

Actual License Plate Images Clarity Classification via Sparse Representation

Yudong Cheng, Feng Liu^(✉), Zongliang Gan, and Ziguan Cui

Jiangsu Key Laboratory of Image Processing and Image Communication,
Nanjing University of Posts and Telecommunications,
Nanjing, China
liuf@njupt.edu.cn

Abstract. The quality of the license plate image has a great influence on the license plate recognition algorithm. Predicting the clarity of license plate image in advance will help the license plate recognition algorithm set appropriate parameters to improve the accuracy of the recognition. In this paper, we propose a classification algorithm based on sparse representation and reconstruction error to divide license plate images into two categories: high-clarity and low-clarity. We produced over complete dictionaries of both two categories, and extract the reconstruction error of the license plate image that to be classified through the two dictionaries as the feature vector. Finally we send the feature vector to SVM classifier. Our Algorithm is tested by the license plate image database, reaching over 90% accuracy.

Keywords: Blind image quality assessment · Licenses plate
Dictionary learning · Sparse representation · Feature extraction

1 Introduction

In recent years, many problems in the industry, if supplemented by high-performance quality evaluation to do the pretreatment, the system will be stable and efficient. License plate image recognition system is the case. License plate recognition system (LPRS) which currently developing gradually, is one of the most popular research direction. LPRS is generally divided into the following sections [1]: license plate location, license plate segmentation, and license plate recognition (LPR) which is the main part of LPRS. At present, there are many LPR algorithms with the combination of the neural network, and the training samples of the algorithm has a very high demand [2–4]. If the license plate image training samples can be classified according to the clarity, thus making different network models, the final accuracy of LPR algorithm could be improved. Based on this, we propose a classification algorithm, dividing the license plate image into high-clarity and low-clarity two categories, to assist the LPRS.

2 Relate Work

The clarity classification of license plate image is a brand new problem in the study of non-reference image quality assessment (NRIQA), and there are no research papers before. Compared with the general NRIQA algorithms, the clarity classification algorithm of the license plate images does not evaluate the quality score, but the image is classified by clarity according to the industrial demand. However, both are essentially the features that can accurately describe the image quality. Therefore, the recent study of NRIQA algorithm can give us a lot of help.

Early NRIQA algorithms generally assume the types of distortion model that affect the image quality are known [5–12]. Based on the presumed distortion types, these approaches extract distortion specific features to predict the quality. However, there are more types of image distortions in reality, and distortions may affect each other. The assumption limits the application of these methods.

Recent numbers of studies on NRIQA take similar architecture. First in the training stage, the data containing the distorted image and the associated subjective evaluation are trained [13] to extract feature vectors. Then a regression model is learned to map these feature vectors to subjective human scores. In the test stage, feature vectors are extracted from test images, and then send into the regression model to predict its quality model [14–18]. The strategy to classify the quality of license plate images is almost the same, but a classification model will replace the regression model.

Moorthy and Bovik proposed a two-step framework of BIQA, called BIQI [14]. Scene statistics that extracted from given distorted image firstly, are used to decide the distortion type this image belongs to. The statistics are also used to evaluate the quality of the distortion type. By the same strategy, Moorthy and Bovik proposed DIIVINE, which use a richer natural scene feature [15]. However, completely different to the actual application, both BIQI and DIIVINE assume that the distortion type in the test image is represented in the training data set.

Saad et al. assumed that the statistics of the DCT characteristics can be varied in a predictable manner as the image quality changes [16]. According to it, a probability model called BLIINDS is trained by the contrast and structural features extracted in the DCT domain. BLIINDS is extended to BLIINDS-II by using more complex NSS-based DCT functionality [17]. Another approach not only extract the DCT feature, but also the wavelet and curvelet [18]. Although this model achieve certain effects in different distortion types, it still cannot work on every type.

As there is a big difference between general natural images and license plate (LP) images, the general NRIQA pursuing universal can not be used on LP images directly. Statistical information of LP image, such as DCT domain, wavelet domain and gradient statistics, not only affected by the distortion, also affected by the LP characters. This makes it difficult for the general NRIQA algorithm to work on the LP image. However, the advantage of the LP image is that the image does not appear in addition to the license plate characters other than the image. This very useful priori information can help us solve the problem of clarity classification of LP images. Base on this, we can produce different large enough over complete dictionaries to represent all types of LP images and to classify them by extracting the LP images from different dictionary reconstruction errors as valid features.

3 Framework of License Plate Image Classification

Here are the six uniform size gray-scale LP images in Fig. 1. We can see that for the distinction between the obvious high-clarity images and low-clarity, their DCT transformation are no rules to follow, gradient statistics also mixed together indistinguishable, which makes the vast majority of known NRIQA algorithms cannot effectively extract the feature of the LP image quality can be described.

Here, we propose a method based on sparse representation algorithm to extract the appropriate feature vector, and then put the feature vector into the appropriate classification model.

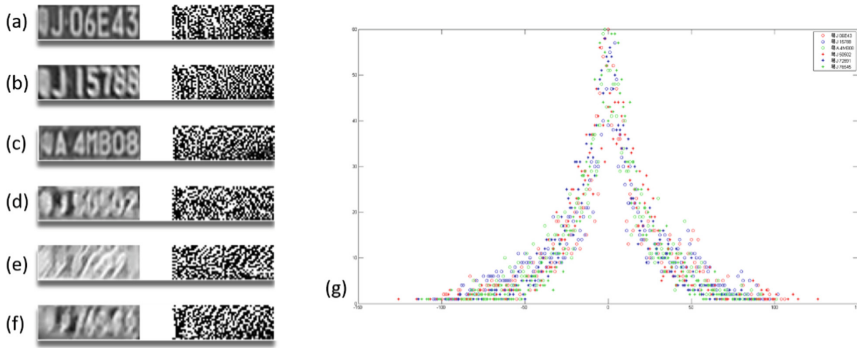


Fig. 1. Different clarity LP images and their DCT transform (a) (b) (c) are high-clarity images (d) (e) (f) are low-clarity images (g) is their gradient histogram

3.1 Classification Principle

Familiar with human subjective scoring that often used in general NRIQA algorithms as the final output, the clarity classification principle for LP images also needs to be artificially developed. In this algorithm, LP images are divided into two categories:

1. High-clarity LP image: the LP image that can be recognized all the last five characters by human.
2. Low-clarity LP image: the LP image that cannot be recognized all the last five characters by human.

Successful recognition of all the last five characters or not is one of the most intuitive manifestations of human eye's assessment. It is worth noting that we believe that the high/low clarity binary classification of images is a prerequisite for faithfully reflecting the subjective evaluation of image quality. Likewise, if the feature we extract can accurately distinguish between high and low clarity images, this feature can also be used to describe the image's quality score. Therefore, the classification principle that we propose is not unique. LP images could be classified into different categories by different principles.

3.2 The Algorithm of Feature Extraction

Sparse representation-based algorithms have been widely applied to computer vision and image processing like image denoising [19] and face recognition [20], especially the famous sparse representation-based classifier [21]. We follow this idea and describe in detail below.

Given sufficient samples of both high-clarity LP images and low, $A_1 = [v_{1,1}, v_{1,2}, \dots, v_{1,n_1}] \in R^{m \times n_1}$, $A_2 = [v_{2,1}, v_{2,2}, \dots, v_{2,n_2}] \in R^{m \times n_2}$, and test sample $y \in R^m$ from the same class would approximately lie in the subspace spanned by the training samples associated with either class:

$$y = a_{1,1} * v_{1,1} + a_{1,2} * v_{1,2} + \dots + a_{1,n_1} * v_{1,n_1} \quad (1)$$

$$y = a_{2,1} * v_{2,1} + a_{2,2} * v_{2,2} + \dots + a_{2,n_2} * v_{2,n_2} \quad (2)$$

and $a_{i,j} \in \mathbb{R}$, $j = 1, 2, \dots, n_i$ is a scalar.

Then we form a dictionary \mathcal{D} by grouping all the samples from both classes.

$$\mathcal{D} = [A_1, A_2] = [v_{1,1}, v_{1,2}, \dots, v_{1,n_1}, v_{2,1}, \dots, v_{2,n_2}] \quad (3)$$

and the linear representation of y can be written as:

$$y = \mathcal{D} * x_0 = a_{1,1} * v_{1,1} + a_{1,2} * v_{1,2} + \dots + a_{2,n_2} * v_{2,n_2} \quad (4)$$

here x_0 is a coefficient vector whose entries are zero except those associated with the first class or second.

So to determine the class of clarity a test LP image is, it would be reconstructed and extracted the reconstruct error:

$$y_1 = \mathcal{D} * \alpha_0(1) \quad (5)$$

$$y_2 = \mathcal{D} * \alpha_0(2) \quad (6)$$

$$e(1) = \|y_{test} - y_1\|_2 \quad (7)$$

$$e(2) = \|y_{test} - y_2\|_2 \quad (8)$$

By compare these two error, the small one would be the clarity class the test image belongs to.

To learn the over-complete dictionary, K-SVD algorithm [22] would be used to solve the next optimization problem:

$$\langle \mathcal{D}, \alpha \rangle = \operatorname{argmin}_{\mathcal{D}, \alpha} \|Y - \mathcal{D} * \alpha\|_2 \text{ s.t. } \|\alpha\|_0 \leq L \quad (9)$$

However, the hypothesis of SRC algorithm is too strong to classify clarity of LP image directly. One of the prerequisites for the SRC algorithm is that the subspace of different classes are the same. To verify this, we create over complete dictionaries of the same parameters for the high/low clarity LP image training samples $D_1 D_2$.

$$\langle D_1, \alpha_1 \rangle = \operatorname{argmin}_{D_1, \alpha_1} \|Y_1 - D_1 * \alpha_1\|_2 \text{ s.t. } \|\alpha_1\|_0 \leq L \quad (10)$$

$$\langle D_2, \alpha_2 \rangle = \operatorname{argmin}_{D_2, \alpha_2} \|Y_2 - D_2 * \alpha_2\|_2 \text{ s.t. } \|\alpha_2\|_0 \leq L \quad (11)$$

In the process of making the dictionary, we found that for different image segmentation size, training error of dictionary produced by high-clarity image samples $E_1 = \|Y_1 - D_1 * \alpha_{1, \text{training}}\|_2$ is always larger than the error of low-clarity dictionary E_2 as shown in Fig. 2.

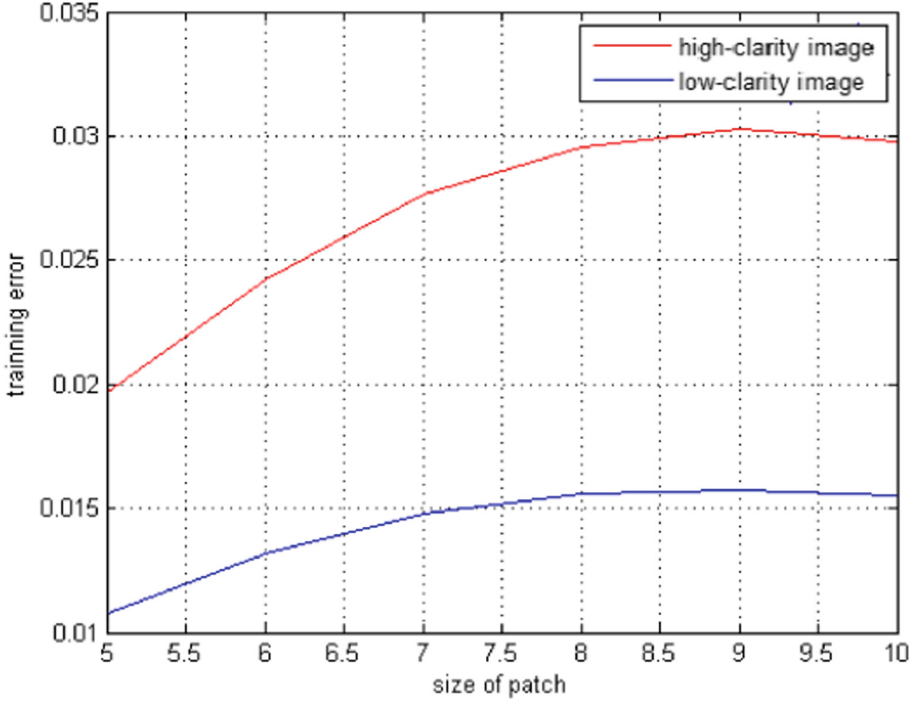


Fig. 2. Training error of different clarity LP images.

This means that high-clarity LP images can provide more information and are not more completely represented than low-clarity LP images that provide lower information. Although the frequency domain features that reflect the amount of LP image information cannot be effectively extracted, the same indirect reconstruction error can also represent the amount of information of the LP image.

The experimental results show that high-clarity LP images require more dictionary atoms to represent, and the dictionary's reconstruction error is related to the amount of information contained in the image. Although the SRC algorithm directly to the reconstruction error comparison method is not effective, the reconstruction error can still be extracted as LP image quality statistical feature.

So here we form two over complete dictionaries by high/low-clarity LP images with the same parameters. Then we get the reconstruction error $e(1)$ and $e(2)$ by Eqs. 5–8 as a two-dimensional feature vector of LP image.

In this algorithm, we utilize a support vector machine (SVM) for classification. While the feature vector was extracted, any classifier can be chosen to map it onto classes. The choice of SVM was motivated by the well performance [23] (Fig. 3).

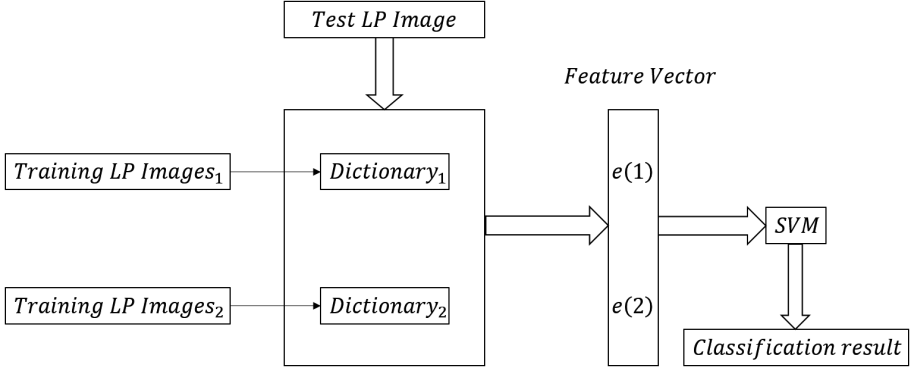


Fig. 3. Algorithm flowchart

4 Experiments and Analysis

4.1 Database Establishment

First of all, there is no public database of clarity classified LP images, so we have to build one. The size of LP images from the surveillance video should be uniformed to $20 * 60$ pixels. Then, all the LP images will be evaluated artificially through the principle above. Based on this principle, we establishment a LP image database containing 500 high-clarity images and 500 low-clarity images. Low clarity LP images include a variety of common distortion types like motion blur, Gaussian blur, defocus blur, compression and noise.

4.2 Algorithm Performance and Analysis

There are a few parameters in the algorithm. We set L (the sparse prior in Eq. 9) to 2, so that we can get obvious reconstruction error with a small sparse prior. The kernel used in the SVM part is the radial basis function (RBF) kernel, whose parameters estimated by using cross-validation. In order to prevent over-fitting, we take random 30% of both clarity LP images, divide them into pieces with the size of $n * n$ by sliding, and use the largest variance of 60% of each LP image to learn the dictionary. The other 70% images used as test samples. The figure reported here are the correct rate of classification by different algorithms.

We test our algorithm on the database we made above. Since we were the first to study this problem, there was no comparison of clarity classification algorithms for LP images. Here we test the performance of No-reference PSNR [24], BIQI [15], NIQE [14] and SSEQ [25]. We record the assessment score of No-reference PSNR, NIQE, SSEQ for the random selection of 150 high-clarity LP images and 150 low-clarity LP images in the database (Figs. 4, 5 and 6).

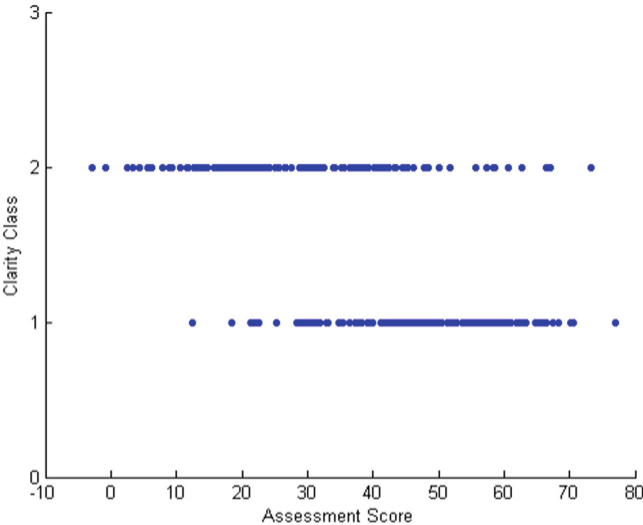


Fig. 4. Assessment score of no-reference PSNR

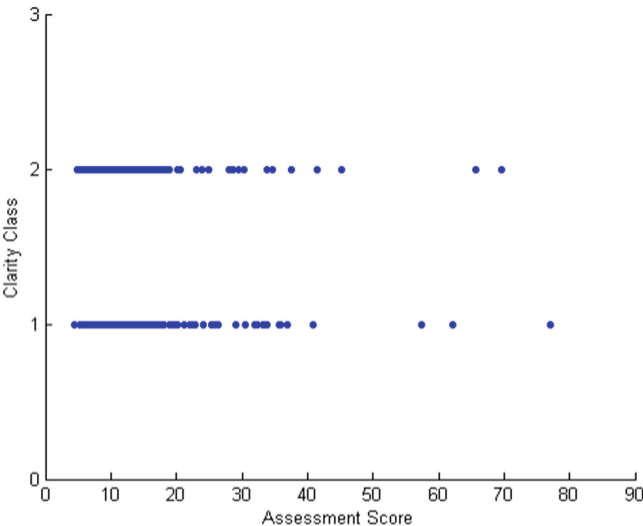


Fig. 5. Assessment score of NIQE

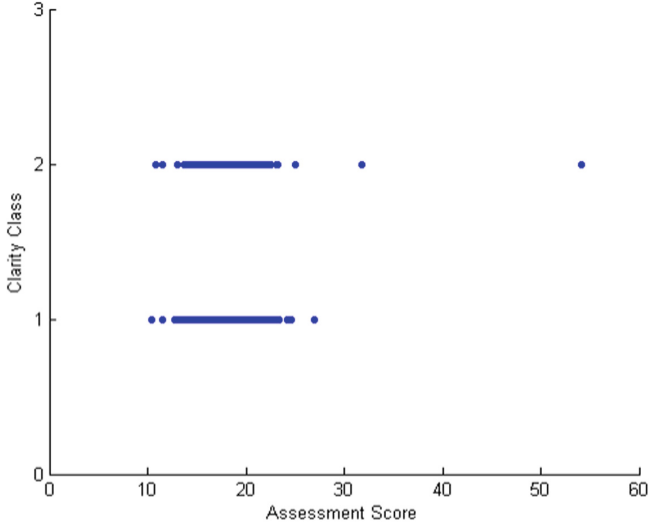


Fig. 6. Assessment score of SSEQ

It can be seen that these assessment scores are very close to the different LP image clarity, and cannot effectively distinguish between high-clarity images and low. Of course, these algorithms cannot evaluate the LP image quality.

For the BIQI algorithm, we tested the algorithm in another way. We did not directly calculate the assessment score of LP images, but sent the feature vector extracted from 9 wavelet transform into the SVM classifier. We compared this feature with ours (Table 1):

Table 1. Classification performance of algorithms on two clarity LP image database

		Class 1: high-clarity	Class 2: low-clarity
BIQI		38.00%	63.00%
Our algorithm	Patch size: 5	68.80%	95.80%
	Patch size: 6	81.80%	97.20%
	Patch size: 7	86.40%	99.20%
	Patch size: 8	88.20%	97.20%
	Patch size: 9	88.80%	97.80%
	Patch size: 10	94.40%	98.00%

It can be seen that the features extracted by BIQI also do not apply to the quality of LP image.

The result shows that our algorithm performs well especially with the large size of patch the LP image divided. Due to the lack of relevant databases and algorithms, we did not have more experiments. Based on SRC, our algorithm considered the

expression ability of different dictionaries. By extracting different reconstruction error, this classification algorithm could be extended to the assessment of LP image quality. Familiar with the information extracted from DCT domain and gradient, the reconstruction error could also be used as feature in regression model to map to the associated human subjective scores.

5 Conclusion

We have proposed a well-performed clarity classification method that improves on the database we built. Not the same as traditional ones, this model extract the different reconstruction error as feature. This method can be widely applied to the quality assessment algorithm for fixed category objects. First determine the image category, and then determine the image quality, this process in more in line with human visual perception.

Acknowledgment. This research was supported in part by the National Nature Science Foundation, P.R. China (Nos. 61471201, 61501260), Jiangsu Province Universities Natural Science Research Key Grant Project (No. 13KJA510004), the Six Kinds Peak Talents Plan Project of Jiangsu Province (2014-DZXX-008), and the 1311 Talent Plan of NUPT.

References

1. Sharma, J., Mishra, A., Saxena, K., et al.: A hybrid technique for license plate recognition based on feature selection of wavelet transform and artificial neural network. In: 2014 International Conference on Optimization, Reliability, and Information Technology (ICROIT), pp. 347–352. IEEE (2014)
2. Nagare, A.P.: License plate character recognition system using neural network. *Int. J. Comput. Appl.* **25**(10), 36–39 (2011). ISSN 0975-8887
3. Akoum, A., Daya, B., Chauvet, P.: Two neural networks for license number plate recognition. *J. Theoret. Appl. Inf. Technol.* (2005–2009)
4. Masood, S.Z., Shu, G., Dehghan, A., et al.: License plate detection and recognition using deeply learned convolutional neural networks. *arXiv preprint [arXiv:1703.07330](https://arxiv.org/abs/1703.07330)* (2017)
5. Ferzli, R., Karam, L.J.: A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). *IEEE Trans. Image Process.* **18**(4), 717–728 (2009)
6. Narvekar, N.D., Karam, L.J.: A no-reference perceptual image sharpness metric based on a cumulative probability of blur detection. In: *Proceedings of the IEEE International Workshop on Quality Multimedia Experience*, July 2009, pp. 87–91 (2009)
7. Varadarajan, S., Karam, L.J.: An improved perception-based no-reference objective image sharpness metric using iterative edge refinement. In: *Proceedings of the IEEE International Conference on Image Processing*, October 2008, pp. 401–404 (2008)
8. Sadaka, N.G., Karam, L.J., Ferzli, R., Abousleman, G.P.: A no-reference perceptual image sharpness metric based on saliency-weighted foveal pooling. In: *Proceedings of the IEEE International Conference on Image Processing*, October 2008, pp. 369–372 (2008)
9. Sheikh, H.R., Bovik, A.C., Cormack, L.K.: No-reference quality assessment using natural scene statistics: JPEG2000. *IEEE Trans. Image Process.* **14**(11), 1918–1927 (2005)

10. Chen, J., Zhang, Y., Liang, L., Ma, S., Wang, R., Gao, W.: A no-reference blocking artifacts metric using selective gradient and plainness measures. In: Huang, Y.-M.R., Xu, C., Cheng, K.-S., Yang, J.-F.K., Swamy, M.N.S., Li, S., Ding, J.-W. (eds.) PCM 2008. LNCS, vol. 5353, pp. 894–897. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89796-5_108
11. Suthaharan, S.: No-reference visually significant blocking artifact metric for natural scene images. *J. Signal Process.* **89**(8), 1647–1652 (2009)
12. Barland, R., Saadane, A.: Reference free quality metric using a region-based attention model for JPEG-2000 compressed images. In: *Proceedings of SPIE*, vol. 6059, pp. 605905-1–605905-10, January 2006
13. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a ‘completely blind’ image quality analyzer. *IEEE Signal Process. Lett.* **20**(3), 209–212 (2013)
14. Moorthy, A.K., Bovik, A.C.: A two-step framework for constructing blind image quality indices. *IEEE Signal Process. Lett.* **17**(5), 513–516 (2010)
15. Moorthy, A.K., Bovik, A.C.: Blind image quality assessment: from natural scene statistics to perceptual quality. *IEEE Trans. Image Process.* **20**(12), 3350–3364 (2011)
16. Saad, M.A., Bovik, A.C., Charrier, C.: A DCT statistics-based blind image quality index. *IEEE Signal Process. Lett.* **17**(6), 583–586 (2010)
17. Saad, M.A., Bovik, A.C., Charrier, C.: Blind image quality assessment: a natural scene statistics approach in the DCT domain. *IEEE Trans. Image Process.* **21**(8), 3339–3352 (2012)
18. Shen, J., Li, Q., Erlebacher, G.: Hybrid no-reference natural image quality assessment of noisy, blurry, JPEG2000, and JPEG images. *IEEE Trans. Image Process.* **20**(8), 2089–2098 (2011)
19. Elad, M., Aharon, M.: Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006)
20. Zhang, Q., Li, B.: Discriminative K-SVD for dictionary learning in face recognition. In: *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE (2010)
21. Wright, J., et al.: Robust face recognition via sparse representation. *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009)
22. Aharon, M., Elad, M., Bruckstein, A.: K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **54**(11), 4311–4322 (2006)
23. Burges, C.: A tutorial on support vector machines for pattern recognition. *Data Min. Knowl. Discov.* **2**(2), 121–167 (1998)
24. Wang, Z., Sheikh, H.R., Bovik, A.C.: No-reference perceptual quality assessment of JPEG compressed images. In: *Proceedings of the 2002 International Conference on Image Processing*, vol. 1, p. I. IEEE (2002)
25. Liu, L., Liu, B., Huang, H., et al.: No-reference image quality assessment based on spatial and spectral entropies. *Signal Process. Image Commun.* **29**(8), 856–863 (2014)

Image and Graphics

9th International Conference, ICIG 2017, Shanghai,
China, September 13-15, 2017, Revised Selected
Papers, Part II

Zhao, Y.; Kong, X.; Taubmann, D. (Eds.)

2017, XVII, 629 p. 268 illus., Softcover

ISBN: 978-3-319-71588-9