those yielded by classical MAP-based SR methods, while it has a small computational cost and can be employed for real time processing. Our method can also be used to improve recognition rate of other video-based OCR systems.

6. References

[1] A.K. Katsaggelos, *Super Resolution of Images and Video*, Mprgan & Claypool Publisher, 2007.

[2] K.V. Suresh, G.M. Kumar, and A.N. Rajagopalan, "Superresolution of License Plates in Real Traffic Videos", *IEEE Trans. Intelligent transportation systems*, June 2007, Vol. 8, No. 2.

[3] M. Tanaka, M. Okutomi, "A fast MAP-based super-resolution algorithm for general motion", *Proc. SPIE Electronic Imaging 2006*, Vol. 6065, pp. 404-415.

[4] R.R. Schultz, R.L. Stevenson, "A bayesian approach to image expansion for improved definition", *IEEE Trans. Image Processing*, May. 1994, Vol. 3, No. 3, pp. 233-242.

[5] J. Xu, S. Li, and M. Yu, "Car License Plate Extraction Using Color and Edge Information", *Machine Learning and Cybernetics*, 2004, Vol. 6, pp. 3904-3907.

[6] Y. Altunbasak, R.M. Mersereau, and A.J. Patti, "A Fast Parametric Motion Estimation Algorithm with Illumination and Lens Distortion Correction", *IEEE Trans. Image Processing*, 2003, Vol. 12, pp. 395-408.

[7] D. Robinson, S. Farsiu, and P. Milanfar, "Optimal Registration of Aliased Images Using Variable Projection with Applications to Superresolution", *The ComputerJournal,* 2007.

[8] X. Zhu, J. Yuan, and S. Du, "Super-Resolution Reconstruction of Image Sequences Compressed with DWT-Based Techniques", *Proc. International Conference on Wavelet Analysis and Pattern Recognition*, Beijing, 2007, pp. 555-560.