Image Super Resolution via Combination of Two Dimensional Quaternion Valued Singular Spectrum Analysis Based Denoising, Empirical Mode Decomposition Based Denoising and Discrete Cosine Transform Based Denoising Methods

Yingdan Cheng
Telephone: +86 20 3932 2246 Fax: +86 20 3932 2252 Email: 1783675881@qq.com
Faculty of Information Engineering, Guangdong University of Technology, Guangzhou, 510006, China.
*Bingo Wing-Kuen Ling
Telephone: +86 20 3932 2246 Fax: +86 20 3932 2252 Email: yongquanling@gdut.edu.cn
Faculty of Information Engineering, Guangdong University of Technology, Guangzhou, 510006, China.
Yuxin Lin
Telephone: +86 20 3932 2246 Fax: +86 20 3932 2252 Email: a19860075121@163.com
Faculty of Information Engineering, Guangdong University of Technology, Guangzhou, 510006, China.
Ziyin Huang
Telephone: +86 20 3932 2246 Fax: +86 20 3932 2252 Email: shmillehzy@gmail.com
Faculty of Information Engineering, Guangdong University of Technology, Guangzhou, 510006, China.
Yui-Lam Chan
Telephone: +852 2766 6213 Fax: +852 2362 8439 Email: enylchan@polyu.edu.hk
Department of Electronic and Information Engineering, The Hong Kong Polytechnic University.

Abstract: This paper formulates the image super resolution problem as various denoising problems. In particular, the discrete cosine transform zero padding approach is used to generate an initial high resolution image. Then, three different time frequency analysis based denoising methods are applied iteratively to improve the quality of the super resolution image. In particular, the two dimensional quaternion valued singular spectrum analysis (2DQSSA) based denoising method, the empirical mode decomposition (EMD) based denoising method and the discrete cosine transform based denoised method are applied. For the 2DQSSA based denoising method, the luminance plane is used as the real part of the quaternion valued image. Since different color planes in the quaternion valued image are fused together via the quaternion valued operation, some high frequency information missing in one color plane can be generated using those in other color planes. On the other hand, for the EMD based denoising method, the selection of the intrinsic mode functions (IMFs) is formulated as a binary linear programming problem. Here, the high frequency components generated by the aliasing are removed by discarding some IMFs. The computer numerical simulation results show that our proposed method can achieve the super resolution performance better than those without performing any one of the above three time frequency analysis based denoising.

Keywords: Super resolution image, two dimensional quaternion valued singular spectrum analysis, empirical mode decomposition, discrete cosine transform, binary linear programming.

1. Introduction

For some applications such as the consumer electronic applications, the overall costs of the products are limited by the affordability of the consumers. Hence, the qualities of the image acquisition devices are low. Only the low resolution images can be obtained [2]. In this case, it is required to perform the image super resolution. That is, to construct the high resolution images from the low resolution images [1]. However, since the low resolution images are suffered from the aliasing, the information in the images is lost. Hence, performing the image super resolution is very challenging [3].

Traditional image super resolution mainly employs the interpolation based methods such as the nearest neighbor interpolation based methods, the bilinear interpolation based methods and the bicubic interpolation based methods [4] to generate the super resolution images. Although these methods are widely used in many science and engineering applications due to their low computational complexity, the blurred edges and the unclear textures are usually found in the obtained high resolution images due to the lack of high frequency details.

The sparse coding [5], [6] is one of the representative image super resolution methods. The overlapping patches are first obtained by cropping from the original image. Then, the cropped images are normalized. Next, the normalized patches are then encoded using a low resolution dictionary. The high resolution patches are then constructed using the high resolution dictionary based on the corresponding

sparse coefficients. In particular, the sparse coefficients are obtained via finding the solution of an optimization problem based on the given dictionary. Then, the efficient mapping functions are constructed [5]-[7].

The convolutional neural network based super resolution (SRCNN) is a deep learning based single image super resolution. The neural network mainly consists of three convolutional layers, a nonlinear layer mapping the high dimensional vectors to the new high dimensional vectors. Finally, a high resolution image is reconstructed [8]. Although the deep learning methods have also been recently proposed [9], the required computational power is very high. Also, the overfitting usually occurs. Hence, this approach is not robust.

To address the above problem, first the discrete cosine transform zero padding approach is used to generate an initial high resolution image. Now, the size of the image is increased. However, as the high frequency components are set to zero, the image quality is poor. To enhance the image quality, it becomes a denoising problem. This paper proposes three different time frequency analysis based denoising methods for performing the color image super resolution. In particular, the 2DQSSA [11], [12] based denoising method, the EMD [13] based denoising method and the discrete cosine transform [14] based denoising method are applied iteratively to obtain the final high quality super resolution images.

For the conventional quaternion valued image processing, the RGB color planes are usually taken as the imaginary parts of the quaternion valued image and the zero matrix is usually taken as the real part of the quaternion valued image. In fact, they are other color spaces such as the HIS color space, the HSV color space and the CMY color space. Since the HSV color space is mainly used for performing the image segmentation via separating the chromaticity and the intensity information from the image for enhancing the image contrast [10], this paper takes the luminance plane as the real part of the quaternion valued image.

This paper aims to generate the super resolution images with the high quality via formulating the super resolution problem as the denoising problem. The outline of this paper is as follows. Section 2 presents our proposed method. Section 3 shows the computer numerical simulation results. Finally, a conclusion is drawn Section 4.

2. Our proposed method

The flowchart of our proposed method is shown in Figure 1. The details of these procedures are discussed in the following sub-sessions:

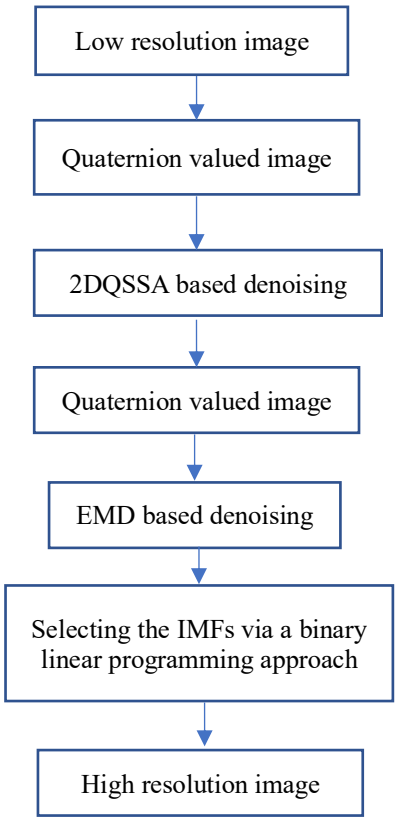


Figure 1. The flowchart of our proposed method.

2.1. Quaternion valued image

Let H be the set of quaternion valued number. Let $q \in H$ be a quaternion valued number. It has one real part and three imaginary parts. Let q_1 be its real component as well as q_2, q_3 and q_4 be its i imaginary component, its j imaginary component and its k imaginary component, respectively. That is,

$$q = q_1 + q_2 i + q_3 j + q_4 k. \quad (1)$$

Here, i, j and k obey the following rules of algebra:

$$i^2 = j^2 = k^2 = ijk = -1,$$
$$ij = -ji = k,$$
$$jk = -kj = i$$

and

$$ki = -ik = j. \quad (2)$$

A color image consists of three color planes. For the representation using the RGB space, they are the red color plane, the green color plane and the blue color plane. The luminance plane of the color image is a weighted sum of these three color planes. Let (x, y) be the location index of a pixel. Here, the pixel is located at the x^{th} row

and the y th column of the image. Let $l(x,y)$, $r(x,y)$, $g(x,y)$ and $b(x,y)$ be the luminance pixel value, the red color pixel value, the green color pixel value and the blue color pixel value at the location (x,y). In this paper, let I be a low resolution image. Let $L_1 \times L_2$ be the total number of pixels in I. That is, $I \in R^{L_1 \times L_2 \times 3}$. Let Y be the initial high resolution image obtained by performing the discrete cosine transform zero padding approach. In particular, the discrete cosine transform is applied to each column of each color plane of the low resolution image. Then, the zeros are padded in the discrete cosine transform domain and the inverse discrete cosine transform is applied to obtain the initial high resolution image. Let *S* be the ratio of the total number of columns in the high resolution image to that of the low resolution image. Obviously, the total number of pixels in Y is $SL_1 \times L_2$. That is, $Y \in R^{SL_1 \times L_2 \times 3}$.

Now, a quaternion valued image is formed such that its real part is the luminance plane, as well as these imaginary parts are the red color plane, the green color plane and the blue color plane. Let $q(x,y)$ be the quaternion valued pixel at the location (x,y). That is,

$$q(x,y) = l(x,y) + r(x,y)i + g(x,y)j + b(x,y)k. \quad (3)$$

Let $H^{SL_1 \times L_2}$ be the set of the $SL_1 \times L_2$ quaternion valued matrices. Let $Z \in H^{SL_1 \times L_2}$ be this quaternion valued image. That is,

$$Z = \begin{bmatrix} q(0,0) & q(0,1) & \cdots & q(0,L_2-1) \\ q(1,0) & q(1,1) & \cdots & q(1,L_2-1) \\ \cdots & \cdots & \cdots & \cdots \\ q(SL_1-1,0) & q(SL_1-1,1) & \cdots & q(LS_1-1,L_2-1) \end{bmatrix}. \quad (4)$$

2.2. 2DQSSA based denoising

2.2.1. Decomposition phase of the 2DQSSA

Z is decomposed by the 2DQSSA. The first step is to perform the embedding. That is to generate the augmented quaternion valued trajectory matrix. Let $u \times v$ be the size of the window. Here, $1 \le u \le SL_1$ and $1 \le v \le L_2$. Let

$$W_{s,t} = \begin{bmatrix} q(s,t) & q(s,t+1) & \cdots & q(s,t+v-1) \\ q(s+1,t) & q(s+1,t+1) & \cdots & q(s+1,t+v-1) \\ \cdots & \cdots & \cdots & \cdots \\ q(s+u-1,t) & q(s+u-1,t+1) & \cdots & q(s+u-1,t+v-1) \end{bmatrix} \quad (5)$$

be the windowed quaternion valued image. Here, $W_{s,t} \in H^{u \times v}$. Then, the elements of $W_{s,t}$ are rearranged into a column vector. Let $\overrightarrow{W}_{s,t}$ are be this column vector. That is:

$$\overrightarrow{W}_{s,t} = [q(s,t) \ \cdots \ q(s+u-1,t) \ \cdots \ q(s,t+v-1) \ \cdots \ q(s+u-1,t+v-1)]^T. \quad (6)$$

Here, $\overrightarrow{W}_{s,t} \in H^{uv \times 1}$. Next, the window is moved from left to right and from top to bottom. Finally, the quaternion valued trajectory matrix is constructed as follows:

$$W = [\overrightarrow{W}_{0,0} \ \cdots \ \overrightarrow{W}_{0,L_2-v} \ \cdots \ \overrightarrow{W}_{SL_1-u,0} \ \cdots \ \overrightarrow{W}_{SL_1-u,L_2-v}]. \quad (7)$$

Here, $W \in H^{uv \times (SL_1-u+1)(L_2-v+1)}$. Let W^{lT}, W^{rT}, W^{gT} and W^{bT} be the involutions of W about the real axis, the i imaginary axis, the j imaginary axis and the k imaginary axis, respectively. Let

$$W^a = [W^{lT} \ \ W^{rT} \ \ W^{gT} \ \ W^{bT}]^T \quad (8)$$

be augmented quaternion valued trajectory matrix of W. Here, $W^a \in H^{4uv \times (SL_1-u+1)(L_2-v+1)}$.

The second step is to perform the quaternion valued singular value decomposition to generate the 2DQSSA components. Let C^a be the covariance matrix of W^a. That is, $C^a = E(W^a W^{aH})$. Then, the quaternion valued singular value decomposition is performed on W^a. Let λ_i be the i^{th} singular value. Let u_i and v_i be the i^{th} left singular vector and the i^{th} right singular vector, respectively. Let $\sqrt{\lambda_i} u_i v_i^H$ be the i^{th} 2DQSSA component of *Z*. Let r be the total number of the positive singular values. That is, $r = max\{i : \lambda_i > 0\}$. Let $\{1,2,\dots,r\}$ be the index set of $\sqrt{\lambda_i} u_i v_i^H$.

2.2.2. Reconstruction phase of the 2DQSSA

In this stage, first $\sqrt{\lambda_i} u_i v_i^H$ are categorized into a finite number of groups. Let *M* be the total number of groups and I_m for $m = 1,\dots,M$ be the index set of these groups. Now, putting $\sqrt{\lambda_i} u_i v_i^H$ into one of these groups is equivalent to perform the partition on $\{1,2,\dots,r\}$. Let

$$\widehat{W}_m^a = \sum_{i \in I_m} \sqrt{\lambda_i} u_i v_i^H. \quad (9)$$

It can be seen that

$$W^a = \sum_{m=1}^{M} \widehat{W}_m^a. \quad (10)$$

Second, the de-Hankelization is performed to obtain the grouped 2DQSSA components. In particular, let W_m be the matrix containing the first *uv* rows of $\widehat{W}_m^a$. W_m is divided into $v \times (L_2 - v + 1)$ blocks with the size of each block being $u \times (SL_1 - u + 1)$. Let $\widehat{W}_{m,s,t}^a$ be these blocks. That is:

$$W_m = \begin{bmatrix} \widehat{W}_{m,1,1}^a & \widehat{W}_{m,1,2}^a & \cdots & \widehat{W}_{m,1,L_2-v+1}^a \\ \widehat{W}_{m,2,1}^a & \widehat{W}_{m,2,2}^a & \cdots & \widehat{W}_{m,2,L_2-v+1}^a \\ \cdots & \cdots & \cdots & \cdots \\ \widehat{W}_{m,v,1}^a & \widehat{W}_{m,uv,2}^a & \cdots & \widehat{W}_{m,v,L_2-v+1}^a \end{bmatrix}. \quad (11)$$

Let $\widetilde{W}_{m,s,t}^a$ be the block obtained by performing the averaging among the off diagonal elements within $\widehat{W}_{m,s,t}^a$. That is, all the elements in the same off diagonal of $\widetilde{W}_{m,s,t}^a$ are the same. Let

$$\widetilde{W}_m = \begin{bmatrix} \widetilde{W}_{m,1,1}^a & \widetilde{W}_{m,1,2}^a & \cdots & \widetilde{W}_{m,1,L_2-v+1}^a \\ \widetilde{W}_{m,2,1}^a & \widetilde{W}_{m,2,2}^a & \cdots & \widetilde{W}_{m,2,L_2-v+1}^a \\ \cdots & \cdots & \cdots & \cdots \\ \widetilde{W}_{m,v,1}^a & \widetilde{W}_{m,uv,2}^a & \cdots & \widetilde{W}_{m,v,L_2-v+1}^a \end{bmatrix}. \quad (12)$$

Let $\widetilde{W}_{m,n}^a$ be the blocks obtained by performing the averaging among the off diagonal blocks in $\widetilde{W}_m$. Let

$$\check{Z}_m = \begin{bmatrix} \check{W}^a_{m,1} & \check{W}^a_{m,2} & \dots & \check{W}^a_{m,L_2-v+1} \\ \check{W}^a_{m,2} & \check{W}^a_{m,3} & \dots & \check{W}^a_{m,2,L_2-v+2} \\ \dots & & \dots & \dots \\ \check{W}^a_{m,v} & \check{W}^a_{m,v+1} & \dots & \check{W}^a_{m,L_2} \end{bmatrix}. \quad (13)$$

Here, all the blocks in the same off diagonal blocks in $\check{Z}_m$ are the same. Finally, the de-Hankelization is performed on $\check{Z}_m$. Let $\hat{Z}_m$ be the obtained de-Hankelized matrix. Here, the n^{th} column of $\hat{Z}_m$ is obtained by performing the de-Hankelization on $\check{W}^a_{m,n}$. Obviously, the dimension of $\hat{Z}_m$ is the same as that of Z. Let $\hat{Z}_{m,r}$, $\hat{Z}_{m,i}$, $\hat{Z}_{m,j}$ and $\hat{Z}_{m,k}$ be the real component, the *i* imaginary component, the *j* imaginary component and the k imaginary component of $\hat{Z}_m$, respectively. That is,

$$\hat{Z}_m = \hat{Z}_{m,r} + i\hat{Z}_{m,i} + j\hat{Z}_{m,j} + k\hat{Z}_{m,k}. \quad (14)$$

Since different color planes in the quaternion valued image are fused together via the quaternion valued operation, some high frequency information missing in one color plane can be generated using those in other color planes.

The last step is to generate the denoised image. Let $\hat{Z}$ be the reconstructed quaternion valued image. Likewise, let $\hat{Z}_r$, $\hat{Z}_i$, $\hat{Z}_j$ and $\hat{Z}_k$ be the real component, the *i* imaginary component, the *j* imaginary component and the *k* imaginary component of $\hat{Z}$, respectively. That is,

$$\hat{Z} = \hat{Z}_r + i\hat{Z}_{,i} + j\hat{Z}_j + k\hat{Z}_k. \quad (15)$$

Here, $\hat{Z}_r$ is the luminance plane while $\hat{Z}_i$, $\hat{Z}_j$ and $\hat{Z}_k$ are the red color plane, the green color plane and the blue color plane of the reconstructed image, respectively. In this paper, $\hat{Z}_r$ is ignored because the further processing is no longer based on the quaternion valued image. As the further processing only based on the color image, $\hat{Z}_i$, $\hat{Z}_j$ and $\hat{Z}_k$ are selected as follows:

$$\hat{Z}_i = \hat{Z}_{1,i} + \hat{Z}_{2,i} + \hat{Z}_{4,i} + \hat{Z}_{8,i},$$
$$\hat{Z}_j = \hat{Z}_{1,j} + \hat{Z}_{2,j} + \hat{Z}_{3,j} + \hat{Z}_{8,j} + \hat{Z}_{10,j} + \hat{Z}_{15,j} + \hat{Z}_{16,j}$$
and
$$\hat{Z}_k = \hat{Z}_{1,k} + \hat{Z}_{2,k} + \hat{Z}_{3,k} + \hat{Z}_{5,k} + \hat{Z}_{7,k} + \hat{Z}_{8,k}. \quad (16)$$

Here, the selections of the grouped 2DQSSA components are based on the expert knowledge. Since the high frequency components generated by the aliasing are removed by discarding some grouped 2DQSSA components, the 2DQSSA based denoising can enhance the image quality and yield a good super resolution performance.

2.3. EMD based denoising

2.3.1. EMD

For each color plane except the luminance plane of the reconstructed image, let $u(x,y)$ be the pixel value at the location (x,y) of this color plane. The EMD represents $u(x,y)$ as the sum of a finite number of components called the IMFs and the residual. The details procedures for performing the EMD are as follows:

Step 0: Initialize $k = 0$ as the iteration index, $l = 0$ as the IMF index and $\varepsilon > 0$ as a parameter for terminating the algorithm. Initialize

$$h_k(x,y) = u(x,y). \quad (17)$$

Step 1: Find out all edges corresponding to the local minima and the local maxima of $h_k(x,y)$.

Step 2: Let $\mathrm{e}_{\max}(\mathrm{x},\mathrm{y})$ and $\mathrm{e}_{\min}(\mathrm{x},\mathrm{y})$ be the planes corresponding to the upper envelope and the lower envelope interpolated based on the edges corresponding to the local maxima and the local minima of $h_k(x,y)$, respectively.

Step 3: Let $\mathrm{e}_{\mathrm{average}}(\mathrm{x},\mathrm{y})$ be the mean of the planes corresponding to the upper envelope and the lower envelope. That is:

$$e_{average}(x,y) = \frac{e_{min}(x,y) + e_{max}(x,y)}{2}. \quad (18)$$

Step 4: Subtract $h_k(x,y)$ from $\mathrm{e}_{\mathrm{average}}(x,y)$. Let $h_{k+1}(x,y)$ be the obtained signal. That is,

$$h_{k+1}(x,y) = h_k(x,y) - \mathrm{e}_{\mathrm{average}}(x,y). \quad (19)$$

Step 5: In order to determine whether $h_{k+1}(x,y)$ is an IMF or not, the following criterion is defined:

$$\frac{\sum_{x=1}^{SL_1}\sum_{y=1}^{L_2}|h_{k+1}(x,y)-h_k(x,y)|^2}{\sum_{x=1}^{SL_1}\sum_{y=1}^{L_2}|h_k(x,y)|^2} \le \varepsilon. \quad (20)$$

If the above criterion is satisfied, then $h_{k+1}(x,y)$ is used to approximate the l^{th} IMF. Define

$$IMF_l(x,y) = h_{k+1}(x,y). \quad (21)$$

Otherwise, increment the value of k and go back to Step 1.

Step 6: Compute the residual signal as follow:

$$res(x,y) = u(x,y) - \sum_{i=0}^{l} IMF_i(x,y) \quad . \quad (22)$$

Reset $k = 0$, increment the value of l and set $h_k(x,y) = res(x,y)$. Go back to Step 1 until there is no edge in the residual for interpolating the envelope.

Let $K-1$ be the total number of the obtained IMFs. Then, we have

$$u(x,y) = \sum_{i=1}^{K-1} IMF_i(x,y) + res(x,y) \quad . \quad (23)$$

Denote $IMF_K(x,y) = res(x,y)$. Hence, there are in total K components.

2.3.2. Selection of the IMFs via a binary linear programming approach

It is worth noting that the low resolution

images are obtained based on the devices with the low sampling rates. Hence, they are suffered from the aliasing. To address this issue, some high frequency components should be discarded in order to minimize the aliasing effect. On the other hand, as the whole frequency band of $u(x,y)$ are covered by different IMFs and different IMFs have different time frequency characteristics, the aliasing effect can be suppressed by an appropriate choice of the IMFs.

It is worth noting that the sizes of the IMFs are $SL_1 \times L_2$. However, the sizes of the corresponding color planes of the original low resolution image are $L_1 \times L_2$. Hence, it is difficult to define the objective function in the pixel domain to select the IMFs. To address this difficulty, this paper proposes an IMF selection method in the frequency domain so that the comparison between the color planes of the low resolution image and the IMFs of the color planes of the 2DQSSA processed image can be performed.

Now, each IMF of each color plane of the 2DQSSA processed image and each color plane of the low resolution image are divided into blocks. Here, the block sizes of the IMFs are different from those of the corresponding color planes of the low resolution image. However, the total number of the blocks in each IMF is the same as that in the corresponding color plane of the low resolution image. Let G_{m_1,n_1} be the $(m_1,n_1)^{\text{th}}$ block of a particular color plane of the low resolution image. Likewise, let $B_{m_1,n_1,l}$ be the corresponding block in the l^{th} IMF of the corresponding color plane of the 2DQSSA processed image. Let $b_{m_1,n_1,l}$ be the coefficient representing whether $B_{m_1,n_1,l}$ is selected or not. Here, "1" represents that $B_{m_1,n_1,l}$ has been selected. On the other hand, "0" represents that $B_{m_1,n_1,l}$ has been discarded. Let

$$b_{m_1,n_1} = \left[b_{m_1,n_1,1} \quad \dots \quad b_{m_1,n_1,K}\right]^T. \quad (24)$$

Let $J\left(b_{m_1,n_1}\right)$ be the objective function for selecting the IMFs. Let $\xi = [1,\dots,1]^T$ be a K dimensional column vector. Since the low resolution image should be visually indistinguishable with the high resolution image, most of $B_{m_1,n_1,l}$ should be selected. Hence, our objective is to retain as much as these blocks of the IMFs as possible. Therefore, we define

$$J\left(b_{m_1,n_1}\right) = \left\|\xi - b_{m_1,n_1}\right\|_1. \quad (25)$$

Let $\hat{B}_{m_1,n_1,l}(\omega_1,\omega_2)$ be the two dimensional discrete time Fourier transform of $B_{m_1,n_1,l}$. Likewise, let $\hat{G}_{m_1,n_1}(\omega_1,\omega_2)$ be the two dimensional discrete time Fourier transform of G_{m_1,n_1}. That is,

$$\hat{B}_{m_1,n_1,l}(\omega_1,\omega_2) = \sum_p \sum_q B_{m_1,n_1,l}(p,q)e^{-j(\omega_1 p+\omega_2 q)} \quad (26)$$

and

$$\hat{G}_{m_1,n_1}(\omega_1,\omega_2) = \sum_p \sum_q G_{m_1,n_1}(p,q)e^{-j(\omega_1 p+\omega_2 q)}. \quad (27)$$

Define

$$\check{B}_{m_1,n_1,l}(\omega_1,\omega_2) = \left[\hat{B}_{m_1,n_1,1}(\omega_1,\omega_2) \quad \dots \quad \hat{B}_{m_1,n_1,K}(\omega_1,\omega_2)\right]^T. \quad (28)$$

Then, $\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T b_{m_1,n_1}$ is the two dimensional discrete time Fourier transform of the block in the corresponding color plane of the image after performing the IMF selection. Here, $\left|\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T b_{m_1,n_1} - \hat{G}_{m_1,n_1}(\omega_1,\omega_2)\right|$ is the modulus error between the block in the corresponding color plane of the image after performing the IMF selection and that in the original low resolution image. In order to enforce the image after performing the IMF selection closed to the original low resolution image, this modulus error should be bounded for all the frequencies. Define δ be the specification on the upper bound of both the real part and the imaginary part of the absolute value of $\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T b_{m_1,n_1} - \hat{G}_{m_1,n_1}(\omega_1,\omega_2)$. That is:

$$\left|real(\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T b_{m_1,n_1} - \hat{G}_{m_1,n_1}(\omega_1,\omega_2))\right| \le \delta. \quad (29)$$

and

$$\left|imag(\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T b_{m_1,n_1} - \hat{G}_{m_1,n_1}(\omega_1,\omega_2))\right| \le \delta. \quad (30)$$

That is equivalent to

$$\begin{bmatrix} \text{real}(\check{\text{B}}_{m_1,n_1}(\omega_1,\omega_2)^{\text{T}}) \\ -\text{real}(\check{\text{B}}_{m_1,n_1}(\omega_1,\omega_2)^{\text{T}}) \\ \text{imag}(\check{\text{B}}_{m_1,n_1}(\omega_1,\omega_2)^{\text{T}}) \\ -\text{imag}(\check{\text{B}}_{m_1,n_1}(\omega_1,\omega_2)^{\text{T}}) \end{bmatrix} \text{b}_{m_1,n_1} \le \begin{bmatrix} \text{real}(\hat{\text{G}}_{m_1,n_1}(\omega_1,\omega_2)) \\ -\text{real}(\hat{\text{G}}_{m_1,n_1}(\omega_1,\omega_2)) \\ \text{imag}(\hat{\text{G}}_{m_1,n_1}(\omega_1,\omega_2)) \\ -\text{imag}(\hat{\text{G}}_{m_1,n_1}(\omega_1,\omega_2)) \end{bmatrix} + \delta \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \quad (31)$$

Let

$$A(\omega_1,\omega_2) = \begin{bmatrix} real(\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T) \\ -real(\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T) \\ imag(\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T) \\ -imag(\check{B}_{m_1,n_1}(\omega_1,\omega_2)^T) \end{bmatrix}$$

and

$$p(\omega_1,\omega_2) = \begin{bmatrix} real(\hat{G}_{m_1,n_1}(\omega_1,\omega_2)) \\ -real(\hat{G}_{m_1,n_1}(\omega_1,\omega_2)) \\ imag(\hat{G}_{m_1,n_1}(\omega_1,\omega_2)) \\ -imag(\hat{G}_{m_1,n_1}(\omega_1,\omega_2)) \end{bmatrix} + \delta \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \quad (32)$$

Then, the IMF selection problem can be formulated as the following optimization problem:

Problem (P_{m_1,n_1})

$$\min_{b_{m_1,n_1}} \quad J\left(b_{m_1,n_1}\right) = \left\|\xi - b_{m_1,n_1}\right\|_1,$$

$$\text{subject to } \text{A}(\omega_1,\omega_2)\text{b}_{\text{m}_1,\text{n}_1} \le \text{p}(\omega_1,\omega_2). \quad (33)$$

It is worth noting that the frequency domain is a continuous set. Therefore, Problem (P_{m_1,n_1}) is an infinite constrained optimization problem. To address this problem, the frequency domain is

sampled into a set of sampling points. Now, the IMF selection problem becomes a standard linear binary programming problem. There are many existing efficient methods for finding the solution of the linear binary programming problem. After finding the solution of Problem (P_{m_1,n_1}) for all the blocks in all the color planes, the super resolution image is constructed.

2.4. Discrete cosine transform based denoising

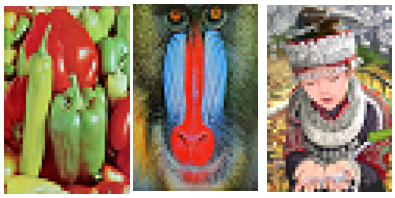
This paper also employs the discrete cosine transform based denoising method to further improve the super resolution performance. First, the error between the initial high resolution image obtained by performing the discrete cosine transform zero padding approach and the EMD processed image is computed. Next, this error image is downsampled to obtain the corresponding low resolution error image. After that, the bi-cubic interpolation method is applied to obtain a new high resolution error image. Now, a new image is reconstructed by summing up the EMD processed image and the new high resolution error image. Finally, the 2DQSSA based denoising method, the EMD based denoising method and the discrete cosine transform based denoising method are repeated until the algorithm is converged.

3. Computer numerical simulation results

3.1. Dataset

This paper takes the images from the ImageNet image library (ILSVRC2013) for demonstrating the effectiveness of our proposed method. This library consists of 395,909 color images with the bitmap format. The camera quality in terms of the sizes of the images is to $L_1 \times L_2 = 60 \times 40$ [15], [6]. However, only those representative low resolution images with different types of contents are demonstrated in this paper because of the page limit. In particular, they are the images of a bird, a baby, a human face, a bunch of flowers, a pepper, a monkey and a woman. Figure 2 shows these images.

(a) (b) (c) (d)



(e) (f) (g)

Figure 2. The low resolution images. (a) A bird. (b) A baby. (c) A human face. (d) A bunch of flowers. (e) A pepper (f) A monkey. (g) A woman.

3.2. Parameter selections

The objective of the image super resolution is to double the total number of the rows of the low resolution image. Hence, $S = 2$. In the 2DQSSA based denoising, a large window size requires a very heavy computation. On the other hand, a small window size does not exploit the correlation among the pixels in the same color plane. To tradeoff between these two factors, this paper sets the window size at 4×4. That is, $u = 4$ and $v = 4$. Besides, as the window size is 4×4, the maximum value of r is 32. This is not a very large number. In order to simplify the grouping procedures in the 2DQSSA based denoising, the total number of the SSA groups is set to the value of r. That is, $M = r$. In this case, no grouping is required. In the EMD based denoising, ε is set at 0.001. This is because this is the typical value used in the EMD. In the binary linear programing, the total number of the sampling points in the frequency domain is set at 225. Since the sampling step size is $\frac{2\pi}{225}$, it is small enough for the majority natural images. This is because the majority natural images are narrow band signals. On the other hand, it is worth noting that a very small value of δ will result to the feasible set of Problem (P_{m_1,n_1}) to be empty. Whereas, a very large value of δ will result to all the decision vectors being in the feasible set of Problem (P_{m_1,n_1}). In this case, the constraint functions are not meaningful and the optimization problem becomes the corresponding unconstrained optimization problem. To tradeoff between these two factors, this paper sets δ at 70.

3.3. Results

Figure 3 show the high resolution images obtained via our proposed method. It can be seen that the images have not been blurred even though the total number of the rows is doubled. Besides, Table 1 show the peak signal to noise ratio (PSNR) between the original high resolution images before performing the downsampling and the high resolution images

obtained by our proposed method. It can be seen that our proposed method can achieve the PSNR between 25.2800dB and 30.9981dB, which is high enough for the majority science and engineering applications.

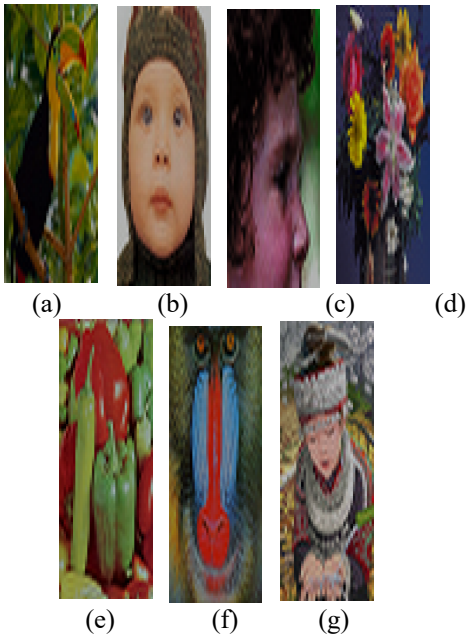


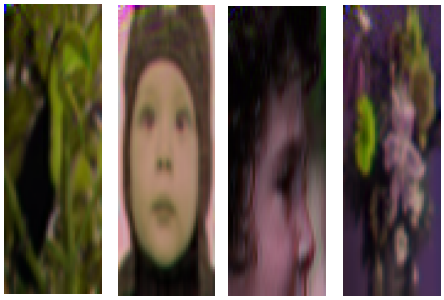
Figure 3. The high resolution images obtained via our proposed method. (a) A bird. (b) A baby. (c) A human face. (d) A bunch of flowers. (e) A pepper (f) A monkey. (g) A woman.

In order to demonstrate the effectiveness of our proposed method, our proposed method without performing the discrete cosine transform based denoising, the EMD based denoising and the 2DQSSA based denoising are compared. The processed images are shown in Figure 4, Figure 5 and Figure 6, respectively. It can be seen that our proposed method outperforms the comparison methods in terms of the quality of the processed images. Besides, Table 2 shows the PSNRs between the original high resolution images before performing the downsampling and the high resolution images obtained by our proposed method without performing various time frequency analysis based denoising. It can be seen easily from Table 1 and Table 2 that the differences of the PSNRs yielded by our proposed method with performing all these three time frequency analysis based denoising and that without performing the discrete cosine transform based denoising are between 5.0186dB and 9.6046dB, that without performing the 2DQSSA based denoising are between 5.2869dB and 13.4406dB, as well as that without performing the EMD based denoising are between 2.1828dB and 9.1725dB. As the improvements are very significant, this demonstrates the effectiveness of our proposed method.

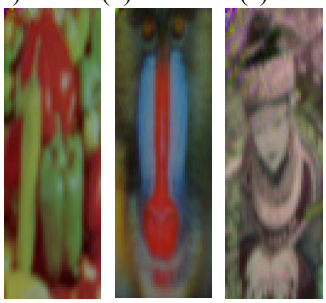
In order to compare the effects of different color spaces on our proposed method, the quaternion valued images are generated using the following methods. First, the real part of the quaternion valued image is set to the luminance plane and these three imaginary parts of the quaternion valued image are set to the red color plane, the green color plane and the blue color plane of the RGB color space. This quaternion valued image is denoted LRGB. Second, the real part of the quaternion valued image is set to the zero matrix and these three imaginary parts of the quaternion valued image are set to the red color plane, the green color plane and the blue color plane of the RGB color space. This quaternion valued image is denoted 0RGB. Third, the real part of the quaternion valued image is set to the V plane of the HSV color space and these three imaginary parts of the quaternion valued image are set to the red color plane, the green color plane and the blue color plane of the RGB color space. This quaternion valued image is denoted VRGB. Fourth, the real part of the quaternion valued image is set to the luminance plane and these three imaginary parts of the quaternion valued image are set to the cyan color plane, the magenta color plane and the yellow color plane of the CMY color space. This quaternion valued image is denoted LCMY. Table 3 shows the PSNRs between the original high resolution images before performing the downsampling and the high resolution images obtained by our proposed method using different planes of different color spaces for generating the quaternion valued images. It can be seen that the performances yielded by our proposed method using different planes of different color spaces for generating the quaternion valued images are similar. Hence, the effects of using different planes of different color spaces for generating the quaternion valued images on our proposed method are very minor. This demonstrates the robustness of our proposed method.

Besides, our proposed method is compared to the state of the art methods. In particular, Table 4 shows the PSNRs between the original high resolution images before performing the downsampling and the high resolution images obtained by the sparse coding based method [5], [6], the SRCNN based method [9] and our proposed method. It can be seen that our proposed method can yield the range of the PSNRs between 25.2800dB and 30.9981dB,

while the sparse coding based method can only yield the range of the PSNRs between 23.0321B and 25.6414dB while the SRCNN based method can only yield the range of the PSNRs between 19.3421dB and 26.1917dB. Obviously, our proposed method significantly outperforms the state of the art method.



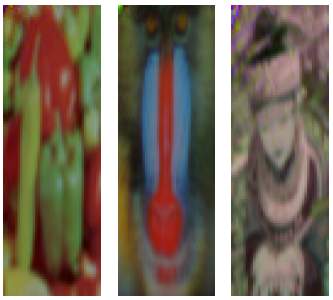
(a) (b) (c) (d)



(e) (f) (g)

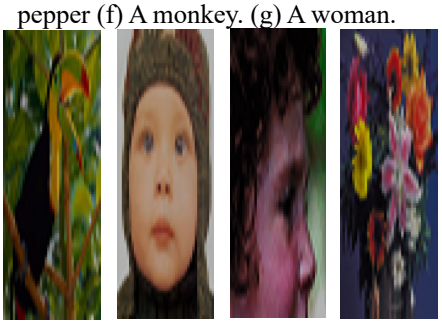
Figure 4. The high resolution images obtained via our proposed method without performing the discrete cosine transform based denoising. (a) A bird. (b) A baby. (c) A human face. (d) A bunch of flowers. (e) A pepper (f) A monkey. (g) A woman.

(a) (b) (c) (d)

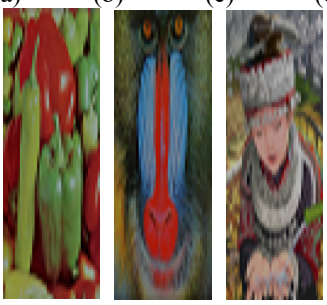


(e) (f) (g)

Figure 5. The high resolution images obtained via our proposed method without performing the EMD based denoising. (a) A bird. (b) A baby. (c) A human face. (d) A bunch of flowers. (e) A pepper (f) A monkey. (g) A woman.



(a) (b) (c) (d)



(e) (f) (g)

Figure 6. The high resolution images obtained via our proposed method without performing the 2DQSSA based denoising. (a) A bird. (b) A baby. (c) A human face. (d) A bunch of flowers. (e) A pepper (f) A monkey. (g) A woman.

Table 1. The PSNRs (dB) between the original high resolution images before performing the downsampling and the high resolution images obtained by our proposed method.

Images	Bird	Baby	Face	Flowers	Pepper	Monkey	Woman
PSNR	27.6150	29.7081	30.9981	25.2800	28.5679	30.2468	28.0936

Table 2. The PSNRs (dB) between the original high resolution images before performing the downsampling and the high resolution images obtained by our proposed method without performing various time frequency analysis based denoising.

Images	Bird	Baby	Face	Flowers	Pepper	Monkey	Woman
PSNRs yielded by our proposed method	21.1191	20.1035	21.9325	20.2614	22.5079	20.2392	20.7238

without performing the discrete cosine transform based denoising							
PSNRs yielded by our proposed method without performing the 2DQSSA based denoising	14.1744	17.8415	13.9146	19.9931	13.8643	20.9623	21.9458
PSNRs yielded by our proposed method without performing the EMD based denoising	23.0354	20.5356	24.3679	23.0972	24.2605	24.2972	20.8714

Table 3. The PSNRs (dB) between the original high resolution images before performing the downsampling and the high resolution images obtained by our proposed method using different planes of different color spaces for generating the quaternion valued images.

Images	Bird	Baby	Face	Flowers	Pepper	Monkey	Woman
LRGB	27.6150	**29.7081**	30.9981	25.2800	28.5679	30.2468	28.0936
0RGB	27.8256	29.6889	29.5977	25.6624	**29.0880**	28.7163	26.0936
VRGB	**28.7885**	27.8242	**32.7102**	**26.2846**	27.9180	28.9870	**30.0038**
LCMY	26.2566	29.1948	31.5315	24.3684	26.2628	**30.2718**	27.9197

Table 4. The PSNRs (dB) between the original high resolution images before performing the downsampling and the high resolution images obtained by the state of the art methods.

Methods	Bird	Baby	Face	Flowers	Pepper	Monkey	Woman
Sparse coding	23.5463	24.6067	25.6414	23.0321	23.6081	24.8456	23.5646
SRCNN	21.7331	23.7308	26.1917	19.3421	20.6534	24.1868	20.2636
Our proposed method	**27.6150**	**29.7081**	**30.9981**	**25.2800**	**28.5679**	**30.2468**	**28.0936**

4. Conclusion

This paper formulates the super resolution problem as the denoising problems. In particular, a combination of three different types of time frequency analysis based denoising methods is proposed. First, the initial high resolution images are generated using the discrete cosine transform zero padding approach. Then, a combination of the 2DQSSA based denoising method, the EMD based denoising method and the discrete cosine transform based denoising method is applied iteratively. For the 2DQSSA based denoising method, as the red color plane, the green color plane and the blue color plane are fused together via the quaternion valued operation, some high frequency information missing in one color plane can be generated using those in other color planes. On the other hand, some distorted high frequency components are removed by discarding some IMFs. Therefore, our proposed method outperforms the methods without performing any one of these three time frequency analysis based denoising.

It is worth noting that the existing deep learning methods require a very huge computation power. Also, it is highly dependent on the given training images. These drawbacks result to the difficulty of applying the deep learning methods to perform the video super resolution. In future, the feasibility of applying our proposed method to perform the video super resolution will be investigated.

Declarations

Conflict of interests: There is no conflict of interest.

Funding: This paper was supported partly by the National Nature Science Foundation of China (no. U1701266, no. 61671163 and no. 62071128), the Team Project of the Education Ministry of the Guangdong Province (no. 2017KCXTD011), the Guangdong Higher Education Engineering Technology Research Center for Big Data on Manufacturing Knowledge Patent (no. 501130144) and the Hong Kong Innovation and

Technology Commission, Enterprise Support Scheme (no. S/E/070/17).
Data availability statement: The data is downloaded from the public database.

References

1. A. Buccini, M. Pasha and L. Reichel, "Generalized singular value decomposition with iterated Tikhonov regularization," Journal of Computational and Applied Mathematics ,2019.
2. Z. Li, Q. Li, W. Wu, J. Yang and X. Yang, "Deep recursive up-down sampling networks for single image super-resolution," Neurocomputing ,2019.
3. C. Y. F. Ho, B. W. K. Ling and P. K. S. Tam, "Representations of linear dual-rate system via single SISO LTI filter, conventional sampler and block sampler," IEEE Transactions on Circuits and Systems-II: Express Briefs, vol. 55, no. 2, pp. 168-172, 2008.
4. H. Wu, J. Zhang and Z. Wei, "High resolution similarity directed adjusted anchored neighborhood regression for single image super-resolution," IEEE Access, vol. 6, pp. 25240-25247, 2018.
5. J. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), pp. 1-8, 2008.
6. J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image super-resolution via sparse representation," IEEE transactions image processing, vol.19, no. 11, pp. 2861–2873, 2010.
7. M. Bevilacqua, A. Roumy, C.Guillemot, and M. L. A. Morel,"Low-complexity single-image super-resolution based on nonnegative neighbor embedding," British Machine Vision Conference BMVA Press, pp. 1–10, 2012.
8. C. Dong, C. C. Loy, K. He, "Image Super-Resolution Using Deep Convolutional Networks," IEEE Trans Pattern Anal Mach Intell, vol.38, no.2, pp. 295-307, 2016.
9. D. Qiu, S. Zhang, Y. Liu, J. Zhu and L. Zheng, "Super-resolution reconstruction of knee magnetic resonance imaging based on deep learning," Computer Methods and Programs in Biomedicine, 2019.
10. H.Junkang, J.Qiuping, C. Runmin, G.Wei, S. Feng, "Two-branch deep neural network for underwater image enhancement in HSV color space," IEEE Signal Processing Letters, vol.28, pp.2152-2156, 2021.
11. S. Enshaeifar, S. Kouchaki, C. C. Took and S. Sanei, "Quaternion singular spectrum analysis of electroencephalogram with application in sleep analysis," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 24, no. 1, pp. 57-67, 2016.
12. Y. Lin, W. K. Ling, L. Hu, Y. Zheng and X. Wang, "Hyperspectral image enhancement by two dimensional quaternion valued singular spectrum analysis for object recognition," Remote Sensing, 2021.
13. A. Ertürk, M. K. Güllü and S. Ertürk, "Hyperspectral image classification with spectral gradient enhancement for empirical mode decomposition," 2012 IEEE International Geoscience and Remote Sensing Symposium, pp. 4162-4165, 2012.
14. V. Patel, C. K. Modi, C. N. Paunwala and S. Patnaik, "Hybrid approach for single image super resolution using ISEF and IBP," International Conference on Communication Systems and Network Technologies, pp. 495-499, 2011.
15. R. Timofte, V. De Smet, and L. Van Gool, "Anchored neighbor hood regression for fast example-based super-resolution," IEEE International Conference on Computer Vision IEEE, pp. 1920–1927, 2014.