Hazardous Gas Emission Monitoring Based on High-Resolution Images

Wenjian Chen,1 Yi Wang,2,3 Xuan Li,2 Wei Gao,1 Shiwei Ma,1 Yuanyuan Duan,1 and Xiaopeng Shao2

1Xi’an Institute of Applied Optics, Xi’an 710065, China
2School of Physics and Optoelectronic Engineering, Xidian University, Xi’an 710071, China
3Chinese Flight Test Establishment, Xi’an 710089, China

Correspondence should be addressed to Xiaopeng Shao; xpshao@xidian.edu.cn

Received 28 September 2017; Accepted 21 November 2017; Published 4 March 2018

Air pollution presents unprecedentedly severe challenges to humans today. Various measures have been taken to monitor pollution from gas emissions and the changing atmosphere, of which imaging is of crucial importance. By images of target scenes, intuitional judgments and in-depth data are achievable. However, due to the limitations of imaging devices, effective and efficient monitoring work is often hindered by low-resolution target images. To deal with this problem, a superresolution reconstruction method was proposed in this study for high-resolution monitoring images. It was based on the idea of sparse representation. Particularly, multiple dictionary pairs were trained according to the gradient features of samples, and one optimal pair of dictionaries was chosen to reconstruct by judging the weighting of the information in different directions. Furthermore, the K-means singular value decomposition algorithm was used to train the dictionaries and the orthogonal matching pursuit algorithm was employed to calculate the sparse coding coefficients. Finally, the experiment’s results demonstrated its advantages in both visual fidelity and numerical measures.

1. Introduction

Today, humans are facing the most severe situation of air pollution due to industrial emission, fuel combustion, motor vehicle exhaust emission, and so on [1–4]. It adversely affects human health and daily life. Various monitoring measures have been taken to provide real-time information of the environment [5]. For example, the monitoring process focuses on the number of chimneys that are working and the size of the emissions. As an important method, imaging technology has made the monitoring task easy and efficient, although it often suffers from low-resolution problems originating from the imaging devices. Details are not available in the captured images, which further limits the judgment of the conditions [6–8].

To improve the resolution of monitoring images, charge-coupled devices (CCDs) with high resolution are exploited, which will certainly lead to a higher cost. To improve the effect of monitoring images without replacing the original detection equipment, people gradually turn to image-processing methods, of which the superresolution reconstruction method is an effective way to improve the quality of images including hyperspectral, visible light, and infrared images. Many significant superresolution (SR) methods have been proposed to deal with SR reconstruction problems in past decades [9–12]. Among them, the conventional interpolation-restoration approach is computational. One improved SR algorithm is the generalized nonlocal means algorithm, which generalized the successful video denoising nonlocal means (NLM) technique [13, 14].

In addition, in the study of improving image resolution, single-frame image SR reconstruction algorithms based on dictionary learning show their unique advantage in combining the prior information of images in this research area.
so these reconstruct images with higher resolution than other methods. Inspired by compressed sensing (CS) theory, Yang et al. firstly proposed the SR reconstruction method based on sparse coding [17, 18], named sparse coding-based superresolution (SCSR), which could reconstruct a HR image by global dictionary learning. Zeyde et al. made an improvement to Yang et al.’s method and proposed the single image scale-up method using sparse representation (SISR) [19]. The SCSR algorithm proposed by Yang et al. and the SISR method proposed by Zeyde et al. are both effective learning-based SR reconstruction methods, which try to obtain a HR image by training a high-resolution dictionary and a corresponding low-resolution (LR) dictionary.

When reconstructing the LR images of different kinds, the mentioned methods choose the only couple of the dictionary. But the information in different gas observation images or different parts in one image may vary a lot [20]. If we separate a whole LR image into many patches and then choose the dictionary couple with the highest similarity according to the features in the LR patches in the process of reconstructing HR patches, the resolution of the reconstructed monitoring image could be improved by a large margin. Hence, a novel SR reconstruction based on adaptive dictionary learning is proposed in this paper according to the analyses above. In particular, the features of training samples are clustered into several groups by the K-means method, and multidictionary pairs are trained off-line by the K-means singular value decomposition (K-SVD) algorithm [21]. And then, the dictionary pair which has the highest similarity with the LR image is selected in the process of online image reconstruction. Finally, several groups of experiments are completed and the results show that the images reconstructed by the proposed method are excellent on both subjective vision perception and objective evaluation value.

2. Principle of SR Image Reconstruction

Dictionary learning is particularly important in the process of image SR reconstruction. By means of sparse representation, more information in the dictionaries can be contained by fewer atoms, which leads to the higher quality of reconstructed images and less time cost. In this section, the calculations of image sparse representation and dictionary pair learning are described in detail.

In the process of imaging, if the original ideal HR image is $\tilde{x}$, which is affected by the blurring of the optical imaging systems and the downsampling of the CCD when displayed on the device, the traditional imaging process can be modeled as [22]

$$
x = SH\tilde{x} + v,
$$

where $x$ is the observed image, $S$ represents the downsampling operator, $H$ is a blurring filter, and $v$ is the additive noise caused by the system.

Based on the theory of CS, to obtain an ideal HR image using a fast SR reconstruction algorithm, both the methods of SISR and SCSR require that the ideal HR image $\tilde{x}$ be sparsely represented by the HR dictionary $D_h$, which can be written as [23]

$$
\tilde{x} = D_h\alpha.
$$

The observed LR image can be expressed by

$$
x = SHD_h\alpha, \quad \|\alpha\|_0 \ll n.
$$

As (2) and (3) show, $\alpha \in R$ is a sparse matrix with $n$ elements. The sparse representation of the HR ideal image is shown in Figure 1. But this equation is not satisfied by many images. The natural image statistics prove that usually the areas of image primitives have very low internal dimensions and thus they can be represented by a small amount of training samples [24]. The primitives mentioned here refer to the high-frequency feature areas such as the edge and inflection points in images.

In this study, an assumption was made that the image primitives can be represented in the sparse form by the dictionary which is trained with a large number of primitive patches. A HR image can be obtained by the following steps, and Figure 2 shows the process of SR reconstruction.

(i) The primitive areas of training samples are extracted and divided into four subsets according to the direction of their gradient features. The four directions are $0^\circ$, $45^\circ$, $90^\circ$, and $135^\circ$, which indicate the angles with respect to the vertical direction.

(ii) Divide the primitive areas into four subsets according to the gradient features, which are used as the data source to train four subdictionary sets, each including a HR dictionary and a corresponding LR dictionary by the K-SVD algorithm.

(iii) According to the weightings of the gradient feature information in each direction within a LR patch, the certain HR subdictionary and the corresponding LR subdictionary are chosen to reconstruct a HR subpatch.

(iv) The HR subpatches reconstructed by different dictionaries are combined into the final HR image according to the weight of the features in four directions.

Typically, the primitive areas in the images can be obtained by (linear) feature extraction operators. For the training sample sets $x$, the HR primitive areas in one direction $x^1_k$ and the corresponding LR primitive areas $x^2_k$ can be obtained by

$$
\begin{align*}
x^1_k &= F_kx, \\
x^2_k &= F^*_kSx,
\end{align*}
$$

where $k = 1, 2, 3, 4$ and $S$ is the downsampling operator. The features in the horizontal gradient ($0^\circ$) can be obtained after being filtered by the operator $F_1$. The features in the vertical gradient ($90^\circ$) can be obtained after being filtered by $F_2$. The
For a certain LR patch, Equation (9) can be solved by the orthogonal matching pursuit (OMP) method [25] and by calculating \( \beta_l \). The process of SR reconstructing the HR patch \( \hat{y}_i \) can be expressed as

\[
\hat{y}_i = D_h \beta_l^i.
\]  

The symbol \( \beta_l^i \) in the above equations means the sparse representation coefficient of \( y_i \) under a certain subdictionary \( D_h^i \). Equation (9) can be solved by the orthogonal matching pursuit (OMP) method [25] and by calculating \( \beta_l^i \). In the process of SR reconstructing the HR patch \( \hat{y}_i \), we need to combine the weight of the information in four different directions. The HR patch \( \hat{y}_i \) can be written as

\[
\hat{y}_i = \sum_{j=1}^{4} q_j \hat{y}_j.
\]  

where

\[
x = \begin{bmatrix}
{x_h} \\
{x_l}
\end{bmatrix}, \quad D = \begin{bmatrix}
D_h \\
D_l
\end{bmatrix},
\]

The \( N \) and \( M \) in the equation above mean the size of the input image. Given an initial dictionary \( D_h \), the dictionary pair \( D_h^k \) and \( D_l^k \) can be solved by the K-SVD algorithm [21]. Figure 3 shows some parts of the four HR feature patches with the same training samples.

For a LR image \( y \) to be reconstructed, at first, a simple segmentation method was employed to divide it into \( p \) patches \( y_i \in \mathbb{R}^n, n = 1, 2, \ldots, p \). For a certain LR patch \( y_i \), its corresponding HR patch is \( \hat{y}_i \) (\( \hat{y}_i \in \mathbb{R}^N, n = 1, 2, \ldots, p \)). According to (1), the LR patches can be written as

\[
y_i = SHD_k + v, \quad v \text{ refers to the noise}.\]

Our purpose is to get \( \hat{y}_i \) by \( y_i \) and a certain couple of subdictionary \( D_h^k \) and \( D_l^k \), the process can be expressed as

\[
\beta_l^i = \arg\min_{\beta_l^i} \| y_i - SHD_l^k \beta_l^i \|_2^2, \quad \text{s.t. } \| \beta_l^i \| \leq T_0, \quad (9)
\]

\[
\hat{y}_i = D_h^k \beta_l^i. \quad (10)
\]

The HR and LR primitive areas obtained in (3) and (4) can be sparsely represented by the corresponding HR and LR dictionaries, respectively. The process of obtaining the dictionaries can be mathematically expressed by

\[
D_h^k = \arg\min_{D_h^k, \alpha} \left\{ \| x_h - D_h^k \alpha \|_2^2 \right\}, \quad \text{s.t. } \forall i, \| \alpha_i \|_0 < T_0, \quad (6)
\]

\[
D_l^k = \arg\min_{D_l^k, \alpha} \left\{ \| x_l - D_l^k \alpha \|_2^2 \right\}, \quad \text{s.t. } \forall i, \| \alpha_i \|_0 < T_0. \quad (6)
\]

Here, \( D_h^k \) refers to the \( k \)th subdictionary of HR and \( D_l^k \) refers to the corresponding LR dictionary. Because the HR patches and the LR patches have the same sparse representation coefficients, (6) is expressed as

\[
\min_{D_l^k, \alpha} \left\{ \| x_l - D_l^k \alpha \|_2^2 \right\}, \quad \text{s.t. } \forall i, \| \alpha_i \|_0 < T_0. \quad (7)
\]

where

\[
x = \begin{bmatrix}
{x_h} \\
{x_l}
\end{bmatrix}, \quad D = \begin{bmatrix}
D_h \\
D_l
\end{bmatrix},
\]

The features in the 45° and 135° directions can be obtained by \( F_3 \) and \( F_4 \), respectively. The four filters are described in

\[
F_1 = \begin{bmatrix}
-1 & 0 & 1 \\
-1 & 0 & 1 \\
-1 & 1 & 1
\end{bmatrix}, \quad F_2 = \begin{bmatrix}
0 & 0 & 0 \\
-1 & -1 & -1 \\
-1 & 1 & 1
\end{bmatrix}, \quad F_3 = \begin{bmatrix}
-1 & 0 & 1 \\
-1 & -1 & 0 \\
1 & 1 & 0
\end{bmatrix}, \quad F_4 = \begin{bmatrix}
1 & 0 & -1 \\
0 & -1 & -1
\end{bmatrix}
\]

The symbol \( q_j \) represents the weighting coefficient of the information in four different directions. To guarantee the compatibility between the LR patches and the chosen dictionary, the weighting coefficient is employed which can be expressed by the amount of the information in different directions, which is

\[
q_j = \frac{F_j y_i}{\sum_{k=1}^{4} F_k y_i}, \quad j = 1, 2, 3, \ldots, 4. \quad (12)
\]

Derived from (9), the final HR image \( \hat{y}_i \) is

\[
\hat{y}_i = \sum_{j=1}^{4} q_j \hat{y}_j. \quad (13)
\]
The whole HR image $\hat{y}$ is finally spliced by all the HR patches $\hat{y}_i$, as shown in

$$\hat{y} = \sum_{i=1}^{p} \hat{y}_i.$$  \hspace{1cm} (14)

3. Experiment and Analysis

According to the analysis above, experiments have been conducted to demonstrate the effectiveness of the proposed method for superresolution image reconstruction. The number of training samples will affect the quality of the
reconstructed HR image, but too many samples lead to a decrease in the reconstruction efficiency. In our experiments, the peak signal-to-noise ratio (PSNR) and the mean square error (MSE) were chosen as the objective evaluation parameters of image quality, which compare the original high-resolution image with each corresponding pixel of the reconstructed image. Because they can reflect the similarity of each pixel of two images, PSNR and MSE are widely used objective evaluation parameters of image quality in the field of image processing. Above all, PSNR and MSE reflect the degree of similarity or fidelity of two images. PSNR and MSE can be expressed as

\[
\text{MSE} = \frac{1}{M \times N} \sum_{i=1}^{M} \sum_{j=1}^{N} (x_{ij} - \hat{x}_{ij})^2, \tag{15}
\]

\[
\text{PSNR} = 10 \log_{10} \left[ \frac{\text{MAX}_I^2}{\text{MSE}} \right]. \tag{16}
\]

The symbols \( M \) and \( N \) in (15) represent the width and height of the image, respectively. \( x_{ij} \) and \( \hat{x}_{ij} \) mean the pixel values corresponding to two images. The result of \( x_{ij} - \hat{x}_{ij} \) is close to zero when these two images are similar. In other words, a small MSE value indicates two analogous images. In addition, the parameter \( \text{MAX}_I \) in (16) means the value of the signal peak. For example, the value equals 255, when we use an 8-bit sample image. Furthermore, MSE is an important parameter for PSNR. A better reconstruction image always has a higher value of PSNR.

In our experiments, the growth of the value of PSNR of the reconstructed HR image is close to zero when the sample number increases to a degree. In order to guarantee the comprehensiveness of the dictionary information and the efficiency of the algorithm, 100 pictures of different kinds were chosen as the training samples. For the LR color images, we transform the RGB images to a YCbCr space. Since the human visual system is more sensitive to the luminance channel, we use the superresolution reconstruction method based on dictionary learning proposed in this study to reconstruct the images in the luminance channel, and the images in the Cb and Cr chrominance channels are simply magnified by the bicubic interpolation method.

The original LR images in these experiments are all downsampled from HR images by a factor of 1/3 in this experiment. The number of atoms in each subdictionary is 1000, and the iteration number of the dictionary training is set to 40. For a better comparison, the LR images are also reconstructed by the bicubic interpolation method and Zeyde et al.’s method, respectively. Figures 4 and 5 show the results of reconstructed gas monitoring images.

Figure 4(a) shows the original LR chimney image, and Figures 4(b)–4(d) are the reconstructed images based on the bicubic method, Zeyde et al.’s method, and the proposed method, respectively. For the enlarged parts of these images, the construction of the chimney in Figures 4(b)–4(d) is more clear and smooth than that in Figure 4(a). The image in Figure 4(d) is smoother and shows better quality including better contrast, improved resolution, clearer texture, and detailed information.

Figure 5 takes an image of a scene full of smog as a sample. With the reconstructed HR images, it is more practical to determine the exact pollution sources and their locations. As for the detailed information in the images, all of the three reconstructed HR images show a better human visual effect than the LR image, which shows obvious mosaic blocks. Furthermore, the HR image reconstructed by
the bicubic method looks blurred compared with the other two HR images. As for the texture, stripe information and other details in Figure 5(d) are more abundant and clearer.

In general, because four different dictionaries contain different features of the training samples and show different similarities with the LR images, the reconstructed images by the proposed SR algorithm of adaptive dictionary learning all present a better image quality than raw LR images do. While the HR reconstruction image will be obtained from the LR image, the HR patches are utilized based on the weight of the features in different directions, which consider the differences in various LR patches and guarantee the best matching results with the dictionaries in the process of SR reconstruction.

To evaluate the reconstruction results, the objective evaluation parameters PSNR and MSE are analyzed and calculated for reconstructed images in both Figures 4 and 5. The results are shown in Table 1. Indeed, PSNR and MSE are two basic evaluation standards. Higher PSNR and lower MSE mean better image quality. As Table 1 shows, the PSNR of two groups of images reconstructed by the proposed method are both higher than the PSNR of images reconstructed by the other two methods while the value of MSE are significantly reduced. Compared with the images reconstructed by Zeyde et al.’s method, the PSNR of the reconstructed chimney image by the proposed method increases by 0.8602 dB and the PSNR of the reconstructed smog image by the proposed method increases by 0.3369 dB. Different LR images will end with different reconstruction results due to the association with the image content. The images with more obvious gradient changes are more possible to be better reconstructed, benefitting from the higher matching accuracy with the various gradient dictionaries.

4. Conclusion

In conclusion, the proposed superresolution reconstruction method based on adaptive dictionary learning of gradient features can be used to obtain HR monitoring images of gas emission with more detailed information. It starts by training four couples of different dictionaries according to the directions of gradient information and reconstructing a HR gas monitoring image by combining the HR patches according to the weight of the features in different directions. Finally, actual images of pollution sources are tested by the proposed method. Experimental results show the effectiveness of the proposed method in reconstructing HR images. Furthermore, our further study will focus on increasing the efficiency of the whole SR reconstruction procedure and improving the dictionaries for some particular images with pertinence to make the monitoring task more effective and efficient.

Conflicts of Interest

The authors declare that they have no conflicts of interest.
Acknowledgments

Grateful acknowledgement is made to the authors’ friends Dr. Fei Liu and Ms. Pingli Han, who gave them considerable help by means of suggestion, comments, and criticism. Thanks are due to Dr. Fei Liu for his encouragement and support. Thanks are also due to Ms. Pingli Han for the guidance of the paper. The authors deeply appreciate the contribution made to this thesis in various ways by their leaders and colleagues. This work was supported by the National Natural Science Foundation of China (no. 61505156); the Fundamental Research Funds for the Central Universities (JB170503); the State Key Laboratory of Applied Optics (CS16017050001); and the Young Scientists Fund of the National Natural Science Foundation of China (61705175).

References


