Journal of Southeast University (English Edition) Vol. 26, No. 3, pp. 457 – 460 Sept. 2010 ISSN 1003—7985

Super-resolution reconstruction for license plate images of moving vehicles

Lu Xiaobo[1] Zeng Weili[2]

([1] School of Automation, Southeast University, Nanjing 210096, China)
([2] School of Transportation, Southeast University, Nanjing 210096, China)

Abstract: A novel reconstruction method to improve the recognition of license plate texts of moving vehicles in real traffic videos is proposed, which fuses complimentary information among low resolution (LR) images to yield a high resolution (HR) image. Based on the regularization super-resolution (SR) reconstruction schemes, this paper first introduces a residual gradient (RG) term as a new regularization term to improve the quality of the reconstructed image. Moreover, L_1 norm is used to measure the residual data (RD) term and the RG term in order to improve the robustness of the proposed method. Finally, the steepest descent method is exploited to solve the energy functional. Simulated and real acquired video sequence experiments show the effectiveness and practicability of the proposed method and demonstrate its superiority over the bi-cubic interpolation and discontinuity adaptive Markov random field (DAMRF) SR method in both signal to noise ratios (SNR) and visual effects.

Key words: super-resolution; residual gradient term; residual data term; license plate; regularization

Super-resolution (SR) techniques aim at estimating a high-resolution (HR) image by utilizing a sequence of low-resolution (LR) frames. In many civilian and military applications, such as medical imaging, traffic video sequences using low cost sensors, HR images are always required. In this paper, we are interested in traffic imaging applications. If a video surveillance system captured a hit-and-run accident on public routes, a post facto analysis of the stored video may be required. Though very high spatial resolution has been used, however, due to the fact that the heavily compressed images or the capture environment is not desirable, information about the identity of the vehicles involved in the accident may not be easily derivable. Moreover, to the best of our knowledge, license plate recognition (LPR) systems remain the principal vehicle identifier despite the fact that they can be deliberately altered in fraud situations or replaced. Although the LPR system has made great progress in the last decade, there is still much work to be done, such as a robust LPR system should be adaptable to the variety of the environment. Assuming that the license plate region is detected at very low resolution, the current LPR system does not effectively work well[1]. The problem addressed in this paper uses image processing techniques to improve the readability of license plate text and some other signs of moving vehicles.

There are many reasons leading to the degradation of the required video quality, such as motion blur, camera lens blur, atmospheric blur, and noise. From a restoration point of view, the idea is to increase the information content in the final image by using the additional spatial-temporal information that is available in each of the LR images. A variety of the SR algorithms[2-6] have been proposed since the multi-frame SR problem was first addressed in Ref. [7]. It is impossible for us to introduce all the related methods due to the limitation of the content of this paper, therefore only those of great importance will be discussed to give readers an overview of the recent papers. In Ref. [8], a method for generating a high resolution slow-motion sequence from a given video is introduced, in which an area of interest such as the license plate is slowed down and super-resolved. Matan et al. [9] used a probabilistic and crude motion estimation to reconstruct a high resolution image.

More recently, a new super-resolution algorithm was proposed by Suresh et al. [10], which modeled the HR image as a Markov random field (MRF) with a discontinuity adaptive regularization. This approach obtained promising results in some cases when tested on real traffic video sequences. In this paper, a residual gradient (RG) term is introduced as a gradient homogeneity constraint term to the regularization SR reconstruction model [10]. At the same time, in order to improve the robustness of our method, the residual data (RD) term and the RG term are measured with L_1 norm.

1 Observation Model

It is important to set up a suitable imaging model, which can truly reflect the real physical process of imaging degradation, to implement the super-resolution reconstruction. In the practical process of image sampling, there exist some degrading factors, i. e., atmospheric blur, motion blur, camera lens blur, and sampling devices.

Since we are interested in traffic imaging applications, camera blur often has a greater effect on images than atmospheric blur (which is very important for remote sensing images) in conventional imaging systems, and thus only camera blur is considered in this paper. Supposing that we observe n LR images $\boldsymbol{Z}_1, \boldsymbol{Z}_2, \ldots, \boldsymbol{Z}_n$, we use the notation in Ref. [10] to formulate the general super-resolution model. The degraded sequence of the original image is assumed to be translational with scaling-variant motions. The blur comes from camera defocus and motion blur. The general vector super-resolution formulation is defined as

$$\boldsymbol{Z}_k = \boldsymbol{E}_k \boldsymbol{B}_k \boldsymbol{H}_k \boldsymbol{X} + \boldsymbol{I}_k \qquad 1 \leqslant k \leqslant n \tag{1}$$

Received 2009-12-30.

Biography: Lu Xiaobo (1965—), male, doctor, professor, xblu@seu.edu.cn.

Foundation items: The National Natural Science Foundation of China (No. 60972001), the National Key Technology R&D Program of China during the 11th Five-Year Plan Period (No. 2009BAG13A06).

Citation: Lu Xiaobo, Zeng Weili. Super-resolution reconstruction for license plate images of moving vehicles[J]. Journal of Southeast University (English Edition), 2010, 26(3): 457 – 460.

where $\boldsymbol{X}$ represents the HR image of size $N_1N_2 \times 1$ which is rearranged in a lexicographic order; $\boldsymbol{Z}_k$ and $\boldsymbol{I}_k$ are two columns of size M_1M_2 representing the observed LR frames and the additive noise, respectively; $\boldsymbol{B}_k$ is the camera lens blur matrix of size $N_1N_2 \times N_1N_2$; $\boldsymbol{H}_k$ is an $N_1N_2 \times N_1N_2$ matrix that models the motion corresponding to the reference frame; $\boldsymbol{E}_k$ is the decimation matrix of size $M_1M_2 \times N_1N_2$.

The recovery of the image $\boldsymbol{X}$ from $\{\boldsymbol{Z}_k\}_{k=1}^n$ relies on the knowledge of the involved operators. By setting the center frame as a reference image, the operator $\boldsymbol{H}_k$ is obtained through the motion estimation between the reference image and the image $\boldsymbol{Z}_k$. Since all the images in the sequence go through the same camera and atmosphere, we assume the noise, the decimation operator, and the blur operator are identical for all the images. Therefore, $\forall\, k$, $\boldsymbol{I}_k = \boldsymbol{I}$, $\boldsymbol{E}_k = \boldsymbol{E}$, and $\boldsymbol{B}_k = \boldsymbol{B}$.

2 Proposed Regularization Framework

The following expression formulates a generalized minimization cost function[4],

$$f(\boldsymbol{X}) = \arg\min_{\boldsymbol{X}} \left\{ \sum_{k=1}^{n} \rho(\boldsymbol{Z}_k, \boldsymbol{EBH}_k\boldsymbol{X}) + \lambda g(\boldsymbol{X}) \right\} \quad (2)$$

where λ is the regularization parameter which provides a tradeoff between the constraint regularized (CR) term $g(\cdot)$ and the RD term $\rho(\cdot)$.

Suresh et al. [10] defined RD and RG terms as

$$\rho(\boldsymbol{Z}_k, \boldsymbol{EBH}_k\boldsymbol{X}) = \| \boldsymbol{Z}_k - \boldsymbol{EBH}_k\boldsymbol{X} \|_2^2 \quad (3)$$

and

$$\begin{aligned} g(\boldsymbol{X}) = \sum_{i=1}^{N_1}\sum_{j=1}^{N_2} 4\gamma - \gamma\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i,j-1)]^2}{\gamma}\right\} - \\ \gamma\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i,j+1)]^2}{\gamma}\right\} - \\ \gamma\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i-1,j)]^2}{\gamma}\right\} - \\ \gamma\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i+1,j)]^2}{\gamma}\right\} \end{aligned} \quad (4)$$

where γ is the annealing parameter.

According to Ref. [4], L_1 norm leads to a more robust result in error estimation than L_2 norm, so we adopt this result in this paper. Therefore, the RD term (3) can be rewritten as

$$\sum_{k=1}^{n} \rho(\boldsymbol{Z}_k, \boldsymbol{EBH}_k\boldsymbol{X}) = \sum_{k=1}^{n} \| \boldsymbol{Z}_k - \boldsymbol{EBH}_k\boldsymbol{X} \|_1 \quad (5)$$

To better preserve edges and guarantee the gradient consistency between the estimated high resolution image and the original LR images, a gradient-based constraint, namely the RG term, is introduced to Eq. (2).

$$\begin{aligned} f(\boldsymbol{X}) = \arg\min_{\boldsymbol{X}} \left\{ \sum_{k=1}^{n} \rho(\boldsymbol{Z}_k, \boldsymbol{EBH}_k\boldsymbol{X}) + \lambda_1 g(\boldsymbol{X}) + \right. \\ \left. \lambda_2 h(\nabla \boldsymbol{Z}_k, \boldsymbol{EH}_k \nabla \boldsymbol{X}_k) \right\} \end{aligned} \quad (6)$$

where $h(\cdot)$, namely the RG term, measures the Euclidean distance between the estimated HR image and the gradient of the observed LR images. We define the RG term as

$$h(\nabla_1\boldsymbol{Z}_k, \boldsymbol{EH}_k(\nabla_l\boldsymbol{X})) = \sum_{k=1}^{n}\sum_{l=1}^{4} \| \nabla_l\boldsymbol{Z}_k - \boldsymbol{EH}_k(\nabla_l\boldsymbol{X}) \|_1 \quad (7)$$

where the L_1 norm is used to measure the difference between the gradient maps. The gradients $\nabla_1\boldsymbol{X}$, $\nabla_2\boldsymbol{X}$, $\nabla_3\boldsymbol{X}$ and $\nabla_4\boldsymbol{X}$ are given by

$$\left.\begin{aligned} \nabla_1\boldsymbol{X} &= (\boldsymbol{S}_y^1 + \boldsymbol{S}_x^{-1}\boldsymbol{S}_y^1 + \boldsymbol{S}_x^1\boldsymbol{S}_y^1)\boldsymbol{X} - (\boldsymbol{S}_y^{-1} + \boldsymbol{S}_x^{-1}\boldsymbol{S}_y^{-1} + \boldsymbol{S}_x^1\boldsymbol{S}_y^{-1})\boldsymbol{X} \\ \nabla_2\boldsymbol{X} &= (\boldsymbol{S}_x^1 + \boldsymbol{S}_x^1\boldsymbol{S}_y^{-1} + \boldsymbol{S}_x^1\boldsymbol{S}_y^1)\boldsymbol{X} - (\boldsymbol{S}_x^{-1} + \boldsymbol{S}_x^{-1}\boldsymbol{S}_y^{-1} + \boldsymbol{S}_x^{-1}\boldsymbol{S}_y^1)\boldsymbol{X} \\ \nabla_3\boldsymbol{X} &= (\boldsymbol{S}_x^1 + \boldsymbol{S}_x^1\boldsymbol{S}_y^{-1} + \boldsymbol{S}_y^{-1})\boldsymbol{X} - (\boldsymbol{S}_x^{-1} + \boldsymbol{S}_x^{-1}\boldsymbol{S}_y^1 + \boldsymbol{S}_y^1)\boldsymbol{X} \\ \nabla_4\boldsymbol{X} &= (\boldsymbol{S}_x^1 + \boldsymbol{S}_x^1\boldsymbol{S}_y^1 + \boldsymbol{S}_y^1)\boldsymbol{X} - (\boldsymbol{S}_x^{-1} + \boldsymbol{S}_x^{-1}\boldsymbol{S}_y^{-1} + \boldsymbol{S}_y^{-1})\boldsymbol{X} \end{aligned}\right\} \quad (8)$$

where $\boldsymbol{S}_x^{\pm 1}$ and $\boldsymbol{S}_y^{\pm 1}$ are horizontal and vertical shift matrices by ± 1 pixel, respectively.

3 Super Resolution Implementation

By using the RD term (5), the regularization term (4), and the RG term (7), we propose the following super-resolution reconstruction model,

$$\begin{aligned} f(\boldsymbol{X}) = \arg\min_{\boldsymbol{X}} \left\{ \sum_{k=1}^{n} \| \boldsymbol{Z}_k - \boldsymbol{EBH}_k\boldsymbol{X} \|_1 + \lambda_1 g(\boldsymbol{X}) + \right. \\ \left. \lambda_2 \sum_{k=1}^{n}\sum_{l=1}^{4} \| \nabla_l\boldsymbol{Z}_k - \boldsymbol{EH}_k(\nabla_l\boldsymbol{X}) \|_1 \right\} \end{aligned} \quad (9)$$

We use steepest descent (SD) optimization procedure to find the solution to minimize problem (9). To derive the SD update procedure for the image estimation, we first compute the gradient of the cost function in Eq. (9) with respect to $\boldsymbol{X}$. Using Eqs. (4) and (8), the gradient of the cost function in Eq. (9) with respect to $\boldsymbol{X}$ is given by

$$\begin{aligned} \nabla f(\boldsymbol{X}) = \left\{ \sum_{k=1}^{n} \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{BE}(\mathrm{sign}(\boldsymbol{EH}_k(\nabla_l\boldsymbol{X}) - \nabla_l\boldsymbol{Z}_k)) + \right. \\ \lambda_1\nabla(g(\boldsymbol{X})) + \lambda_2 \sum_{k=1}^{n}\sum_{l=1}^{4} \boldsymbol{H}_k^{\mathrm{T}}\boldsymbol{E}(\mathrm{sign}(\boldsymbol{EH}_k(\nabla_l\boldsymbol{X}) - \\ \left. \nabla_l\boldsymbol{Z}_k)) \frac{\partial(\nabla_l\boldsymbol{X})}{\partial\boldsymbol{X}} \right\} \end{aligned} \quad (10)$$

where

$$\begin{aligned} \nabla(f(\boldsymbol{X})) = \\ 2[\boldsymbol{X}(i,j) - \boldsymbol{X}(i,j-1)]\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i,j-1)]^2}{\gamma}\right\} + \\ 2[\boldsymbol{X}(i,j) - \boldsymbol{X}(i,j+1)]\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i,j+1)]^2}{\gamma}\right\} + \\ 2[\boldsymbol{X}(i,j) - \boldsymbol{X}(i-1,j)]\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i-1,j)]^2}{\gamma}\right\} + \\ 2[\boldsymbol{X}(i,j) - \boldsymbol{X}(i+1,j)]\exp\left\{-\frac{[\boldsymbol{X}(i,j) - \boldsymbol{X}(i+1,j)]^2}{\gamma}\right\} \end{aligned} \quad (11)$$

The SD optimization iteration procedure can be summarized briefly as follows:

1) Choose a proper value of γ and an iteration number N;

2) Use the bi-cubic interpolation to calculate $\boldsymbol{X}^{(0)}$ as the initial guess of the HR image;

3) Update the state $\boldsymbol{X}^{(n+1)} = \boldsymbol{X}^{(n)} - \beta^n \nabla f(\boldsymbol{X}^{(n)})$;

4) If $\| \boldsymbol{X}^{(n+1)} - \boldsymbol{X}^{(n)} \| / \| \boldsymbol{X}^{(n)} \| \leqslant \varepsilon$ or $n \leqslant N$, then $\boldsymbol{X} = \boldsymbol{X}^{(n)}$. Where ε is a small positive constant and β^n presents the step size at the n-th iteration.

4 Experiments

In order to show the effectiveness of SR reconstruction with the proposed algorithm based on the proposed L_1 norm, the discontinuity adaptive Markov random field (DAMRF) and residual gradient, we test it in both simulated and real acquired sequences. For the simulated experiment, the performance of the reconstruction algorithm is evaluated by measuring the improvement in signal to noise ratio (SNR)[10], which is defined by

$$S = 10\lg \frac{\| \bar{\boldsymbol{X}} - \boldsymbol{X}^{(0)} \|^2}{\| \boldsymbol{X} - \bar{\boldsymbol{X}} \|^2} \tag{12}$$

where $\bar{\boldsymbol{X}}$ is the original image; $\boldsymbol{X}^{(0)}$ is the initial estimation of the super-resolution image which can be obtained by interpolating the first low-resolution frame using standard bilinear or bi-cubic interpolation; $\boldsymbol{X}$ is the super-resolution image. To verify the effectiveness of the proposed algorithm, denoted as L_1 + DAMRF + RG, for SR reconstruction, we evaluate the proposed algorithm subjectively and objectively in comparison with bi-cubic interpolation and the DAMRF SR model[10]. The criterion for selecting the parameter in each experiment of each algorithm is to choose parameters to produce the visually most appealing results. For fair evaluation, we use the trial-and-error method for determining the optimal parameter and the best result is chosen as the output of the algorithm for this experiment.

4.1 Evaluation of simulated sequence

For the simulation experiment, we use an HR license plate of size 28 × 91 pixels for the simulated data (see Fig. 1 (a)). First, we shift the HR image by a pixel in the horizontal and vertical directions. Then the shifted image is corrupted by the independent white Gaussian noise of size 5 × 5 with a standard deviation equal to 5 and salt-and-pepper noise with 0.02 noise density. The resulting image is downsampled by a factor of 2 in both horizontal and vertical directions to generate six LR images from the original HR image. Fig. 1(b) shows one representative LR example. Fig. 1 (c) shows the SR reconstruction obtained by using the bi-cubic interpolation, which does not help to improve the legibility. Fig. 1(d) shows the result obtained by using the DAMRF method[10]. Fig. 1(e) shows the SR reconstruction of the proposed L_1 + DAMRF + RG. The SNR of the DAMRF method[10] and the proposed method L_1 + DAMRF + RG are 9.15 dB and 9.45 dB, respectively. We observe that the proposed method looks better because L_1 norm leads to a more robust result and CG is helpful to suppress noise and preserve more edge information than the algorithm without this constraint.

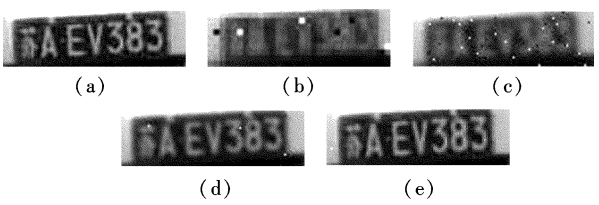

Fig. 1 Simulation results of different super-resolution methods. (a) Original HR frame; (b) One of six LR frames; (c) Bi-cubic interpolation; (d) DAMRF model[10]; (e) L_1 + DAMRF + RG

4.2 Evaluation of real acquired sequence

To further evaluate the performance of the proposed method, two real acquired LR sequences are used. In Fig. 2, we show the effectiveness of the proposed SR method in comparison with bi-cubic interpolation and the DAMRF method[10]. Fig. 2(a) shows six consecutive LR frames of a real traffic video. Fig. 2(b) shows the result of the SR reconstruction result obtained by using bi-cubic interpolation. The output of the bi-cubic interpolation is quite blurred, and some of the numbers are not at all discernible. The result of the SR reconstruction by using the DAMRF method[10] is shown in Fig. 2(c), which performs relatively better, but some of the letters are not easily legible. For example, the letter "C" and the last number "2" are confused. For comparison, the proposed L_1 + DAMRF + RG is shown in Fig. 2 (d). Note that the license plate can be read clearly. The experiments described above demonstrate the practical utility and the potential of the proposed method in enhancing LR frames captured from real traffic conditions.

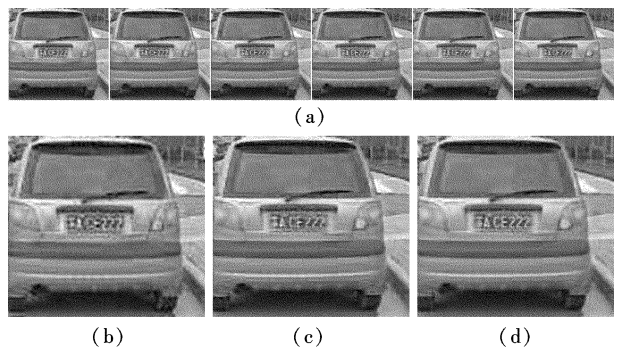

Fig. 2 SR reconstruction results from real LR video. (a) Six frames of the acquired LR sequence; (b) Bi-cubic interpolation; (c) DAMRF[10]; (d) L_1 + DAMRF + RG

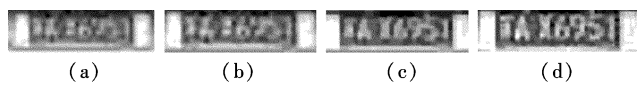

Fig. 3 Results of different methods applied to license plate. (a) One of five LR frames; (b) Bi-cubic interpolation; (c) DAMRF model[10]; (d) L_1 + DAMRF + RG

In Fig. 3, one license plate sequence which is cropped from a real video frame captured using a camera at the overbridge is used. This paper considers five frames of the low resolution sequence, one of which for our captured sequence is shown in Fig. 3(a). Note that this license plate is quite blurred and almost all of the numbers are difficult to read. The outputs corresponding to different super-resolution methods are given in Figs. 3(b) to (d). We again notice that the reconstructed image using the proposed method is significantly better compared with the existing methods.

5 Conclusion

We present a new super-resolution image reconstruction

algorithm to enhance the quality of a set of blurred image frames of moving vehicles and produce an HR image with fewer blur effects. The new method, based on the DAMRF model, uses L_1 norm to measure the RD term and introduces the RG term to improve the quality of the reconstructed image. Experimental results show the effectiveness of the proposed method, especially when compared with the DAMRF method.

References

[1] Anagnostopoulos C N E, Anagnostopoulos I E, Psoroulas I D, et al. License plate recognition from still images and video sequences: a survey [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2008, **9**(3): 377 - 391.

[2] Takeda H, Milanfar P, Protter M, et al. Super-resolution without explicit subpixel motion estimation [J]. *IEEE Transactions on Image Processing*, 2009, **18**(9): 1958 - 1975.

[3] Baker S, Kanade T. Limits on super-resolution and how to break them [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, **24**(9): 1167 - 1183.

[4] Farsiu S, Robinson D, Elad M, et al. Fast and robust multi-frame super-resolution [J]. *IEEE Transactions on Image Processing*, 2004, **3**(10): 1327 - 1344.

[5] Farsiu S, Robinson D, Elad M, et al. Advances and challenges in super-resolution [J]. *International Journal of Imaging Systems and Technology*, 2004, **14**(2): 47 - 57.

[6] Park S C, Park M K, Kang M G.. Super-resolution image reconstruction: a technical review[J]. *IEEE Signal Processing Magazine*, 2003, **1**(5): 21 - 36.

[7] Huang T S, Tsai R Y. Multi-frame image restoration and registration [J]. *Advances in Computer Vision and Image Processing*, 1984, **1**(2): 317 - 339.

[8] Chaudhuri S, Taur D R. High-resolution slow-motion sequencing—how to generate a slow-motion sequence from a bit stream [J]. *IEEE Signal Processing Magazine*, 2005, **22**(2): 16 - 24.

[9] Matan P, Elad M. Super resolution with probabilistic motion estimation [J]. *IEEE Transactions on Image Processing*, 2009, **18**(8): 1899 - 1904.

[10] Suresh K V, Kumar G M, Rajagopalan H N. Super-resolution of license plates in real traffic videos [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2007, **8**(2): 321 - 331.

运动车辆车牌图像的超分辨率重建

路小波[1] 曾维理[2]

([1] 东南大学自动化学院,南京 210096)

([2] 东南大学交通学院, 南京 210096)

摘要:为了改善实际交通环境中运动车辆车牌图像的质量,提出一种新的超分辨率重建方法,即通过融合低分辨率图像间的互补信息得到一幅高分辨率车牌图像. 首先,在超分辨率重建正则化框架下引入梯度残差项作为一个梯度强制项来改善重建图像的质量. 其次,为了提高重建算法的鲁棒性,用 L_1 范数度量数据残差项和梯度残差项. 最后,用最速下降法求解相应的最小能量泛函. 模拟和实际视频图像序列的实验结果验证了所提方法的有效性和实用性,所提方法在重建图像的信噪比指标和视觉效果方面均优于双三次插值和 DAMRF 法.

关键词:超分辨率; 梯度残差项; 数据残差项; 车牌; 正则性

中图分类号:TP391. 41