

# Optical Engineering

[SPIDigitalLibrary.org/oe](http://SPIDigitalLibrary.org/oe)

## **Toward high-quality image communications: inverse problems in image processing**

Cheolkon Jung  
Licheng Jiao  
Bing Liu  
Hongtao Qi  
Tian Sun

# Toward high-quality image communications: inverse problems in image processing

Cheolkon Jung

Licheng Jiao

Bing Liu

Hongtao Qi

Tian Sun

Xidian University

Key Lab of Intelligent Perception and Image

Understanding of Ministry of Education of  
China

Xi'an 710071, China

E-mail: [ckjung@ece.skku.ac.kr](mailto:ckjung@ece.skku.ac.kr)

**Abstract.** Recently, image communications are becoming increasingly popular, and there is a growing need for consumers to be provided with high-quality services. Although the image communication services already exist over third-generation wireless networks, there are still obstacles that prevent high-quality image communications because of limited bandwidth. Thus, more research is required to overcome the limited bandwidth of current communications systems and achieve high-quality image reconstruction in real applications. From the point of view of image processing, core technologies for high-quality image reconstruction are face hallucination and compression artifact reduction. The main interests of consumers are facial regions and several compression artifacts inevitably occur by compression; these two technologies are closely related to inverse problems in image processing. We review recent studies on face hallucination and compression artifact reduction, and provide an outline of current research. Furthermore, we discuss practical considerations and possible solutions to implement these two technologies in real mobile applications. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.OE.51.10.100901]

Subject terms: compression artifact reduction; face hallucination; image communications; image reconstruction; inverse problems in image processing..

Paper 120750V received May 24, 2012; revised manuscript received Aug. 27, 2012; accepted for publication Aug. 30, 2012; published online Sep. 27, 2012.

## 1 Introduction

With advancements in mobile communication devices, technology now allows people to communicate while looking at each other's face. This technology is also referred to as videoconferencing and basically transmits images to a display system so users can see each other while talking, as shown in Fig. 1(a). Many market analysts predict the number of subscribers to image communication services grows exponentially every year because of lower mobile device prices and aggressive marketing of communication companies, as shown in Fig. 1(b). As image communication services come into wide use, consumers want high-quality services. Although image communication services already exist over third-generation (3G) wireless networks, such as the high-speed downlink packet access (HSDPA), there are still obstacles that prevent high-quality communications because of limited bandwidth (maximum uploading and downloading speeds are 14.4 and 5.76 Mbps, respectively). Consequently, more research is required to overcome the limited bandwidth of current communications systems and achieve high-quality image reconstruction in mobile devices. In terms of image processing, core technologies for high-quality image reconstruction are face hallucination and compression artifact reduction.

Face hallucination technology, which is also referred as face super-resolution (SR), is very important for image communications because the main interests of consumers are facial regions, as shown in Fig. 2. A number of related face hallucination methods have been proposed in recent years. Among them, learning-based methods have received

much attention because they can achieve a high magnification factor and produce good SR results compared with other methods. Baker and Kanade<sup>1,2</sup> first introduced a face hallucination method which constructs the high frequency components from a parent-structure resorting to the training set. Wang and Tang<sup>3</sup> presented a principal component analysis (PCA)-based face hallucination algorithm to globally infer the high-resolution face image. Liu et al.<sup>4</sup> developed a two-step statistical modeling approach which integrates a global model and a local model corresponding to the common and specific face characteristics, respectively. Although complicated probabilistic models are required in Liu et al.'s method,<sup>4</sup> the idea of the two-step approach became more and more popular since then. Recently, a novel face hallucination method based on position-patch has been proposed. The position-patch based method hallucinates the high resolution (HR) image patch using the same position image patches of training images.<sup>5-7</sup> Thus, it is able to save computational time and produce high-quality SR results compared to manifold learning-based methods.

With respect to the compression artifact reduction, several compression artifacts inevitably occur because of the loss of high frequency components caused by lossy compression techniques such as H.264 or MPEG-4 (representative artifact: blocking artifact). They seriously degrade the picture quality and are annoying to viewers of the reconstructed images as shown in Fig. 3.<sup>8,9</sup> Accordingly, compression artifact reduction is also very important for image communications. Blocking artifacts appear as grid noise along the block boundaries because each block is transformed and quantized independently. Blocking artifacts occur because of the independent transform and quantization of each block without

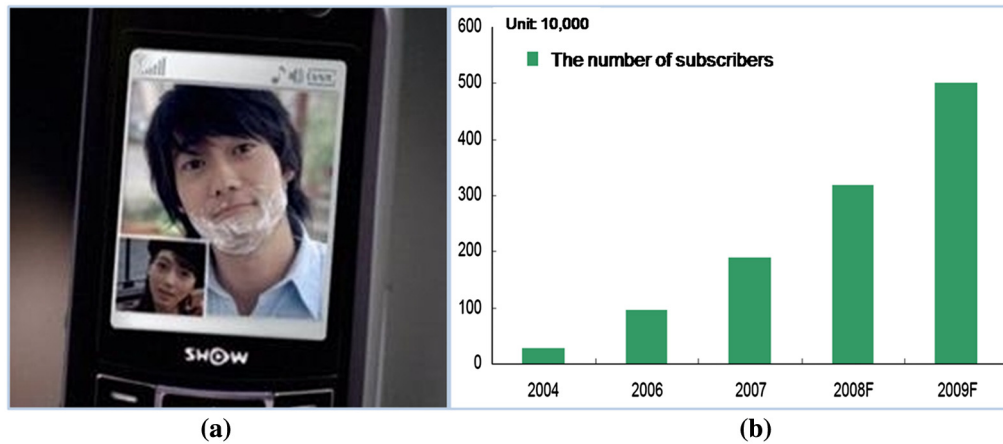


Fig. 1 (a) Image communications in mobile devices. (b) The number of subscribers (from Atlas research).

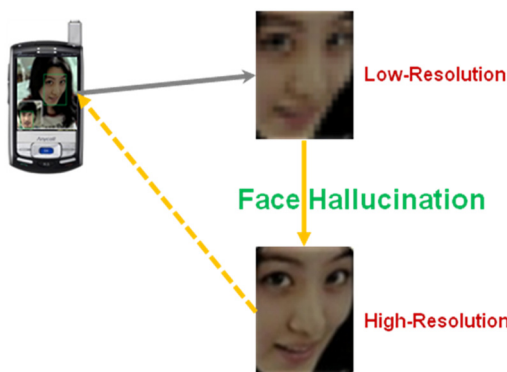


Fig. 2 Face hallucination in image communications.

considering inter-block correlations. Up to now, many studies have been conducted to reduce blocking artifacts from compressed images. Among them, image restoration techniques are commonly used to reduce blocking artifacts and recover the original image;<sup>10</sup> projection-onto-convex-sets (POCS)-based methods are representative research results of such techniques. In the POCS-based methods, prior information was represented as convex sets for reconstruction, and blocking artifacts were reduced by iteration

procedures.<sup>11</sup> POCS-based methods are very effective in reducing blocking artifacts because they are easy to impose smoothness constraints around block boundaries. Total variation (TV)-based methods are actively studied for image deblocking.<sup>12,13</sup> TV provides an effective criterion for image restoration, and thus can be successfully used as prior information for image deblocking. Alter et al.<sup>13</sup> proposed a constrained TV minimization method to reduce blocking artifacts without removing perceptual features. By the TV minimization, edge information was effectively preserved while reducing blocking artifacts. Moreover, a field of experts (FoE) prior was successfully applied to image deblocking.<sup>10</sup> In this method, the image deblocking problem was solved by the maximum a posteriori (MAP) estimation based on the FOE prior. The two technologies are associated with inverse problems in image processing. In this article, we provide an outline of recent studies on face hallucination and compression artifact reduction.

The rest of this article is organized as follows. In Sec. 2, we describe the inverse problems in image processing. In Sec. 3, we explain recent research trends and results related to face hallucination, and we address them related to compression artifact reduction in Sec. 4. In Sec. 5, we discuss practical considerations and possible solutions to implement two technologies in mobile applications. Finally, conclusions are made in Sec. 6.



Fig. 3 Example of compression artifacts in the Ballet sequences. (a) Original image. (b) Compressed image.

## 2 Inverse Problems in Image Processing

Inverse problems involve estimating parameters or data from inadequate observations; the observations are often noisy and contain incomplete information about the target parameter or data due to physical limitations of the measurement devices. Due to lack of sufficient information in the indirect observations, solutions to inverse problems are usually nonunique and challenging. That is, they are ill-posed problems, and thus, some other reconstruction technologies are required to solve them including machine learning, Bayesian inference, convex optimization, sparse representation, and so on.<sup>14–16</sup>

Indeed, many problems in image processing can be represented as inverse problems. They are modeled by relating the observed image  $g(r)$  to the unknown original image  $f(r)$ . A general form for the relation is as follows:<sup>14</sup>

$$g(r) = [Hf](r) + n(r), \quad r \in R, \quad (1)$$

where  $r$  represents the pixel position,  $R$  represents the whole surface of  $g(r)$ ,  $H$  is an operator representing the forward problem, and  $n(r)$  represents the errors (modeling uncertainty and observation errors). If we assume operator  $H$  is linear, we can write the observation model in a vector-matrix form as follows:

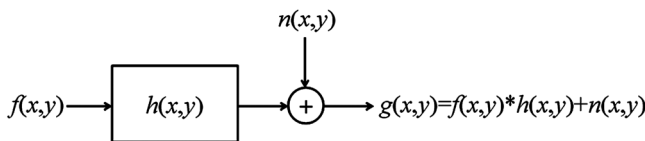
$$\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}, \quad (2)$$

where  $\mathbf{g} = \{g(r), r \in R\}$ ,  $\mathbf{f} = \{f(r), r \in R\}$  and  $\mathbf{n} = \{n(r), r \in R\}$  are vectors containing the observed image pixel values, unknown original image pixel values, and observation errors, respectively; and  $\mathbf{H}$  is a huge dimensional matrix whose elements are defined from  $H$ .

Figure 4 shows the observation model in image processing which can be formulated as inverse problems. In image processing, there are many inverse problems such as image denoising, image SR, image deblurring, image decompression, and so on. Above all, we inevitably meet several inverse problems in image communications because transmission bandwidth is strictly limited in a mobile communication environment. Consequently, image sequences are compressed and transmitted using lossy compression techniques such as H.264 and MPEG-4, and thus, undesired image distortions also occur because of compression artifacts resulting from lossy compression techniques. In this article, we deal with two representative inverse problems in image processing: face hallucination and image deblocking.

## 3 Face Hallucination

Since the concept of face hallucination is introduced by Baker and Kanade,<sup>1,2</sup> a number of related face hallucination methods have been proposed during the past decade. In general, there are two classes of SR techniques: multiframe SR



Observation model:  $\mathbf{g} = \mathbf{H}\mathbf{f} + \mathbf{n}$

Fig. 4 Observation model in image processing, redrawn from Ref. 14.

(from inputs images only) and single-frame SR (from other training images). From a methodological viewpoint, it can be widely divided into interpolation-based,<sup>17,18</sup> reconstruction-based,<sup>19–24</sup> and learning-based<sup>3,6,7,25–30</sup> methods.

First, the basic interpolation methods include nearest-neighbor interpolation, bilinear interpolation and bicubic interpolation, etc.<sup>17,18</sup> Given one low resolution (LR) image, they only use the information of the original pixel and several pixels around it to estimate the missing pixels. It is simple and fast and can get some results when the interpolation factor is small. However, when the interpolation factor is large, the performance is not good because the high frequency information is missed. Second, reconstruction-based methods firstly build an observation model to connect the original HR image and realistic LR image, simulating the process to get a LR image from a HR image. There are many reconstruction-based methods, such as POCS,<sup>19</sup> MAP method,<sup>20</sup> iterative back-projection method,<sup>21,22</sup> regular method,<sup>23</sup> and mixed method,<sup>24</sup> etc. All of them need some locality prior assumptions, and can make the blur and saw-tooth effects to a certain extent. Since the prior knowledge is somewhat little, the information provided by LR images may not satisfy with the demand for HR images. Third, learning-based methods have received much attention in recent years because they can achieve a high magnification factor and produce good SR results compared with other methods. The basic idea is to compute the neighborhood between the patch of test images and the patches of training images set, and construct the optimal coefficients to approximate the HR image using the learned prior knowledge. In this article, we focus on learning-based face hallucination methods and introduce some representative works and our research results.

### 3.1 Example-Based Image SR

In 2001, example-based image SR was proposed by Freeman et al. Its core idea was to learn the fine details from HR images of training datasets, and use the learned relationships between LR and HR to predict fine details of a test image. Above all, Freeman et al. employed a nonparametric patch-based prior along with the Markov random field (MRF) model to generate the desired HR images. A large dataset of HR and LR patch pairs was generated and used for seeking the nearest neighbors of the LR input patches. The selected HR patch neighbors were treated as the candidates for the target HR patch. The block diagram of the method is shown in Fig. 5. As shown in the figure, the key procedure of

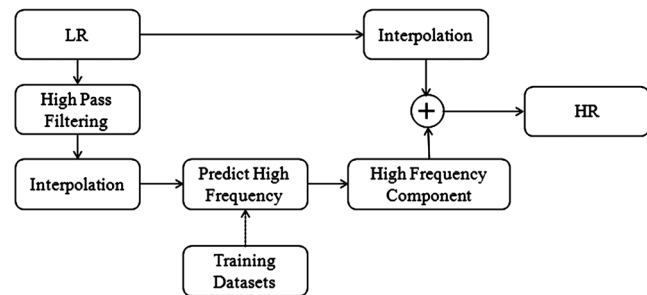


Fig. 5 Block diagram of the example-based super resolution (SR) method.<sup>25</sup>

this method is to predict the missing high frequencies using the training datasets.

### 3.2 Neighbor-Embedding Based Image SR

In 2004, Chang et al. proposed a novel method for solving single-image SR problems. In this method, given an LR image as input, a set of training examples were used to recover its HR counterpart. While this formulation resembled other learning-based methods for SR, this method was inspired by manifold learning-based methods, particularly locally linear embedding (LLE). More specifically, small image patches in LR and HR images formed manifolds with similar local geometry in two distinct feature spaces. Then, multiple nearest neighbors were selected in the feature space, and SR images were reconstructed by the corresponding HR patches of the nearest neighbors. Since then, this method has been extensively applied to solving image SR problems including face hallucination.

### 3.3 PCA-Based Face Hallucination

In 2005, a new face hallucination method using eigen-transformation was proposed by Wang et al. In contrast to conventional methods based on probabilistic models, this method viewed face hallucination as a transformation between different image styles. PCA was used to fit the input face image as a linear combination of the LR face images in the training dataset. The HR image was rendered by replacing the LR training images with HR ones, while retaining the same combination coefficients. Since face images were well structured and had similar appearances, they spanned a small subset in the high dimensional image space. In the work of Penev and Sirovich,<sup>31</sup> face images were shown to be well reconstructed by PCA representation with 300 to 500 dimensions. The system diagram of this method is shown in Fig. 6. As shown in the figure, this method first employed PCA to extract useful information as much as possible from an LR face image, and then rendered an HR face image by eigen-transformation.

### 3.4 Sparse Coding Based Face Hallucination

In 2008, a new approach to single-image SR based on sparse signal representation was proposed by Yang et al. This method was motivated by the image statistics that image patches could be well-represented as a sparse linear combination of elements from an appropriately chosen overcomplete dictionary. They found sparse representation for each patch of the LR input, and then used the coefficients of

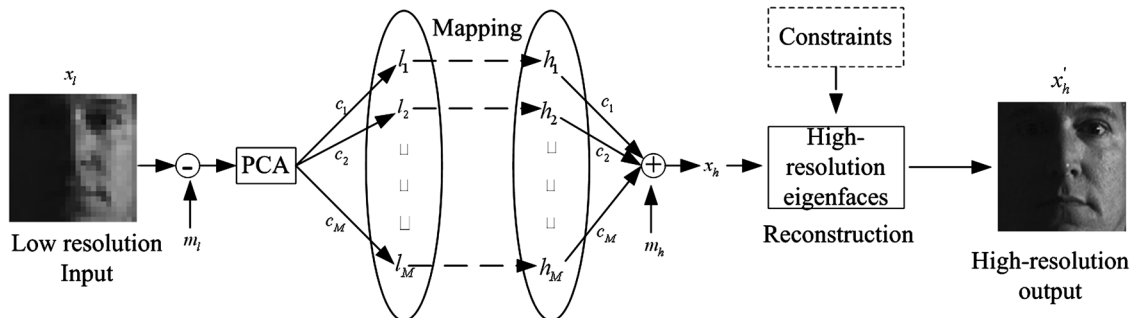
this representation to generate the HR output. Theoretical results from compressed sensing suggested that under mild conditions, the sparse representation could be correctly recovered from the down-sampled signals. By jointly training two dictionaries for the LR and HR image patches, they made the similarity of sparse representations between the LR and HR pairs with respect to their own dictionaries. Therefore, the sparse representation of an LR patch was applied to the reconstruction of SR images with the HR patch dictionary. The learned dictionary pair was a more compact representation of the patch pair compared to previous approaches, and simply sampled a large amount of image patch pairs reducing the computational cost effectively.

### 3.5 Position-Patch Based Face Hallucination

In 2010, a novel face hallucination approach was proposed by Ma et al. In contrast to most of the conventional methods based on probabilistic models or manifold learning, the position-patch based method hallucinated the HR image patch using the same position image patches of each training images. The optimal weights of the training image position-patches were estimated and the hallucinated patches were reconstructed using the same weights. The final SR face images were formed by integrating the hallucinated patches. It was able to save computational time and produce high-quality SR results compared to conventional manifold learning based methods. The position-patch based face hallucination method is briefly described in Algorithm 1.

**Algorithm 1** Position-patch based face hallucination.<sup>5</sup>

- 
- Step 1:** Denote the input LR image, LR training image, and HR training image in overlapping patches as  $\{X_L^P(i, j)\}_{p=1}^N$ ,  $\{Y_L^{mP}(i, j)\}_{p=1}^N$ , and  $\{Y_H^{mP}(i, j)\}_{p=1}^N$ , respectively, for  $m = 1, 2, \dots, M$ .
- Step 2:** For each patch  $X_L^P(i, j)$ :
- (a) Compute the reconstruction weights  $w(i, j)$  by least square estimation
  - (b) Synthesize the HR patch  $X_H^P(i, j)$
- Step 3:** Concatenate and integrate the hallucinated HR patches to form a facial image, which is the target HR facial image  $\{X_H^P(i, j)\}_{p=1}^N$ .
- 



**Fig. 6** System diagram of the principal component analysis (PCA)-based face hallucination.<sup>3</sup>

### 3.6 Convex-Optimization-Based Face Hallucination

Inspired by the position-patch based face hallucination method, a new convex optimization based face hallucination method is proposed. The position-patch based method has employed least square estimation to get the optimal weights for face hallucination; however, the least square estimation approach can provide biased solutions when the number of the training position-patches is much larger than the dimension of the patch. To overcome this problem, we make use of constrained convex optimization instead of least square estimation to obtain the optimal weights for face hallucination. The optimal weights ( $w$ ) are computed by solving the following convex optimization problem:

$$\min_w \|w\|_1 \quad \text{subject to } \|X_L^P - Y_L^P \cdot w\|_2^2 \leq \varepsilon, \quad (3)$$

where  $Y_L^P$  is a column matrix of the training patches  $Y_L^{mP}(i, j)$  for  $m = 1, 2, \dots, M$ ; and  $\varepsilon$  is a error tolerance. Consequently, the hallucinated HR patch  $X_H^P(i, j)$  is obtained by:

$$X_H^P(i, j) = \sum_{m=1}^M Y_H^{mP}(i, j) \cdot w_m(i, j). \quad (4)$$

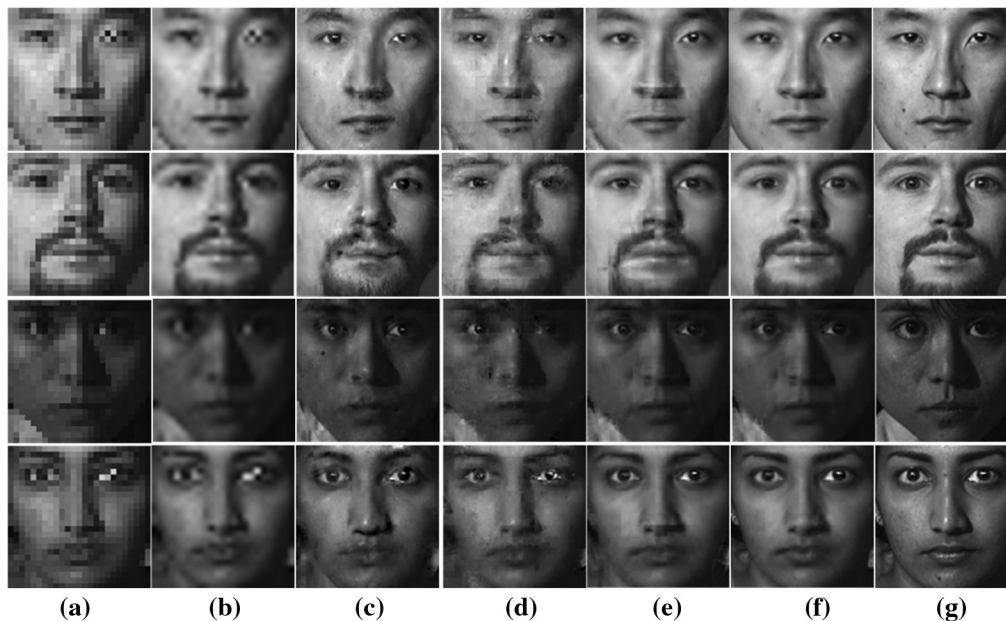
By Eqs. (3) and (4), we can get more stable reconstruction weights for face hallucination because  $l_1$ -norm is more suitable for this problem, and because each patch can be approximated with a smaller subset of patches than  $l_2$ -norm. In contrast,  $l_2$ -norm provides nonzero weights for all patches. Figure 7 shows the face hallucination results by bi-cubic interpolation, example-based image SR,<sup>25</sup> neighbor-embedding based image SR,<sup>26</sup> position-patch based face hallucination,<sup>5</sup> and convex optimization based face hallucination.<sup>7</sup> We performed experiments on the CMU-PIE face database which contains 41,368 images obtained from 68 subjects.

We took the frontal face images with 21 different illumination conditions. Thus, the total number of images was 1,428. Among them, 630 images of 30 subjects were used in the training stage, and the rest were used in the synthesis stage. In the neighbor-embedding method, the HR patch size of  $Y_H^m$  was  $12 \times 12$  pixels, while the corresponding LR patch size of  $Y_L^m$  was  $3 \times 3$  pixels. In addition, the number of the neighbor patches for reconstruction was 5. The size of the image patches in position-patch and convex optimization methods was  $3 \times 3$  pixels. The size of LR images for training and synthesis was  $25 \times 25$  pixels, while that of hallucinated results was  $100 \times 100$  pixels. That is, interpolation factor was 4. As shown in the figure, learning based methods generally produce better face hallucination results than traditional bicubic interpolation. Above all, the hallucinated results of Refs. 25 and 26 are somewhat blurred and with some artifacts; however, results of Refs. 5 and 7 produce more natural looking facial images. Further examination of the results reveals that Ref. 7 is more effective in preserving the edge and image details in the nose and mouth areas than Ref. 5.

For a more quantitative test, average peak-to-noise ratio (PSNR) and structural similarity (SSIM) values of the face hallucination results are provided in Table 1. The SSIM is a complementary measure of the PSNR, which gives an indication of image quality based on known characteristics of the human visual system.<sup>32</sup> Here, the unit of PSNR is dB. As shown in the table, our method achieves the best hallucination performances in terms of the PSNR and SSIM. Here, the bold numbers represent the best PSNR and SSIM values.

### 4 Compression Artifact Reduction

Block-based discrete cosine transform (BDCT) has been widely used in image and video compression due to its energy compacting property and relative ease of



**Fig. 7** Face hallucination results. (a) Input LR faces ( $25 \times 25$  pixels). (b) Bicubic interpolated images. (c) Example-based image super resolution (SR).<sup>25</sup> (d) Neighbor-embedding based image SR.<sup>26</sup> (e) Position-patch based face hallucination.<sup>5</sup> (f) Convex optimization based face hallucination.<sup>7</sup> (g) Original HR faces ( $100 \times 100$  pixels).

**Table 1** Average PSNR and SSIM values of different methods.

Measure	Bicubic	Example-based (Ref. 25)	Neighbor embedding (Ref. 26)	Position-patch (Ref. 5)	Convex optimization (Ref. 7)
PSMR	24.5388	26.0954	26.3758	28.1613	<b>28.2437</b>
SSIM	0.7278	0.7544	0.7444	0.8146	<b>0.8178</b>

implementation.<sup>33–36</sup> Thus, BDCT has been adopted in most image/video compression standards including JPEG (joint photographic experts group) and MPEG (motion picture experts group). However, BDCT has a major drawback, which is usually referred as blocking artifacts. Blocking artifacts appear as grid noise along the block boundaries because each block is transformed and quantized independently. Usually, the lower the bit rate is, the more serious the blocking artifacts are. Blocking artifacts occur because of the independent transform and quantization of each block without considering inter-block correlations.

#### 4.1 Main Techniques for Image Deblocking

There are two main techniques to deal with the blocking artifacts: in-loop filtering and postprocessing methods. The in-loop filters operate within coding loop while the postprocessing methods are applied after the decoder and make use of decoded parameters. Table 2 lists the deblocking filters employed by current video coding standards.<sup>37</sup> As listed in the table, in-loop filters have been optionally or not used because of the need of changing the encoder structure. Thus, postprocessing methods are promising solutions to this problem and comparable results have been achieved by researchers.

#### 4.2 Postprocessing Methods For Image Deblocking

Since early 1980s, postprocessing of low bit-rate BDCT coded images has a lot of research attention. Postprocessing methods are classified into three main groups: filtering-based, denoising-based, and restoration-based methods.<sup>10</sup>

First, some researchers viewed the distortions around the block boundaries as spatial, high-frequency components.

**Table 2** Deblocking filters for video coding standards.<sup>37</sup>

Standard	Deblocking filter
H.261	Optional in-loop filter
MPEG-1	None
MPEG-2	None, post-processing often used
H.263	None
MPEG-4	Optional in-loop filter, post-processing suggested
H.264	Mandatory in-loop filter, post-processing suggested

Thus, many filtering-based methods have been proposed to reduce them. In 1984, Lim and Reeve<sup>38</sup> first applied low-pass filtering to the pixels along the boundary to remove the blocking artifacts. Then, in 1986, Ramamurthi and Gersho<sup>39</sup> proposed a nonlinear space-variant filter to perform filtering in parallel with the edges. Since then, many filtering-based methods have been presented, and the representative work is the adaptive deblocking filter, which has been used in the H.264/MPEG-4 advanced video coding (AVC) standards to reduce the distortions.<sup>40</sup>

Second, some researchers viewed deblocking as a denoising problem. They proposed some efficient noise models and some deblocking methods based on the wavelet technique. In 1997, Xiong et al.<sup>41</sup> exploited cross-scale correlation by the overcomplete wavelet transform, and used the thresholds to reduce the distortions. In 2004, Liew and Yan<sup>34</sup> made a theoretical analysis of the blocking artifacts, and used the three-scale overcomplete wavelet scheme to reduce them.

Third, many researchers viewed deblocking as a restoration problem, and proposed restoration-based deblocking methods. The POCS-based method was a representative approach of the restoration-based methods for deblocking.<sup>42</sup> In the POCS-based methods, prior information was represented as convex sets for reconstruction, and blocking artifacts were reduced by iteration procedures. The POCS based methods were very effective for reducing blocking artifacts because they were easy to impose smoothness constraint around block boundaries. In 2003, Kim et al.<sup>11</sup> proposed a new smoothness constraint set (SCS) and an improved QCS to improve performances of the POCS-based methods. Furthermore, the TV-based methods were actively studied for image deblocking. TV provided an effective criterion for image restoration, and thus could be successfully used as prior information for image deblocking.<sup>13,43</sup> In 2004, Alter et al. proposed a constrained TV minimization method to reduce blocking artifacts without removing perceptual features. In 2010, a human visual system (HVS)-based TV method using a new weighted regularization parameter was proposed by Do et al.<sup>44</sup> In 2007, a FoE prior<sup>45,46</sup> was successfully applied to image deblocking by Sun and Cham.<sup>10</sup> In this method, the image deblocking problem was solved by the MAP estimation, based on the FOE prior. In addition, they employed the narrow quantization constraint set (NQCS) for further PSNR gain.<sup>47</sup> Consequently, this method achieved a high PSNR gain and produced state-of-the-art results on deblocking.

#### 4.3 Sparse Representation Based Image Deblocking

Recently, sparse representation has been actively studied to solve various restoration problems in image processing.<sup>48–52</sup> Some researchers have made significant contributions to

image denoising, restoration and SR using sparse representation. Sparse representation assumes that original signals can be accurately recovered by several elementary signals called atoms.<sup>50,53</sup> Thus, it has been proven very effective for image restoration tasks. Inspired by recent results of sparse representation, we provided a novel deblocking method based on sparse representation.<sup>48</sup> To remove blocking artifacts, we obtain a general dictionary from a set of training images using K-singular value decomposition (K-SVD) algorithm, which can effectively describe the content of an image. Then, an error threshold for orthogonal matching pursuit (OMP) is automatically estimated to use the dictionary for image deblocking by the quality of compressed image. Our deblocking method is comprised of two main procedures: generation of a deblocking dictionary using K-SVD algorithm, and image deblocking by the deblocking dictionary. That is, the deblocking dictionary is generated in the training stage, and blocking artifact reduction is performed in the testing stage.

### 4.3.1 Deblocking dictionary design using K-SVD algorithm

In the training stage, image patches are selected to generate a dictionary for image deblocking. From the image patches, a deblocking dictionary is trained by the K-SVD algorithm. Here, to solve the optimization problem, the batch-OMP method is used.<sup>54</sup> The K-SVD algorithm is an iterative method to generate an overcomplete dictionary that fits training examples well. It is simple and designed to be truly direct generalization of the K-Means algorithm.<sup>52-56</sup> In general, it alternates between sparse coding and dictionary update while training.

Let  $\bar{\mathbf{X}} = [x_1; \dots; x_p]$  be an  $n \times P$  matrix of  $P$  training patches of  $n$ -length pixels, used to train an overcomplete dictionary  $\mathbf{D}$  of size  $n \times K$  with  $P \gg K$  and  $K > n$ . For generating  $\mathbf{D}$ , the objective function of the K-SVD algorithm is defined as follows:<sup>55,57</sup>

$$\min_{\mathbf{D}, \Theta} \|\bar{\mathbf{X}} - \mathbf{D} \cdot \Theta\|_F^2 \quad \text{subject to } \|\theta_i\|_0 \leq S, \quad (5)$$

where  $S$  is a given sparsity level,  $\Theta = [\theta_1 \dots \theta_p]$ , and  $\theta_i$  is the sparse vector of coefficients representing the  $i$ 'th patch in terms of the columns of  $\mathbf{D} = [d_1 \dots d_K]$ . The K-SVD algorithm progressively creates the deblocking dictionary  $\mathbf{D}$  from an initial dictionary by solving Eq. (5). The full steps of dictionary generation are described in Algorithm 2.

### 4.3.2 Automatic estimation of error threshold

The deblocking dictionary  $\mathbf{D}$  is employed to reduce blocking artifacts. The objective function for image deblocking is as follows:

$$\min_{\Theta} \|\Theta\|_1 \quad \text{subject to } \|\mathbf{Y} - \mathbf{D} \cdot \Theta\|_2 \leq T, \quad (6)$$

where  $\mathbf{Y}$  is the corrupted image by blocking artifacts and  $T$  is an error threshold for OMP. Blocking artifacts are reduced by optimizing Eq. (6), and we can reconstruct the original image. As can be expected, an error threshold  $T$  of Eq. (6) should be estimated to use the deblocking dictionary in

### Algorithm 2 Dictionary generation by the K-SVD algorithm.

**Step 1:** Initialize a dictionary  $\mathbf{D}$  (an overcomplete DCT dictionary)

**Step 2:** Repeat  $n$  times ( $n$ : number of training iterations)

- a) Sparse coding stage: compute  $\theta_i$  using OMP for  $i = 1, 2, \dots, P$

$$\min_{\mathbf{D}, \Theta} \|\bar{\mathbf{X}} - \mathbf{D} \cdot \Theta\|_F^2 \quad \text{subject to } \|\theta_i\|_0 \leq S$$

- b) Dictionary update stage: update the dictionary atom  $d_k$  and coefficient  $\theta_k$  for  $k = 1, 2, \dots, K$

- b-1) Obtain the set of all indices corresponding to the training patches that use  $d_k$  and  $\theta_k$ .

- b-2) Compute the matrix of residuals  $\mathbf{E}_k$ :

$$\mathbf{E}_k = \bar{\mathbf{X}} - \sum_{j \neq k} d_j \theta_j$$

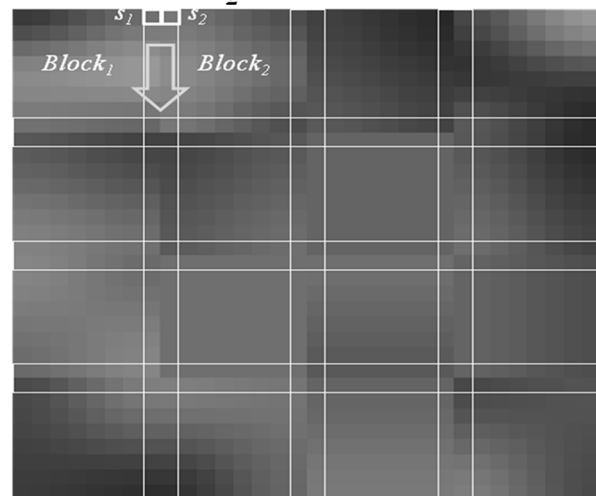
- b-3) Restrict  $\mathbf{E}_k$  by selecting only the columns corresponding to those elements that initially used  $d_k$  in their representation, and obtain  $\mathbf{E}_k^R$ .

- b-4) Apply SVD decomposition  $\mathbf{E}_k^R = \mathbf{U} \Delta \mathbf{V}^T$ , and update  $d_k = \mathbf{u}_1$ ,  $\theta_k^R = \Delta(1, 1) \cdot \mathbf{v}_1$  where  $\Delta(1, 1)$  is the largest singular value of  $\mathbf{E}_k^R$ ; and  $\mathbf{u}_1$  and  $\mathbf{v}_1$  are the corresponding left and right singular vectors, respectively.

reducing blocking artifacts. We can estimate  $T$  for OMP automatically using quality information of JPEG compressed images. The procedures of estimating  $T$  are summarized as follows.

First, the standard deviation of the quantization noise,  $\sigma_N$ , is estimated as shown in Fig. 8. Since the blocking artifacts mostly occur around the block boundaries,  $\sigma_N$  is computed from the intensity difference *Diff* between two boundary pixels on both sides of a boundary between two blocks as follows:

$$Diff = \frac{|I(s_1) - I(s_2)|}{2}$$



**Fig. 8** Block discontinuity estimation: *Diff* is the absolute value of one-half the intensity difference between two pixels; two pixels  $s_1$  and  $s_2$  belong to  $Block_1$  and  $Block_2$ , respectively; and  $I(s)$  is the intensity of a pixel  $s$ . Here, *Diff* is computed between two boundary pixels on both sides of a boundary between two blocks.



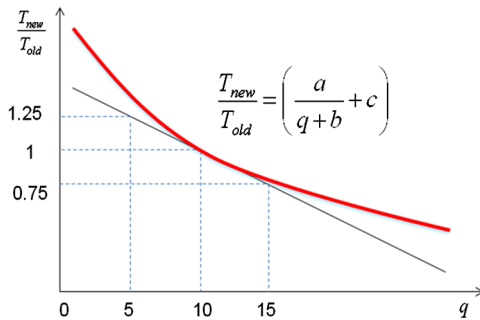
$$\text{Diff} = \frac{|I(s_1) - I(s_2)|}{2}, \quad (7)$$

where *Diff* is the absolute value of one-half the intensity difference between two pixels,  $s_1$  and  $s_2$ . In computing *Diff*, only horizontal or vertical block discontinuities are considered as mentioned in Ref. 34. In the figure, pixels  $s_1$  and  $s_2$  belong to Block<sub>1</sub> and Block<sub>2</sub>, respectively;  $I(s)$  is the intensity of a pixel  $s$ . Accordingly, we compute  $\sigma_N$  of the compressed blocky image from *Diff*.

Then,  $T$  is computed based on  $\sigma_N$ . In the previous works for image denoising,<sup>34,53,57</sup>  $T_{\text{old}}$  is obtained by the following equation:

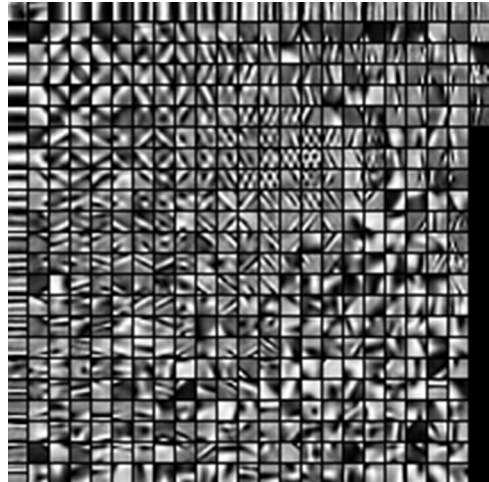
$$T_{\text{old}} = C \cdot \sigma_N. \quad (8)$$

Here, the noise gain  $C$  is set to 1.15. In the JPEG coding standard, the most important parameter is the quality  $q$ ,

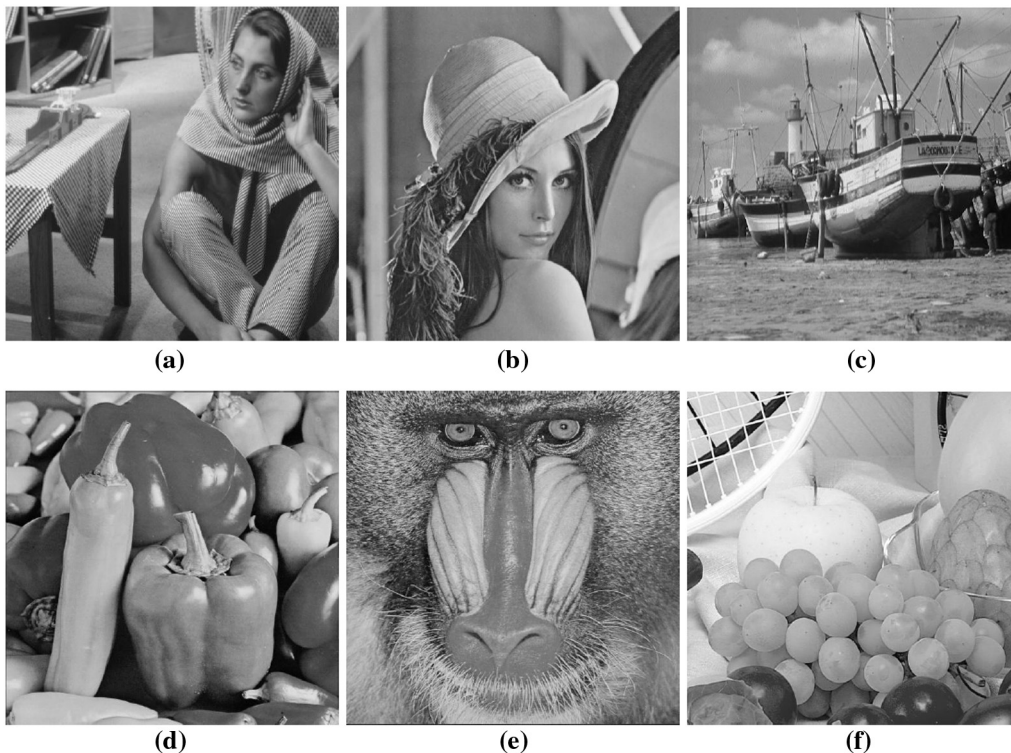


**Fig. 9** Distribution of the error threshold according to quality: the red line is the actual distribution of the error threshold.

which contains a value between 0 and 100. The higher  $q$  is, the less image degradation due to compression is; however, when  $q$  is high, the resulting file size is large. For image deblocking, we found that  $T_{\text{old}}$  fits well when  $T_{\text{old}}$  is only 10 by various experiments. In other cases,  $T_{\text{old}}$  do not follow the distribution of the error threshold  $T$  of Eqs. (6) by (8). Instead, we found that  $T_{\text{new}}/T_{\text{old}}$  follows nonlinear distribution according to a given quality  $q$  as shown in Fig. 9. Thus, we modify Eq. (8) as follows:



**Fig. 11** The general dictionary by K-singular value decomposition (K-SVD) using 100,000 image patches (total 512 atoms are learned with each atom of size  $8 \times 8$  pixels).



**Fig. 10** Test images: (a) *Barbara*, (b) *Lena*, (c) *Boat*, (d) *Peppers*, (e) *Baboon*, and (f) *Fruits*.



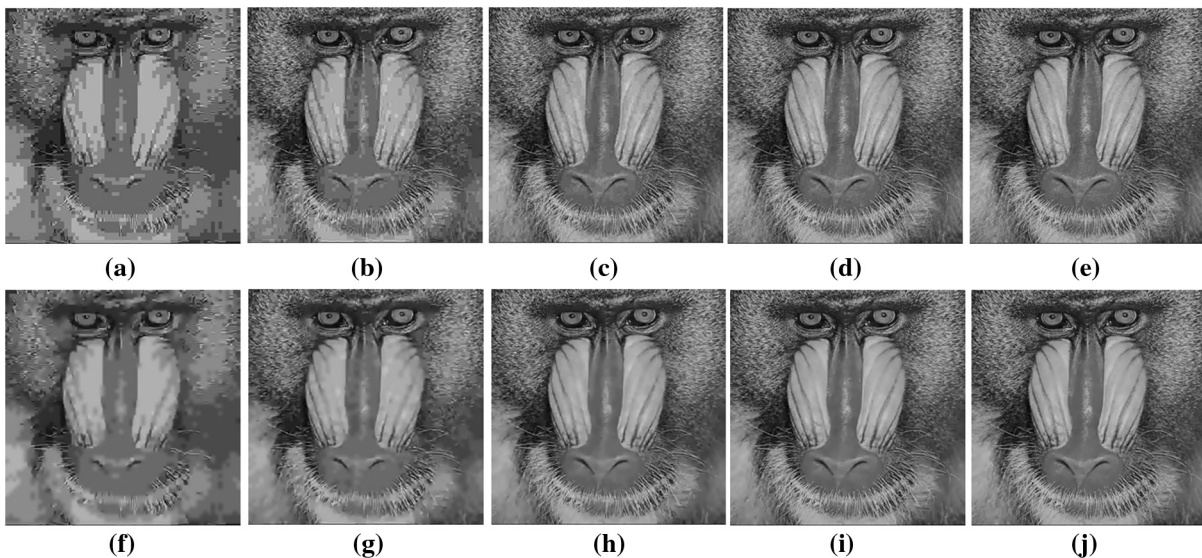
**Fig. 12** JPEG compressed images and their deblocked results of the *Barbara* image according to different quality values: (a) quality = 1, (b) quality = 5, (c) quality = 10, (d) quality = 15, (e) quality = 20, (f) the deblocked result of (a), (g) the deblocked result of (b), (h) the deblocked result of (c), (i) the deblocked result of (d), and (j) the deblocked result of (e).

$$T_{\text{new}} = T_{\text{old}} \cdot \left( \frac{a}{q+b} + c \right) = C \cdot \sigma_N \cdot \left( \frac{a}{q+b} + c \right), \quad (9)$$

where  $a$ ,  $b$ , and  $c$  are the control parameters, and their appropriate values are adjusted by experiments. Here,  $a$ ,  $b$ , and  $c$  are set to be 20, 10, and 0, respectively. Consequently, the error threshold for OMP is computed by  $T_{\text{new}}$  of Eq. (6), and used to solve Eq. (3). As a result, we get deblocked results of JPEG compressed images by the learned dictionary  $\mathbf{D}$ .

As shown in Fig. 10, six typical images were used for the tests, *Barbara*, *Lena*, *Boat*, *Peppers*, *Baboon*, and *Fruits*, whose sizes were  $512 \times 512$  pixels. In the training stage, total 91 natural images provided by the Yang et al.'s work<sup>51</sup> were used to generate a general dictionary. Dictionary

size and all parameters including  $C$ ,  $a$ ,  $b$ , and  $c$  of Eqs. (8) and (9) are determined on the training data set. In addition, the dictionary was trained from randomly sampled 100,000-image patches using K-SVD, i.e., the size of each patch is  $8 \times 8$  pixels. Thus, the size of the training data was  $64 \times 100,000$  pixels. We performed the experiments until  $q$  was 20 because the blocking effects mainly occur when  $q$  was from 0 to 20.<sup>36</sup> The dictionary with the 512 atoms is used in our experiments. Figure 11 shows the generated dictionary from the training data. Figures 12 and 13 show the JPEG compressed images and their deblocked results of the *Barbara* and *Baboon* images, respectively, according to different quality values, i.e.,  $q$  is 1, 5, 10, 15, or 20. It can be observed that the lower  $q$  is, the more blocking artifacts



**Fig. 13** JPEG compressed images and their deblocked results of the *Baboon* image according to different quality values: (a) quality = 1, (b) quality = 5, (c) quality = 10, (d) quality = 15, (e) quality = 20, (f) the deblocked result of (a), (g) the deblocked result of (b), (h) the deblocked result of (c), (i) the deblocked result of (d), and (j) the deblocked result of (e).

**Table 3** Performance evaluation results from test images using the proposed and FoE prior-based methods.<sup>a</sup>

Image	Quality	Metric	JPEG	FoE-based method (Ref. 10)	Our method	Ours + NQCS (Ref. 47)
<i>Barbara</i>	$q = 11$	PSNR	26.0311	26.7018	26.8194	<b>26.9108</b>
		SSIM	0.7761	0.7998	0.7966	<b>0.8081</b>
	$q = 9$	PSNR	25.5054	26.2071	26.3213	<b>26.4032</b>
		SSIM	0.7466	0.7780	0.7732	<b>0.7846</b>
	$q = 5$	PSNR	24.0165	24.4042	24.981	<b>25.0092</b>
		SSIM	0.6579	0.6751	0.7121	<b>0.7172</b>
<i>Lena</i>	$q = 11$	PSNR	30.7633	31.9666	31.9513	<b>31.9696</b>
		SSIM	0.8271	0.8626	0.8627	<b>0.8641</b>
	$q = 9$	PSNR	29.9766	<b>31.3018</b>	31.2704	31.2902
		SSIM	0.8069	0.8515	0.8506	<b>0.8521</b>
	$q = 5$	PSNR	27.319	27.7019	28.8602	<b>28.8650</b>
		SSIM	0.7394	0.7620	0.8065	<b>0.8070</b>
<i>Boat</i>	$q = 11$	PSNR	28.4561	<b>29.4076</b>	29.3438	29.3956
		SSIM	0.77	0.7979	0.7937	<b>0.7997</b>
	$q = 9$	PSNR	27.7544	<b>28.7647</b>	28.6988	28.7522
		SSIM	0.7441	0.7789	0.7721	<b>0.7794</b>
	$q = 5$	PSNR	25.4801	25.8192	26.6330	<b>26.6683</b>
		SSIM	0.6514	0.6708	0.6976	<b>0.7030</b>
<i>Peppers</i>	$q = 11$	PSNR	30.7451	<b>32.0341</b>	31.8679	31.8787
		SSIM	0.7951	<b>0.8356</b>	0.8324	0.8328
	$q = 9$	PSNR	30.011	<b>31.4735</b>	31.3072	31.3119
		SSIM	0.7761	<b>0.8276</b>	0.8239	0.8238
	$q = 5$	PSNR	27.4385	27.8965	<b>29.1197</b>	29.1138
		SSIM	0.7074	0.7339	<b>0.7864</b>	0.7854
<i>Baboon</i>	$q = 11$	PSNR	24.5851	24.9784	25.0323	<b>25.0877</b>
		SSIM	0.6891	0.6833	0.6789	<b>0.6949</b>
	$q = 9$	PSNR	24.048	24.4971	24.5315	<b>24.5883</b>
		SSIM	0.6535	0.6517	0.6424	<b>0.6604</b>
	$q = 5$	PSNR	22.3936	22.5909	23.0026	<b>23.0460</b>
		SSIM	0.5245	0.5356	0.5217	<b>0.5389</b>

**Table 3** (Continued).

Image	Quality	Metric	JPEG	FoE-based method (Ref. 10)	Our method	Ours + NQCS (Ref. 47)
Fruits	$q = 11$	PSNR	30.1973	<b>31.4000</b>	31.3322	31.3977
		SSIM	0.7961	0.8391	0.8378	<b>0.8414</b>
	$q = 9$	PSNR	29.4625	30.7641	30.7147	<b>30.7725</b>
		SSIM	0.7758	0.8275	0.8262	<b>0.8294</b>
	$q = 5$	PSNR	27.0479	27.5133	28.5934	<b>28.623</b>
		SSIM	0.7043	0.7297	0.7819	<b>0.7829</b>

<sup>a</sup>In the FoE prior-based method,<sup>10</sup> the results combined with the NQCS method<sup>47</sup> are reported. The **bold** numbers represent the best PSNR and SSIM values of each image at each quality. The unit of PSNR is dB.

occur along block boundaries in the compressed images. This is because transform coefficients of blocks are quantized independently in BDCT based image compression. As can be seen in (a)-(e) of the figures, the blocking artifacts are degrading the quality of picture seriously. In addition, the blocking artifacts are remarkably reduced as the quality increases. In the figures, (f)-(j) show the reduction results of the blocking artifacts by the proposed method. It can be observed that the proposed method suppresses the blocking artifacts efficiently and improves the picture quality, especially along block boundaries where the block discontinuities are severe.

To provide more reliable performance evaluation of the results, we compare our method with the latest state-of-the-art one which is based on the FoE prior.<sup>10</sup> It has been reported that the method has achieved the best deblocked results in terms of PSNR. As evaluation metrics, the PSNR and SSIM are considered to measure the quality of the estimated images. To simulate various types of BDCT compression, three quantization tables, usually denoted as Q1, Q2, and Q3, have been commonly used by many researchers.<sup>10,34</sup> The Q1, Q2, and Q3 tables correspond to a medium to high compression level, similar to what can

be obtained by using JPEG with  $q = 11$ ,  $q = 9$ , and  $q = 5$ , respectively.<sup>9</sup> Accordingly, in our experiments, the values of  $q$  are used instead of the quantization tables when the performance of our method is evaluated because our method is based on the quality information. Table 3 lists the PSNR and SSIM values of the deblocked results obtained by the FoE prior-based method and ours. In the FOE prior-based method,<sup>10</sup> the FoE prior captures the statistics of natural images, and thus, has been effectively employed for image denoising and inpainting.<sup>45,46</sup> The FOE prior has been successfully applied to deblocking of BDCT compressed images.<sup>10</sup> We have obtained the corresponding software for evaluation at <http://www.cs.brown.edu/~dqsun/research/software.html>. In the experiments, the FoE filter size is  $5 \times 5$  pixels and the maximum number of iterations is 200. In the FoE prior-based method,<sup>10</sup> the narrow quantization constraint set (NQCS)<sup>47</sup> have been used for the higher PSNR gain of deblocked results, and thus we also report the improved PSNR values by NQCS (see the 7th column). Combined with the NQCS method,<sup>47</sup> our method generally achieves the best PSNR and SSIM results about the test images. In the table, the bold numbers represent the best PSNR and SSIM values of each image at each quality.

**Table 4** Mobile phone trends in 5-year intervals.<sup>58</sup>

Year	1995	2000	2005	2010	2015
Cellular generation	2 G	2.5-3 G	3.5 G	Pre-4 G	4 G
Cellular standards	GSM	GPRSUMTS	HSPA	HSPALTE	LTELTE-A
Downlink bitrate (Mb/s)	0.01	0.1	1	10	100
Battery capacity (Wh)	1	2	3	4	5
Phone CPU Clock (MHz)	20	100	200	500	1000
Phone CPU Power (W)	0.05	0.05	0.1	0.2	0.3
Workload (GOPS)	0.1	1	10	100	1000
#Programmable cores	1	2	4	8	16

## 5 Practical Considerations for Mobile Applications

Currently, high-end mobile phones, which are usually referred to as smartphones, support multiple radio standards and a rich suite of applications including advanced radio, audio, video, and graphics processing. They provide more advanced computing ability and connectivity than contemporary feature phones using multiple chips such as a base-band processor and an application processor. Moreover, it is expected that new functionalities are being added to smartphones at an increasing rate; however, the increases in battery capacity have not matched increases in functionality.<sup>58–62</sup> In fact, battery capacities have not been growing more than 10% every year, whereas the number of features and applications.<sup>59</sup> Thus, the needs for low power and high performance are growing at a significantly higher rate. As listed in Table 4, the present workload of a 3.5 G smartphone amounts to nearly 100 giga operations per second (GOPS). This workload increases at a steady rate, roughly by an order of magnitude every 5 years. The workload is partitioned by application processing, radio processing, media processing, and 3D graphics. Among them, about 60% of the workload is used for radio and application processing. More than 30% of the workload is assigned to media processing including the functions such as display processing, camera processing, video decoding, and encoding. Here, video encoding requires the most amount of operations, i.e., 17 GOPS. In the workload for media processing, 10 GOPS is available, and thus two new functions (e.g., face hallucination and image deblocking) can be realized using it. Recently, the multicore architecture for mobile applications has been proposed to support a workload of 100 GOPS with 1 W.<sup>58</sup> We believe the multicore architecture can be effectively employed for implementing the new functions.

Another way to implement them is to use the graphics processing units (GPU)-based parallelization technology. Fortunately, due to the strong computational locality of video processing algorithms, video processing is highly amenable to parallel processing. Such locality makes it possible to divide video processing tasks into smaller, weakly interacting pieces for parallel computing.<sup>63</sup> The GPU-based parallelization technology drastically reduces the amount of operations, and thus, effective parallel architectures and programming also can be used to implement the new functions for mobile applications.

## 6 Conclusions

In this article, we provided two core technologies for high-quality image communications from the point of view of image processing: face hallucination and compression artifact reduction. The technologies have a close relation to inverse problems in image processing, and thus, we have described recent studies and our related research results to deal with the inverse problems effectively. When image data are transmitted over mobile communication networks, data loss inevitably occurs in the high frequency components of images because of lossy compression techniques. Thus, the quality of facial regions (i.e., main interests of image communications) is reduced and several compression artifacts inevitably occur. We have demonstrated that convex optimization and sparse representation can be effectively employed for solving the inverse problems and achieving high-quality image communications. In addition, to

implement the technologies in actual mobile devices, power management is a critical issue due to the limited capacity of batteries. Therefore, this article also discusses practical considerations and possible solutions to implement two technologies in mobile applications.

Nowadays, displays of many different sizes, including mobile displays, have come into wide use. They also have the same problems of high-quality image reconstruction. We believe the two technologies can be effectively employed for enhancing image quality in various displays.

## Acknowledgments

The authors would like to thank all the anonymous reviewers for their valuable comments and useful suggestions on this paper. This work was supported by the National Natural Science Foundation of China (Nos. 61050110144, 60803097, 60972148, 60971128, 60970066, 61072106, 61075041, 61003198, 61001206, and 61077009), the National Research Foundation for the Doctoral Program of Higher Education of China (No. 200807010003 and 20100203120005), the National Science and Technology Ministry of China (Nos. 9140A07011810DZ0107 and 9140A07021010DZ0131), the Key Project of Ministry of Education of China (No. 108115), and the Fundamental Research Funds for the Central Universities (Nos. JY10000902001, K50510020001, and JY10000902045).

## References

1. S. Baker and T. Kanade, "Hallucinating faces," in *Proc. IEEE Int. Conf. Automatic Face and Gesture Recogn.*, pp. 83–88, IEEE, Grenoble (2000).
2. S. Baker and T. Kanade, "Limits on super-resolution and how to break them," *IEEE Trans. Pattern Anal. Machine Intell.* **24**(9), 1167–1183 (2002).
3. X. G. Wang and X. O. Tang, "Hallucinating face by eigen-transformation," *IEEE Trans. Sys. Man Cybernetics- C* **35**(3), 425–434 (2005).
4. C. Liu, H. Y. Shum, and W. T. Freeman, "Face hallucination: theory and practice," *Int. J. Computer Vis.* **75**(1), 115–134 (2007).
5. X. Ma, J. Zhang, and C. Qi, "Hallucinating face by position-patch," *Pattern. Recogn.* **43**(6), 2224–2236 (2010).
6. X. Ma, J. Zhang, and C. Qi, "Position-based face hallucination method," in *Proc. IEEE Conf. Multimedia and Expo.*, pp. 290–293, IEEE, New York, NY (2009).
7. C. Jung et al., "Position-patch based face hallucination using convex optimization," *IEEE Signal Process. Lett.* **18**(6), 367–370 (2011).
8. C. Jung and L. C. Jiao, "Novel Bayesian deringing method in image interpolation and compression using a SGLI prior," *Opt. Express* **18**(7), 7138–7149 (2010).
9. A. Foi, V. Katkovnik, and K. Egiazarian, "Pointwise shape-adaptive DCT for high-quality denoising and deblocking of grayscale and color images," *IEEE Trans. Image Process.* **16**(5), 1395–1411 (2007).
10. D. Sun and W. K. Cham, "Postprocessing of low bit-rate block DCT coded images based on a fields of experts prior," *IEEE Trans. Image Process.* **16**(11), 2743–2751 (2007).
11. Y. Kim, C. S. Park, and S. J. Ko, "Fast POCS based postprocessing technique for HDTV," *IEEE Trans. Consumer Electron.* **49**(4), 1438–1447 (2003).
12. A. Gothandaraman, R. T. Whitaker, and J. Gregor, "Total variation for the removal of blocking effects in DCT based encoding," in *Proc. IEEE Conf. Image Process.*, Vol. 2, pp. 455–458, IEEE, Thessaloniki (2001).
13. F. Alter, S. Y. Durand, and J. Froment, "Adapted total variation for artifact free decomposition of JPEG images," *J. Math. Imaging Vis.* **23**(2), 199–211 (2005).
14. A. Mohammad-Djafari, "Bayesian inference for inverse problems in signal and image processing and applications," *Int. J. Imaging Sys. Appl.* **16**(5), 209–214 (2006).
15. H. H. Szu, "Inverse problem of image processing," *J. Math. Phys.* **25**(9), 2767–2772 (1984).
16. G. Wang, J. Zhang, and G. W. Pan, "Solution of inverse problems in image processing by wavelet expansion," *IEEE Trans. Image Process.* **4**(5), 579–593 (1995).
17. D. Rajan and S. Chaudhuri, "Generalized interpolation and its application in super-resolution imaging," *Image Vis. Comput.* **19**(13), 957–969 (2001).

18. S. Lertrattanapanich and N. K. Bose, "High resolution image formation from low resolution frames using Delaunay triangulation," *IEEE Trans. Image Process.* **11**(12), 1427–1441 (2002).
19. H. Stark and P. Oskoui, "High-resolution image recovery from image-plane arrays, using convex projections," *J. Opt. Soc. Am.-A* **6**(11), 1715–1726 (1989).
20. R. R. Schulz and R. L. Stevenson, "Extraction of high-resolution frames from video sequences," *IEEE Trans. Image Process.* **5**(6), 996–1011 (1996).
21. M. Irani and S. Peleg, "Super resolution from image sequences," in *Proc. Int. Conf. Pattern Recogn.*, Vol. 2, pp. 115–120, IEEE, Atlantic City, NJ (1990).
22. M. Irani and S. Peleg, "Improving resolution by image registration," *CVGIP Graphical Models Image Process.* **53**(3), 231–239 (1991).
23. N. Nguyen, P. Milanfar, and G. Golub, "Efficient generalized cross-validation with applications to parametric image restoration and resolution enhancement," *IEEE Trans. Image Process.* **10**(9), 1299–1308 (2001).
24. M. Elad and A. Feuer, "Restoration of a single super resolution image from several blurred, noisy, and undersampled measured images," *IEEE Trans. Image Process.* **6**(12), 1646–1658 (1997).
25. W. T. Freeman, T. R. Jones, and E. C. Pasztor, "Example-based super-resolution," *IEEE Comput. Graph. Appl.* **22**(2), 56–65 (2002).
26. H. Chang, D. Y. Yeung, and Y. Xiong, "Super-resolution through neighbor embedding," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, Vol. 1, pp. 1-275–1-282, IEEE (2004).
27. C. Liu, H. Shum, and C. Zhang, "A two-step approach to hallucinating faces: Global parametric model and local nonparametric model," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, Vol. 1, pp. 1-192–1-198, IEEE (2001).
28. K. Jia and S. G. Gong, "Generalized face super-resolution," *IEEE Trans. Image Process.* **17**(6), 873–886 (2008).
29. J. Yang et al., "Face hallucination via sparse coding," in *Proc. IEEE Conf. Image Process.*, pp. 1264–1267, IEEE, San Diego, CA (2008).
30. J. Yang et al., "Image super-resolution as sparse representation of raw image patches," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, pp. 1–8, IEEE, Anchorage, AK (2008).
31. P. S. Penev and L. Sirovich, "The global dimensionality of face space," in *Proc. IEEE Conf. Automatic Face and Gesture Recogn.*, pp. 264–270, IEEE, Grenoble (2000).
32. Z. Wang et al., "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.* **13**(4), 600–612 (2004).
33. Y. Luo and R. K. Ward, "Removing the blocking artifacts of block-based DCT compressed images," *IEEE Trans. Image Process.* **12**(7), 838–842 (2003).
34. A. W. C. Liew and H. Yan, "Blocking artifacts suppression in block-coded images using overcomplete wavelet representation," *IEEE Trans. Circuits Sys. Vid. Technol.* **14**(4), 450–461 (2004).
35. S. Singh, V. Kuamr, and H. K. Verma, "Reduction of blocking artifacts in JPEG compressed images," *Dig. Sign. Process.* **17**(1), 225–243 (2007).
36. B. Jeon and J. Jeong, "Blocking artifacts reduction in image compression with block boundary discontinuity criterion," *IEEE Trans. Circuits Sys. Vid. Technol.* **8**(3), 345–357 (1999).
37. G. Raja and M. J. Mirza, "In-loop deblocking filter for H.264/AVC video," in *Proc. Int. Sym. Commun., Control Sign. Process.*, EURASIP (2006).
38. H. C. Reeve and J. S. Lim, "Reduction of blocking effect in image coding," *Opt. Eng.* **23**(1), 34–37 (1984).
39. B. Ramamurthi and A. Gersho, "Nonlinear space-variant postprocessing of block coded images," *IEEE Trans. Acoustics, Speech, Sign. Process.* **34**(5), 1258–1268 (1986).
40. P. List et al., "Adaptive deblocking filter," *IEEE Trans. Circuits Sys. Vid. Technol.* **13**(7), 614–619 (2003).
41. Z. Xiong, M. Orchard, and Y. Q. Zhang, "A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations," *IEEE Trans. Circuits Sys. Vid. Technol.* **7**(2), 433–437 (1997).
42. R. E. Rosenholtz and A. Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding," *Proc. SPIE* **1452**, 116–126 (1991).
43. F. Alter, S. Y. Durand, and J. Froment, "Deblocking DCT-based compressed images with weighted total variation," in *Proc. IEEE Conf. Acoustics, Speech, Sign. Process.*, pp. 221–224, IEEE (2004).
44. Q. B. Do, A. Beghdadi, and M. Luong, "A new adaptive image post-treatment for deblocking and deringing based on total variation method," in *Proc. ISSPA*, pp. 464–467, IEEE, Kuala Lumpur (2010).
45. S. Roth and M. J. Black, "Field of experts: Aa framework for learning image priors," in *Proc. IEEE Conf. Comput. Vis. Pattern Recogn.*, Vol. 2, pp. 860–867, IEEE (2005).
46. S. Roth and M. J. Black, "Fields of experts," *Int. J. Comput. Vis.* **82**(2), 205–229 (2009).
47. S. H. Park and D. S. Kim, "Theory of projection onto the narrow quantization constraint set and its application," *IEEE Trans. Image Process.* **8**(10), 1361–1373, (1999).
48. C. Jung et al., "Image deblocking via sparse representation," *Sign. Process. Image Commun.* **27**(6), 663–677 (2012).
49. J. Wright et al., "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.* **31**(2), 210–227 (2009).
50. K. Huang and S. Aviyente, "Sparse representation for signal classification," *Adv. Neur. Info. Process. Sys.* **19**, 609–616 (2006).
51. J. Yang et al., "Image super-resolution via sparse representation," *IEEE Trans. Image Process.* **19**(11), 2861–2873 (2010).
52. M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.* **15**(12), 3736–3745 (2006).
53. M. Aharon, M. Elad, and A. Bruckstein, "The KSVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. Sign. Process.* **54**(11), 4311–4322 (2006).
54. R. Rubinstein, M. Zibulevsky, and M. Elad, "Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit," *CS Technion* (2008).
55. J. M. D. Carvajalino and G. Sapiro, "Learning to sense sparse signals: Simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.* **18**(7), 1395–1408 (2009).
56. M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: Design of dictionaries for sparse representation," *Proc. SPARSE*, 9–12 (2005).
57. R. Yang and M. Ren, "Learning overcomplete dictionaries with application to image Denoising," in *Proc. Int. Sym. Photon. Optoelectron.*, pp. 1–4, IEEE, Chengdu (2000).
58. C. H. Berkel, "Multi-core for mobile phones," in *Proc. Conf. Design, Automation and Test in Europe*, ACM (2009).
59. H. Falaki, R. Govindan, and D. Estrin, "Smart screen management on mobile phones," *Tech. Rep. Center Embedded Networked Sensing* (2009).
60. H. Kim and I. C. Park, "High-performance and low-power memory-interface architecture for video processing applications," *IEEE Trans. Circuits Sys. Vid. Technol.* **11**(11), 1160–1170 (2011).
61. T. H. Meng et al., "Low-power signal processing system design for wireless applications," *IEEE Personal Commun.* **5**(3), 20–31 (1998).
62. T. C. Chen et al., "Fast algorithm and architecture design of low-power integer motion estimation for H.264/AVC," *IEEE Trans. Circuits Sys. Vid. Technol.* **17**(5), 568–577 (2007).
63. D. Lin et al., "Parallelization of video processing," *IEEE Sign. Process. Mag.* **26**(6), 103–112 (2009).



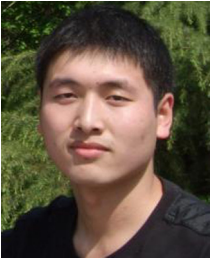
**Cheolkon Jung** received the BS, MS, and PhD degrees in electronic engineering from Sungkyunkwan University, Republic of Korea, in 1995, 1997, and 2002, respectively. He is currently a professor at Xidian University, China. His main research interests include computer vision, pattern recognition, image and video processing, multimedia content analysis and management, and 3D TV.



**Licheng Jiao** received the BS degree from Shanghai Jiao Tong University, China, in 1982, and the MS and PhD degrees from Xian Jiao Tong University, China, in 1984 and 1990, respectively. From 1990 to 1991, he was a postdoctoral fellow in the National Key Lab for Radar Signal Processing at Xidian University, China. Since 1992, he has been with the School of Electronic Engineering at Xidian University, China, where he is currently a distinguished professor. He is the dean of the School of Electronic Engineering and the Institute of Intelligent Information Processing at Xidian University, China. His current research interests include signal and image processing, nonlinear circuit and systems theory, learning theory and algorithms, computational vision, computational neuroscience, optimization problems, wavelet theory, and data mining.



**Bing Liu** received the BS degree in electronic engineering from Henan Polytechnic University, China, in 2009. He is currently pursuing the MS degree in Xidian University, China. His research interests include image processing and machine learning.



**Hongtao Qi** received the BS degree in electronic engineering from Xidian University, China, in 2009. He is currently pursuing the MS degree in the same university. His research interests include image processing and 3D TV.



**Tian Sun** received the BS degree in electronic engineering from Xidian University, China, in 2009. He is currently pursuing the MS degree in the same university. His research interests include computer vision and pattern recognition.