

# A Single Image Compression Framework Combined with Sparse Representation-Based Super-Resolution

He Xiaohai, He Jingbo, Huang Jianqiu

College of Electronics and Information Engineering  
Sichuan University  
Chengdu, China  
nic5602@scu.edu.cn

Wu Di

Information Engineering Department  
Police officer college of CAPF  
Chengdu, China  
wudi1945@163.com

**Abstract**—This paper addresses the problem of single image compression by using sparse representation-based super-resolution modeling. The proposed single image compression framework (SICF) is mainly composed of two components, including image encoder and decoder. Given an image, the framework takes down-sampling processing. Then, the down-sampling image is fed into JPEG and reconstructed by the sparse representation-based super-resolution, which is using the proposed targeted dictionary based on soft-decision adaptive interpolation (TD-SAI) and targeted residual dictionary (TRD). For further improving the quality of the final decoded image, the feedback is designed to obtain residual assistance image which can be compensated for the loss of high-frequency details in process of super-resolution reconstruction. Supported by the targeted dictionary, our method achieved an improvement in objective quality assessment and competitive performance in visual quality when comparing with other SR algorithms. More importantly, the proposed SICF achieved significant bit-rate saving under the same PSNR compared with JPEG standard. Extensive experiments manifest the viability and efficiency of the proposed image compression framework.

**Keywords**—image compression; targeted dictionary training; super-resolution

## I. INTRODUCTION

With the accelerative growth of computer technology and high-performance display devices, it is desirable to develop efficient image processing techniques for acquiring high-quality image with limited bit-rate. Under the circumstance, the combination of super-resolution reconstruction and conventional image compression has become the hot topic among the scholars.

In the past several years, the correlational research in video has acquired some continuous development due to the existence of great correlation and redundant information between two adjacent frames in the video sequence. [1] proposes a different framework of down-sampling based coding with super resolution technique, where a super-resolution technique is employed to restore the downsampled frames to their original resolutions. [2] extends a multi-

resolution approach to example-based super-resolution and discuss codebook construction for video sequences. In [3], the authors apply the example-based SR algorithm to restore the down-sampled frames. [4] adopts a novel compression algorithm using selective key-frame detection to encode video and patch-based super-resolution to decode.

By contrast, the corresponding research developed slowly in the field of single image, for which it is hard to break through the limitation of single image and provide enough priori information for reconstruction. However, sparsity-based regularization has been achieved great success in the past several years. We refer the reader to the following papers [5-10]. The sparse model assumes that image patches can be well represented as a sparse linear combination of elements chosen from over-complete dictionary, which ensures that linear relationships among high-resolution signals can be recovered from the corresponding low-dimensional projections. [5-7] extend the approach to learn dictionaries for coupled feature spaces by using single sparse coding to learn the coupled dictionaries in concatenated spaces. In [8], the authors propose an in-scale self-learning framework for single image super resolution, which advance the learning of support vector regression (SVR) and image sparse representation in super resolution method. The paper[10] adopts a beta process joint dictionary learning method (BP-JDL) for solving the dictionary learning problem in coupled feature spaces, which a Bayesian method using a beta process prior is applied to learn the over-complete dictionaries.

Inspired by above mentioned approach, we propose a single image compression framework combining the conventional image compression with sparse representation-based super-resolution algorithm. Aiming at the specific process of proposed image compression framework, we put forward the targeted dictionary based on soft-decision adaptive interpolation (TD-SAI) and the targeted residual dictionary (TRD), which the training process is directed at the corresponding design of the framework. The targeted training contributes significantly to the super resolution.

The remainder of this paper is organized as follows. Section II elaborates the design of our proposed single image

---

This work is supported by National Natural Science Foundation of China (No.61471248).

compression framework (SICF). In section III, we discuss how to train the targeted dictionary for single-image sparse representation-based super-resolution. Extensive experimental results are reported and discussed in section IV. Finally, section V concludes the paper.

## II. SINGLE IMAGE COMPRESSION FRAMEWORK

In this section, we give the detail description of our proposed single image compression framework (SICF). The image compression framework consists of two significant components, including image encoder and decoder. By applying our proposed targeted dictionary based on soft-decision adaptive interpolation (TD-SAI) and target residual dictionary training (TRD) to the present framework, we achieve a great improvement in image compression efficiency which benefits from the combination of sparse representations-based super-resolution and image compression processing.

### A. Encoder of the Framework

For the convenience of expression, we will take the 2:1 downsampling processing as the standard of our framework in this paper. The sketch map of image encoder in proposed framework is illustrated in Fig. 1. From this sketch map, it is clear to see the production of two components in image compression encoder. The two different encoded images will have different effects in next processing of targeted dictionary reconstruction.

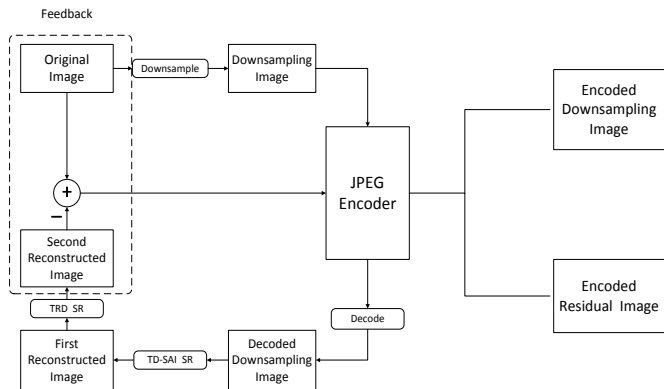


Figure 1. Image encoder of the proposed single image compression framework (SICF)

Firstly, given an image, the proposed framework takes down-sampling processing and then the down-sampling image is fed into JPEG. Numerous experimental results demonstrate the necessity of the denoising process to the decompressed low resolution image. Owing to the JPEG encoder, the encoded downsampling image has quantization noise which can be reduced by BM3D [11] algorithm. Consequently, before the process of super-resolution reconstruction, the framework will adopt BM3D-denoising to handle the encoded downsampling image in the case of different quantitative parameters.

The decoder first decompresses the low resolution downsampling image and then up-converts it to the original resolution by the sparse representation-based super-resolution

using targeted dictionary based on soft-decision adaptive interpolation (TD-SAI). To get more high-frequency details for the first reconstructed image, we reconstruct the images with the assistance of the targeted residual dictionary (TRD). As a joint result of the TD-SAI and TRD, the final reconstructed image gets satisfactory visual quality without help of the feedback.

For further improving the visual quality of the final decoded image, the feedback part is particularly designed to acquire the residual assistance image which can be compensated for the loss of high-frequency details in process of super-resolution reconstruction. The quality of decoded image can be improved significantly under limited bit-rate of residual assistance image. The image subtraction between original test image with the second reconstructed image is just feedback of our framework which is also the other important component needed to feed into the JPEG encoder. Consequently, the final encoder of the framework is made up of the encoded downsampling image and encoded residual image.

### B. Decoder of the Framework

The proposed single image compression framework (SICF) is mainly composed of the image encoder and decoder. As shown in Fig.1, it has given a presentation of the sketch map about the image encoder of our framework.

After the above treatment, we acquire two important images, including the encoded downsampling image and the encoded residual image. These two essential images can provide convenience for next processing, which is the sparse representation-based super resolution and final decoded image. As illustrated in Fig.2, the presented image decoder of compression framework is aimed at the image encoder.

It should be noted that in Fig.2, through the image decoder, we obtain not only the decoded downsampling image, but also the decoded residual image. It is obvious to find that decoded downsampling image is the key part for super-resolution reconstruction in framework. In addition, the decoded residual image can generate a complementation of high-frequency detail information for our reconstructed image. On the contrary, if the proposed framework doesn't have the design of the feedback, worse image visual quality and the loss of more image detail information can be observed in the final image.

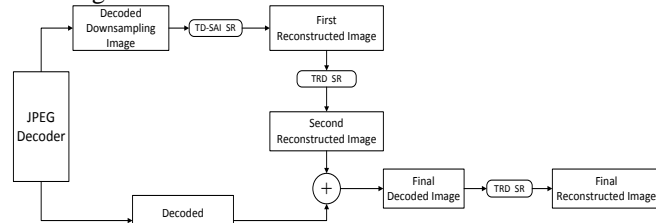


Figure 2. Image decoder of the proposed single image compression framework

In decoder of the framework, we can obtain two corresponding decoded downsampling image and decoded residual image. The targeted dictionary based on soft-decision adaptive interpolation (TD-SAI) is used in reconstruction of

the decoded downsampling image to get the first reconstructed image. With the help of the targeted residual dictionary (TRD), more detail information is added to the first reconstructed image, so that we can acquire more satisfied reconstructed image. Due to the excellent design of feedback in encoder, the final decoded image will be the joint effect of second reconstructed image and the decoded residual image. It must be said that residual image is compressed by JPEG. Thus, the resulting decoded residual image just has limited bit-rate which can be hugely helpful for the bit-rate saving of the framework.

While, in order to obtain a more satisfactory improving both in subjective and objective quality, it is desirable to regain more detail information during the decoded image reconstruction. Consequently, the proposed framework is aiming at the final decoded image to train new targeted residual dictionary and then apply it to the reconstruction for the final decoded image.

Because of the targeted dictionary training, our super-resolution method produce excellent reconstructed results on a variety of images, and meanwhile we also achieve a more competitive performance both in objective and subjective quality comparing with other super-resolution approaches. But above all, the proposed SICF acquires meaningful bit-rate saving under the same PSNR compared with JPEG standard. Owing to the adaptation of JPEG compression standard in image encoder and decoder of proposed framework, so the comparison with the JPEG standard is directly encoded and decoded test image under the same PSNR results.

### III. THE TARGETED DICTIONARY TRAINING FOR SUPER-RESOLUTION

#### A. The Training of Targeted Dictionary Based on Soft-decision Adaptive Interpolation(TD-SAI)

For this proposed SICF, the sparse representation-based super-resolution plays an important role in it. An effective dictionary training method will becomes the core of the super resolution. In this section, we focus on training a specific dictionary pair which directs at the proposed SICF.

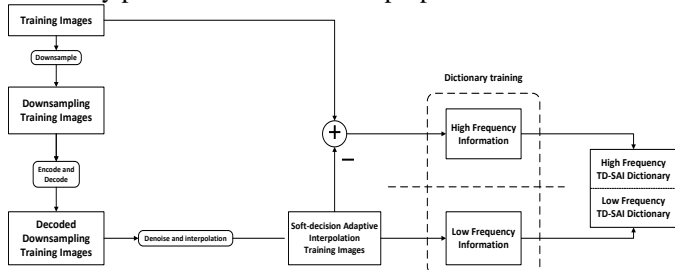


Figure 3. The training of targeted dictionary based on soft-decision adaptive interpolation (TD-SAI)

As illustrated in Figure 3, it shows the training of targeted dictionary based on soft-decision adaptive interpolation (TD-SAI). More details about soft-decision adaptive interpolation can be found in [12]. Comparing with other state-of-the-art SR approaches, it is obviously to realize the superior of our

dictionary training method which aims at the processing of the image compression in our proposed framework.

Firstly, we take the 2:1 downsampling processing for the training images, and then the downsampling training images are encoded and decoded by the JPEG. Owing to the existence of quantization noise, we decided to use the BM3D method for denoising to acquire better image with less quantization noise. The denoising parameter is associated with the quantitative parameters (QP), which corresponding denoising parameter is decreasing with the increasing of quantitative parameters (QP). Then, we up-convert the training images to the original resolution using the soft-decision adaptive interpolation which could be provide more prior information for next dictionary training. A set of high-resolution training images are collected, decoded and denoised low-resolution images are constructed using soft-decision adaptive interpolation and pairs of matching patches that form the training database are extracted. Each of these patch-pairs undergoes a pre-processing stage that removes the low-frequencies from the training images and extracts features from soft-decision adaptive interpolation images. At this moment, a dictionary is trained for the low-resolution patches, such that they can be represented sparsely, and a corresponding dictionary is constructed for the high-resolution patches that it matches the low-resolution one. Finally, we obtain the high-frequency and low-frequency TD-SAI dictionary from this training database. In this paper, we adopt the method published in [7] for the super-resolution reconstruction.

#### B. The Training of Targeted Residual Dictionary(TRD)

In order to further improve performance of our proposed framework, we decided to take full advantage of the residual information. By using the targeted dictionary based on soft-decision adaptive interpolation, we get the first reconstructed training images. The residual high-frequency information, which is subtraction between training images with the first reconstructed training images, is the key of the targeted residual dictionary (TRD).

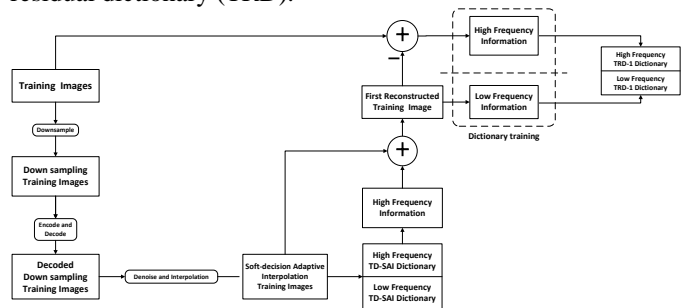


Figure 4. The training of targeted residual dictionary (TRD-1)

As illustrated in Figure 4, the high-resolution training images are collected, first reconstructed images are acquired by the TD-SAI and pairs of matching patches that form the training database are extracted. Each of these patch-pairs undergoes a pre-processing stage that removes the low-frequencies which is just the first reconstructed images from the training images and extracts features from the first reconstructed images. We get the low-frequency and high-frequency targeted residual dictionary from the low and high

frequency information, providing by the first reconstructed training images and original training images. The TRD-1 is used in the encoder of framework to improve the first reconstructed images both objective and subjective, and then we can get the second reconstructed training images by using the TRD-1.

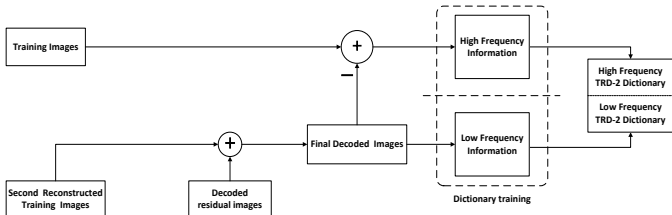


Figure 5. The training of targeted residual dictionary (TRD-2)

As shown in Figure 5, the second reconstructed training image is added to the decoded residual training image from feedback, and then we will acquire the final decoded training image. The high-resolution training images are collected, second reconstructed images are acquired by the TRD-1 and pairs of matching patches that form the training database are extracted. Each of these patch-pairs undergoes a pre-processing stage that removes the low-frequencies which is just the final decoded images from the training images and extracts features from the final decoded images. The TRD-2 in the decoder of framework is acquired in the end.

#### IV. EXPERIMENT RESULTS

##### A. The Results of Single Image Super Resolution via TD-SAI and TRD

During the encoding and decoding of the compression framework, there is a quantitative parameters (QP = 1~100) needed to set. With the increasing of the QP, the quality of image is much better both in objective and subjective quality. In our experiment, we pick up three different quantitative parameters (QP=40, 50, 60) to demonstrate the effectiveness and robustness of the targeted dictionary based on soft-decision adaptive interpolation and the targeted residual dictionary.

Extensive experiments were carried out to evaluate the proposed targeted dictionary based on soft-decision adaptive interpolation (TD-SAI) and targeted residual dictionary (TRD) algorithm, in signal-to-noise ratio (PSNR), structural similarity (SSIM) and visual quality. The representative image super-resolution methods, including ScSR [6] and BPJDL [10], are employed to compare with the proposed TD-SR (TD-SAI and TRD) method.

Table 1. PSNR (dB) and SSIM comparisons of different QP by different methods for super-resolution

Image (QP=40)	ScSR		BPJDL		Proposed	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Flower	33.55	0.949	33.78	0.950	<b>34.81</b>	<b>0.959</b>
Cablecar	28.46	0.898	28.92	0.906	<b>29.61</b>	<b>0.912</b>
Soccer	28.51	0.907	28.79	0.912	<b>29.75</b>	<b>0.921</b>
Parrot	31.38	0.945	31.78	0.947	<b>32.66</b>	<b>0.953</b>
Butterfly	26.71	0.944	27.39	0.953	<b>28.62</b>	<b>0.967</b>

Image (QP=50)	ScSR		BPJDL		Proposed	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Flower	34.29	0.957	34.53	0.959	<b>35.39</b>	<b>0.961</b>
Cablecar	29.06	0.913	29.48	0.921	<b>30.28</b>	<b>0.926</b>
Soccer	29.22	0.921	29.49	0.926	<b>30.44</b>	<b>0.934</b>
Parrot	31.94	0.952	32.29	0.954	<b>33.15</b>	<b>0.957</b>
Butterfly	27.28	0.951	27.97	0.960	<b>28.87</b>	<b>0.971</b>

Image (QP=60)	ScSR		BPJDL		Proposed	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
Flower	34.95	0.963	35.17	0.964	<b>36.03</b>	<b>0.967</b>
Cablecar	29.65	0.924	30.08	0.931	<b>30.84</b>	<b>0.936</b>
Soccer	29.89	0.932	30.14	0.936	<b>31.03</b>	<b>0.941</b>
Parrot	32.46	0.958	32.83	0.960	<b>33.41</b>	<b>0.954</b>
Butterfly	27.83	0.959	28.53	0.966	<b>29.56</b>	<b>0.972</b>

The PSNR and SSIM comparisons for five gray images in the case of QP=40, 50, 60 measurements are provided in Table 1. From the PSNR and SSIM comparison results, we notice that the most recent beta process joint dictionary learning method (BP-JDL) outperforms the coupled learned dictionaries (ScSR). The proposed TD-SR (TD-SAI and TRD) method is able to provide high-res images with the best recovery accuracy.

Here, we pick up two representative images to report the super resolution experimental results without the assistance of feedback. The visual comparisons of the super resolution for downsampling and compressed image results are shown in Figure 6~7. From the visual comparison results, we notice the improvement of proposed TD-SR (TD-SAI and TRD) method in terms of artifact on the edges compared to the other methods. It is clear to observe that our proposed algorithm generates shaper edges, and the reconstructed image is much more faithful to the original images. The proposed method not only eliminates the ringing effects, but also preserves sharper edges and more high-frequency details, showing much clearer and better subjective results than the other competing super resolution algorithms for the downsampling and compressed image in framework.



Figure 6. Visual quality comparison of different methods for gray image Parrot (QP=40).

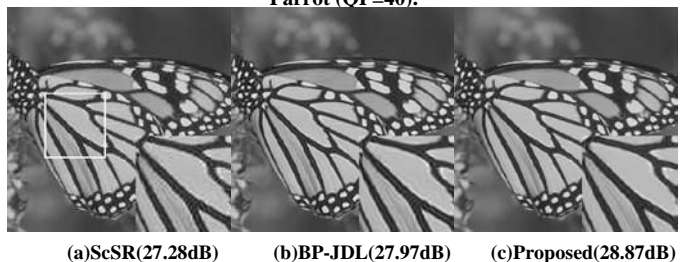


Figure 7. Visual quality comparison of different methods for gray image Butterfly (QP=50).

## B. Analysis of our Proposed Entire Framework including the Feedback

Here, the proposed single image compression framework (SICF) is composed of six components: (1) downsampling; (2) JPEG encoder; (3) JPEG decoder; (4) denoising; (5) super-resolution via TD (TD-SAI and TRD-1); (6) feedback and TRD-2 super-resolution.

A set of images are selected to test the efficiency of the proposed SICF. We compare the image processed by our framework with the image encoded directly by JPEG without downsampling under the same PSNR. It should be noted that the efficiency of our framework can be reflected by the rate-distortion performance. Here, we pick up two representative images to report the bit-rate saving results with the assistance of the feedback and the final TRD-2 super-resolution.

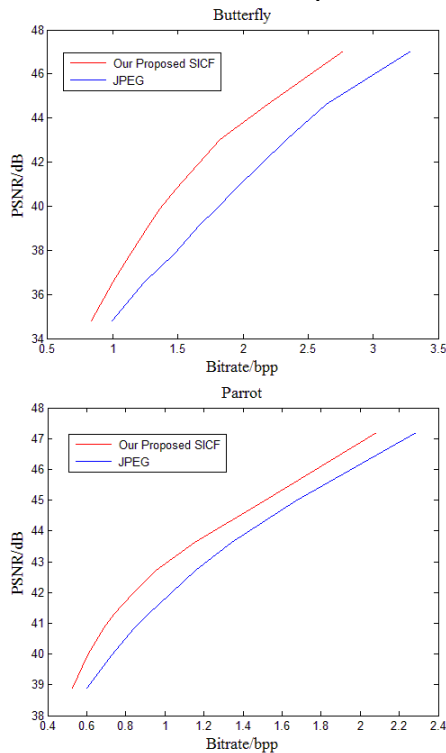


Figure 8. Rate-distortion Performance of Butterfly and Parrot

As shown in Figure.8, the rate-distortion performance of image, which was processed by our SICF, outperforms the JPEG standard under the same PSNR and visual quality. Obviously, our proposed SICF achieves significant bit-rate saving under the same PSNR compared with JPEG standard, for which the proposed SICF have unique feedback design for image compression and the targeted residual dictionary for image super-resolution.

## V. CONCLUSIONS

In this paper, we establish a single image compression framework (SICF) combined with sparse representation-based super-resolution. An effectual targeted dictionary based on soft-decision adaptive interpolation (TD-SAI) and targeted residual dictionary (TRD) method was proposed to achieve higher reconstruction quality. For further improving the visual

quality of the final decoded image, the feedback is designed to obtain the residual assistance image which can be compensated for the loss of high-frequency details in super-resolution process at decoder side. The proposed framework achieves significant bit-rate saving under the same PSNR compared with JPEG standard. Our research group is now conducting experiments which focus on the JPEG-2000 image compression standard, and it has made the preliminary achievement.

## REFERENCES

- [1] M. Shen, P. Xue, and C. Wang, "Down-Sampling Based Video Coding Using Super-Resolution Technique," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 6, pp.755-765, Jun. 2011.
- [2] E. M. Hung, R. L. de Queiroz, F. Brandi, K. F. de Oliveira, and D. Mukherjee, "Video Super-Resolution Using Codebooks Derived From Key-Frames," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 9, pp.1321-1331, Sep. 2012.
- [3] F. Brandi, R. de Queiroz, and D. Mukherjee, "Super resolution of video using key frames," in *Proc. IEEE ISCAS*, pp. 321-324, Oct. 2008.
- [4] J. Glaister, C. Chan, M. Frankovich, A. Tang, A. Wang, "Hybrid Video Compression Using Selective Keyframe Identification and Patch-Based Super-Resolution," 2011 IEEE International Symposium on Multimedia, pp105-110, Dec.2011.
- [5] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE CVPR*, pp. 1-8. Jun.2008,
- [6] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," *IEEE Trans. Image Process.*, vol. 19, no. 11, pp. 2861-2873, Nov. 2010.
- [7] R. Zeyde, M. Elad, and M. Protter, "On single image scale-up using sparse-representations," in *Proc. Curves Surfaces*, Jun. 2010, pp. 711-730.
- [8] S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-coupled dictionary learning with applications to image super-resolution and photo-sketch synthesis," in *Proc. IEEE CVPR*, pp2216-2223, Jun.2012
- [9] M.C. Yang, and Y.C. F. Wang, "A Self-Learning Approach to Single Image Super-Resolution," *IEEE Trans. on Multimedia*, vol. 15, no. 3, pp. 498 - 508, Apr. 2013.
- [10] H. Li, H. Qi, and R. Zaretzki, "Beta process joint dictionary learning for coupled feature spaces with application to single image super-resolution," in *Proc. IEEE Int. Conf. Computer Vision and Pattern Recognition*, 2013, pp. 345-352.
- [11] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080-2095, Aug. 2007.
- [12] X. Zhang and X. Wu, "Image interpolation by adaptive 2-D autoregressive modeling and soft- decision estimation," *IEEE Trans. Image Process.*, vol. 17, no. 6, pp. 887-896, Jun. 2008.