SINGLE COLOR IMAGE SUPER-RESOLUTION USING QUATERNION-BASED SPARSE REPRESENTATION

Mengqi Yu, Yi Xu, Peng Sun

Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China
Shanghai Key Lab of Digital Media Processing and Transmations
xuyi@sjtu.edu.cn

ABSTRACT

In current color image super-resolution methods, super-resolution based on sparse representation achieves state-of-the-art performance. However, the exploited sparse representation models deal with the color images as independent channel planes. Consequently, these approaches process the color pixels as scalar quantity, lacking of accuracy in describing inter-relationship among color channels. In this paper, we propose a quaternion-based online dictionary learning method and solve color image super-resolution by employing a quaternion-based sparse representation model. This sparse representation model implements color image super-resolution in a kind of vectorial reconstruction, effectively accounting for both luminance and chrominance geometry in images. The proposed color image super-resolution method can better describe the inter-channel changes. In the case that changing lighting conditions affect color more than the luminance perception, it can obtain superior performance comparing to the methods based on monochromatic sparse models with 1dB improvement.

Index Terms— Quaternion, super-resolution, sparse representation, dictionary learning, PCA, OMP

1. INTRODUCTION

Image SR problem is a highly ill-posed inverse problem since many HR images may produce the same LR image when blurred and down-sampled. Therefore, some prior knowledge about the decimation model is necessary for the solution of the image SR problem. Typically, there are three kinds of SR approaches, i.e. interpolation-based methods, reconstruction-based methods and learning-based methods. Learning-based methods usually achieve better visual quality than the other two categories of approaches, since learning-based methods have more redundant information available with the help of sample dataset.

Recently, a sparse prior has been employed in the learning-based SR methods[1, 2], which achieves state-of-the-art performance. There are two kinds of color image sparse models used in SR tasks: (1) The color image is separated into three channel images and then the sparse representation is enforced on each image independently. (2) The three channels of the color image are concatenated and then the sparse representation is processed for this generated monochromatic image. Both approaches consider no constraints among the color channels. Therefore, color bias would be introduced[3].

In this paper, we propose a quaternion-based sparse prior model for single color image SR, which formulates a color pixel as a quaternion unit and thus processes multichannel information in a parallel way. In essence, SR image is reconstructed as a vectorial operation between the color atoms in the learned quaternion dictionary and sparse quaternion coefficients. The experimental results demonstrate that this sparse representation model can better describe the inter-channel changes, especially under the cases that changing lighting conditions affects color more than the luminance perception. In such cases, the reconstructed SR image from the quaternion-based sparse model can achieve image quality improvement of 1dB as compared with the images from the monochromatic sparse models.

This work was supported by National Nature Science Foundation of China 60932006, 61025005, the China National Key Technology R&D Program 2012BAH07B01, the National Basic Research Program 2010CB731400 and 111 plan B07022.

2. QUATERNION-BASED SPARSE REPRESENTATION MODEL

We employ quaternion to represent color pixels and process three channels in a parallel way. The proposed sparse representation model is,

$$\dot{\mathbf{p}} = \dot{\mathbf{D}}\dot{\alpha} \tag{1}$$

where $\dot{\mathbf{p}} = \mathbf{p}_r \cdot i + \mathbf{p}_g \cdot j + \mathbf{p}_b \cdot k$ denotes a quaternion represented color image patch, $\dot{\mathbf{D}} = \mathbf{D}_r \cdot i + \mathbf{D}_g \cdot j + \mathbf{D}_b \cdot k$ is the learned quaternion dictionary and $\dot{\alpha} = \alpha_0 + \alpha_1 \cdot i + \alpha_2 \cdot j + \alpha_3 \cdot k$ is the corresponding sparse coefficient vector. We expand (1) using quaternion algebra operation and get,

$$\begin{aligned} 0 &= \mathbf{D}_r \cdot \alpha_1 + \mathbf{D}_g \cdot \alpha_2 + \mathbf{D}_b \cdot \alpha_3 \\ \mathbf{p}_r &= \mathbf{D}_r \cdot \alpha_0 + \mathbf{D}_g \cdot \alpha_3 - \mathbf{D}_b \cdot \alpha_2 \\ \mathbf{p}_g &= -\mathbf{D}_r \cdot \alpha_3 + \mathbf{D}_g \cdot \alpha_0 + \mathbf{D}_b \cdot \alpha_1 \\ \mathbf{p}_b &= \mathbf{D}_r \cdot \alpha_2 - \mathbf{D}_g \cdot \alpha_1 + \mathbf{D}_b \cdot \alpha_0 \end{aligned} \tag{2}$$

From (2), we observe that the three channel images are uniformly represented using three channel dictionaries $\mathbf{D_r}, \mathbf{D_g}$ and $\mathbf{D_b}$, which are linearly related with each other. By training the quaternion dictionary $\dot{\mathbf{D}}$ in a proper way, the interrelationship of the three channels

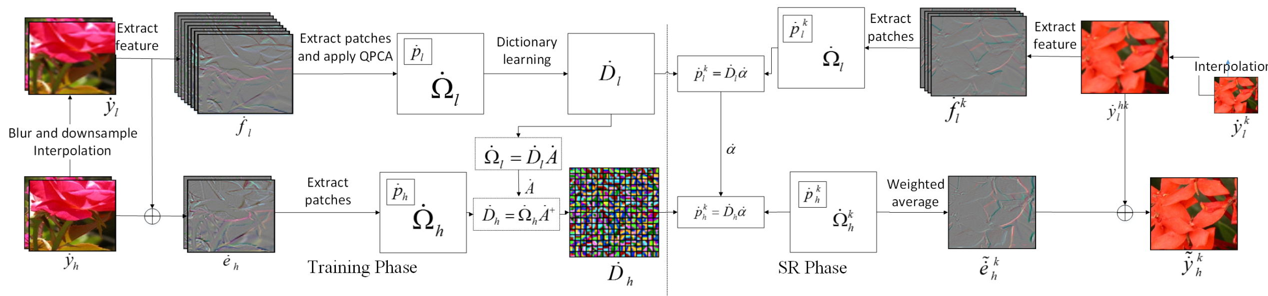

Fig. 1. Flow chart of proposed method

for color patches can be well preserved. In contrast, current color image sparse models represented each channel image using independent dictionaries, providing no assurance of channel correlation in the reconstruction.

The quaternion dictionary $\dot{\mathbf{D}}$ contains a certain amount of prototype atoms as its columns, which is denoted as $\dot{\mathbf{d}}_j$. Then a color image patch $\dot{\mathbf{p}}$ can be represented as a sparse linear combination of these atoms,

$$\dot{\mathbf{p}} = \sum_j \dot{\mathbf{d}}_j \dot{\alpha}_\mathbf{j} \quad j = 1, 2, \ldots, K \tag{3}$$

where K is the size of $\dot{\mathbf{D}}$, and $\dot{\alpha}_j$ is the j^{th} element of vector $\dot{\alpha}$. We formulate $\dot{\alpha}_j$ as the composite of a scalar part and a vector part by writing $\dot{\alpha}_j = (a_0, a_1, a_2, a_3) = [S(\dot{\alpha}_j), V(\dot{\alpha}_j)]$, where $S(\dot{\alpha}) = a_0$ and $V(\dot{\alpha}) = \{a_1, a_2, a_3\}$. Similarly, we formulate element i of color atom $\dot{\mathbf{d}}_j$ as $\dot{d}_j^{(i)} = (0, d_r, d_g, d_b) = \left[0, V\left(\dot{d}_j^{(i)}\right)\right]$, then get

$$\begin{aligned} \mathrm{S}\left(\dot{d}_j^{(i)} \times \dot{\alpha}_j\right) &= -V\left(\dot{d}_j^{(i)}\right) \circ V\left(\dot{\alpha}_j\right) \\ V\left(\dot{d}_j^{(i)} \times \dot{\alpha}_j\right) &= S\left(\dot{\alpha}_j\right) V\left(\dot{d}_j^{(i)}\right) + V\left(\dot{d}_j^{(i)}\right) \otimes V\left(\dot{\alpha}_j\right) \end{aligned} \tag{4}$$

In (4), $S(\cdot)$ and $V(\cdot)$ extract the scalar part and the vector part of a quaternion, respectively. Symbol '$\circ$' denotes dot product operator and '$\otimes$' denotes cross product operator of two vectors. It should be noted that color image patch is sparsely represented as a kind of vectorial operations in a non-commutative way.

3. SUPER-RESOLUTION USING QUATERNION-BASED SPARSE MODEL

3.1. Formation of training data set

Many previous works[4, 5, 6] suggest that distinct features from the LR image are very important to accurately predict the HR image. Similar to the works in [1, 2], we choose the first order and the second order gradient operators to obtain distinct features because of their simplicity and effectiveness. Four gradient operators are involved in the feature extraction,

$$\begin{aligned} f_1 &= [-1, 0, 1] \quad f_2 = f_1^T \\ f_3 &= [1, 0, -2, 0, 1] \quad f_4 = f_3^T \end{aligned} \tag{5}$$

where subscript 'T' denotes the transpose operator of a matrix. As shown in Fig. 1, the implementation details of color image superresolution can be summarized as follows,

1) We convolve these four filters with the three color channels separately and then formulate the extracted gradient maps as quaternion matrices. These four gradient maps are concatenated as one gradient map.
2) We extract image patch pairs from the HR-LR gradient image pair and denote them as $\dot{\mathbf{p}}_l \in \mathbb{Q}^n$ and $\dot{\mathbf{p}}_h \in \mathbb{Q}^m$, $m > n$. After extracting enough training image patches, we obtain a training set pair of $\dot{\Omega}_{l0} \in \mathbb{Q}^{n \times L}$ and $\dot{\Omega}_{h0} \in \mathbb{Q}^{m \times L}$
3) We apply Quaternion Principal Component Analysis (QPCA)[7] to the LR data set $\dot{\Omega}_{l0}$ to obtain a subspace $\dot{\Omega}_l$ which can preserve more than 99.9% of the total variance. First, we multiply $\dot{\Omega}_{l0}$ with $\dot{\Omega}_{l0}^H$, noting that superscript 'H' means conjugate transpose operation. Then a Hermitian matrix $\dot{S} = \dot{\Omega}_{l0}\dot{\Omega}_{l0}^H$ is obtained, $\dot{S} \in \mathbb{Q}^{n \times n}$. Therefore, we can find the eigenvalue matrix $E \in \mathbb{R}^{n \times n}$ with the corresponding eigenvectors $\dot{V} \in \mathbb{Q}^{n \times n}$ which satisfy $\dot{S} = \dot{V}E\dot{V}^H$. Next, we preserve the top q largest eigenvalues and their corresponding eigenvectors $\dot{V}^{QPCA} \in \mathbb{Q}^{n \times q}$. We then obtain $\dot{\Omega}_l \in \mathbb{Q}^{q \times L}$ by

$$\dot{\Omega}_l = \left(\dot{V}^{QPCA}\right)^H \dot{\Omega}_{l0} \tag{6}$$

3.2. Quaternion orthogonal matching pursuit

The QOMP algorithm solves the problem of decomposing signal $\dot{\mathbf{p}} \in \mathbb{Q}^q$ on a quaternion dictionary $\dot{\mathbf{D}} \in \mathbb{Q}^{q \times N}$ satisfying either of the equation (7) or (8)

$$\dot{\alpha} = \arg\min_{\dot{\alpha} \in \mathbb{Q}} \|\dot{\mathbf{p}} - \dot{\mathbf{D}}\dot{\alpha}\|_2^2 \quad s.t. \ \|\dot{\alpha}\|_0 \leq K \tag{7}$$

$$\dot{\alpha} = \arg\min_{\dot{\alpha} \in \mathbb{Q}} \|\dot{\alpha}\|_0 \quad s.t. \ \|\dot{\mathbf{p}} - \dot{\mathbf{D}}\dot{\alpha}\|_2^2 \leq \epsilon \tag{8}$$

where $\dot{\alpha} \in \mathbb{Q}^N$ are the sparse coefficient vector. ε and K are two types of stopping criteria. The implementation details of QOMP can be summarized as,

1) We initialize the residual $\dot{\varepsilon}$ as the signal $\dot{p}$ itself, quaternion dictionary as an empty set.
2) At the k^{th} iteration, QOMP selects the atom that produces the absolute largest decrease in the mean square error $\left\|\dot{\varepsilon}^{(k-1)}\right\|_2^2$. In implementation, we compute the correlation between the residual signal $\dot{\varepsilon}$ and each atom $\dot{\mathbf{d}}_m$ in dictionary, i.e. $C_m^{(k)} = \dot{\varepsilon}^{(k-1)}\dot{\mathbf{d}}_m^H$. Then select the atom which achieves the highest correlation value and record its index as $m^{(k)}$. We add the index $m^{(k)}$ to an index array $\mathbf{M}$ and label $\dot{\mathbf{D}}^{(k)}$ as the active atoms.

Algorithm 1 Quaternion Online Dictionary Learning

Require:
Training set $\dot{\Omega}_l \in \mathbb{Q}^{q\times K}$, error tolerance ε, iteration times T, mini-batch size η, data reduction parameter ρ;
$\dot{\mathbf{A}}_0 \leftarrow 0$,$\dot{\mathbf{B}}_0 \leftarrow 0$
Initialize the dictionary $\dot{\mathbf{D}}_l$ with patched in $\dot{\Omega}$
for $t = 1$ *to* T **do**
  Randomly select a mini-batch $\dot{\psi}$ with η patches from $\dot{\Omega}$ and $\dot{\mathbf{p}}_l \in \dot{\psi}$
  Sparse coding: use QOMP to solve

$$\dot{\alpha}_t = \arg\min_{\dot{\alpha}\in\mathbb{Q}} \|\dot{\alpha}\|_0 \quad s.t.\ \|\dot{\mathbf{p}}_l^{(t)} - \dot{\mathbf{D}}_l^{(t-1)}\dot{\alpha}^{(t)}\|_2^2 \le \varepsilon \tag{9}$$

  $\beta_t = \left(1 - \frac{1}{t}\right)^{\rho}$
  $\dot{\mathbf{A}}^{(t)} \leftarrow \beta^{(t)}\dot{\mathbf{A}}^{(t-1)} + \frac{1}{\eta}\sum_{i=1}^{\eta}\dot{\alpha}^{(t)}\left(\dot{\alpha}^{(t)}\right)^H$
  $\dot{\mathbf{B}}^{(t)} \leftarrow \beta^{(t)}\dot{\mathbf{B}}^{(t-1)} + \frac{1}{\eta}\sum_{i=1}^{\eta}\dot{\mathbf{p}}_l^{(t)}\left(\dot{\alpha}^{(t)}\right)^H$
  repeat
    for each column d_j in $\dot{\mathbf{D}}_l^{(t)}$ **do**

$$\begin{aligned}\dot{\mathbf{u}}_j &\leftarrow \frac{1}{\dot{A}_{jj}}\left(\dot{\mathbf{b}}_j - \dot{\mathbf{D}}_l^{(t)}\dot{\mathbf{a}}_j\right) + \dot{\mathbf{d}}_j \\ \dot{\mathbf{d}}_j &\leftarrow \frac{1}{max(\|\dot{\mathbf{u}}_j\|_2^2, 1)}\dot{\mathbf{u}}_j\end{aligned} \tag{10}$$

    end for
  until convergence
end for
return $\tilde{\dot{\mathbf{D}}}_l$;

Algorithm 2 Quaternion-based Super-resolution

Require:
Dictionary pair $\dot{\mathbf{D}}_h$ and $\dot{\mathbf{D}}_l$, a LR image y_l, high-pass filters G, non-zero coefficients number K
Interpolate the LR image y_l and obtain y_l^h
Extract the feature from LR image channels respectively by $f_l^c = y_l^c * G$, where $c = y, cb, cr$ and then forming f_l by setting the three imaginary parts as f_l^c respective and the real part zeros.
Select patches $\dot{\mathbf{P}}_l$ from f_l
for each patch $\dot{\mathbf{p}}_l^k$ in $\dot{\mathbf{P}}_l$ **do**
  Sparse coding: use QOMP to solve

$$\dot{\alpha}^k = \arg\min_{\dot{\alpha}^k\in\mathbb{Q}} \|\dot{\mathbf{p}}_l^k - \dot{\mathbf{D}}_l\dot{\alpha}^k\|_2^2 \quad s.t.\ \|\dot{\alpha}\|_0 \le K \tag{11}$$

  Synthesize the high frequency component in each patch via

$$\dot{\mathbf{e}}_h = \dot{\mathbf{D}}_h\dot{\alpha}^k \tag{12}$$

end for
Average the overlapping area and obtain $\tilde{\dot{e}}_h$.
Add the interpolated image with $\tilde{\dot{e}}_h$ to obtain HR image $\dot{y}_{h0} = \dot{y}_l^h + \tilde{\dot{e}}_h$.
Find the image $\dot{y}_h$ close to $\dot{y}_{h0}$ using back-projection

$$\tilde{\dot{y}}_h = \arg\min_{\dot{y}_h} \|\dot{y}_h - \dot{y}_{h0}\| \quad s.t.\ \dot{y}_l = SH\dot{y}_h \tag{13}$$

where S is a down-sampling operator and H is a blurring filter.
return $\tilde{\dot{y}}_h$;

3) Compute Coefficients: $\dot{\alpha}^{(k)} = \left(\dot{\mathbf{D}}^{(k)}\right)^+\dot{\mathbf{P}}$, where superscript '$+$' denotes quaternionic pseudo-inverse operation. Then update the residual signal $\dot{\varepsilon}^{(k)} = \dot{\mathbf{p}} - \dot{\mathbf{D}}^{(k)}\dot{\alpha}^{(k)}$.
Iterate step 2)-3) until the residual signal is no greater than a tolerance error bound, i.e. $\left\|\dot{\mathbf{p}} - \dot{\mathbf{D}}\dot{\alpha}\right\|_2^2 \le \varepsilon$ or the number of active atoms is no less than K, i.e. $\|\dot{\alpha}\|_0 \ge K$. Output $\dot{\alpha}^{(k)}$ as the solution of QOMP algorithm.

3.3. Quaternion online dictionary learning

The sparse coding problem in online dictionary learning (ODL) is l_1 norm convex optimization when fixing the dictionary and is solved by LARS[8]. The dictionary updating problem is also a convex optimization when fixing the coefficients and is solved by using a method based on block-coordinate descent. ODL has an advantage on memory usage and computation over ordinary batch methods such as K-SVD[9]. This online method is more suitable to deal with a large number of training samples. Therefore, we employ this method and develop a quaternion version of ODL with a modification of the sparse coding phase, but not relaxing l_0 pseudo-norm to l_1 norm as ODL does. We apply QOMP described in section 3.2 to the following non-convex optimization problems because of its efficiency. Subsequently, dictionary is updated using a similar approach described in [10]. The sparse coding and the dictionary updating step are implemented alternatively. The scheme is formulated in Algorithm 1. Note that $\beta^{(t)}$ is used in iterations to reduce the weight of previous data. $\dot{\mathbf{A}}^{(t)}$ and $\dot{\mathbf{B}}^{(t)}$ are used to carry the information about coefficients in iteration t. $\dot{\mathbf{a}}_j$ and $\dot{\mathbf{b}}_j$ denote the j^{th} column in $\dot{\mathbf{A}}^{(t)}$ and $\dot{\mathbf{B}}^{(t)}$ respectively. $\dot{A}_{jj}$ is the j^{th} element in the leading diagonal of $\dot{\mathbf{A}}^{(t)}$.

After finding an over-complete dictionary $\dot{\mathbf{D}}_l$ for LR training set $\dot{\mathbf{\Omega}}_l$, we apply QOMP to find the final sparse coefficients $\dot{\alpha}$ of $\dot{\Omega}_l$ with respect to $\dot{\mathbf{D}}_l$. The corresponding objective function is the same as (7). The HR dictionary should have the ability to sparse recover the HR training set as accurately as possible. Thus, the objective function can be formulated as

$$\begin{aligned}\dot{\mathbf{D}}_h &= \arg\min_{\dot{\mathbf{D}}_h\in\mathbb{Q}} \sum_k \|\dot{\mathbf{p}}_h^k - \dot{\mathbf{D}}_h\dot{\alpha}^k\|_2^2 \\ &= arg\min_{\dot{\mathbf{D}}_h\in\mathbb{Q}} \|\dot{\mathbf{P}}_h - \dot{\mathbf{D}}_h\dot{\mathbf{A}}\|_F^2\end{aligned} \tag{14}$$

where superscript 'k' locates the position of patch $\dot{\mathbf{p}}_h^k$, $\dot{\mathbf{P}}_h$ is the patch array of patch $\dot{\mathbf{p}}_h^k$ and $\dot{\mathbf{A}}$ is the coefficients array of $\dot{\alpha}$. Similar to the approach in [11], $\dot{\mathbf{D}}_h$ can be solved by

$$\dot{\mathbf{D}}_h = \dot{\mathbf{P}}_h\dot{\mathbf{A}}^+ \tag{15}$$

where $\dot{\mathbf{A}}^+$ means the quaternion pseudo-inverse matrix of $\dot{\mathbf{A}}$.

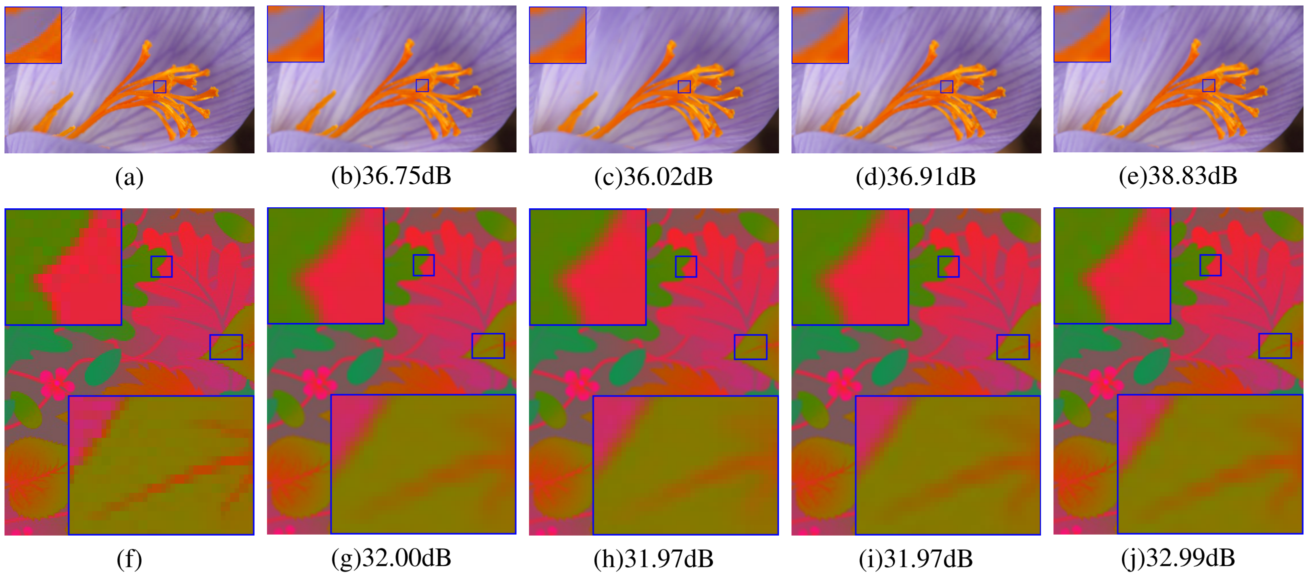

Fig. 2. Columns from left to right:gound truth,bicubic interpolation, Yang et al.[1, 2], Zeyde et al.[11], proposed method

Quaternion pseudo-inverse can be implemented using Quaternion Singular Value Decomposition depicted in [12].

3.4. Color image super-resolution

With an assumption that the sparse coefficient $\dot{\alpha}$ is shared for HR-LR image patch pair, i.e. $\dot{\mathbf{p}}_l^k = \dot{\mathbf{D}}_l\dot{\alpha}^k$ and $\dot{\mathbf{p}}_h^k = \dot{\mathbf{D}}_h\dot{\alpha}^k$, we first compute the sparse coefficients of LR image and then synthesize the high-frequency bands of the HR image. The weighted-averaging is conducted in the overlapping areas to enforce local smoothness on high-frequency bands, which is added with the interpolated LR image to generate an initial estimation of HR image. Finally, we use back-projection to remove the artifacts and further refine the SR results. We summarize our method in Algorithm 2.

4. EXPERIMENTAL RESULTS

The test images cover both generic images whose high-frequency bands are concentrated in luminance channel and full-color images which present significant edges in all color channels.

In all the experiments, the magnification factor is set as 3 and the size of the quaternion dictionary of size 256 is used. We use the training image set in the Yang's package which is available at `http://www.ifp.illinois.edu/~jyang29/`. Some commonly-used parameters of our method in the experiments are listed as follows. We choose patch size as 5×5 considering the effectiveness and the computation efficiency. Empirically, we choose the error tolerance ε of QOMP in the training phase as 0.05. We prune the training samples from 150,000 to about 100,000 by removing those samples with relatively small variance. The iteration number T in Algorithm 1 is set as 100. The sparsity parameter K is set as 15 in algorithm 2 for QOMP during the SR phase.

We compare our results with the typical interpolation method of bicubic and two typical SR methods using image sparse model, including the work of Yang et al.[1, 2], and the work of Zeyde et al.[11]. The parameters set in the works of Yang et al. can be referred to [1, 2]. The parameters set in the work of Zeyde et al. is suggested in [11]. We evaluate our experiment using both subjective visual perception and objective quality measurement.

images	bicubic	Yang	Zeyde	ours
baboon	21.01	21.36	21.38	**21.41**
barbara	24.67	24.86	25.17	**25.19**
lenna	28.27	29.02	29.02	**29.13**
monarch	28.18	29.40	29.47	**29.52**

Table 1. PSNR values of different methods.

Fig. 2 shows the results of abovementioned 4 methods. The works of Yang et al.[1, 2] and Zeyde et al.[11] cannot synthesize the desired sharp color edges, while the latter achieves a greater PSNR value than the former. The proposed SR method using quaternion-based sparse model synthesizes sharper edges and yields fewer artifacts. It outperforms the state-of-the-art methods [1, 2, 11] in effectively preserving both luminance and chrominance geometry in images. Some generic images which presents less color structures are also chosen in the test images. We observe from Table 1 that our method remains a competitive performance to Yang el al.[1, 2] and Zeyde et al.[11].

5. CONCLUSION

We propose a new color image super-resolution approach based on quaternion sparse representation. It implements SR task as a kind of vectorial signal reconstruction and thus avoid color bias problem. More specifically, the proposed method can obtain superior performance in full color images comparing to the methods based on monochromatic sparse models with 1dB improvement.

6. REFERENCES

[1] Jianchao Yang, John Wright, Thomas Huang, and Yi Ma, "Image super-resolution as sparse representation of raw image patches," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.

[2] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, "Image super-resolution via sparse representation," *Image Processing, IEEE Transactions on*, vol. 19, no. 11, pp. 2861–2873, 2010.

[3] Julien Mairal, Michael Elad, and Guillermo Sapiro, "Sparse representation for color image restoration," *Image Processing, IEEE Transactions on*, vol. 17, no. 1, pp. 53–69, 2008.

[4] Jian Sun, Nan-Ning Zheng, Hai Tao, and Heung-Yeung Shum, "Image hallucination with primal sketch priors," in *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*. IEEE, 2003, vol. 2, pp. II–729.

[5] Hong Chang, Dit-Yan Yeung, and Yimin Xiong, "Super-resolution through neighbor embedding," in *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. IEEE, 2004, vol. 1, pp. I–275.

[6] Wei Fan and Dit-Yan Yeung, "Image hallucination using neighbor embedding over visual primitive manifolds," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*. IEEE, 2007, pp. 1–7.

[7] Nicolas Le Bihan and Stephen J Sangwine, "Quaternion principal component analysis of color images," in *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*. IEEE, 2003, vol. 1, pp. I–809.

[8] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani, "Least angle regression," *The Annals of statistics*, vol. 32, no. 2, pp. 407–499, 2004.

[9] Michal Aharon, Michael Elad, and Alfred Bruckstein, "K-svd: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Transactions on*, vol. 54, no. 11, pp. 4311–4322, 2006.

[10] Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro, "Online learning for matrix factorization and sparse coding," *The Journal of Machine Learning Research*, vol. 11, pp. 19–60, 2010.

[11] Roman Zeyde, Michael Elad, and Matan Protter, "On single image scale-up using sparse-representations," in *Curves and Surfaces*, pp. 711–730. Springer, 2012.

[12] Licheng Yu, Yi Xu, Hongteng Xu, and Hao Zhang, "Quaternion-based sparse representation of color image," in *Multimedia and Expo (ICME), 2013 IEEE International Conference on*. IEEE, 2013, pp. 1–7.