

Vehicle license plate super-resolution using soft learning prior

Yushuang Tian · Kim-Hui Yap · Yu He

Published online: 24 May 2011
© Springer Science+Business Media, LLC 2011

Abstract This paper proposes a new algorithm to perform single-frame image super-resolution (SR) of vehicle license plate (VLP) using soft learning prior. Conventional single-frame SR/interpolation methods such as bi-cubic interpolation often experience over-smoothing near the edges and textured regions. Therefore, learning-based methods have been proposed to handle these shortcomings by incorporating a learning term so that the reconstructed high-resolution images can be guided towards these models. However, existing learning-based methods employ a binary hard-decision approach to determine whether the prior models are fully relevant or totally irrelevant. This approach, however, is inconsistent with many practical applications as the degree of relevance for the prior models may vary. In view of this, this paper proposes a new framework that adopts a soft learning approach in license plate super-resolution. The method integrates image SR with optical character recognition (OCR) to perform VLP SR. The importance of the prior models is estimated through relevance scores obtained from the OCR. These are then incorporated as a soft learning term into a new regularized cost function. Experimental results show that the proposed method is effective in handling license plate SR in both simulated and real experiments.

Keywords License plate super-resolution · Soft learning prior · Optical character recognition

Y. Tian · Y. He
Media Technology Lab, School of Electrical and Electronic Engineering, Nanyang Technological University, Nanyang Avenue, Singapore 639798

Y. Tian
e-mail: tian0042@e.ntu.edu.sg

Y. He
e-mail: heyu@ntu.edu.sg

K.-H. Yap (✉)
School of Electrical and Electronic Engineering, Nanyang Technological University, Block S2.2, Nanyang Avenue, Singapore 639798
e-mail: ekhyap@ntu.edu.sg

1 Introduction

Image super-resolution (SR) is a process to reconstruct high-resolution (HR) image from single or multiple low-resolution (LR) observations. An important application of image SR is reconstruction of HR vehicle license plate (VLP) from LR observations. This could be used in surveillance systems or traffic regulation enforcement, amongst others. Image SR generally can be categorized into single-frame or multi-frame SR, where a HR image is reconstructed from a single LR image or multiple LR images, respectively. This paper will focus on single-frame license plate SR as at times, only a single frame of the license plate image can be captured due to various imaging constraints.

Different single-frame image SR or interpolation methods have been proposed over the years. Classical interpolation methods such as bi-cubic interpolation use piecewise polynomials to model smoothness in image intensity of the local spatial image neighborhood [2]. However, these methods tend to produce over-smoothness near the edges and textured regions. Hence, other methods have been developed to handle this issue. Adaptive bi-cubic interpolation methods use local features such as edge direction to improve interpolation. Inverse gradient [10] and warped distance [17] have been employed to determine the weights of bi-cubic interpolation. In [9], the edge information of each local region is extracted by using the discrete cosine transform (DCT). Different edge types are identified and used to determine different interpolation strategies for each area. The basic idea of [13] is to estimate the HR local covariance coefficients from the LR counterpart based on their geometric duality. The edge-directed interpolation is tuned based on the covariance. It is noted that all the methods discussed involve estimation of heuristic threshold values and filter weights. Thus, the results are sensitive to changes in values of these parameters.

Regularization methods formulate the single-frame SR problem into the optimization of a cost function consisting of a data fidelity term and a regularization term. The regularization term incorporates some prior information of the image models to achieve a stable solution to the problem [11]. In [5], a smoothness prior is used and the experimental results show that it is superior to linear minimum mean squared error (LMMSE) interpolation algorithm and maximum entropy technique. In [20], the authors apply a discontinuity adaptive regularization term to preserve the edges in the SR reconstruction. A framework of constrained least squares minimization using singular value decomposition (SVD) was proposed in [3]. However, it is noted that most the regularization terms used in the methods above are centered on smoothness priors. They do not incorporate specific prior information about the targeted types of images that are being reconstructed. To address this issue, a text-based interpolation method using L_1 -norm regularization term is developed in [21]. However, the regularization term becomes increasingly ineffective as the magnification factor of the HR image reconstruction increases [1, 8, 14].

Learning-based SR methods use training samples and models to perform single-frame SR. An advantage of these methods is that even when the magnification ratio of the HR image to the LR image is large, the information obtained from learning examples can still restore the details of the HR images effectively [7]. Through the learning process, a lookup table between the HR patches and their corresponding LR patches is built. Given a representative learning dataset, these methods can work well to produce sharper images. A maximum a posteriori (MAP) model using learning-based prior has been introduced in [4]. The authors use nearest-neighbor (NN) search to find the HR patches. The learning term then integrates the details information based on relationship between the example HR and LR patches. Existing learning-based methods use binary weighting schemes, where the

weights are taken as ‘0’ for totally irrelevant samples and ‘1’ for fully relevant samples. The main shortcoming of these methods is that they are prone to artifacts when unsuitable HR patches are chosen for learning. The binary strategy tends to produce artifacts as the HR patches cannot be integrated partially in the cases when the patches only bear partial resemblance to the target LR patches. The HR patches either have to be incorporated fully or discarded outright. Further, these methods consider the HR reconstruction from a local perspective. They do not utilize the global information of the alpha-numeric existed in the VLP when performing reconstruction.

In view of this, this paper proposes a new single-frame SR framework to address the issues above. The main novelties of the proposed method are threefold: (a) It introduces a new iterative technique that integrates optical character recognition (OCR) and HR reconstruction into license plate SR. (b) It proposes a soft learning prior that estimates the importance of different learning samples, and integrates this information into HR reconstruction of the license plate. This is as opposed to conventional hard-decision methods, which either incorporate the training sample fully, or discard them outright. (c) It introduces a scheme that utilizes both global and local characteristics when performing license plate HR reconstruction. This is in contrast with the existing learning methods that use local information alone. Experimental results show that the proposed method can outperform other well-known methods.

The rest of this paper is organized as follows. The problem formulation is introduced in Section 2. The overview of the proposed framework is explained in Section 3. An iterative HR reconstruction and VLP recognition algorithm using soft learning is developed in Section 4. Experimental results on both simulated and real-life images are presented and discussed in Section 5. A brief conclusion is given in Section 6.

2 Problem formulation

The single-frame LR image can be modeled as applying blurring and down-sampling to the original HR image, followed by addition of noise. The observation model can be expressed in matrix–vector form as:

$$\mathbf{g} = \mathbf{W}\mathbf{f} + \mathbf{n} \quad (1)$$

where $\mathbf{f} = [f_1, f_2, \dots, f_{DN_1 \times DN_2}]^T$, $\mathbf{g} = [g_1, g_2, \dots, g_{N_1 \times N_2}]^T$, and $\mathbf{n} = [n_1, n_2, \dots, n_{N_1 \times N_2}]^T$ represent the HR image of sized $DN_1 \times DN_2$, the observed LR image of sized $N_1 \times N_2$ and the additive noise of sized $N_1 \times N_2$, respectively. The parameter D denotes the decimation factor, and matrix \mathbf{W} with dimension $N_1 N_2 \times DN_1 DN_2$ represents the operator that includes the blurring and down-sampling process.

The objective of single-frame SR is to reconstruct f from the observed g . The probability density functions (pdfs) of g and f are given by $p(g)$ and $p(f)$, respectively. The development of the new soft learning cost function is explained as follows. To estimate the HR image, we will determine the argument f that maximizes the *a posteriori* probability $p(f|g)$. Applying the Bayesian theorem, we have

$$p(f|g) = \frac{p(g|f)p(f)}{p(g)} \quad (2)$$

As the denominator $p(g)$ is a common term that does not affect the maximization of $p(f|g)$, we will drop it in the subsequent algorithm development. Assuming n is additive white

Gaussian with variance σ_n^2 , then the likelihood $p(g|f)$ can be expressed as:

$$p(g|f) \propto \exp\left\{-\frac{\|g - Wf\|_2^2}{2\sigma_n^2}\right\} \tag{3}$$

We represent the *prior* of the HR image $p(f)$ in an exponential form, which is the general format of Gibbs distribution:

$$p(f) \propto \exp\{-A(f)\} \tag{4}$$

The term $A(f)$ is a non-negative energy function, with smaller values for more probable signals, and vice versa. Taking the logarithm of $p(f|g)$, the MAP estimation can be formulated as minimizing the following cost function:

$$L(f) = \|g - Wf\|_2^2 + \lambda A(f) \tag{5}$$

where λ is the regularization parameter that controls the relative contribution between the data fidelity term $\|g - Wf\|_2^2$ and the HR image prior $A(f)$. As this paper focuses on HR reconstruction of license plates in a domain-specific application, we propose a new soft learning cost function that incorporates the characteristics of license plate to perform learning:

$$L(f) = \|g - Wf\|_2^2 + \lambda \sum_{i,j} \sum_{k \in \Omega[i,j]} w_k^{[i,j]} \left\| R_{[i,j]} f - x_k^{[i,j]} \right\|_2^2 \tag{6}$$

The new cost function employs a soft learning prior instead of conventional smoothness prior. The first term in the cost function represents the data fidelity term, while the second term is the learning-based functional which incorporates the learning samples. The operator $R_{[i,j]}$ extracts a $m \times m$ block from the image f , and compares it with the HR training patches $x_k^{[i,j]}$ at location $[i,j]$. The inner summation is conducted over all HR patches $x_k^{[i,j]}$ where $k \in \Omega[i,j]$ is the patch index and $\Omega[i,j]$ is the set of all selected HR patches at location $[i,j]$. The outer summation runs through all the pixel locations of the HR image. λ is the regularization parameter that controls the relative contribution between the data fidelity term and the learning functional. A small λ value will cause the reconstructed HR image to emphasize on the data fidelity term, while a large λ value will provide more emphasis on learning towards the training samples. $w_k^{[i,j]}$ is the soft learning weight for a selected training patch.

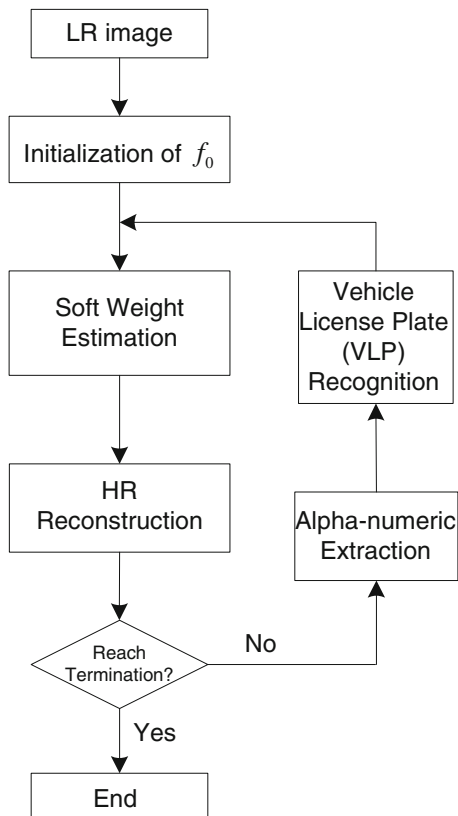
The main feature of the proposed method is that it uses a soft learning prior in the cost function. As opposed to existing learning approaches such as [4] where binary weighting schemes are used (i.e. the weight $w_k^{[i,j]}$ is either 0 or 1), the proposed framework uses a soft learning framework where the weight $w_k^{[i,j]} \in [0, 1]$ is assigned to a continuous value between 0 and 1. The weight is estimated based on the relevance/resemblance of the training patches with respect to the current patch under reconstruction. This will ensure a more meaningful incorporation of information from learning patches, and reduce the number of outliers (dissimilar patches) from being included into the training process. Further, the proposed method incorporates OCR to improve the performance of the system as the license plate is consisted of alphanumeric. To the best of our knowledge, this is the first study where an iterative process of image SR and OCR has been integrated with soft learning to perform license plate SR. The OCR will be used to estimate the importance of learning patches which is reflected in $w_k^{[i,j]}$. An advantage of the proposed method is that it

will take both the global and local information into consideration when performing HR reconstruction. In contrast with the conventional learning-based method which relies on local patch matching alone, the method will combine the global information obtained using OCR with the local patch matching to reach an ideal compromise in performing image SR.

3 Overview of the proposed framework

The overview of the proposed soft learning license image SR is given in Fig. 1. It consists of two major iterative steps of HR image reconstruction and VLP recognition. Firstly, the initial HR image f_0 is estimated using bi-cubic interpolation of the LR image. The initial estimate f_0 will then be passed into the soft weight estimation to determine the importance of the learning patches. It is noted that in the first iteration, a binary weight estimation scheme for the learning patches will be used to reconstruct the HR image. For the subsequent iterations, a soft weight estimation scheme using the inputs from VLP will be employed to determine the importance of the learning patches. The HR reconstruction of the license image using the soft learning will then performed by minimizing the proposed cost function using conjugate gradient optimization. The reconstructed HR image is then passed to alphanumeric extraction when the image is segmented into different characters. These characters are then passed through the VLP recognition which involves OCR of each

Fig. 1 Flowchart for the proposed framework



segmented alphanumeric. The OCR outputs may contain some uncertainties, especially in the beginning of the iterations. They are used to estimate the soft weight estimation of the HR learning patches. The iterative process of HR image reconstruction and VLP recognition will continue so that progressively better HR image and VLP recognition can be obtained. The process will continue until convergence or a maximum number of iterations is reached.

4 Iterative HR reconstruction and VLP recognition using soft learning

4.1 Creation of learning patch database

In many domain-specific applications such as SR of license plates and facial images, the use of training examples has been shown to be effective. This is because the target images are known to possess certain characteristics that can be utilized in HR reconstruction. For example, in the case of VLP HR reconstruction, the VLPs are consisted of 36 alphanumeric (26 letters and 10 digits) in many countries. In order to exploit this information, learning-based approaches will generate sets of training patches for the LR images and the corresponding HR images. This work will follow the standard procedure of generating the LR and HR learning patch database.

Figure 2 shows the creation of learning patch database. First, for each HR alphanumeric image of sized $K \times L$, a set of overlapped HR patches of sized $m \times m$ is sampled. Then, their corresponding LR patches of sized $n \times n$ are obtained as follows: (a) Apply the uniform blur to the HR images to produce a set of overlapped blurred patches of sized $m \times m$, (b) downsample the blurred patches by a factor D , followed by an additive noise. A database containing corresponding HR and LR patch pairs is generated. The pairs of HR-LR patches in the database are partitioned into 36 alphanumeric categories.

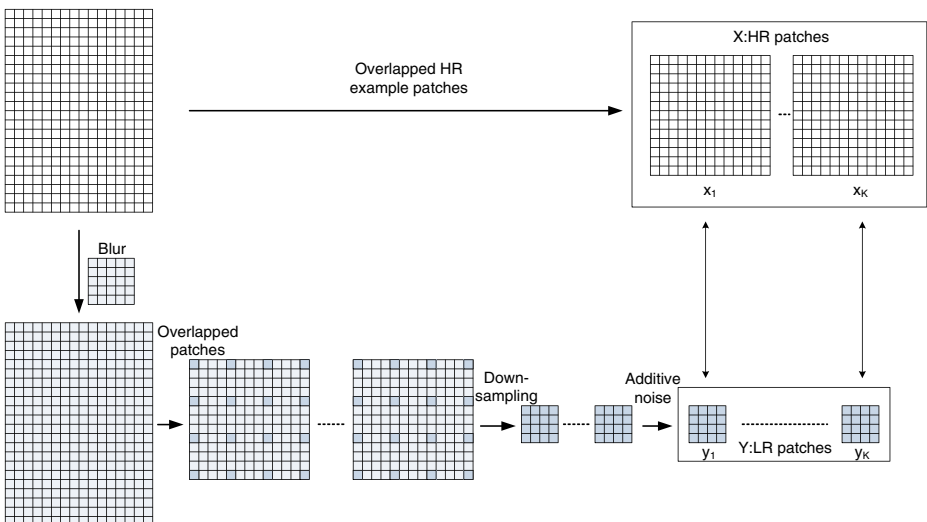


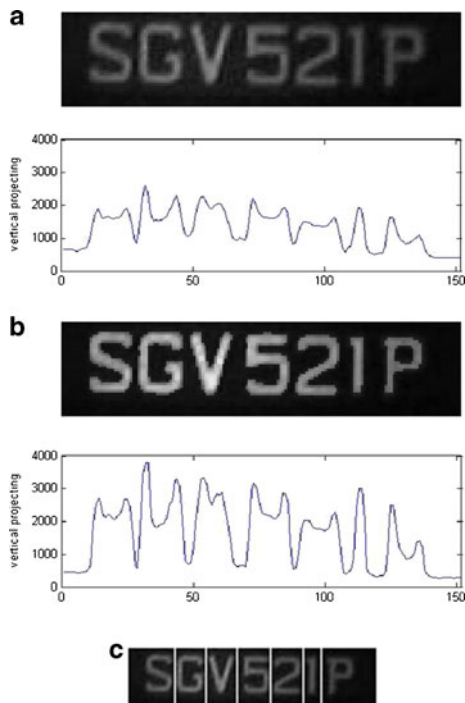
Fig. 2 Creation of the patch learning database

4.2 VLP recognition

The main objective of VLP recognition is to exploit the information that the characters on the license plate are consisted of alphanumerics, hence the outcome of the OCR can be used to provide appropriate soft learning weights for different HR patches during the HR reconstruction. The VLP recognition consists of three steps: image preprocessing, character segmentation and OCR. As the quality of VLP images may vary significantly in different image capturing conditions, preprocessing is a necessary step prior to segmentation. A real-world example of captured VLP image and its vertical projection are given in Fig. 3a. It can be seen that the non-uniform illumination will pose a problem in character segmentation. In order to address the issues of changes in illumination and additive noise, we extend the image enhancement method in [22] to the license plate image. The result after applying the image enhancement technique to Fig. 3a is given in Fig. 3b. It can be seen from Fig. 3a and b that the enhanced image has an intensity profile that will simplify subsequent character segmentation. The vertical projection of the enhanced image also shows more pronounced peaks and valleys when compared to the original image.

Character segmentation is then employed to segment the license image into individual characters using the vertical projection image histogram. A problem with projection analysis is that the vertical profile may sometimes yield spurious boundaries which may cause a character to be segmented into two or more parts erroneously. To address this problem, additional information such as ratio of the height and width of the alphanumeric are also used to enhance the robustness of the segmentation technique. Figure 3c shows the segmentation results after image enhancement and projection analysis.

Fig. 3 VLP segmentation. **a** Original image, **b** Enhanced image, **c** Segmentation result



The next step involves performing OCR on the segmented characters. Different OCR techniques including neural network (NN) techniques [19, 15] and support vector machine (SVM) [16, 12], have been used. However, these methods need extensive training and they are usually computational intensive. As the VLP consists of alphanumeric with rigid format, computationally efficient technique such as template matching is used in this case. The cross-correlation scores between the normalized alphanumeric image and each character prototype are calculated [18]. The normalized cross-correlation score is given as follows:

$$\rho = \frac{(\mathbf{f} - \mu_f)^T (\mathbf{v} - \mu_v)}{\|\mathbf{f} - \mu_f\|_2 \|\mathbf{v} - \mu_v\|_2} \quad (7)$$

where \mathbf{f} represents the individual segmented alphanumerics in the license plate, \mathbf{v} represents each of the 36 character templates used in VLP, μ_v and μ_f are the uniform vectors formed by the average gray level of \mathbf{v} and \mathbf{f} , respectively. The OCR process is centered on the computation of the normalized cross-correlation value ρ for all the templates \mathbf{v} . The value ρ is used as an input to estimate the soft weights $w_k^{[i,j]}$ in the next section.

4.3 Soft learning weight estimation

Conventional learning-based approaches employ crisp binary decision when determining whether a HR patch should be used in learning. In other words, the patch is either considered to be fully relevant or totally irrelevant in guiding the local HR reconstruction. This approach is clearly inconsistent with real-world observations as the HR learning patches often contain only partial resemblance to the HR image under reconstruction.

In view of this, this section will introduce a soft learning weight estimation scheme that determines which HR patches will be chosen for learning and what weight $w_k^{[i,j]} \in [0, 1]$ will be assigned to these learning patches. The main idea is that those HR patches originating from the top-matched character templates will be assigned greater weight during the soft learning process. For example, while performing HR reconstruction of a digit “6”, the OCR template matching may yield the top three matching character templates as “6”, “5” and “S”. Therefore, the HR patches arising from these three character templates should be given higher weights to guide the solution towards these characters.

The process of soft weight estimation is explained as follows. First, the LR license image is partitioned into various overlapping LR patches. The current target LR patch is then compared with all the LR patches in the database. If the difference between the target LR patch with the LR patches in the database is less than a predefined threshold, the corresponding HR patches in the database will be used in the learning. In current binary weighting scheme such as [4], all the selected HR patches will be assigned equal weight. In other words, the binary scheme considers all the shortlisted HR patches to be fully and equally relevant while all the other unselected patches to be totally irrelevant. This approach, clearly, is non-ideal as it tends to include outliers as part of the learning patches. Further, the approach tends to focus on the local matching and ignore the global information, namely which character template is the current patch most likely to belong to.

In this work, the normalized cross-correlation score ρ is used to determine the soft weights. The top three character templates based on ρ will be assigned 90% of total weights (i.e. the combined weights of the top three categories will be 0.9), while the rest of the categories will be assigned 10% of total weights (i.e. the combined weights of the rest of the categories will be 0.1). The relative weighting within each of these two grouping is

Table 1 Summary of CG optimization for HR image reconstruction

Initialization:

$$p_o = r_o, \text{ where } r_o = b - A\hat{f}_o$$

Step 1. At the k -th iteration

Update the k -th iteration conjugate vector:

$$p_k = r_k + \beta_k p_{k-1}, \text{ where } \beta_k = \frac{r_k^T r_k}{r_{k-1}^T r_{k-1}}. \tag{10}$$

Step 2. Update the $(k+1)$ -th iteration HR image:

$$\hat{f}_{k+1} = \hat{f}_k + \alpha_k p_k, \text{ where } \alpha_k = \frac{r_k^T r_k}{p_k^T A p_k} \tag{11}$$

Step 3. Update the $(k+1)$ -th iteration residual:

$$r_{k+1} = r_k - \alpha_k A p_k \tag{12}$$

Step 4. Repeat Steps 1–3 until convergence or a maximum number of iteration is reached.

linearly scaled using the score ρ . The main rationale for this weight assignment scheme is that a misclassified character is typically misclassified into one of the top three categories. For instance, a digit “6” may likely be classified as “6”, “5” or “S”. Thus, the HR patches originating from the top 3 characters templates should be given a large combined weight of 90%, while the rest of the categories will be considered as outliers and assigned a combined weight of 10%.

4.4 HR image reconstruction

After determining the soft learning weights $w_k^{[i,j]} \in [0, 1]$, the HR image can be estimated by minimizing the proposed cost function in (6):

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left(\|g - Wf\|_2^2 + \lambda \sum_{i,j} \sum_{k \in \Omega[i,j]} w_k^{[i,j]} \|R_{[i,j]} f - x_k^{[i,j]}\|_2^2 \right) \tag{8}$$

Table 2 Summary of the proposed method

Initialization:

- a. Initialize estimated image f_o using bi-cubic interpolation
- b. Assign $w_k^{[i,j]} = 1$ for all i, j and k .

1. Extract the overlapped LR patches $g_{[i,j]}$ of sized $n \times n$ from the observed LR image g at location $[i, j]$.
2. Find the corresponding set of $m \times m$ HR patches $x_{[i,j]}$ with respect to the current LR patch $g_{[i,j]}$.
3. Construct the matrix $R_{[i,j]}$ that extract an $m \times m$ block at location $[i, j]$.
4. At the i -th iteration:

- Perform image enhancement and segmentation in Section 4(B)
- Calculate the cross-correlation score ρ in Section 4(C)
- Perform soft weight estimation $w_k^{[i,j]}$ of learning patches based on ρ
- Employ conjugate gradient optimization to estimate the HR image \hat{f} by solving the soft-learning cost function:

$$\hat{\mathbf{f}} = \arg \min_{\mathbf{f}} \left(\|g - Wf\|_2^2 + \lambda \sum_{i,j} \sum_{k \in \Omega[i,j]} w_k^{[i,j]} \|R_{[i,j]} f - x_k^{[i,j]}\|_2^2 \right)$$

5. Repeat step 4 until convergence or a maximum number of iterations is reached.

This is equivalent to solving for \hat{f} using the following linear equation:

$$A\hat{f} = b \tag{9}$$

where $A = W^T W + \lambda \sum_{i,j} \sum_{k \in \Omega[i,j]} w_k^{[i,j]} (R_{[i,j]}^T R_{[i,j]})$ and $b = W^T g + \lambda \sum_{i,j} \sum_{k \in \Omega[i,j]} w_k^{[i,j]} (R_{[i,j]}^T x_{[i,j]})$

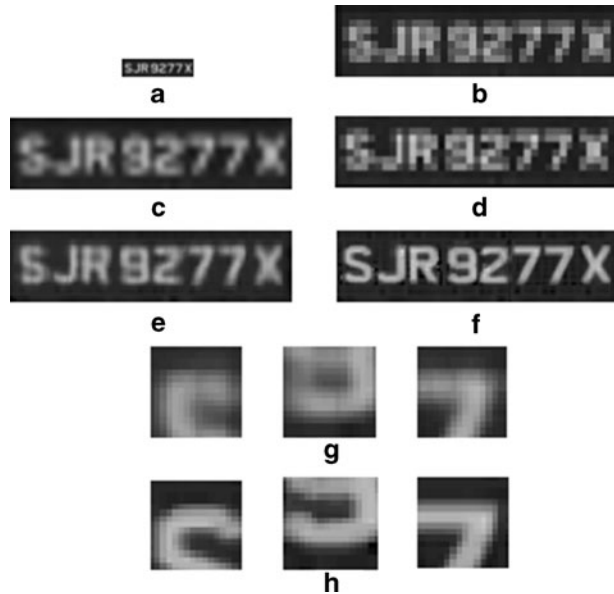
In order to solve the optimization problem in (9), conjugate gradient (CG) optimization is employed. CG utilizes conjugate direction instead of local gradient to search for the minima. Therefore, it can achieve faster convergence when compared with steepest descent method [6]. It also requires less storage requirement and computation complexity when compared with Quasi-Newton method. The mathematical formulation of the HR image reconstruction is derived in Table 1.

The SR algorithm using the soft learning prior is summarized in Table 2.



Fig. 4 Vehicle license plate images

Fig. 5 SR for license plate image. **a** LR image, **b** The scaled-up LR image, **c** Reconstructed image using the bi-cubic interpolation, **d** Reconstructed image using the regularization method in [3], **e** Reconstructed image using binary learning-based method in [4], **f** Reconstructed image using the proposed method, **g** Selected enlarged regions of (e), **h** Selected enlarged regions of (f)



5 Experimental results

In this section, we will demonstrate the performance of the proposed method using both simulated and real-life images. In the patch database, the patch dimension of the LR and HR patches are chosen to be 4×4 and 13×13 respectively. Other patch dimensions can also

Table 3 PSNR of HR image reconstruction at different noise levels

SNR	PSNR			
	Bicubic interpolation	Regularization interpolation	Binary learning-based method	The proposed method
Honda				
45	13.950	15.107	18.787	19.837
35	13.939	15.033	18.753	19.767
25	13.932	15.008	18.560	19.674
Hyundai				
45	16.350	18.779	21.566	22.924
35	16.330	18.681	21.525	22.888
25	16.326	18.641	21.402	22.519
White Toyota				
45	15.018	17.686	20.943	22.622
35	15.015	17.677	20.907	22.481
25	14.978	17.256	20.467	21.703
Black Toyota				
45	14.556	16.744	20.498	22.086
35	14.548	16.646	20.476	21.779
25	14.544	16.611	19.831	21.083

Fig. 6 Image with license plate number ‘SJK5572G’



be chosen. The LR patches are obtained by applying uniform blur on the HR patches, followed by a decimation factor D of 4, and additive white noise.

5.1 HR reconstruction of license images

In the first part of the experiment, the license plate image of the top right image (white Toyota) in Fig. 4 is selected as the test image. To simulate the LR image, the HR image in Fig. 4 is blurred by uniform blur with support of 5×5 . A decimation factor of 4 is chosen and the images are degraded by additive white Gaussian noise (AWGN) to produce a signal-to-noise ratio (SNR) at 35 dB. The proposed algorithm is then run on the LR image. The result is compared with three other methods: bi-cubic interpolation, regularization method in [3] and binary learning-based method in [4]. The reconstructed HR images using the four methods are given in Fig. 5.

The initial LR image and the scaled-up version of it are given in Fig. 5a and b, respectively. The result obtained using the proposed method is given in Fig. 5f. Comparing the result with the original HR image in Fig. 4, it is clear that the proposed algorithm has reconstructed the HR image effectively. The results obtained using the bi-cubic interpolation, the regularization method [3], and the binary learning-based method [4] are given in Fig. 5c–e. It is observed that the bi-cubic interpolation method suffers from significant blurring, and the regularization method produces ragged edges which are unsatisfactory. The enlarged regions of the reconstructed HR images using the method in [4] and the proposed method are given in Fig. 5g and h, respectively. From the figures, it can be seen the proposed method can offer superior reconstruction. This demonstrates that the proposed soft learning framework is better than the binary weighting scheme when performing VLP reconstruction.

Table 4 The license plate recognition results

Iteration number	Recognition results
1	SJK5S7ZG
2	SJK5572G
3	SJK5572G

Fig. 7 Car with license plate number ‘SFJ1848R’



5.2 HR reconstruction under different noise levels

In order to illustrate the effectiveness of the proposed method to handle LR images degraded by different amount of noise, additive noise is added to the test images in Fig. 4 to produce four different sets of LR images with different signal-to-noises (SNRs) at 45, 35 and 25 dB. The same experimental setup as in Section A was used to conduct the experiments. The peak SNR (PSNR) of the reconstructed HR images obtained using the four methods are given in Table 3. From the table, it can be seen that the PSNR of reconstructed image using the proposed method outperforms all the other three methods under different noise levels. The objective performance measures reconfirm the subjective evaluation of the reconstructed images.

Fig. 8 SR for first real-life license plate. **a** The scaled-up LR image, **b** Reconstructed image using the bi-cubic interpolation, **c** Reconstructed image using the regularization method in [3], **d** Reconstructed image using the binary learning-based method in [4], **e** Reconstructed image using the proposed method, **f** Selected enlarged regions of **d**, **g** Selected enlarged regions of **e**

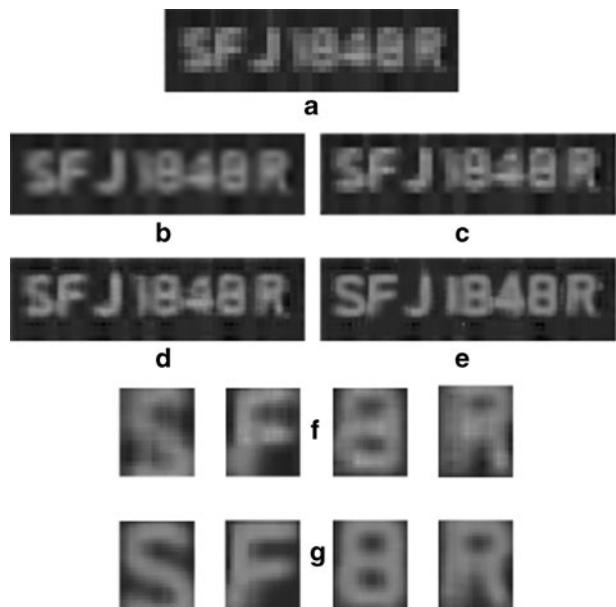


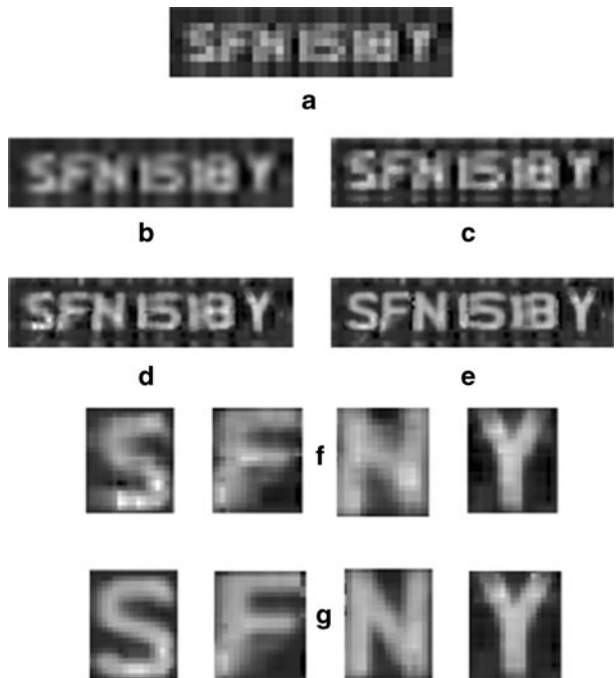
Fig. 9 Car with plate number ‘SFN1518Y’



5.3 Iterative HR reconstruction and VLP recognition

In this section, we will demonstrate the advantages of iterative HR reconstruction and VLP recognition. The VLP image in Fig. 6 is used to evaluate the performance of the proposed algorithm. It can be seen from the image that VLP number cannot be recognized readily due to the low resolution of the image. The proposed iterative process of HR reconstruction and OCR is run on the image. The result of the OCR after each iteration is given in Table 4. From the table, it can be seen the initial recognition contains two errors. This is because the actual characters “5” and “2” are similar to the misclassified “S” and “Z”, respectively. However, as the iteration progresses, the quality of the HR image starts to improve, giving

Fig. 10 SR for second real-life license plate. **a** The scaled-up LR image, **b** Reconstructed image using the bi-cubic interpolation, **c** Reconstructed image using the regularization method in [3], **d** Reconstructed image using the binary learning-based method in [4], **e** Reconstructed image using the proposed method, **f** Selected enlarged regions of (**d**), **g** Selected enlarged regions of (**e**)



rise to better recognition performance, as shown in iterations 2 and 3. This clearly demonstrates the advantages of using OCR to guide HR reconstruction, which in turn is used to improve the performance of OCR iteratively.

5.4 Real-life license plate HR reconstruction

In this section, we will conduct HR reconstruction of real-life license plate images. Two VLP images are captured using a digital camera, as given in Figs. 7 and 9. The size of the reconstructed HR images is set to be four times that of the captured license plate images. The reconstructed HR images using the proposed method are given in Figs. 8 and 10. It can be seen from the figures that the proposed method has achieved satisfactory HR reconstruction where the overall clarity and sharpness of the license plates has been recovered. Significant amount of details has also been restored near the edges. When compared with the regularization method in [3] and binary learning-based method in [4], it is clear that the proposed approach is superior in handling real-life VLP SR reconstruction. The enlarged regions of images obtained using the proposed method also show that it manages to achieve better performance than the binary learning-based method in [4].

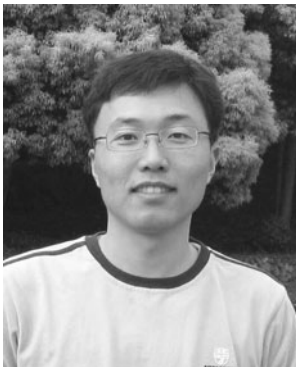
6 Conclusion

This paper presents a new framework using a soft learning approach in single frame VLP super-resolution. As opposed to the conventional binary learning-based techniques, the method estimates the importance of the learning patches through the relevance scores obtained from the OCR. These are then incorporated as a soft prior into a new cost function. The iterative framework of HR reconstruction and VLP recognition enables the quality of the image SR and VLP recognition to be improved progressively. Experimental results show that the proposed method is effective in achieving good HR license plate reconstruction.

References

1. Baker S, Kanade T (2002) Limits on super-resolution and how to break them. *IEEE Trans Pattern Anal Mach Intell* 24:1167–1183
2. Blu T, Thevenaz P, Unser M (2004) Linear interpolation revitalized. *IEEE Trans Image Process* 13:710–719
3. Chen L, Yap K.-H (2005) Regularized interpolation using kronecher product for still images. *IEEE International Conference on image processing*, pp. 1014–1017, Sept
4. Datsenko D, Elad M (2007) Example-based single document image super-resolution: a global map approach with outlier rejection. *J Math Signal Process* 18:103–121
5. El-Khamy SE, Hadhoud MM, Dessouky MI, Salam BM et al (2005) Efficient implementation of image interpolation as an inverse problem. *Digital Signal Process* 15:137–152
6. Farsiu S, Robinson MD, Elad M, Milanfar P (2004) Fast and robust multiframe super resolution. *IEEE trans Image Process* 13:1327–1344
7. Freeman WT, Jones TR, Pasztor EC (2002) Example based super-resolution. *IEEE Comput Graph Appl* 22:56–65
8. He Y, Yap K.-H, Chen L, Chau L.-P (2007) A nonlinear least square technique for simultaneous image registration and super-resolution. *IEEE Transactions on Image Processing*, vol. 16, pp. 2830–2841
9. Hong KP, Paik JK, Kim HJ, Lee CH (1996) An edge-preserving image interpolation system for a digital camcorder. *IEEE Trans Consum Electron* 42:279–284

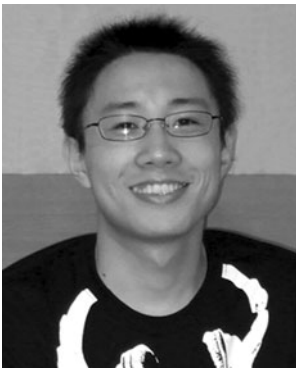
10. Hwang JW, Lee HS (2004) Adaptive image interpolation based on local gradient features. *IEEE Signal Process Lett* 11:359–362
11. Kang MG, Katsagelos AK (1992) Simultaneous iterative image restoration and evaluation of the regularization parameter. *IEEE Trans Signal Process* 40:2329–2334
12. Kazemi FM, Samadi S, Poorreza HR, Akbarzadeh-T M.-R (2007) Vehicle Recognition Using Curvelet Transform and SVM. *Fourth International Conference on Information Technology*, pp. 516–521
13. Li X, Orchard MT (2001) New edge-directed interpolation. *IEEE Trans Image Process* 10:1521–1527
14. Lin Z, Shum HY (2004) Fundamental limits of reconstruction-based super-resolution algorithms under local translation. *IEEE Trans Pattern Anal Mach Intell* 26:83–97
15. Nijhuis JAG, Ter Brugge MH, Helmholt KA, Pluim JPW et al (1995) Car license plate recognition with neural networks and fuzzy logic. *IEEE International Conference on Neural Networks*, pp. 2232–2236
16. Pu H, Wei H, Dong-Feng W, Yong-Jie Z (2003) Car license plate feature extraction and recognition based on multi-stage classifier. *International Conference on Machine Learning and Cybernetics*, pp. 128–132
17. Ramponi G (1999) Warped distance for space-variant linear image interpolation. *IEEE Trans Image Process* 8:629–639
18. Rosenfeld A, Kak CA (1982) *Digital picture processing*, 2nd edn. Academic Press, London
19. Shyang-Lih C, Li-Shien C, Yun-Chung C, Sei-Wan C (2004) Automatic license plate recognition. *IEEE Trans Intell Transp Syst* 5:42–53
20. Suresh KV, Rajagopalan AN (2006) A discontinuity adaptive method for super-resolution of license plates. *Graphics and Image Processing Computer Vision*, vol. 4338/2006, pp. 24–34
21. Yap K-H, He Y, Tian Y, Chau L-P (2009) A nonlinear L1-norm approach for joint image registration and super-resolution. *IEEE Signal Process Lett* 16:981–984
22. Zhang Y, Zhang C (2003) A new algorithm for character segmentation of license plate. *IEEE International Conference on Intelligent Vehicles Symposium*, pp. 106–109



Yushuang Tian received the B.Eng. degree in information engineering from Zhejiang University, Hangzhou, China, in 2005, and the M.Sc. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2007. After that he worked as a Research and Development Engineer in the Huawei Technologies Co. Ltd. for one year. He is currently a Ph.D student at Nanyang Technological University, Singapore. His research interests include image/video processing, computer vision, and image/video super-resolution.



Kim-Hui Yap (S'99-M'03) received the B.Eng. and Ph.D. degrees in electrical engineering from the University of Sydney, Sydney, Australia, in 1998 and 2002, respectively. Currently, he is a faculty member at Nanyang Technological University, Singapore. His main research interests include image/video processing, media content analysis, computer vision, and computational intelligence. He has worked in projects on image processing and computer vision including image restoration, segmentation, super-resolution and reconstruction. He is also the Group Leader in Content-based Analysis for the Center for Signal Processing, Nanyang Technological University, Singapore. His works cover image/video content analysis and understanding, media processing, image/video indexing, retrieval, and summarization. He has numerous publications in various international journals, book chapters and conference proceedings. He is also the editor of a book entitled "Intelligent Multimedia Processing with Soft Computing by Springer-Verlag" in 2005.



Yu He received the B.Eng. degree in precision instrument and optoelectronics engineering from Tianjin University, Tianjin, China, in 2002, and the Ph.D. degree in information engineering from Nanyang Technological University, Singapore, in 2010.