Mixed noise reduction via sparse error constraint representation of high frequency image for wildlife image

Yuan Xu [1] · Yaqin Zhao [1] · Peng Lu [1]

Received: 8 February 2021 / Revised: 23 February 2022 / Accepted: 16 May 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-022-12747-0

Mixed noise reduction via sparse error constraint representation of high frequency image for wildlife image

Yuan Xu [1] · Yaqin Zhao [1] · Peng Lu [1]

Received: 8 February 2021 / Revised: 23 February 2022 / Accepted: 16 May 2022
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

Abstract
Wildlife image noise reduction is a difficult and challenging problem since the images are inevitably corrupted by the mixed noise in the complex field environments. Most of the existing denoising methods focused on the noise removal added manually to pure images. However, the noise of wildlife images taken in the field is random mixed noise; thus, the existing noise reduction algorithms are not suitable for the denoising of wildlife images. In this paper, we propose a novel mixed noise reduction method based on sparse error constraint representation for removing the wildlife image noise. Firstly, we use 2D-DCT method to decompose a noisy image into a high frequency image and a low frequency image, and then rank 2D-DCT coefficients based on Zig-zag sorting algorithm. As we know, the high frequency image contains more noise, so the dictionary learning model of high frequency image is established to recover the images corrupted by the mixed noise. The sparse error term describes the error between the sparse coefficients of the original image and those obtained by the error constraint method. And then we utilize the algorithm proximal alternating linearization minimization to solve the objective function due to the nonconvex and non-smooth minimization problem. In order to update the dictionary, we apply the $(p, l_1$ norm term for sparse coding to obtain the optimal solution of sparse coefficients. The experiment results show that the proposed method has good noise reduction results for both the noisy images recorded in the wild and the images artificially corrupted by mixed noise while retain more details of the wildlife objects in the restored images.

Keywords Sparse error constraint representation · Proximal alternating linearization minimization · Zig-Zag sorting · 2D-DCT

Yaqin Zhao
yaqinzhao@163.com

[1] College of Mechanical and Electronic Engineering, Nanjing Forestry University, Nanjing 210037, China

Published online: 21 May 2022

Springer

Professor : Dr. Mohsen Ebrahimi Moghadam

Presenter : Kian Anvari

Springer Link

What is noise ?

2-D DCT Example

Abstract

Introduction

2-D DCT Dictionary Example

Block-Based 2-D DCT Example

Sparse Representation

Results

Discussion

Springer Link

What is noise ?

Noise is typically defined as a random variation in brightness or color information and it is frequently produced by technical limits of the image collection sensor or by improper environmental circumstances. These difficulties are frequently inevitable in real scenarios, making image noise a common issue that must be addressed with appropriate denoising approaches.

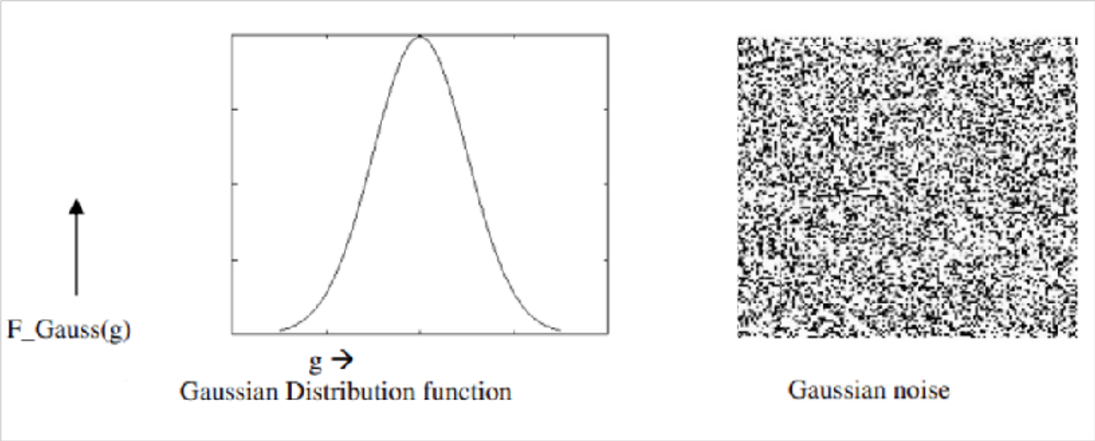
The following are the primary sources of noise in digital images:

Environmental factors may have an impact on the imaging sensor.

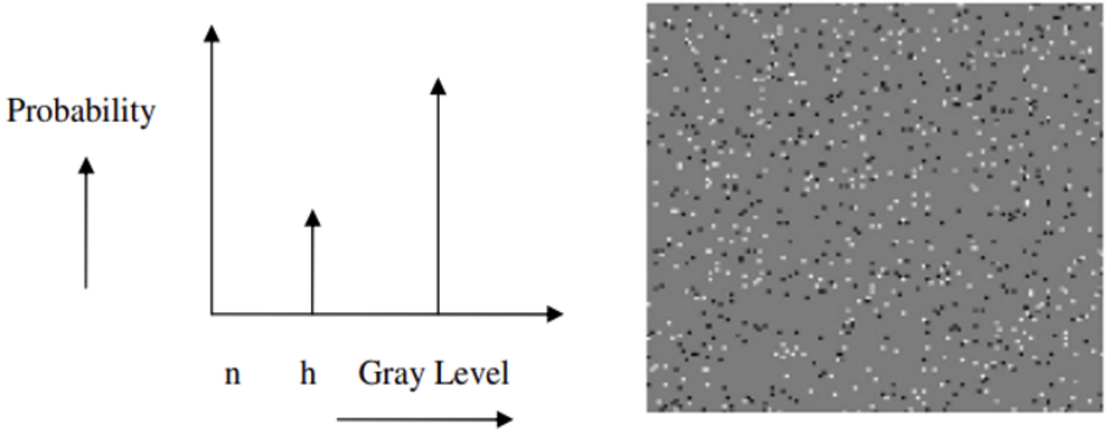
Low light and sensor temperature may cause image noise.

Dust particles in the scanner can cause noise in the digital image.

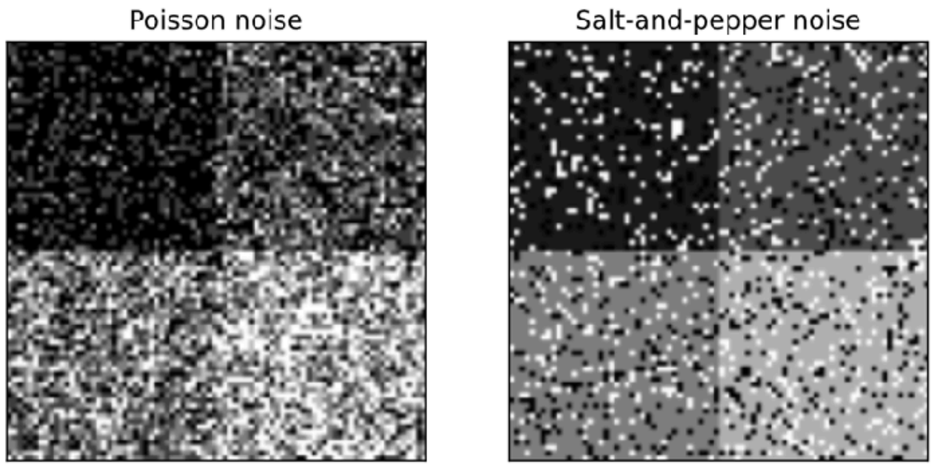
Transmission channel interference.



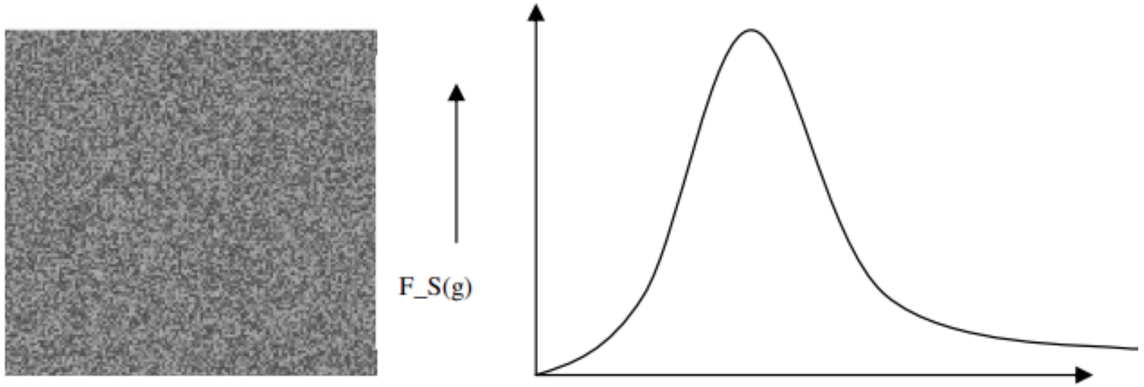
Gaussian noise



Salt and Pepper noise



Poisson noise



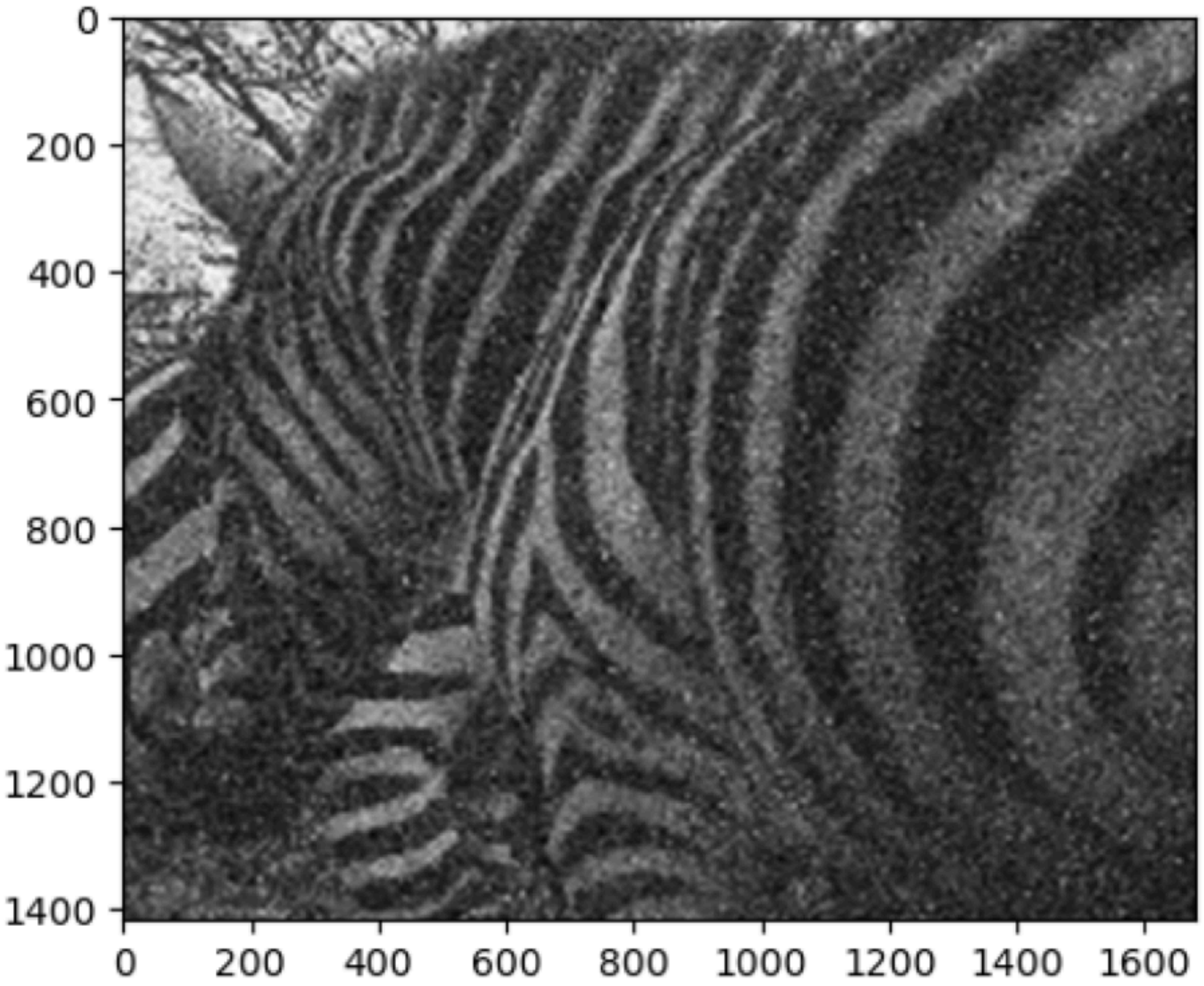
Speckle Noise

2-D DCT

Example

Springer Link

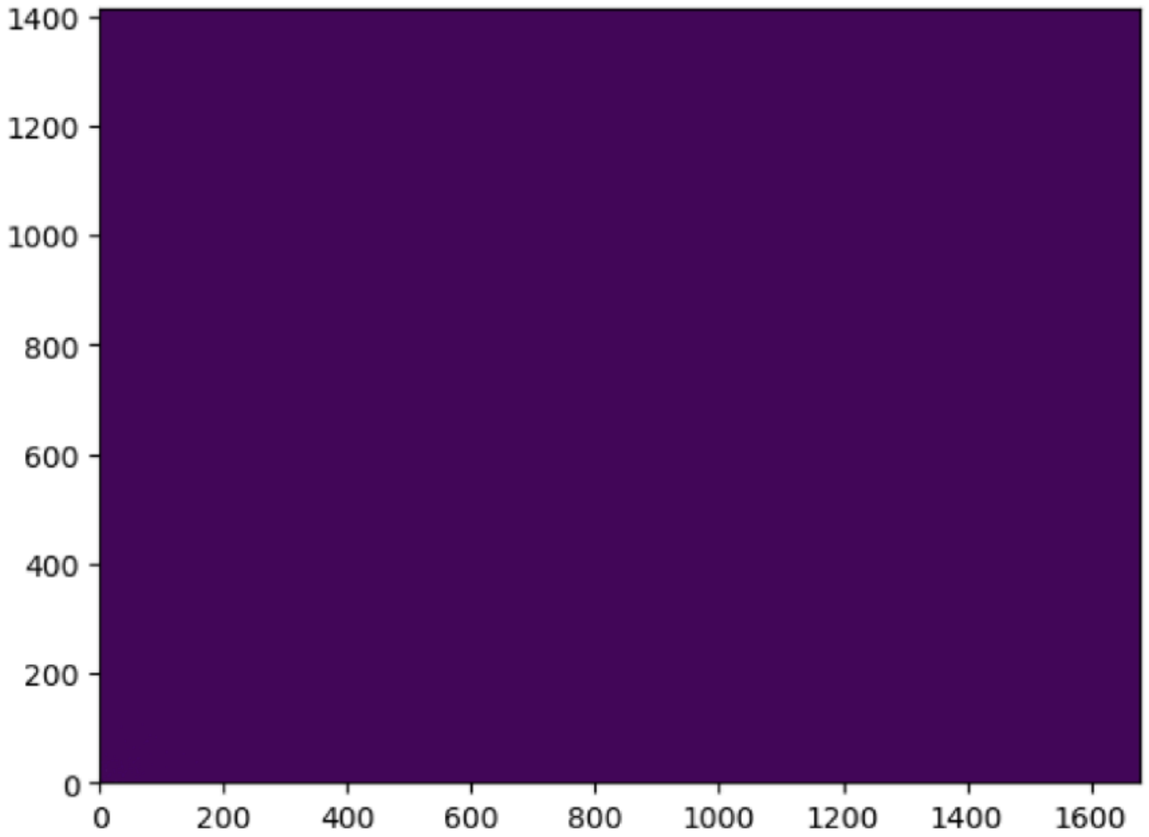
DCT- WORKING (WHOLE IMAGE DCT)



ORGINAL IMAGE

References

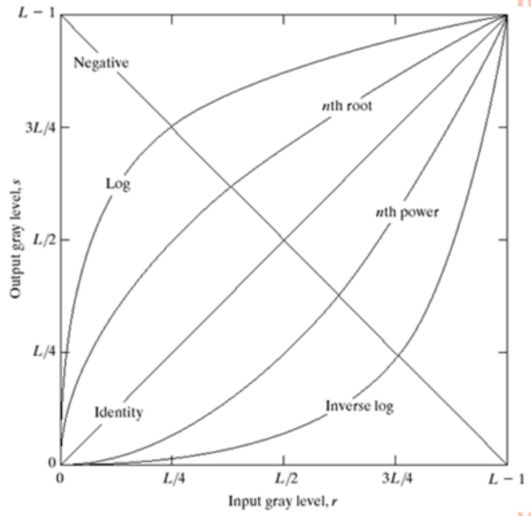
Chapter 3 - Image Enhancement. Dr Mohsen Ebrahimi Moghadam - Shahid Beheshti University

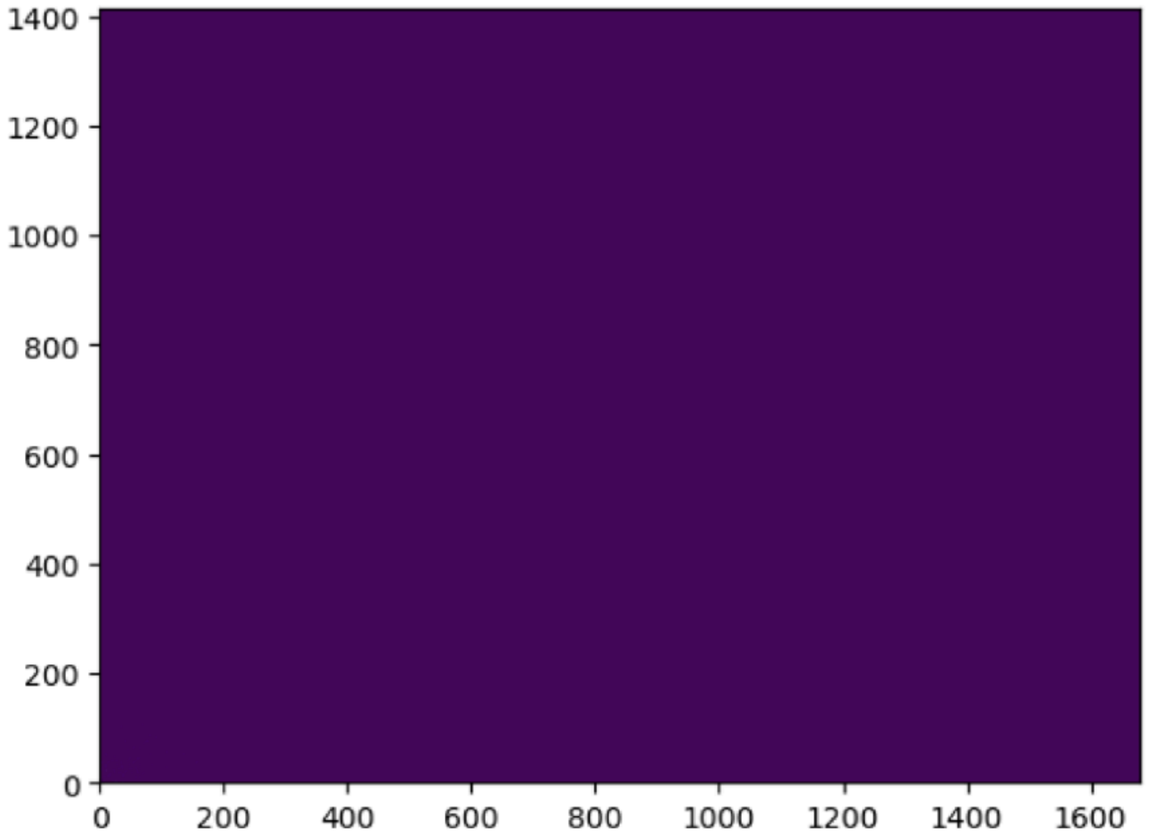


DCT(IMAGE)

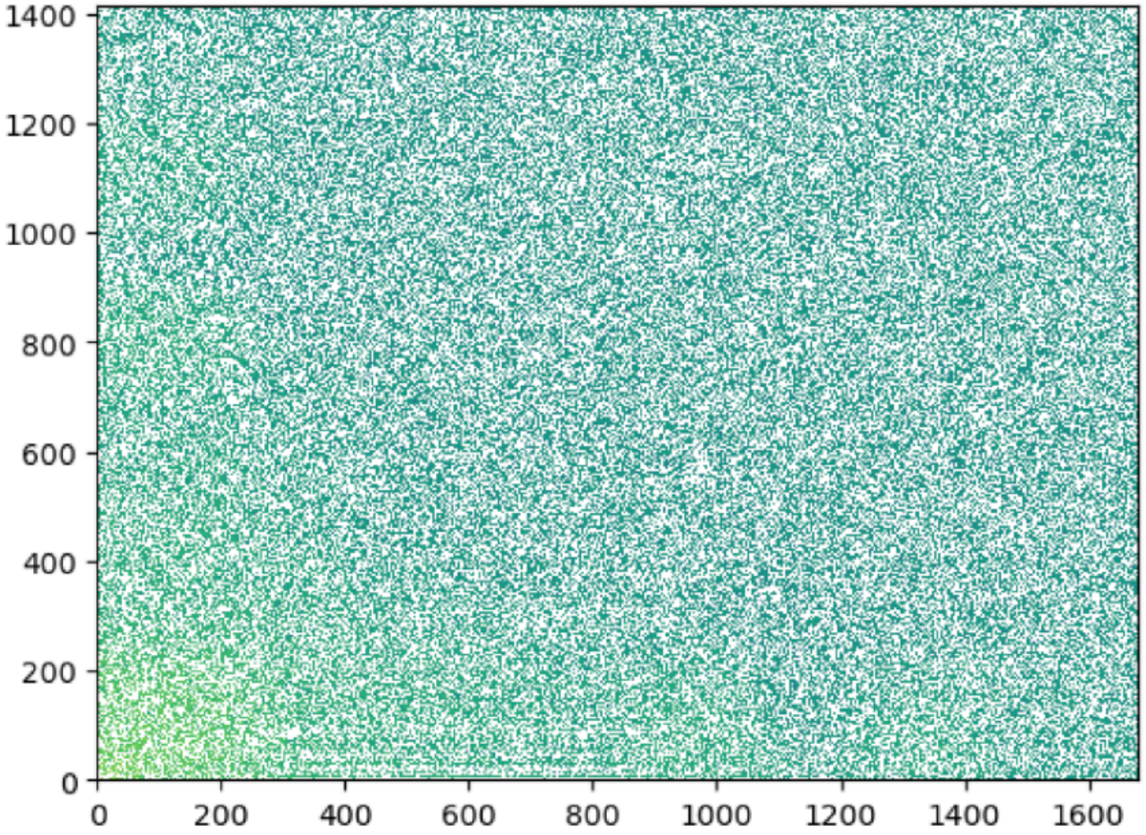
Basic Gray level transformations

- A *point processing* approach!
- Three basic types of functions used
  - Linear (negative and identity transformation)
  - Logarithmic (log and inverse-log)
  - Power-law (nth power and root)

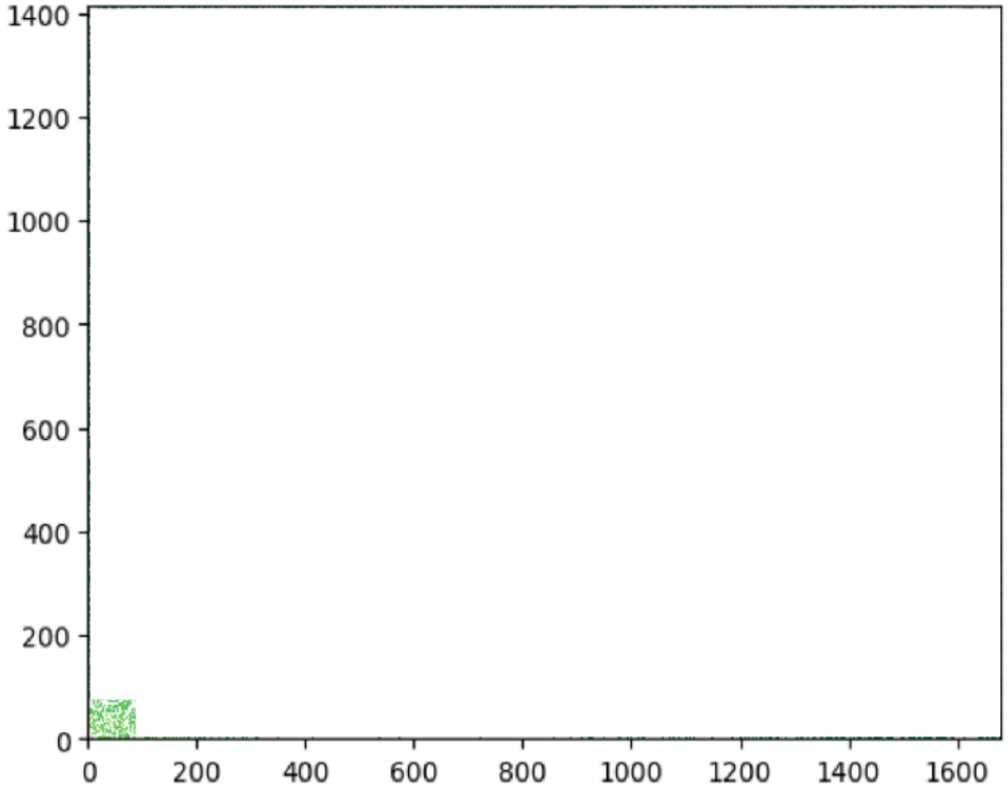




DCT(IMAGE)

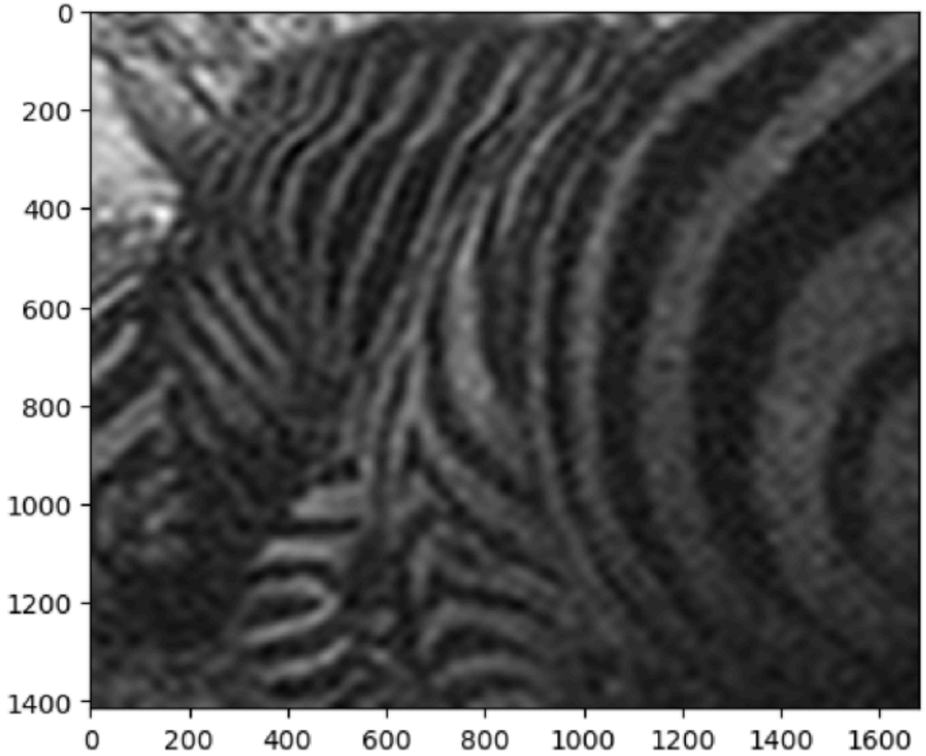


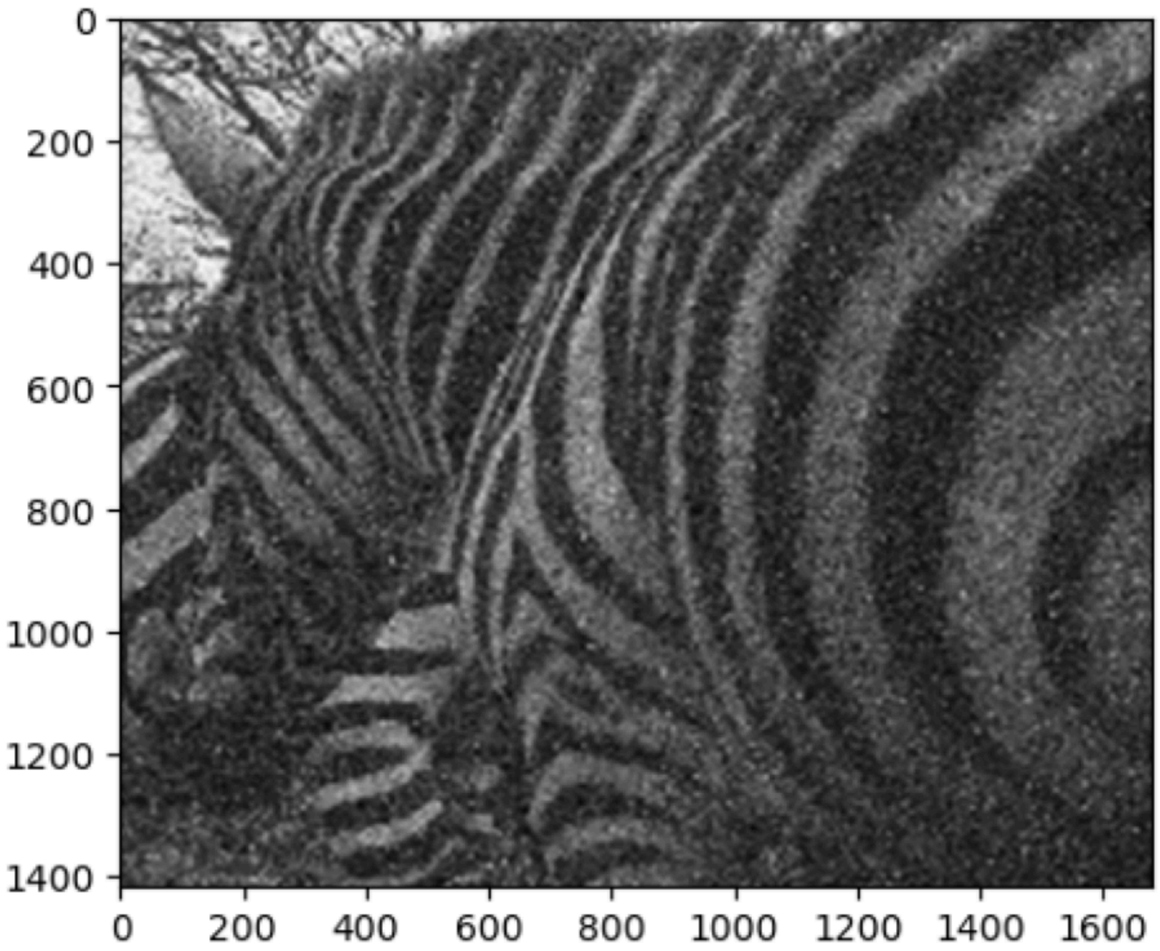
LOG(DCT(IMAGE))



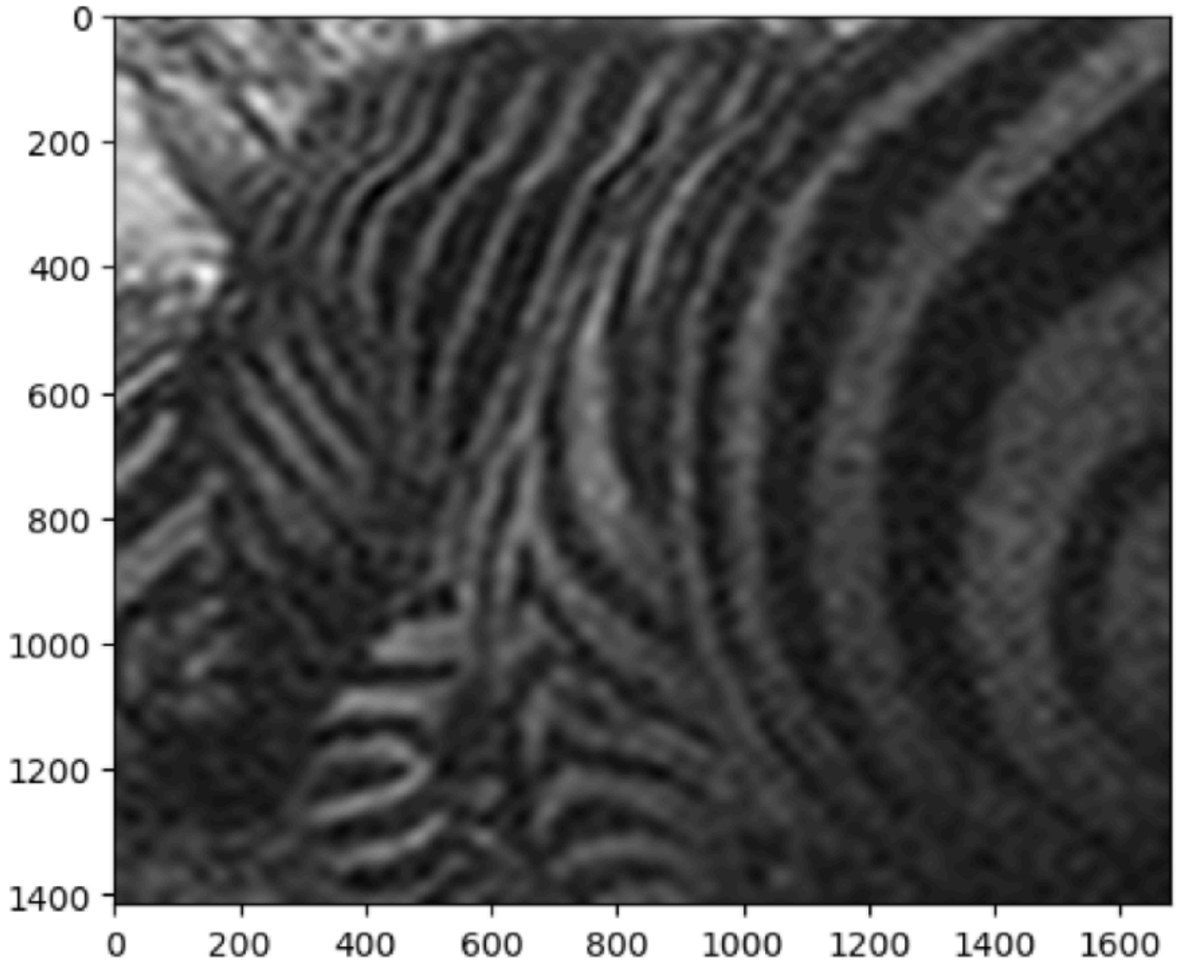

```
In [6]: dct_low = dct_img.copy()
        dct_low[70:] = 0
        dct_low[:,70:] = 0

        img_low = idct2(dct_low)
        plt.imshow(img_low, cmap="gray")
```

```
Out[6]: <matplotlib.image.AxesImage at 0x7f883c953c10>
```

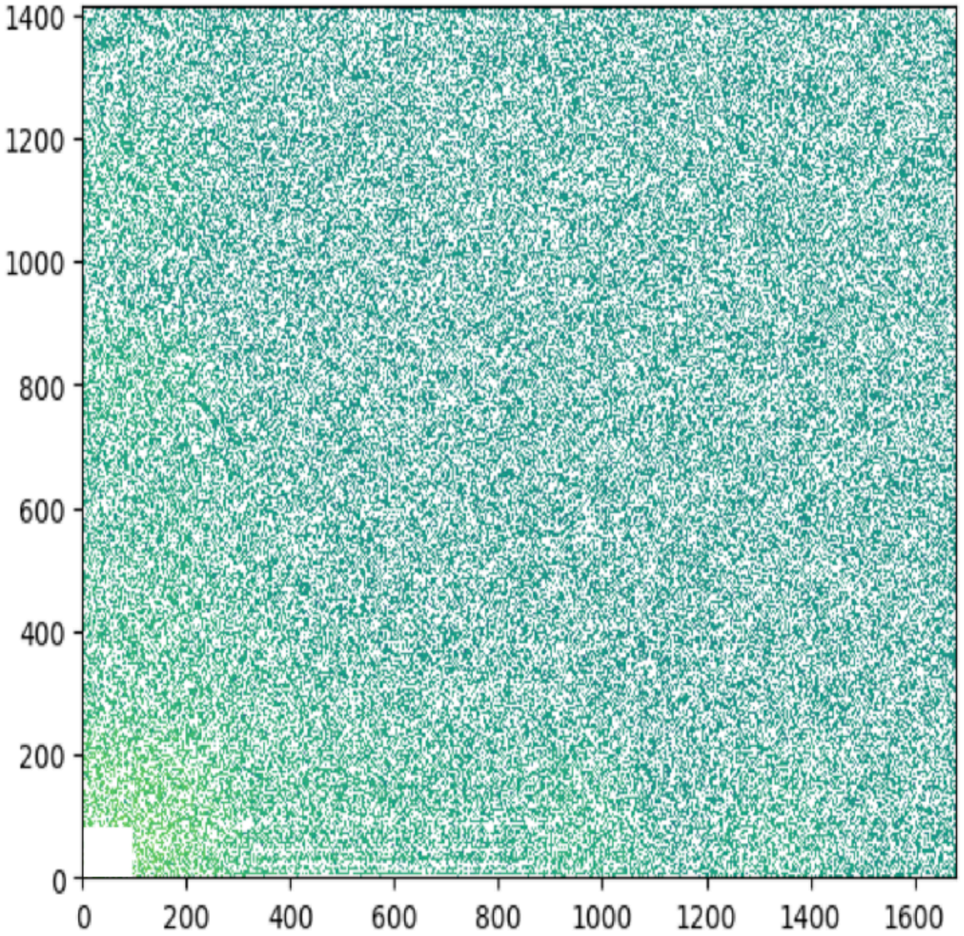




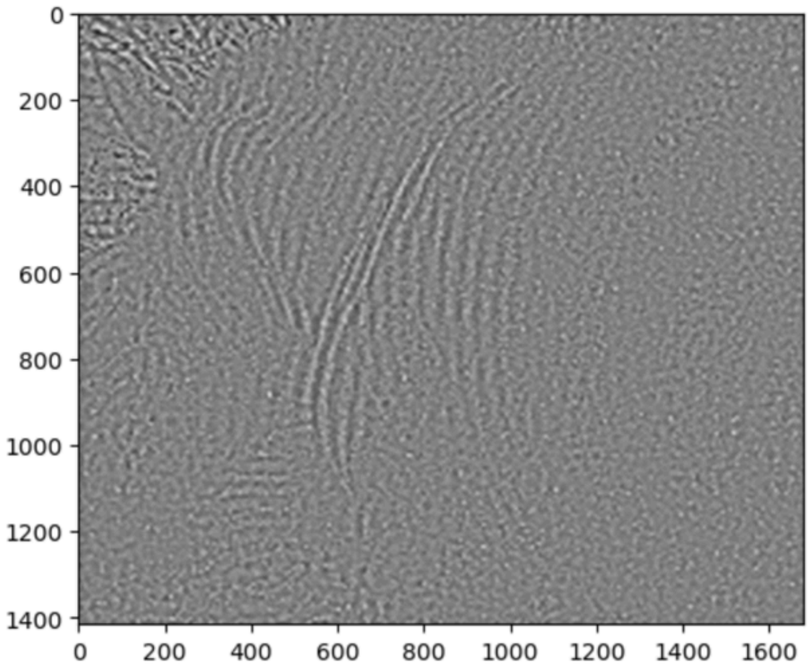
ORGINAL IMAGE

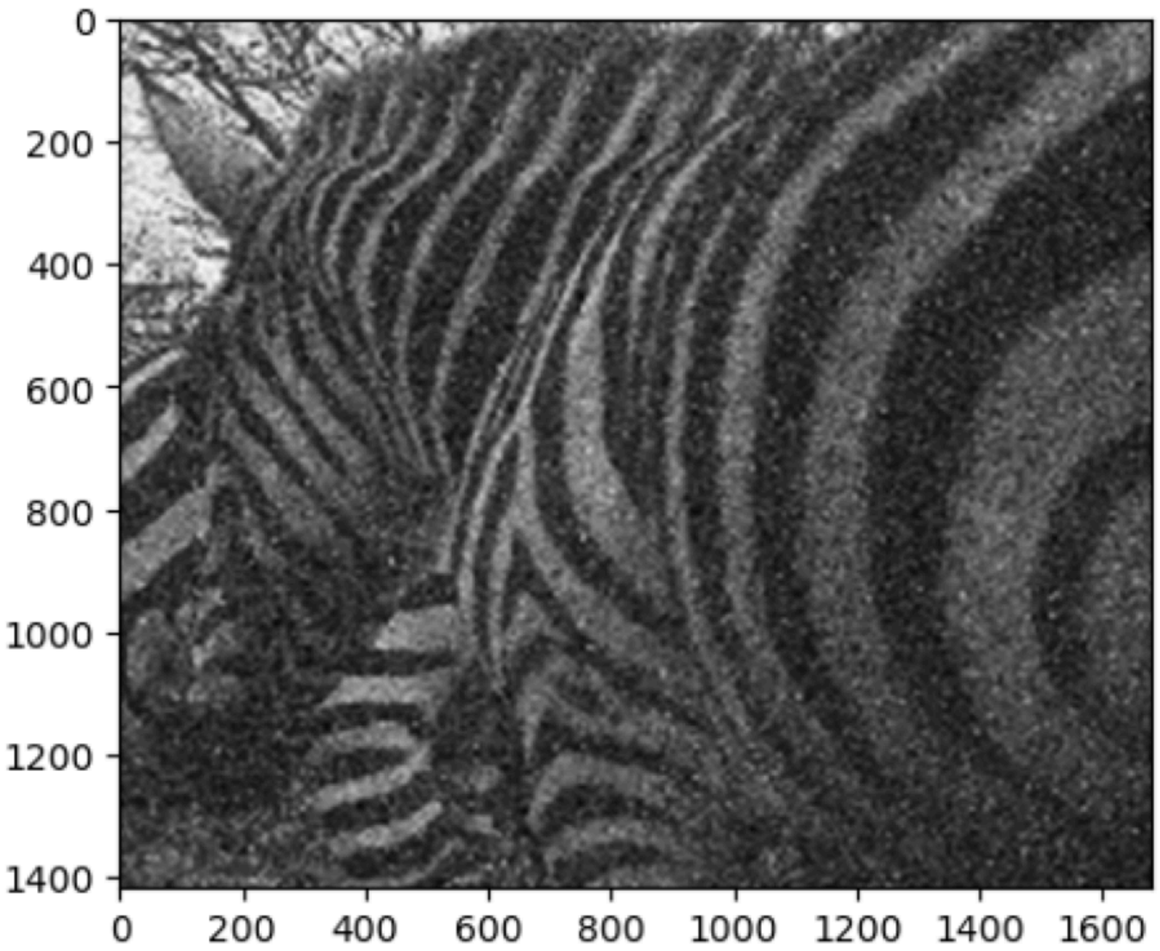


LOW FREQUENCY SUB IMAGE

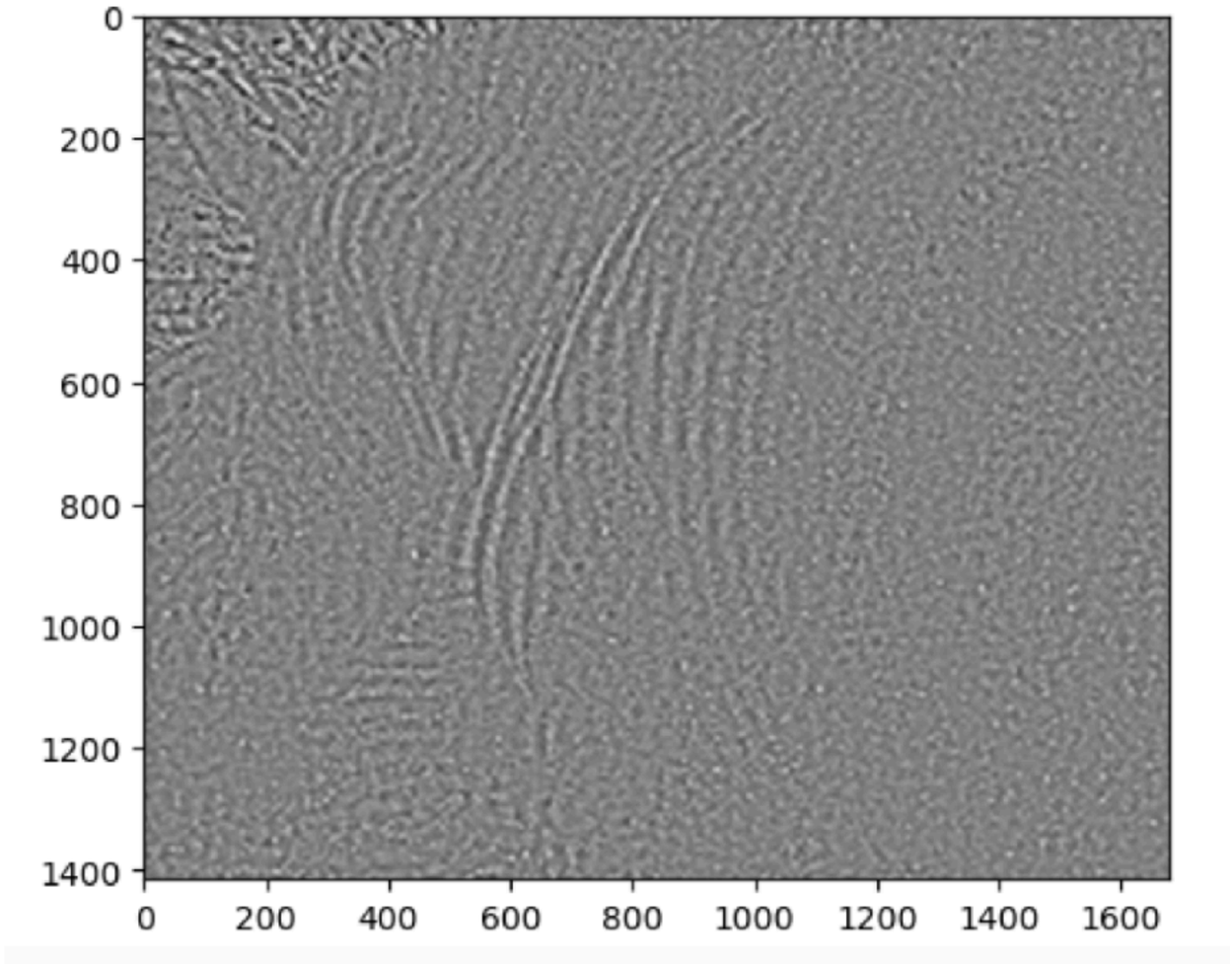



```
In [7]: dct_high = dct_img.copy()

        dct_high[:70,:70] = 0

        img_high = idct2(dct_high)
        plt.imshow(img_high, cmap="gray")

Out[7]: <matplotlib.image.AxesImage at 0x7f883d8519d0>
```





ORGINAL IMAGE



HIGH FREQUENCY SUB IMAGE

Abstract

Abstract

Wildlife image noise reduction is a difficult and challenging problem since the images are inevitably corrupted by the mixed noise in the complex field environments. Most of the existing denoising methods focused on the noise removal added manually to pure images. However, the noise of wildlife images taken in the field is random mixed noise; thus, the existing noise reduction algorithms are not suitable for the denoising of wildlife images. In this paper, we propose a novel mixed noise reduction method based on sparse error constraint representation for removing the wildlife image noise. Firstly, we use 2D-DCT method to decompose a noisy image into a high frequency image and a low frequency image, and then rank 2D-DCT coefficients based on Zig-zag sorting algorithm. As we know, the high frequency image contains more noise, so the dictionary learning model of high frequency image is established to recover the images corrupted by the mixed noise. The sparse error term describes the error between the sparse coefficients of the original image and those obtained by the error constraint method. And then we utilize the algorithm proximal alternating linearization minimization to solve the objective function due to the nonconvex and non-smooth minimization problem. In order to update the dictionary, we apply the *lp*- *l*1- norm term for sparse coding to obtain the optimal solution of sparse coefficients. The experiment results show that the proposed method has good noise reduction results for both the noisy images recorded in the wild and the images artificially corrupted by mixed noise while retain more details of the wildlife objects in the restored images.

Wildlife image noise reduction is a difficult and **challenging** problem since the images are inevitably corrupted by the **mixed noise** in the complex field environments.

Most of the existing denoising methods focused on **the noise removal added manually to pure images.**

However, the noise of **wildlife images** taken in the field is **random mixed noise**; thus, the existing noise reduction algorithms are not suitable for the denoising of wildlife images.

In this paper, we propose a **novel mixed noise reduction** method based on **sparse error constraint representation** for removing the wildlife image noise.

Firstly, we use **2D-DCT** method to decompose a noisy image into **a high frequency image and a low frequency image**, and then rank 2D-DCT coefficients based on **Zig-zag sorting** algorithm.

As we know, **the high frequency image contains more noise**, so the **dictionary learning model of high frequency image is established to recover the images corrupted by the mixed noise.**

The sparse error term describes the error between the sparse coefficients of the **original image** and those obtained by the error constraint method. And then we utilize the algorithm proximal alternating linearization minimization to solve the objective function due to the **non-convex and non-smooth minimization problem.**

In order to update the dictionary, we apply the **lp- l1- norm** term for sparse coding to obtain the optimal solution of sparse coefficients.

The experiment results show that the proposed method has **good noise reduction results for both the noisy images recorded in the wild and the images artificially corrupted by mixed noise** while retain more details of the wildlife objects in the restored images.

Introduction

1 Introduction

Image processing technology has been widely used in wildlife protection. The external factors have a impact on the quality of images when taking images. Noisy images have a great influence on the wildlife image analysis and processing. Many methods were proposed for removing Gaussian noise [3, 6], salt-and-pepper noise [4, 10] and Impulse Noise [1, 9]. Among them, BM3D (block matching and 3D filtering) [4] is considered as a good method to remove noise by using image redundancy and looking for similar blocks in the image for filtering. The detail information can be preserved well after denoising the image. A few researches have focused on mixed Gaussian noise and impulse noise removal based on sparse representation [13, 19]. However, the methods are only shown to be effective to remove the noise added manually to the pure images. Recently, CNN (Convolutional Neural Networks) was also used to image denoising [2, 7, 18, 21]. CNN needs pure images as training samples, which is difficult for wild camera to record pure images. Thus, wildlife noisy images cannot be removed by the methods above.

Sparse representation is the key step in the process of image noise reduction. SVD (singular value decomposition) [11] proposed by Mairal et al. has a good performance in denoising images with additive noise. The combination of generalized K-singular value decomposition algorithm (K-SVD) and orthogonal matching pursuit(OMP) algorithm is also used to calculate the sparse coefficient, but it causes easily blurring or distortion [12, 16]. Some scholars decomposed the source images into low frequency sub-images and high frequency sub-images, and explored sparse representation in frequency domain [5, 14, 15]. Among them, discrete wavelet transform (DWT) is a common transformation models from spatial domain to frequency domain. To remove the multiplicative (speckle) noise, Gibran et al. proposed a novel sparse representation scheme based on DWT and maximum a posteriori (MAP)

Image processing technology has been widely used in wildlife protection.

The external factors have a impact on the quality of images when taking images. **Noisy images have a great influence on the wildlife image analysis and processing.**

Many methods were proposed for removing Gaussian noise, salt-and-pepper noise and Impulse Noise.

References

4. Djurovi I (2016) BM3D filter in salt-and-pepper noise removal [J]. EURASIP J Image Video Process 2016(1):1–11

BM3D (block matching and 3D filtering) is considered as a good method to remove noise by using image redundancy and looking for similar blocks in the image for filtering. The detail information can be preserved well after denoising the image.

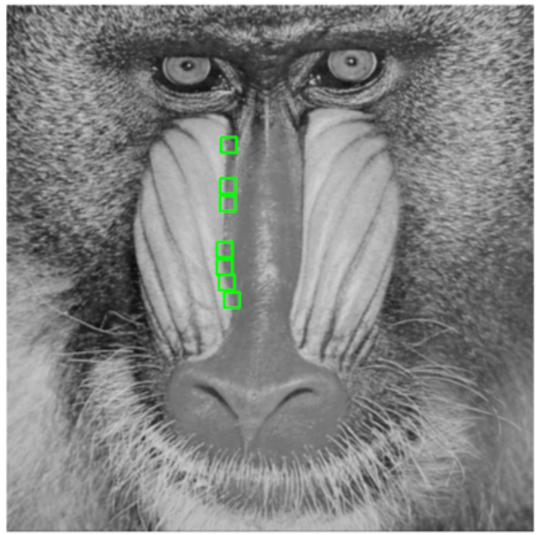


Fig. 1 Test image "baboon" with single set of similar blocks recognized in the BM3D algorithm

Algorithm. *BM3D filter.*
Consider an image region D centered around the pixel $x(n, m)$.

1. Search for regions D' that are similar to the considered region D. Note that this similarity check is performed in the transform domain (commonly Hadamard/Walsh transform). These transform coefficients are compared with threshold and those below the threshold are set to 0. Similar patches are selected to be relatively close to the considered patch in order to simplify search procedure.
2. Similar patches are put into the 3D matrix. 3D discrete linear transform of the 3D matrix is evaluated. Obtained transform coefficients are thresholded and all coefficients below the threshold are removed. Temporary filtered blocks are obtained using inverse 3D transform. Note that in practice, separated 2D/1D transforms are used instead of the 3D transform due to efficiency reasons.
3. These filtered blocks are returned back to the image. The procedure is performed for each pixel. Pixels belong to various number of different blocks. Therefore, an aggregation of the filtered blocks should be performed. Weighted coefficients in aggregation are calculated based on number of transform coefficients that are above the threshold in the 3D transforms. More coefficients above the threshold mean that there is more residual noise in the blocks and vice versa.
4. Similar blocks are found for the filtered image from the previous step.
5. Similar blocks form the 3D matrix.
6. The 3D transform is calculated for the 3D matrix.
7. 3D transform coefficients are filtered using the Wiener filter.
8. Inverse 3D filtering is performed to obtain the final

A few researches have focused on mixed Gaussian noise and impulse noise removal based on sparse representation. However, **the methods are only shown to be effective to remove the noise added manually to the pure images.**

Recently, **CNN** (Convolutional Neural Networks) was also used to image denoising. **CNN needs pure images as training samples, which is difficult for wild camera to record pure images.**

Dictionary learned with pure images without noise and, used for remove noise in noisy images.

Thus, wildlife noisy images cannot be removed by the methods above.

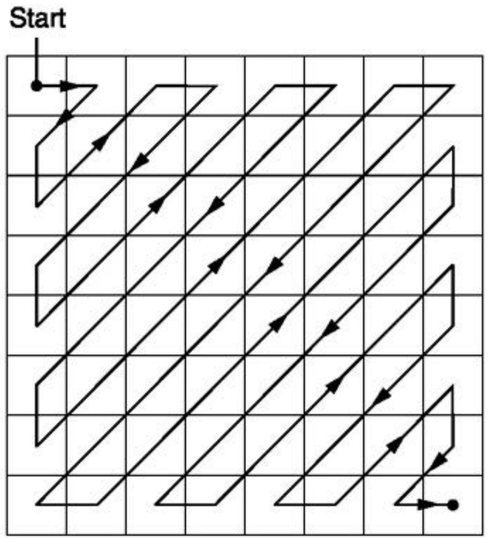
Our contributions The contributions of this paper are three folds.

- We propose a novel noise reduction method for real-world wildlife images. An original image is decomposed into an original image into a high frequency sub-image and a low frequency sub-image by ranking 2D-DCT coefficients based on the Zig-zag algorithm. We mostly denoise the high frequency sub-image and utilizes the low high frequency sub-image to reserve more image details.
- We propose a sparse representation of high frequency sub-image based on sparse coefficients error constraint. We apply the *lp*- *l1*-norm term for sparse coding to obtain the optimal solution of sparse coefficients and utilize proximal alternating linearization minimization algorithm for solving the objective function due to the nonconvex and non-smooth minimization problem.
- Although the proposed method is designed to recover the real-world wildlife images corrupted by noise, it is also effective to remove mixed Gaussian noise and impulse noise added artificially.

The high frequency sub-image is mostly used to remove noise while reducing **time-consuming**, and the low frequency sub-image reserves more image details.

We propose a sparse representation of high frequency sub-image based on sparse coefficients error constraint.

We apply the lp- l1-norm term for sparse coding to obtain the optimal solution of sparse coefficients and utilize proximal alternating linearization minimization algorithm for solving the objective function due to the non-convex and non- smooth minimization problem.

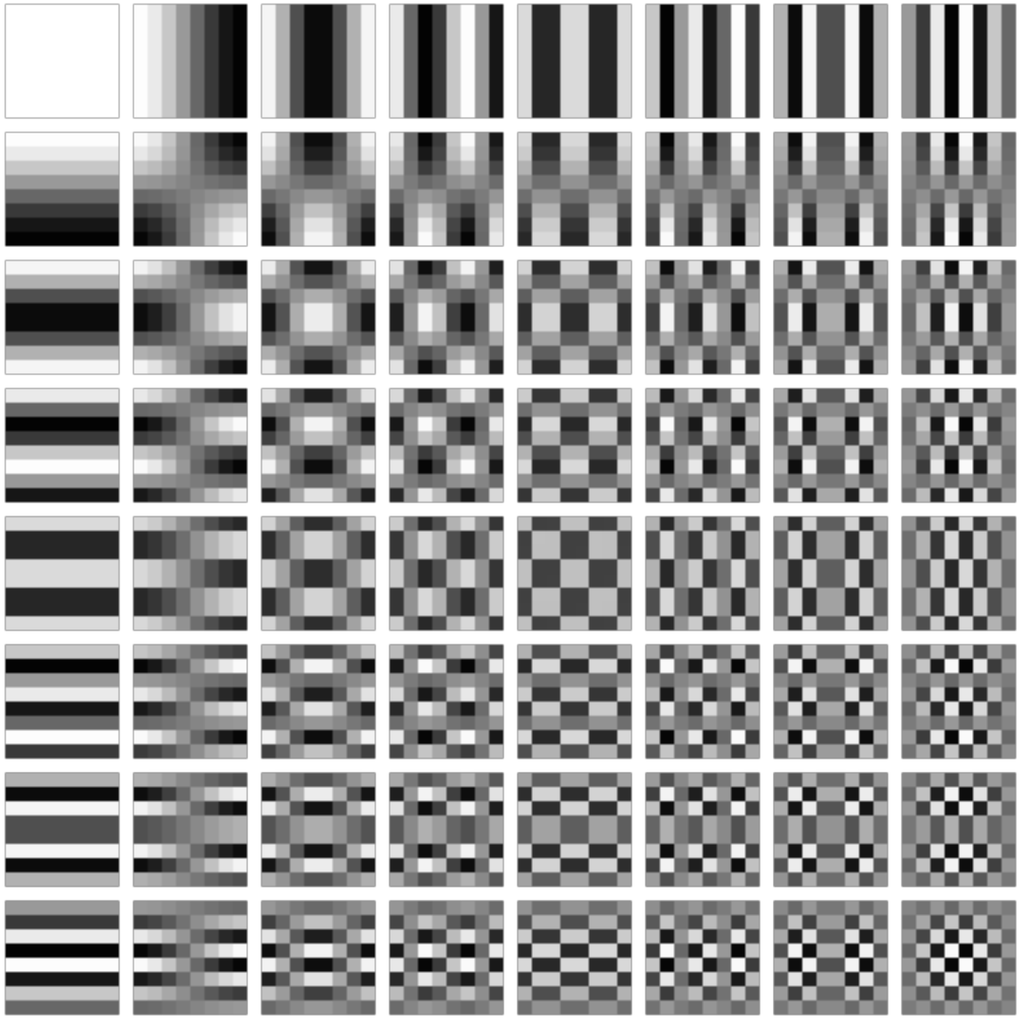


Decomposition of high and low frequency

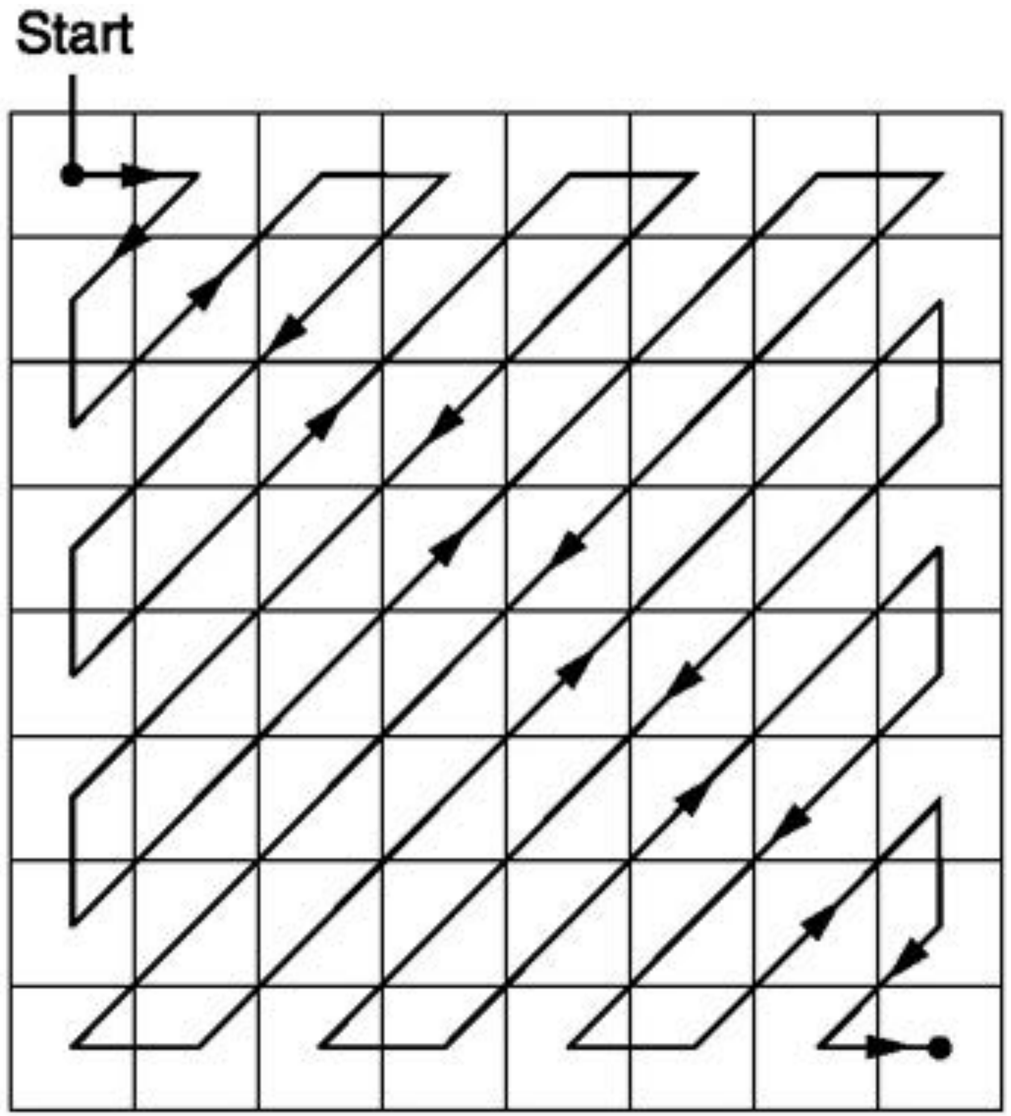
Firstly, the original image is divided into 8 × 8 blocks, and then the block-based DCT transform is carried out. The coefficients ranked in the top (8 × 8)/64 are selected as the low frequency coefficients that are sorted by Zig-zag to construct the low frequency sub-image. The high frequency sub-image is obtained by subtracting the low frequency image from the original image. High and low frequency decomposition of a wildlife image is shown in Fig. 2.

2-D DCT Dictionary

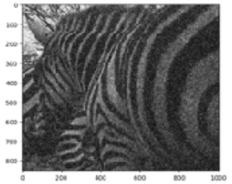
Example

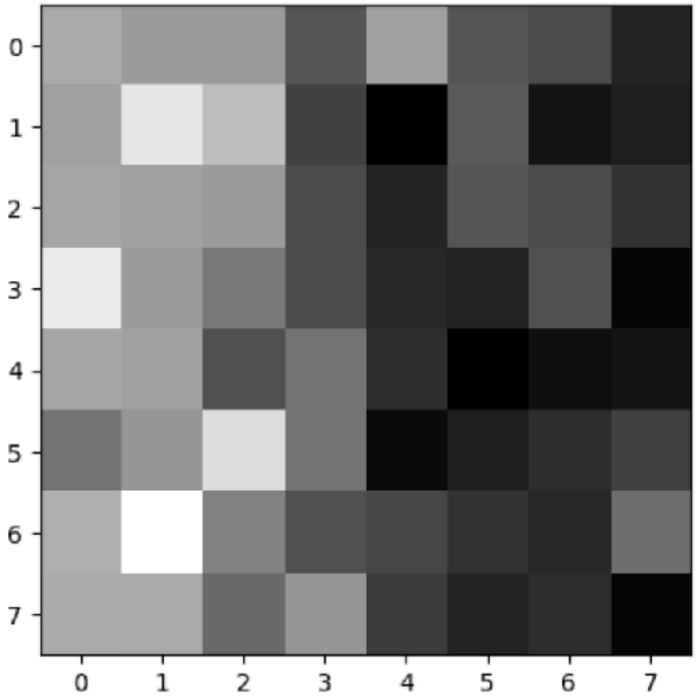


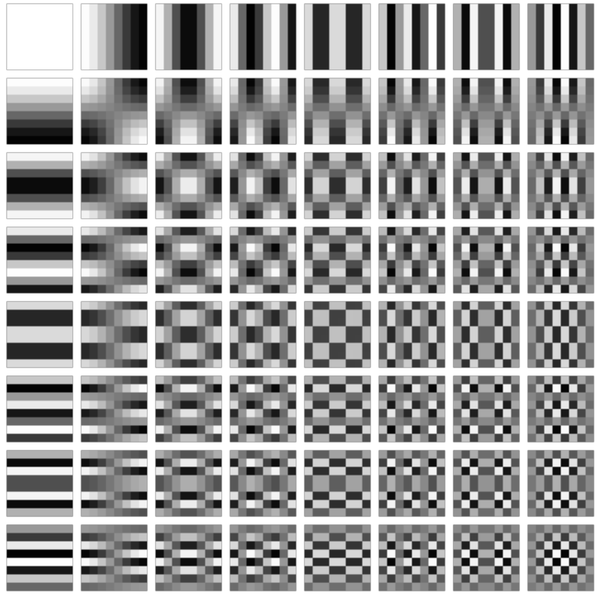
DCT-8*8

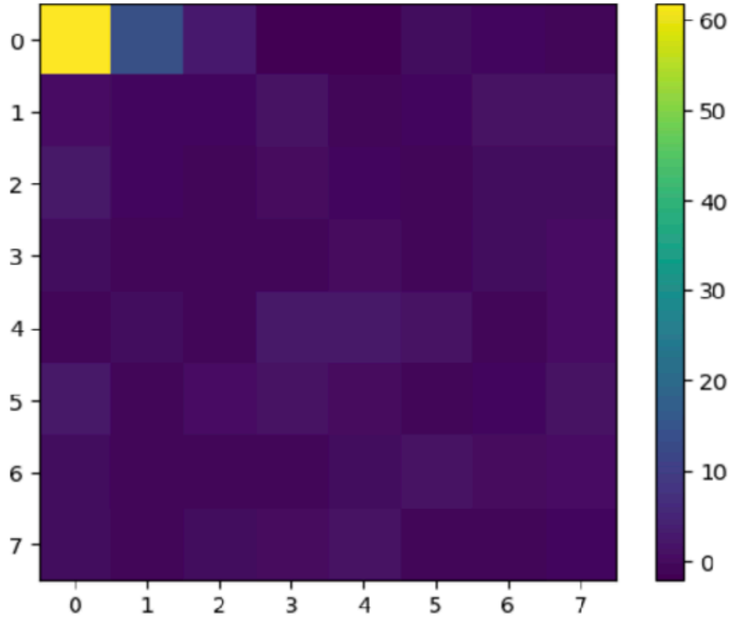


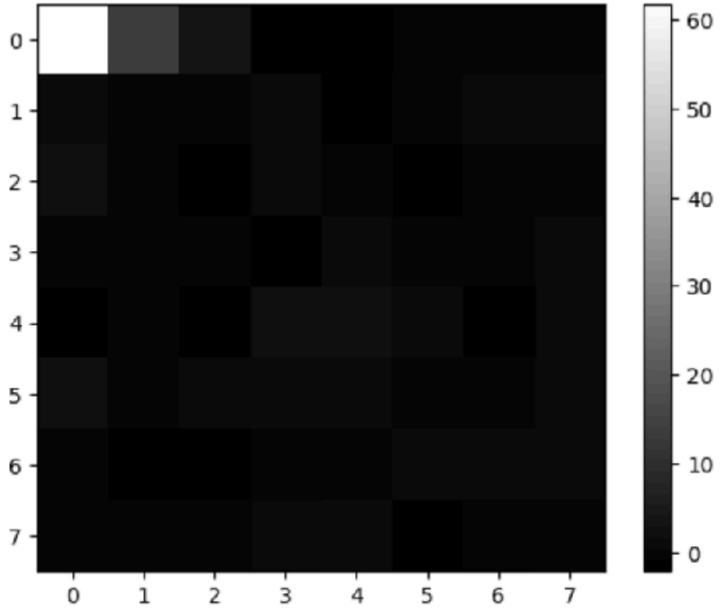
Examples of 8*8 DCT



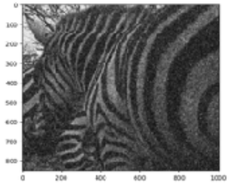


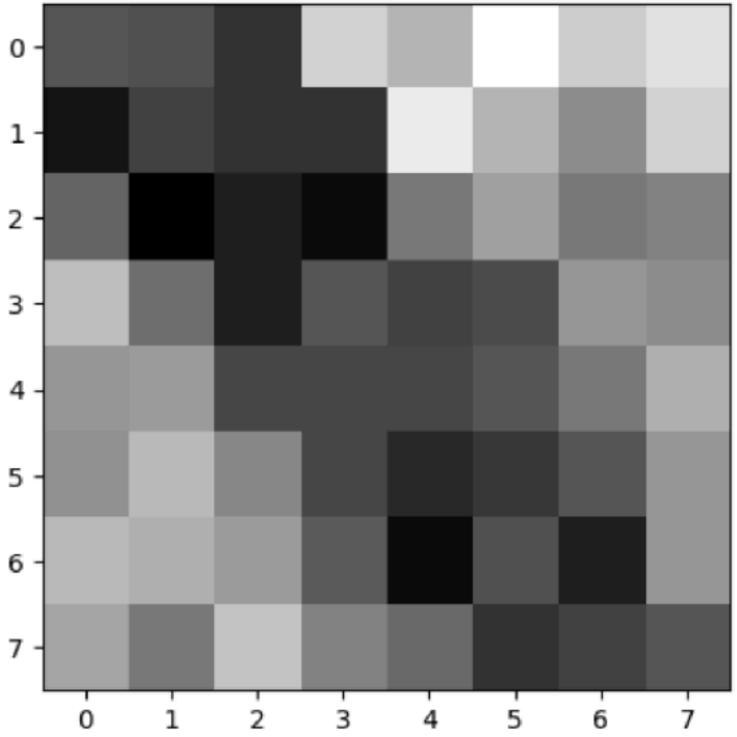


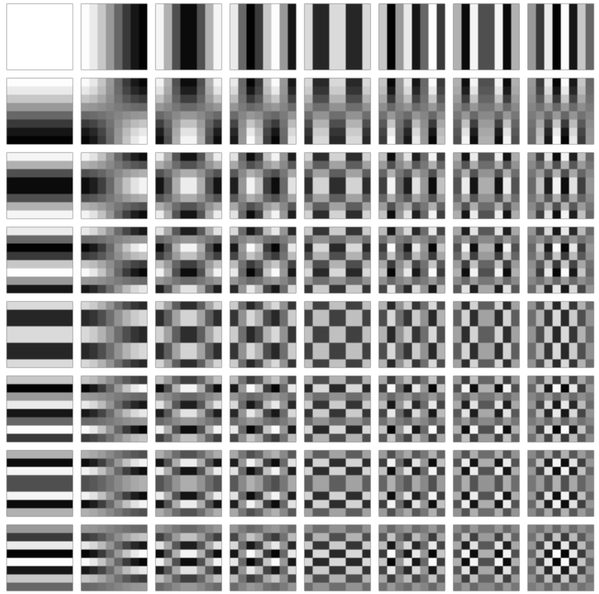


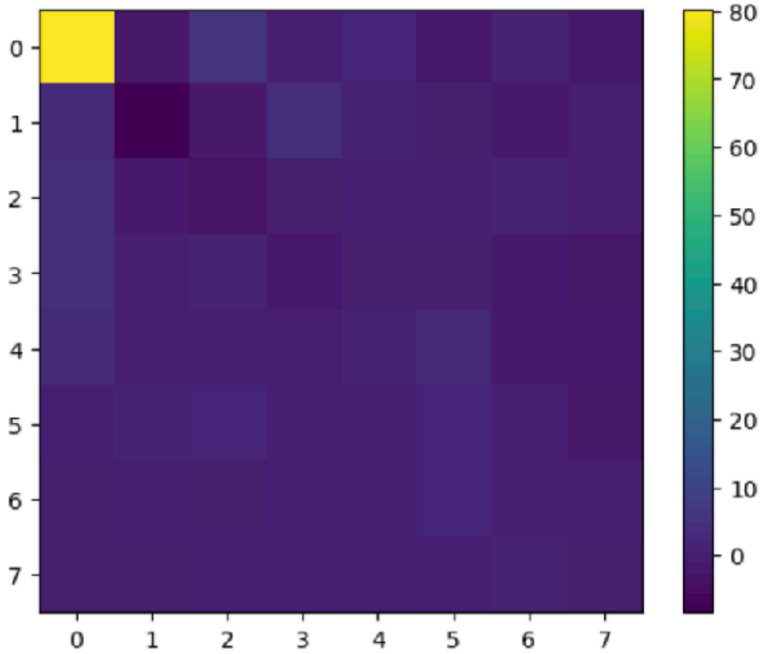


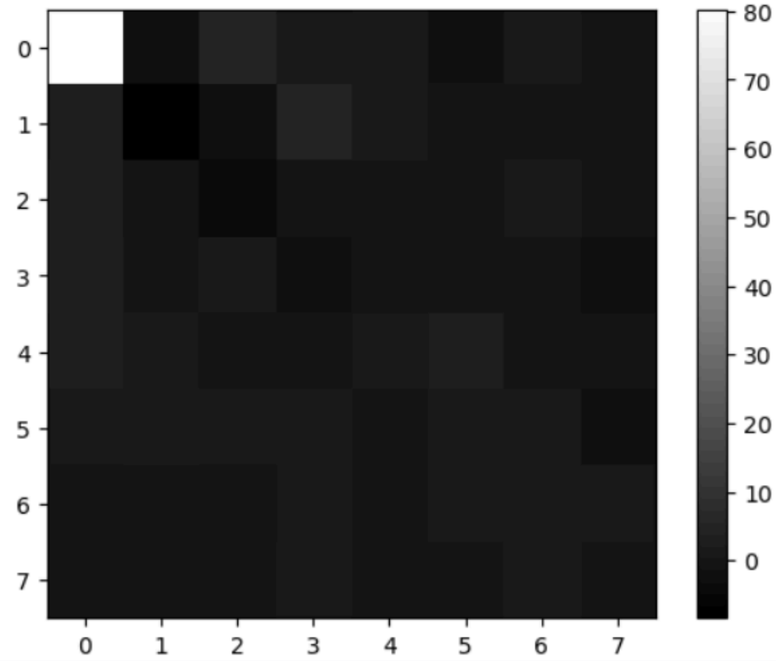
Examples of 8*8 DCT



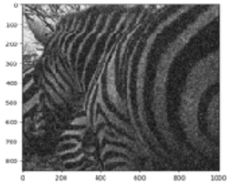


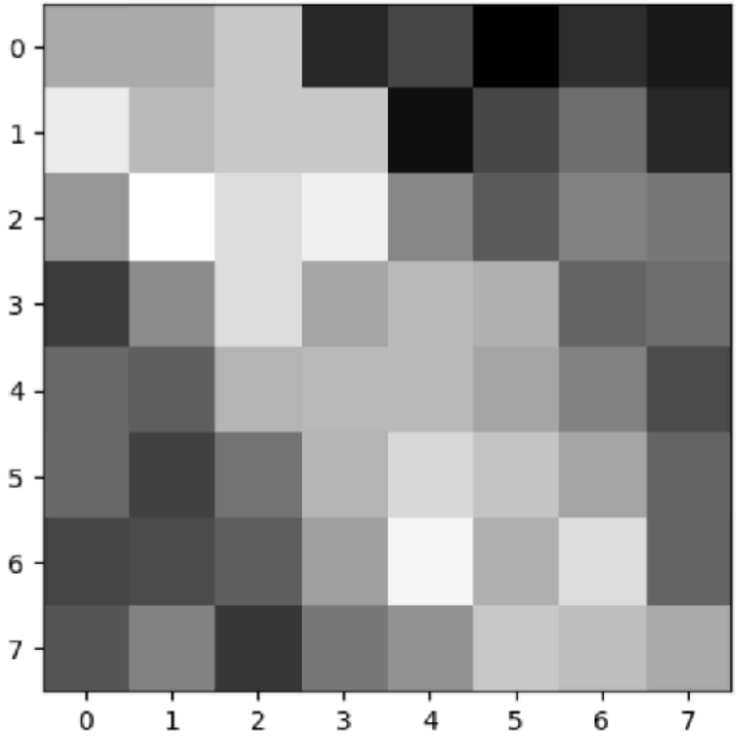


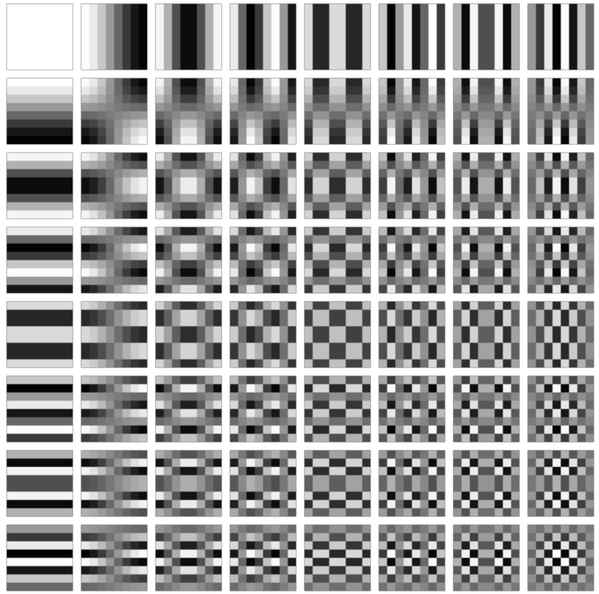


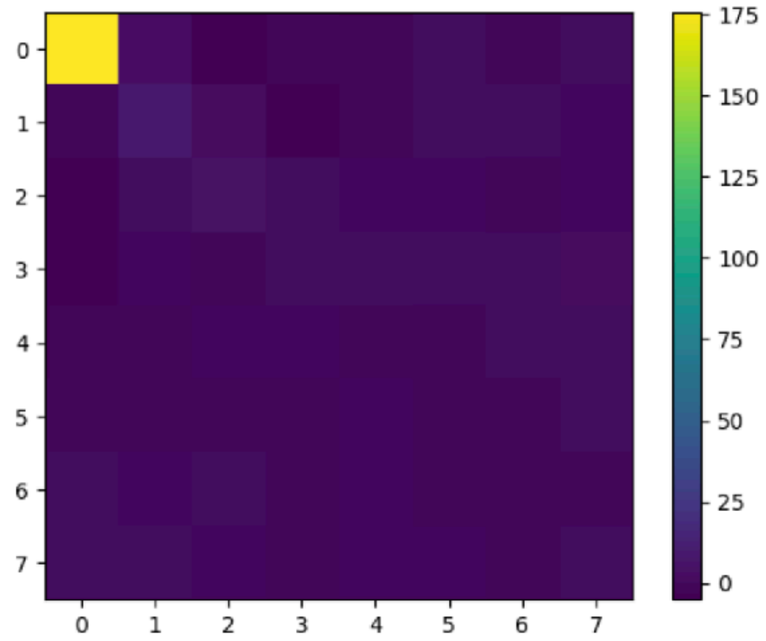


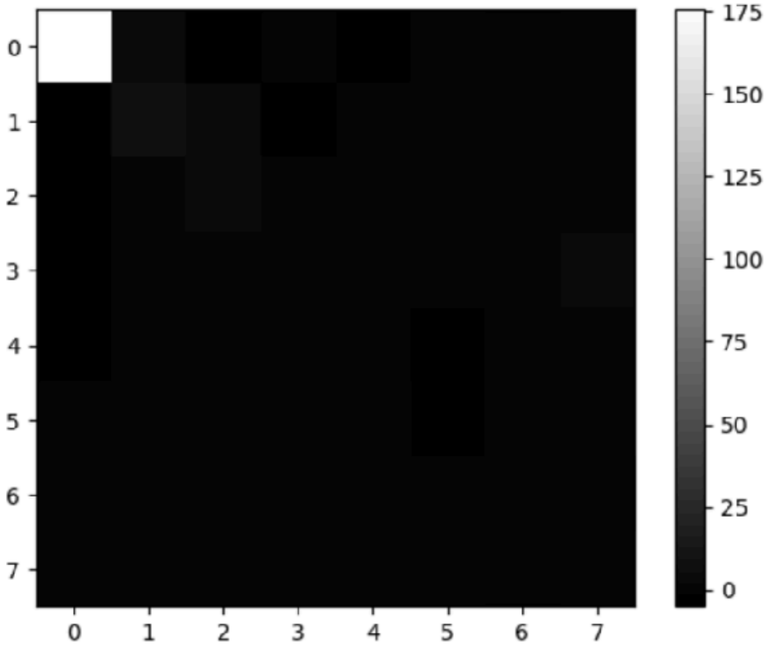
Examples of 8*8 DCT







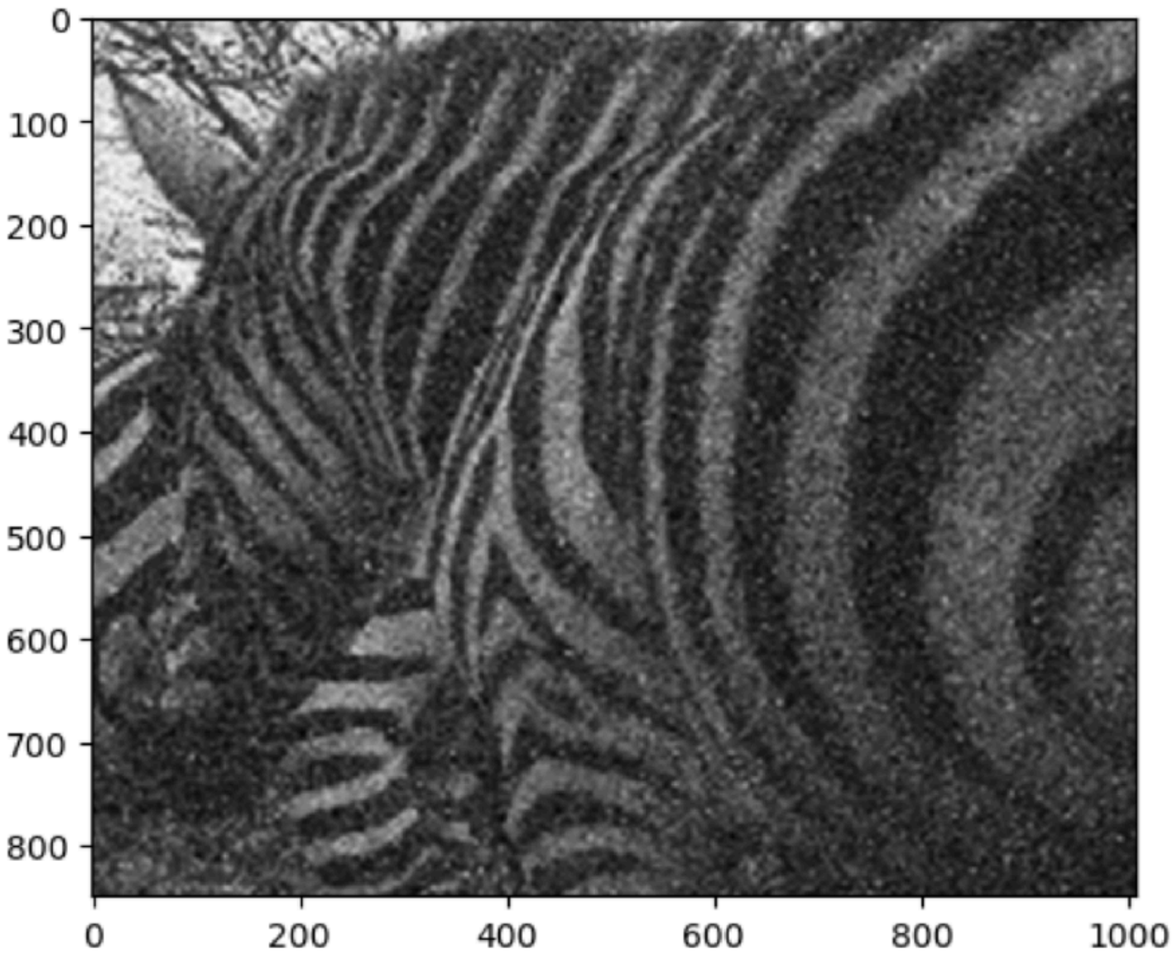




Block-Based 2-D DCT

Example

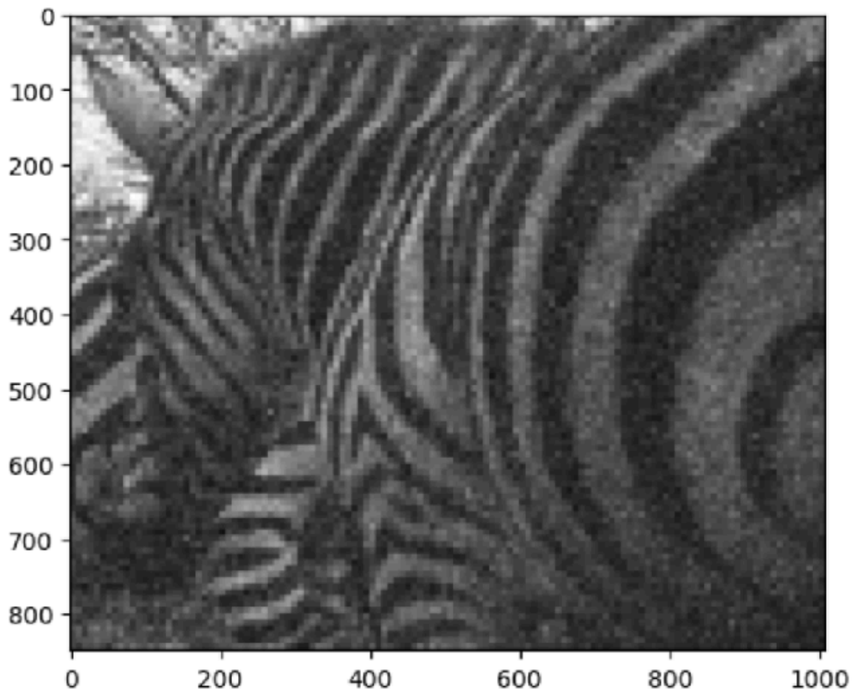
8*8 Block DCT

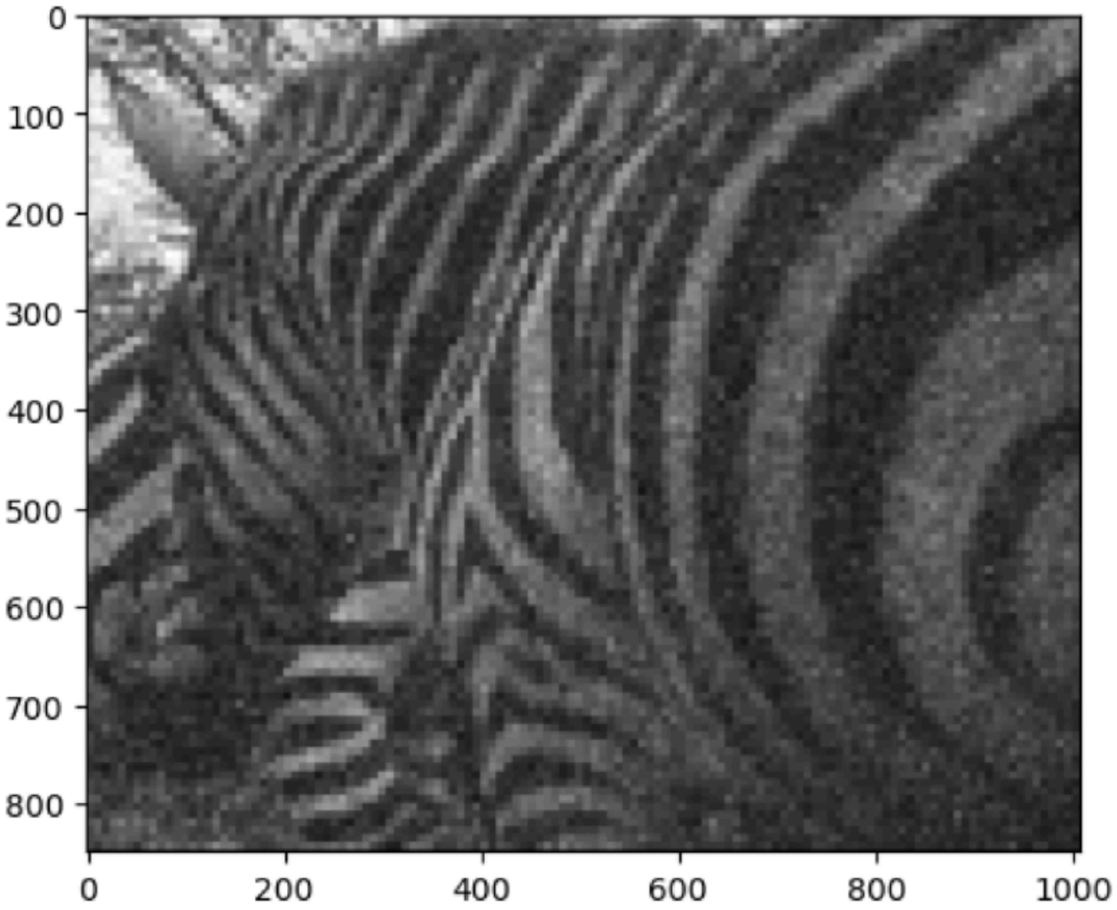


ORGINAL IMAGE

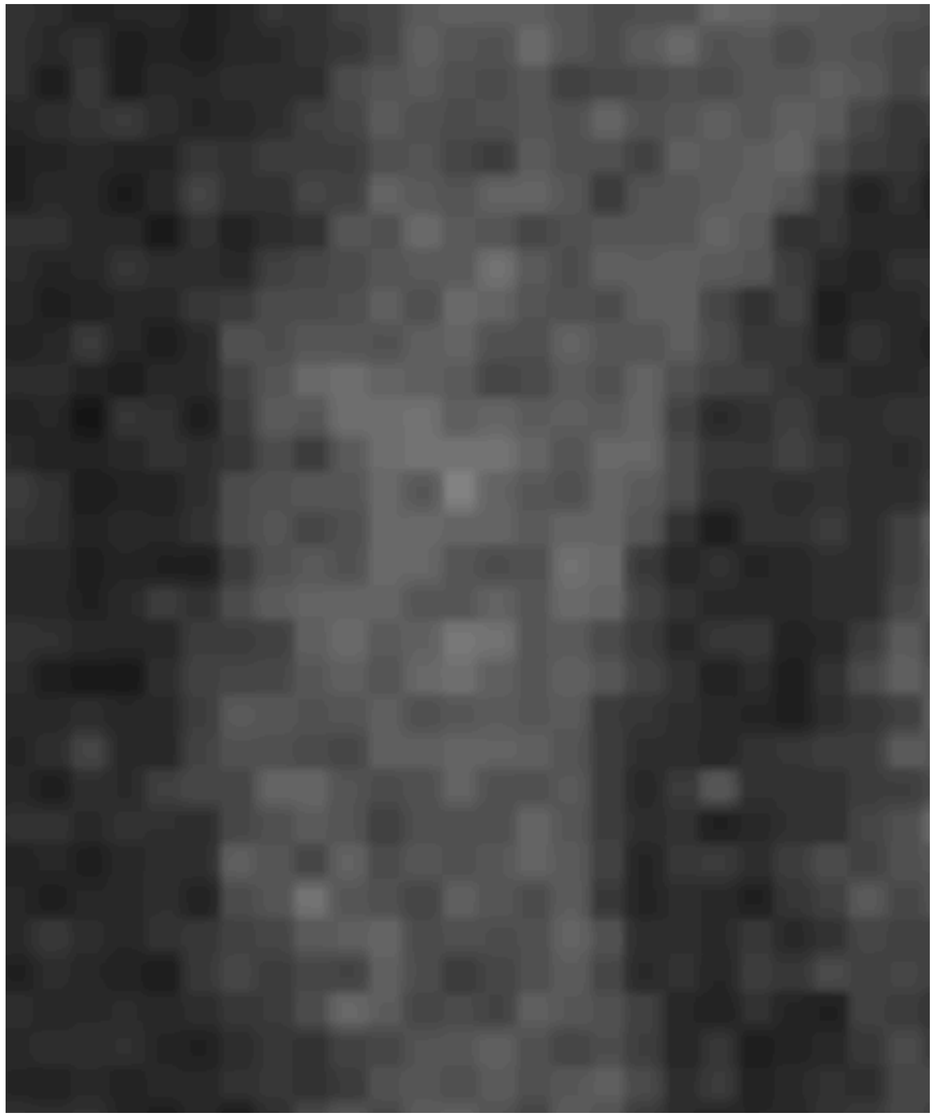
In [2]:
```
img_for_low = img.copy() / 255


for i in range(0,img.shape[0], 8):
    for j in range(0,img.shape[1], 8):

        patch = img_for_low[i:i+8,j:j+8]
        patch = dct2(patch)
        patch[1:] = 0
        patch[:,1:] = 0
        patch_low = idct2(patch)

        img_for_low[i:i+8,j:j+8] = patch_low


plt.imshow(img_for_low, cmap="gray")
```

Out[2]: <matplotlib.image.AxesImage at 0x7fdd5c6254f0>





LOW FREQUENCY SUB IMAGE



References

Chapter 4 - Image Enhancement in Frequency Domain. Dr Mohsen Ebrahimi Moghadam - Shahid Beheshti University

Introduction to fourier transform

- It is common practice to multiply input images with *(-1)* $^{x+y}$ *because*

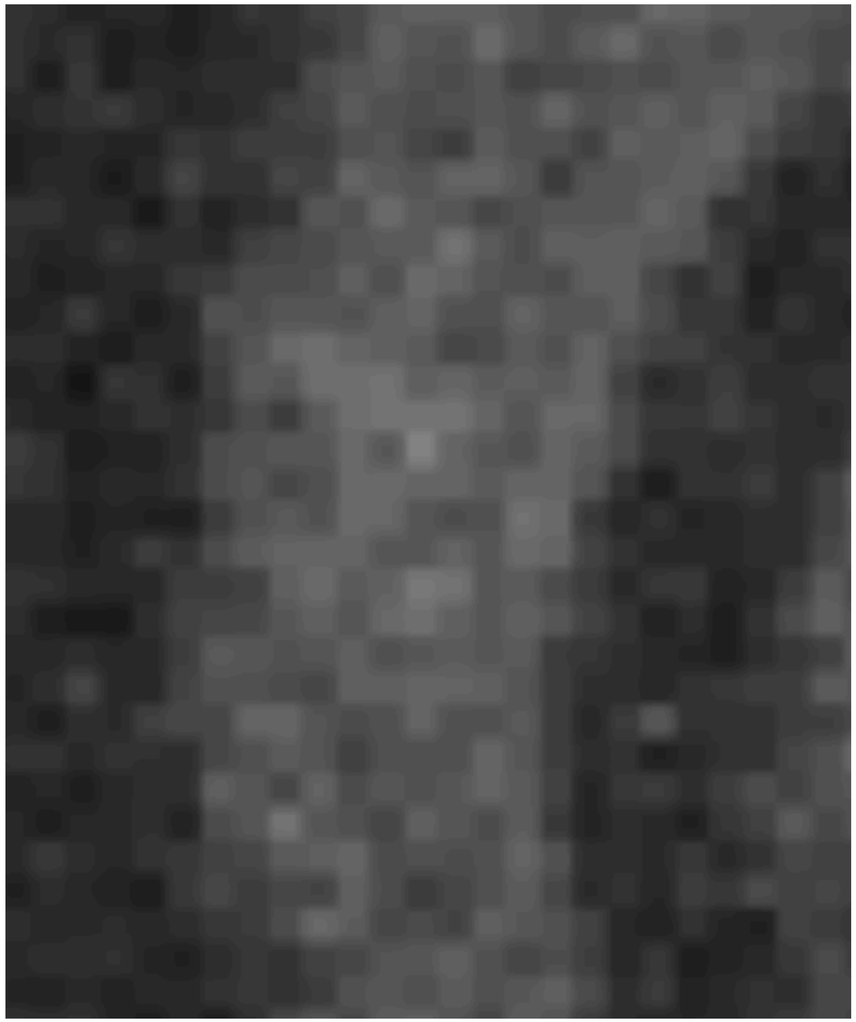
$$\Im\left[f(x,y)(-1)^{x+y}\right] = F(u - M/2, v - N/2)$$

- it shifts the origin of *F(u,v)* to frequency coordinates *(M/2, N/2)*
- In order to guarantee that these shifted coordinates are integers, M and N should be even numbers
- The value of transform at (u,v)=(0,0)

$$F(0,0) = \frac{1}{MN}\sum_{x=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)$$

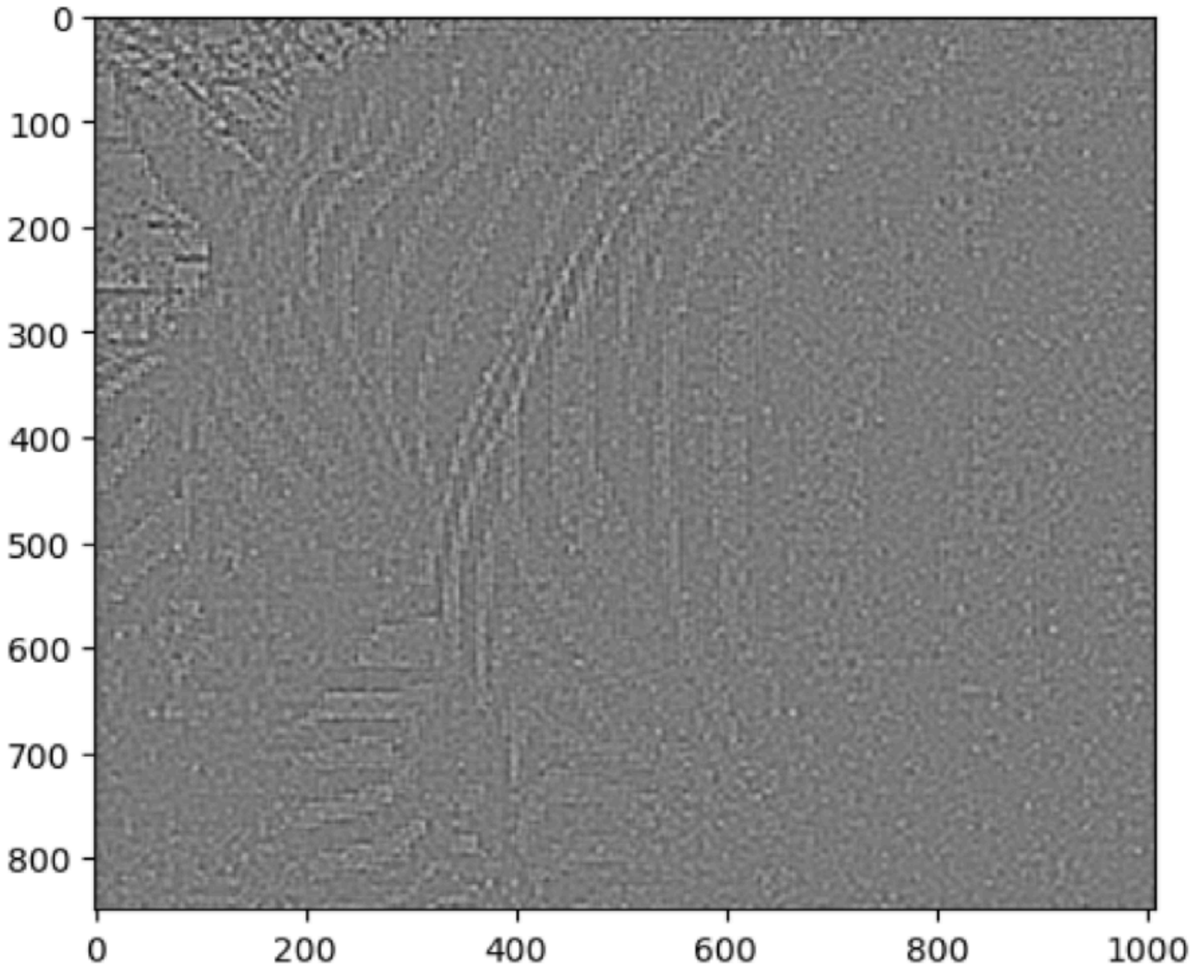
which is the average of *f(x,y)*

it is called dc component



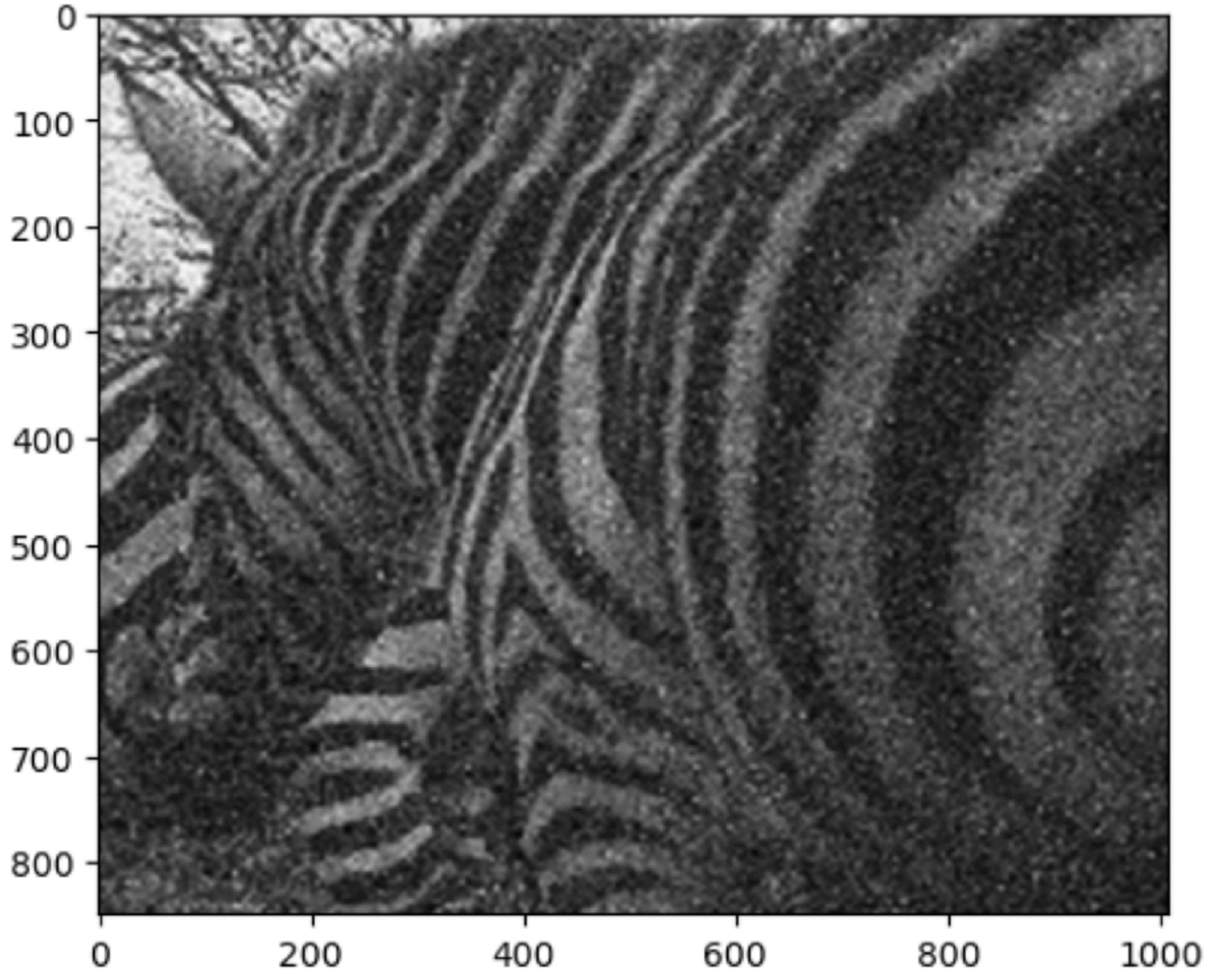
In [4]:
```
img_for_high = img.copy() - img_for_low

plt.imshow(img_for_high, cmap="gray")
```

Out[4]: <matplotlib.image.AxesImage at 0x7fdd5e8482e0>



In [5]:
```
res_img = cv.addWeighted(img_for_low, 0.5, img_for_high, 0.5, 0)

plt.imshow(res_img, cmap="gray")
```

Out[5]: <matplotlib.image.AxesImage at 0x7fdd5c728ee0>



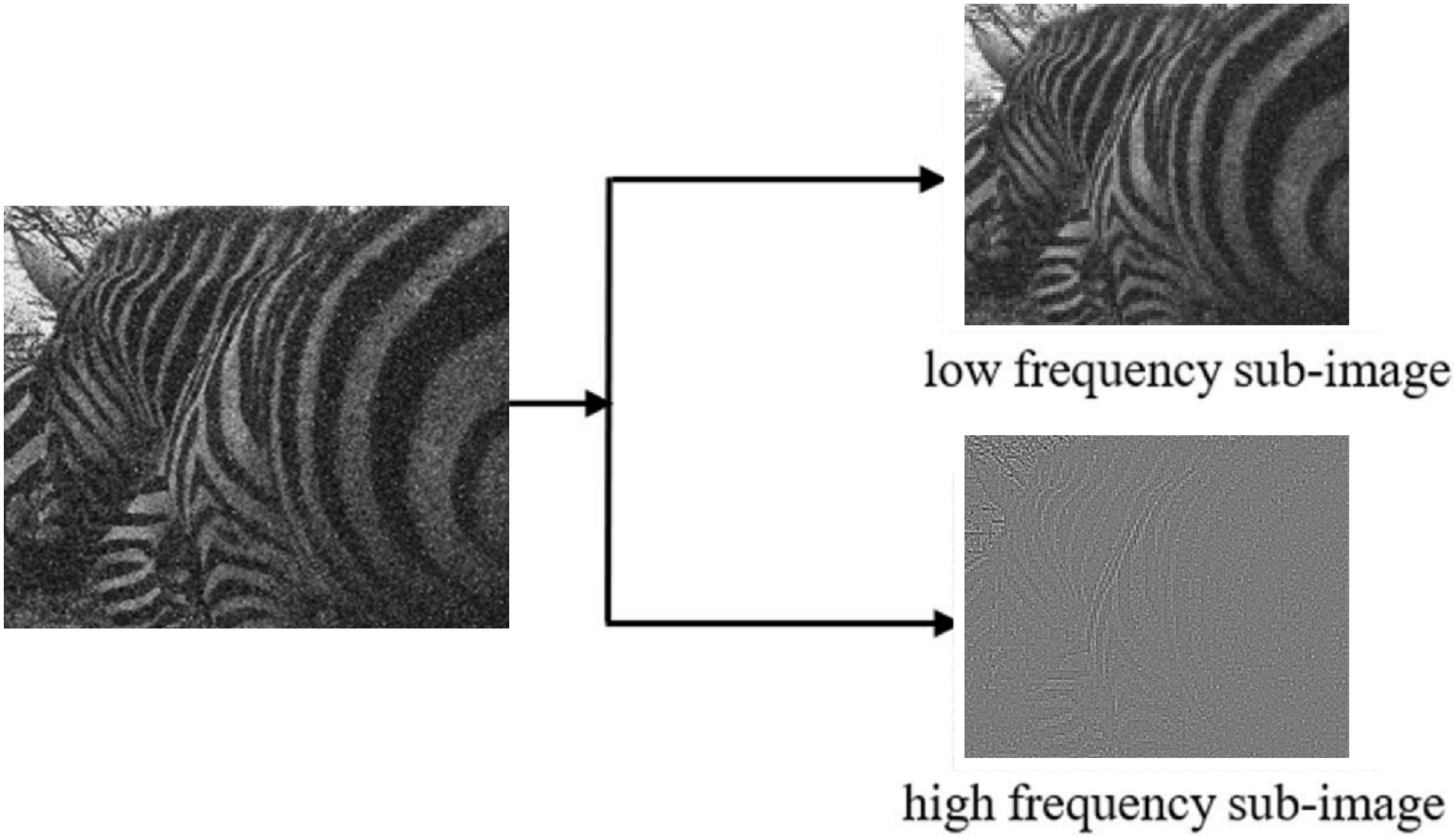


Fig. 2 An original image is separated into high frequency image separation and low frequency

Sparse Representation

Mixed noise reduction via sparse error constraint representation of high frequency image for wildlife image

A. *Problem formulation*

Given a training set of wildlife images $Y = [Y_1, Y_2, \ldots, Y_L]$, the purpose of dictionary learning is to reduce noise by following a new dictionary $D \in \mathbb{R}^{n \times b}$ and sparse coefficients is $X = [X_1, X_2, \ldots, X_b]$, where n denotes the number of the image patches, then the sparse representation model is formulated as follows [13]:

$$\underset{D,X}{argmin} \left\{ \|Y_i - DX_i\|_F^2 \right\} st. \|X_i\|_0 \leq \varepsilon \tag{1}$$

where $\|\cdot\|_0$ is the l^0 norm, counting the nonzero entries of a vector.

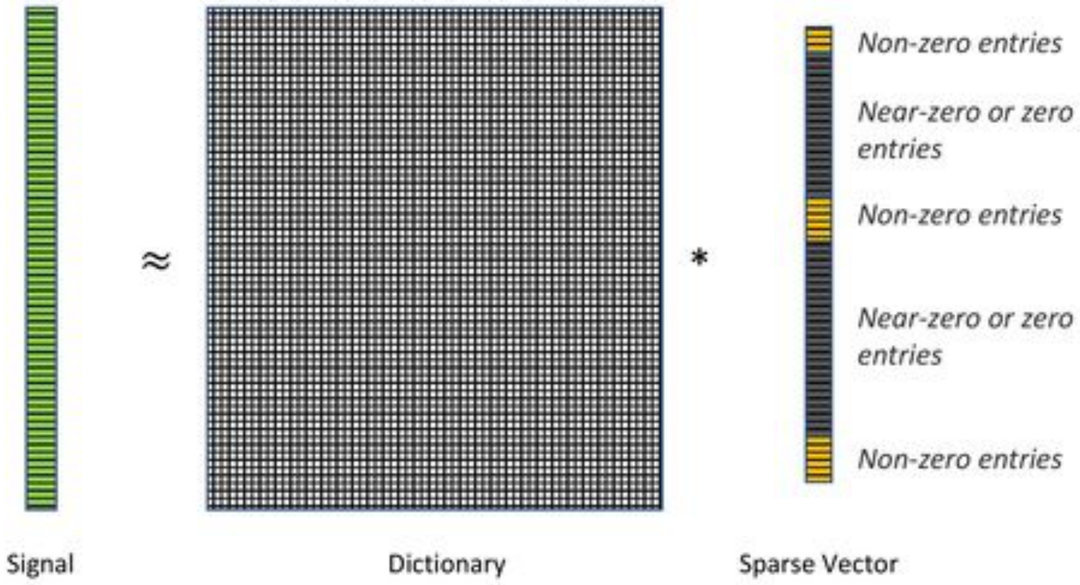
Sparse Representation

$$x \approx D * a$$

$x \in \mathcal{R}^n$
is a signal

$D = [d1, d2..dn] \in \mathcal{R}^{nxm}$
$d^T d = 1$
is a set of normalized "basis vectors" or "atoms"

$a \in \mathcal{R}^m$
is a sparse vector that solves the equation



References

13. Pan H, Jing Z, Qiao L, Li M (2018) Discriminative structured dictionary learning on Grassmann manifolds and its application on image restoration[J]. IEEE Trans Cybern 48(10):2875–2885

Sparse representation has been viewed as a powerful tool for representing high-dimensional signals. This scheme [29]–[31] is a patch-based approach by constructing over-complete dictionaries and implementing sparse coding. Xiao *et al.* [26] reconstructed the latent image by solving ℓ_1-ℓ_0 constraint-based problem, which turns out to be advantageous in the removal of mixed Gaussian-impulse noise. This method involved three steps, including outliers detection [16] by rank-order absolute difference detector, initially approximating [30] by K-singular value decomposition (K-SVD) and enhancing the latent image by minimizing the resulting energy function. Recently, a mixed soft-hard thresholding-based sparse representation method [27] was presented. It can be seen that the construction of over-completed dictionary is important to the success of these methods. Among them, the update of

$$DC \simeq X \tag{1}$$

where D is a dictionary matrix with the size of $m \times k$, and C denotes the corresponding coefficients matrix. The main idea of this scheme is to represent matrix C by a predefined number of blocks. Each block corresponds to a fixed set of columns of D. Subdictionaries of D are selected for stabilizing the encoding procedure, which denoted by $[B_1, \ldots, B_K]$. For a fixed cluster coming from n columns of X, its main idea can be expressed as follows:

$$\min_{D,C} \ \|X - DC\|_F^2$$

References

13. Pan H, Jing Z, Qiao L, Li M (2018) Discriminative structured dictionary learning on Grassmann manifolds and its application on image restoration[J]. IEEE Trans Cybern 48(10):2875–2885

In this paper, we focus on structured sparse model selection approach [15] (SSMS). For a xed width $\sqrt{m}$, SSMS merges the image plane into a data matrix X $(m \times n)$ with the patches by $\sqrt{m} \times \sqrt{m}$. And, n denotes the number of the patches. Then, X can be represented by

$$DC \simeq X \tag{1}$$

where D is a dictionary matrix with the size of $m \times k$, and C denotes the corresponding coef cients matrix. The main idea of this scheme is to represent matrix C by a prede ned number of blocks. Each block corresponds to a xed set of columns of D. Subdictionaries of D are selected for stabilizing the encod-

Mixed noise reduction via sparse error constraint representation of high frequency image for wildlife image

$$\underset{D,X}{argmin}\left\{\|Y_i - DX_i\|_F^2\right\} \; st. \|X_i\|_0 \leq \varepsilon$$

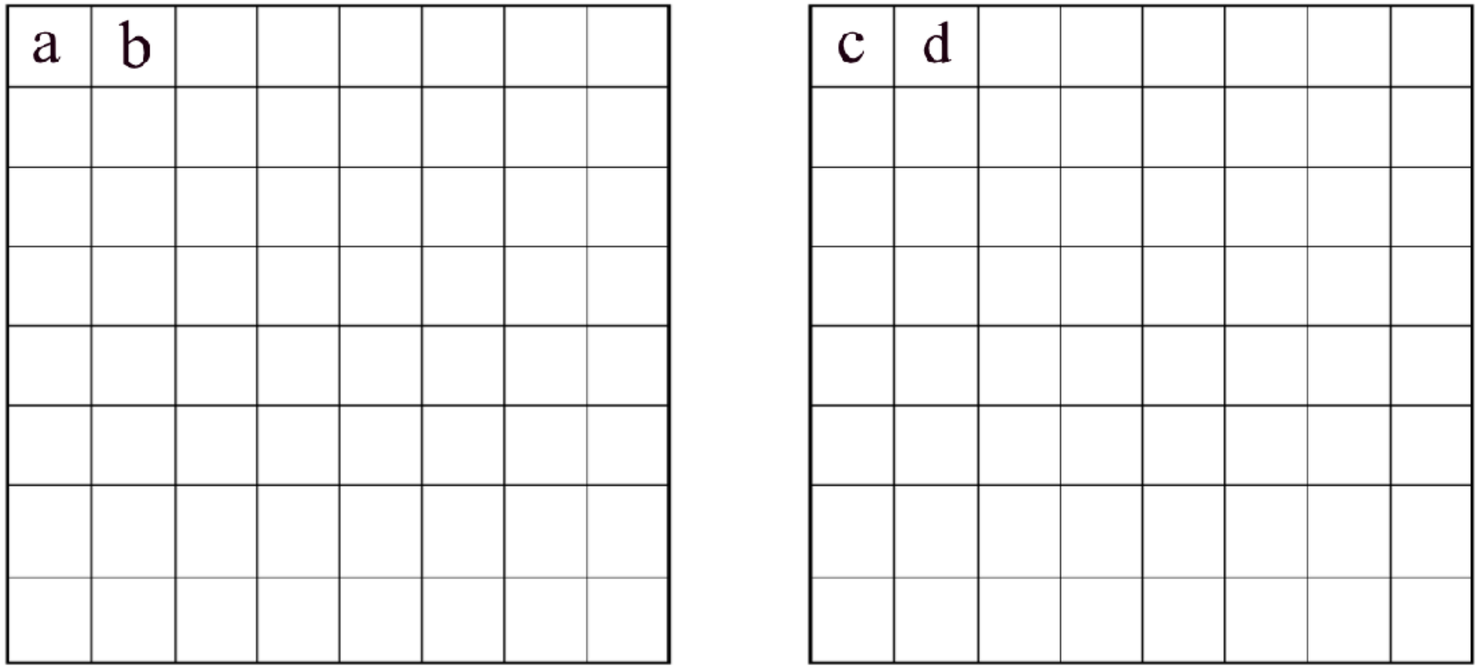
$$\boxed{\underset{D,X}{argmin}\left\{\|Y_i - DX_i\|_F^2\right\}} \; st. \|X_i\|_0 \leq \varepsilon$$

$$\|\mathbf{A}\|_F = \sqrt{\sum_{ij} A_{ij}^2}.$$

References

16. Wei D, Zhou J (2016) Application of K-SVD and OMP algorithm on ultrasonic signal denoising [J]. J Appl Acoust 35(2):95–100

Here we have concatenated all the examples $\mathbf{y}_i$ as columns of the matrix $\mathbf{Y}$ and similarly gathered the representations coefficient vectors $\mathbf{x}_i$ to build the matrix $\mathbf{X}$. The notation $\|\mathbf{A}\|_F$ stands for the Frobenius norm, defined as $\|\mathbf{A}\|_F = \sqrt{\sum_{ij} A_{ij}^2}$.



Y_i 8*8 Blocks D_iX_i

$$\|Y_i - DX_i\|_F^2 = \left(\sqrt[2]{(a-c)^2 + (b-d)^2 + \ldots}\right)^2$$

References

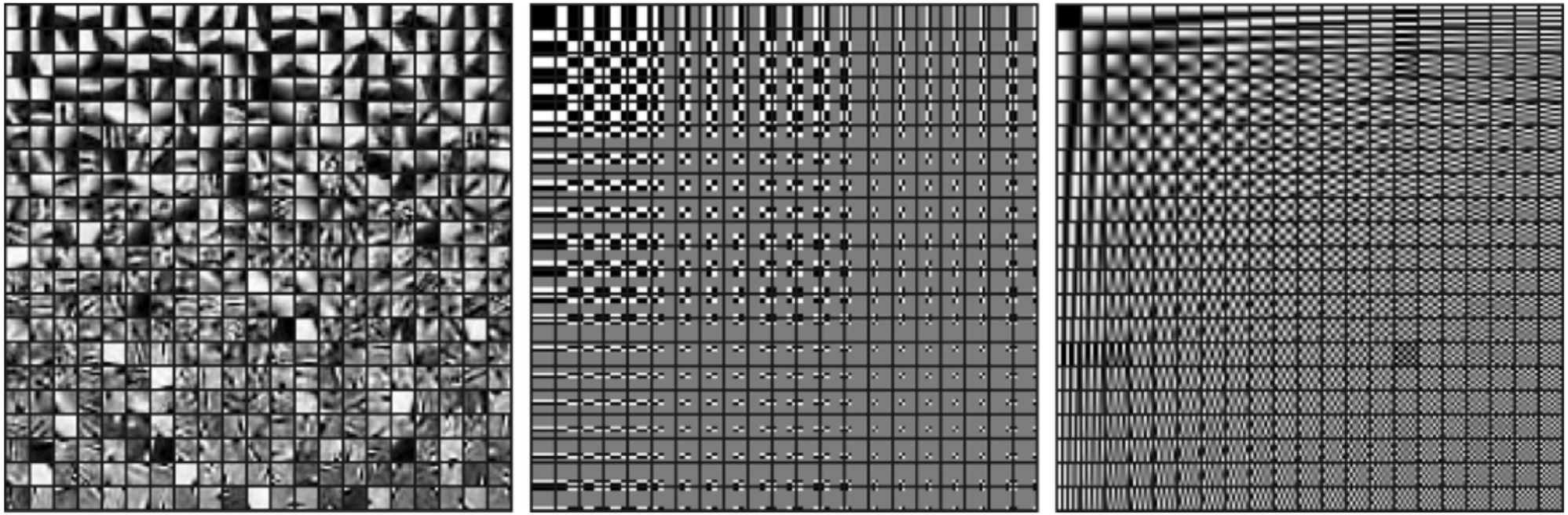
12. Meng Q, Li D, Chen S (2020) Sparse representation and reconstruction of image based on K-SVD dictionary learning[C]. 2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS). pp. 457–462

RECENT years have witnessed a growing interest in the search for sparse representations of signals. Using an overcomplete dictionary matrix $\mathbf{D} \in \mathbb{R}^{n \times K}$ that contains K prototype signal-atoms for columns, $\{\mathbf{d}_j\}_{j=1}^{K}$, a signal $\mathbf{y} \in \mathbb{R}^n$ can be represented as a sparse linear combination of these atoms. The representation of $\mathbf{y}$ may either be exact $\mathbf{y} = \mathbf{D}\mathbf{x}$ or approximate, $\mathbf{y} \approx \mathbf{D}\mathbf{x}$, satisfying $\|\mathbf{y} - \mathbf{D}\mathbf{x}\|_p \leq \epsilon$. The vector $\mathbf{x} \in \mathbb{R}^K$ contains the representation coefficients of the signal $\mathbf{y}$. In approximation methods, typical norms used for measuring the deviation are the ℓ^p-norms for $p = 1, 2$, and ∞. In this paper, we shall concentrate on the case of $p = 2$.

Thus X is Sparse Vector

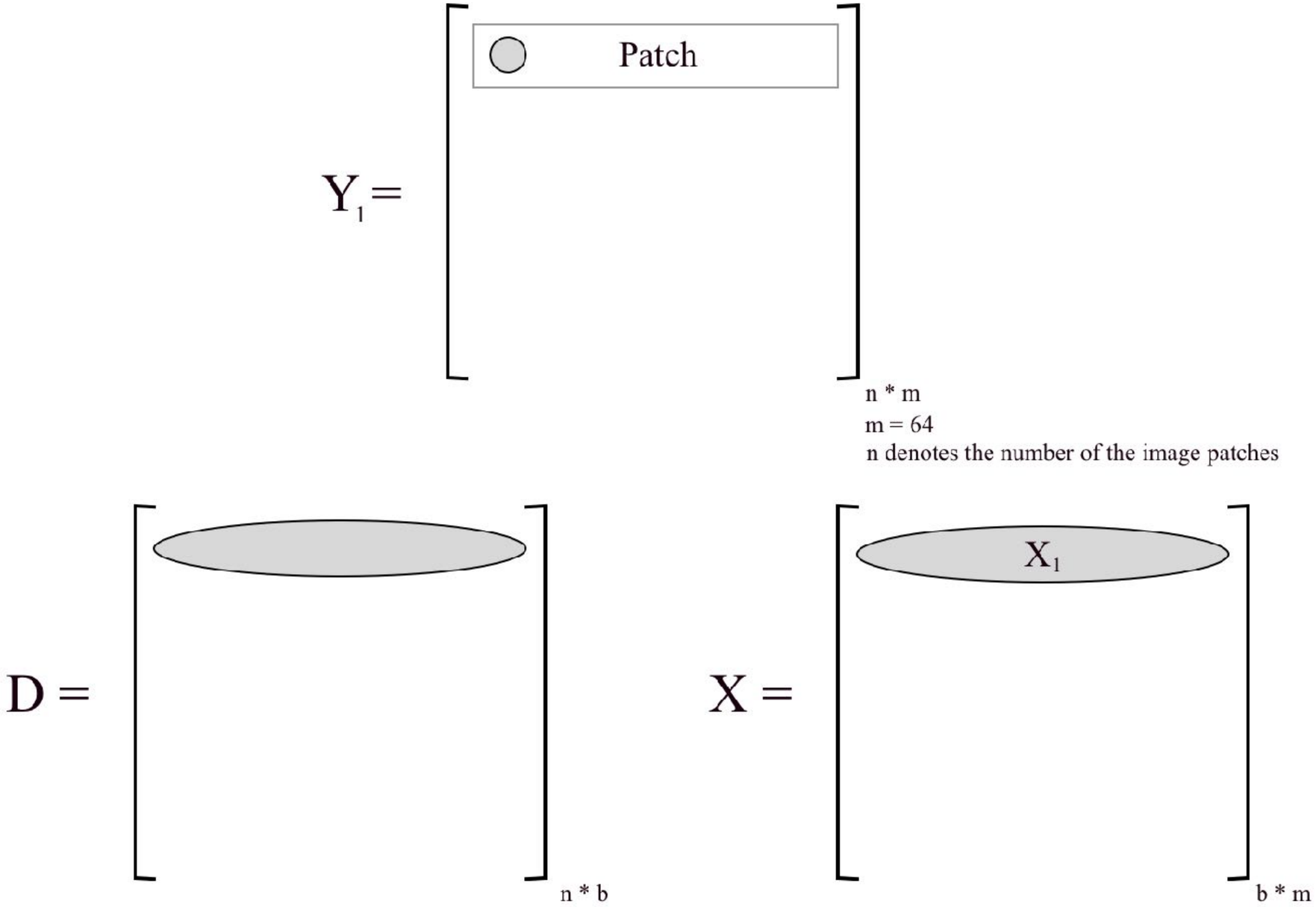
An overcomplete dictionary $\mathbf{D}$ that leads to sparse representations can either be chosen as a prespecified set of functions or designed by adapting its content to fit a given set of signal examples.

Fig. 4. A collection of 500 random blocks that were used for training, sorted by their variance.



(a) (b) (c)

Fig. 5. (a) The learned dictionary. Its elements are sorted in an ascending order of their variance and stretched to maximal range for display purposes. (b) The overcomplete separable Haar dictionary and (c) the overcomplete DCT dictionary are used for comparison.


Patch
$Y_1 =$
n * m
m = 64
n denotes the number of the image patches
D =
n * b
X =
X_1
b * m

$$\underset{D,X}{argmin} \left\{ \| Y_i - DX_i \|_F^2 \right\} \boxed{st. \| X_i \|_0 \leq \varepsilon}$$

where $\|\cdot\|_0$ is the l^0 norm, counting the nonzero entries of a vector.

Sparsity of Sparse matrix

$$\underset{D,X}{argmin} \left\{ \|Y_i - DX_i\|_F^2 \right\} \; st. \|X_i\|_0 \leq \varepsilon$$

L1 Regolarization

We want to optimize the error function and keep sparsity of X

Update the D and X vectore and keep sparsity of X

$$N_y = \underset{\{N_i\},\{D_i\}}{argmin} \sum_{i=1}^{K} \left(\|Y_i - D_i X_i\|_F^2 + \lambda_i \|N_i\|_1 \right)$$

L1 Regolarization

$$\text{L1 norm: } \|\mathbf{w}\|_1 = \sum_{i}^{n} |w_i|$$

Add l-1 norm of sparse vector to loss function

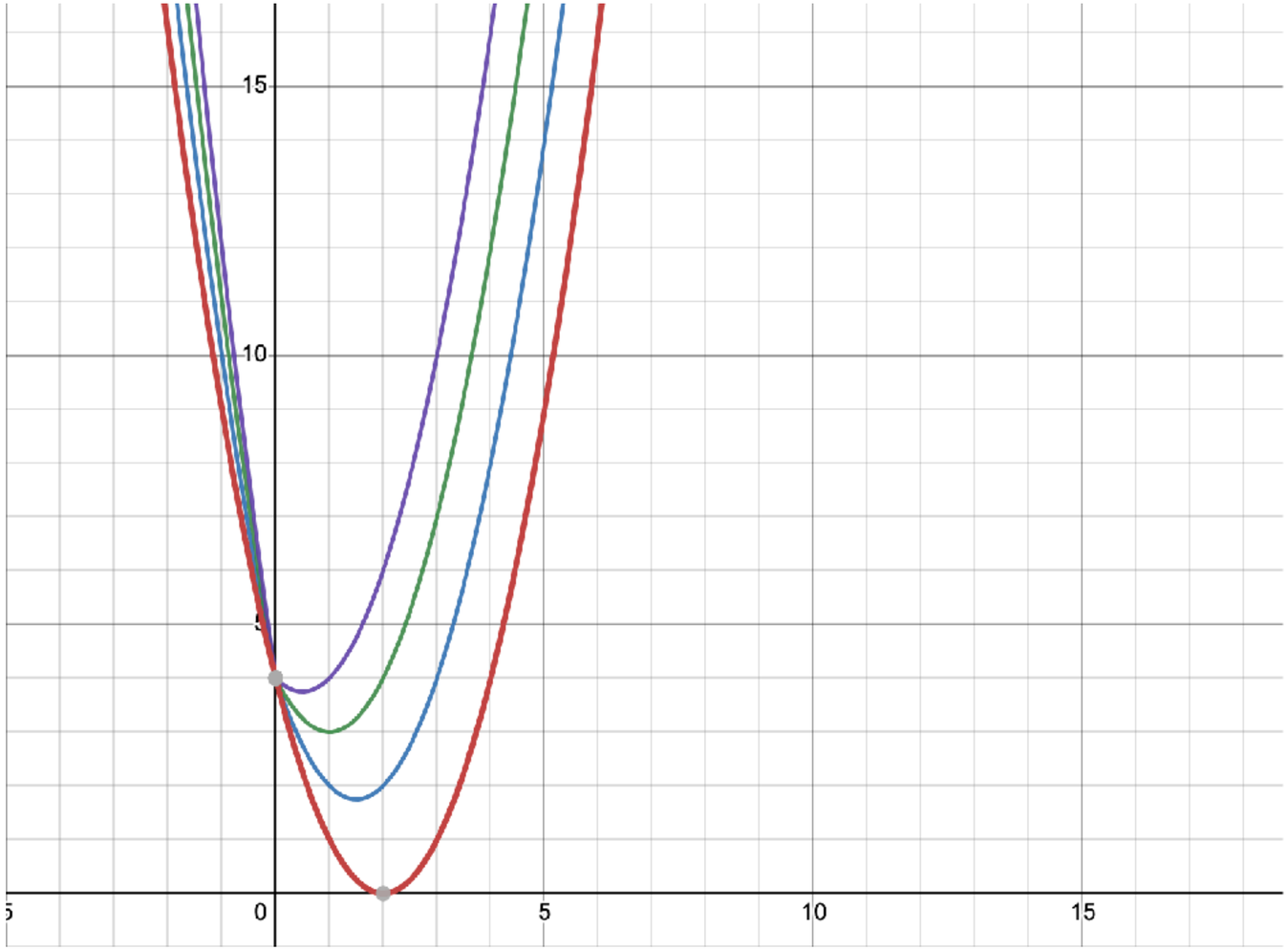
L1 Regolarization Example:

1 $(x-2)^2$

2 $(x-2)^2 + |x|$

3 $(x-2)^2 + 2|x|$

4 $(x-2)^2 + 3|x|$



Lamda : control sparsity of X

A. *Problem formulation*

Given a training set of wildlife images $Y = [Y_1, Y_2, \ldots, Y_L]$, the purpose of dictionary learning is to reduce noise by following a new dictionary $D \in \mathbb{R}^{n \times b}$ and sparse coefficients is $X = [X_1, X_2, \ldots, X_b]$, where n denotes the number of the image patches, then the sparse representation model is formulated as follows [13]:

$$\underset{D,X}{argmin} \left\{ \|Y_i - DX_i\|_F^2 \right\} st. \|X_i\|_0 \leq \varepsilon \tag{1}$$

where ε denotes the sparsity level. Let N_y denote the optimal value of the sparse coefficient N_i by $l2$-$l1$-norm, then N_y can be expressed as [13]:

$$N_y = \underset{\{N_i\},\{D_i\}}{argmin} \sum_{i=1}^{K} \left(\|Y_i - D_i N_i\|_F^2 + \lambda_i \|N_i\|_1 \right) \tag{2}$$

where each column vector X_i of the length L is the sparse representation of the image Y_i over the dictionary D_i, $\lambda_i (\lambda_i > 0)$ is a regularization parameter, and K is the degree of freedom in the model selection. We only use the formula to construct the initial dictionary instead of noise removal due to the lack of the pure image.

$$N_y = \underset{\{N_i\},\{D_i\}}{argmin} \sum_{i=1}^{K} \left(\|Y_i - D_i X_i\|_F^2 + \lambda_i \|N_i\|_1 \right)$$

$$NH_y = \underset{\{NH_i\},\{DH_i\}}{argmin} \sum_{i=1}^{K} \left(\|YH_i - DH_i NH_i\|_F^2 + \lambda_i \|NH_i\|_1 \right)$$

References

12. Meng Q, Li D, Chen S (2020) Sparse representation and reconstruction of image based on K-SVD dictionary learning[C]. 2020 International Conference on Intelligent Computing, Automation and Systems (ICICAS). pp. 457–462

2.2. Weighted l_p Norm Minimization. The norm minimization problem aims to estimate the true sparse result under certain constraint conditions. Traditional patch-based sparse coding is resolved by using the l_1 norm and the weighted l_1 norm, regularly. However, for some image inverse problems, such as image deblurring, image denoising, and other image restoration problems, the convex regularization does not get accurate sparse results. The reason is that the sparsity of an image cannot be measured by any benchmark. In [26], in order to improve the accuracy of sparse representation consequence, the convex l_1 norm was substituted by the nonconvex l_p norm. For image denoising, the weighted l_p $(0 < p < 1)$ penalty function is extended to the sparse representation based patch group, and the weighted l_p norm minimization (WPNM) can be represented as

Inspired by [13, 19], we use the following *lp-l*1-norm(0 < p < 1) to remove the mixed noise of the high-frequency sub-image:

$$NH_y = \underset{\{NH_i\},\{DH_i\}}{argmin} \sum_{i=1}^{K} \left(\|YH_i - DH_i NH_i\|_p^p + \lambda_i \|NH_i\|_1 \right) \tag{4}$$

Some nonconvex but smooth objective functions would be solved using a Riemannian conjugate gradient-like minimization process, which can be expressed as follows [13]:

$$h_\mu^{lp} : \mathbb{R}^{m\times d} \rightarrow \mathbb{R}^+, X \rightarrow \sum_{j=1}^{d} \sum_{i=1}^{m} \left(x_{ij}^2 + \mu \right)^{\frac{p}{2}}, 0 < p < 1, \mu > 0; \tag{5}$$

$(\boldsymbol{x}^2 + \boldsymbol{\mu})^{\frac{p}{2}}$ is approximated to $\|\boldsymbol{x}\|_{\boldsymbol{p}}^{\boldsymbol{p}}$, then

$$\|YH_i - DH_i NH_i\|_p^p \approx h_\mu (YH_i - DH_i NH_i) \tag{6}$$

The eq. (4) can be approximated by

$$NH_y = \underset{\{NH_i\},\{DH_i\}}{argmin} \sum_{i=1}^{K} \left(h_\mu (YH_i - DH_i NH_i) + \lambda_i \|NH_i\|_1 \right) \tag{7}$$

$$NH_y = \underset{\{NH_i\},\{DH_i\}}{argmin} \sum_{i=1}^{K} \left(\|YH_i - DH_i NH_i\|_F^2 + \lambda_i \|NH_i\|_1 \right)$$

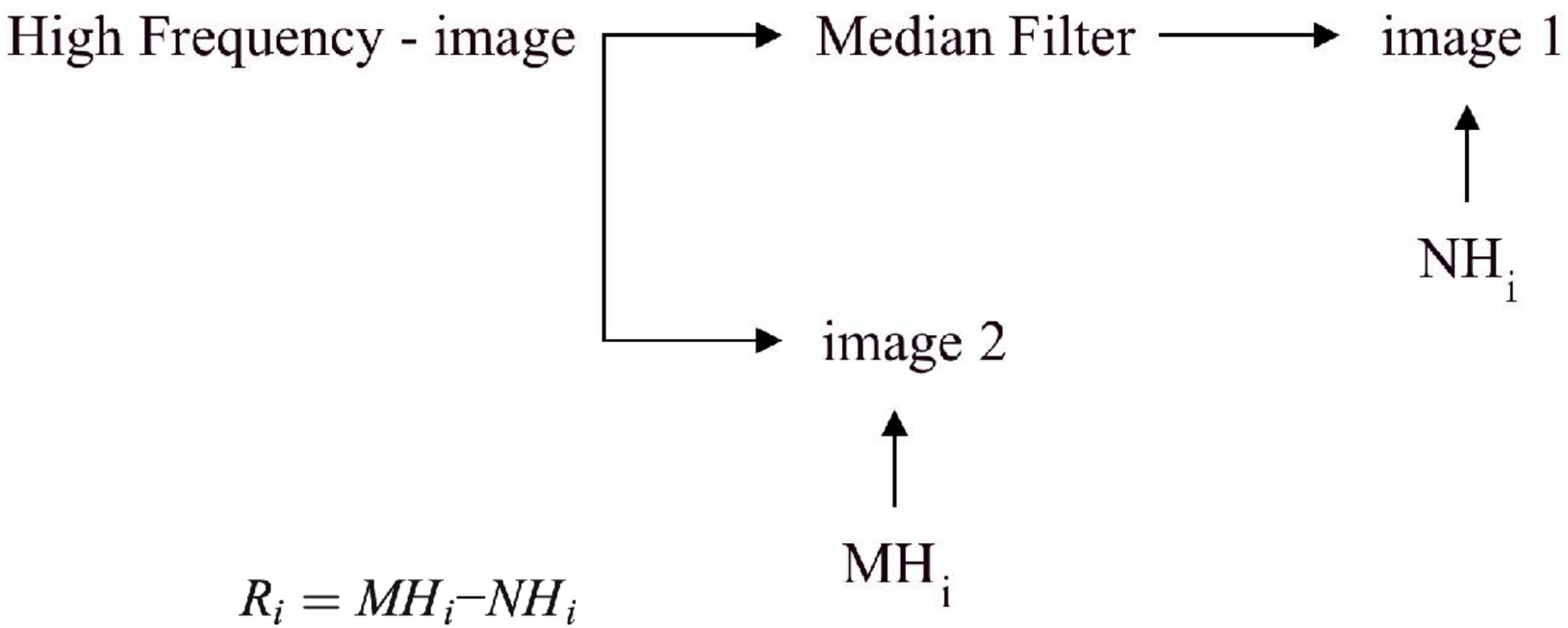
$$NH_y = \underset{\{NH_i\},\{DH_i\}}{argmin} \sum_{i=1}^{K} \left(h_\mu (YH_i - DH_i NH_i) + \lambda_i \|NH_i\|_1 \right)$$

We firstly filter the initial image using median filter in the pre-processing step. The algorithms mostly depend on the accuracy of the sparse coefficients; thus, it is very important to reduce the sparse error. The error between the sparse coefficient NH_i and the sparse coefficients MH_i of original high frequency is

$$R_i = MH_i - NH_i \tag{8}$$

We improve the noise reduction performance by enhancing the error constraint of sparse coefficients, and the formula (7) can be written as:

$$NH_y = \underset{\{NH_i\},\{DH_i\}}{argmin} \sum_{i=1}^{K} \left(h_\mu (YH_i - DH_i NH_i) + \lambda_i \| MH_i - NH_i \|_1 \right) \tag{9}$$



$$R_i = MH_i - NH_i$$

$$R = \left(\begin{bmatrix} 0 \\ 2 \\ 0 \\ 1.5 \end{bmatrix} - \begin{bmatrix} 0.5 \\ 2 \\ 0 \\ 1.5 \end{bmatrix} \right)$$

2.5 Wildlife image denoising algorithm

Basic steps of the wildlife image denoising algorithm are as follows,

Wildlife image denoising algorithm

Input: Wildlife image $\{Y_i\}$, smoothing parameter μ, K: number of subspaces, d: dimension of subspaces, and regularization parameter $\lambda = 10^{-8}$, the approximate parameter γ_1 and γ_2.

Output: A sequence $\left(\{DH_i\}_{i=1}^{K}, \{NH_i\}_{i=1}^{K}\right)$;

1: Decompose an original image into a high frequency sub-image YH_i and a low frequency sub-image YL_i based on Zig-zag sorting of 2D-DCT coefficients;

2: Applying median-based filters on the image YH_i to image patches $\{yh_i\}_{i=1}^{b}$ of the high frequency sub-image;

3: Construct initial dictionary $\{DH_i^0\}_{i=1}^{K}$ for the high frequency sub-image;

4: Update the sparse representation $\{NH_i\}_{i=1}^{K}$ based on $\{DH_k^0\}_{i=1}^{K}$ by the equation (9);

5: For i=1:n do

While $\max_i \left|(A_{MH}^i, A_{MD}^i)\right| > \epsilon$ do

Compute $\{(MH_i, MD_i)\}$ by the equations (12) and (13);

Compute (A_{MH}^i, A_{MD}^i) by the equations (14) and (15);

6: Update the sparse representation $\{NH_i\}_{i=1}^{K}$ by the equation (9);

7: end for

8: return $\left(\{DH_i\}_{i=1}^{K}, \{NH_i\}_{i=1}^{K}\right)$, restoration YH_i.

9: Remove the noise of the low frequency sub-image YL_i based on Wiener filtering;

10:$Y_i = YH_i + YL_i$

$$MH_{i+1} \in prox_{c_i}^{f}\left(MH_i - \frac{1}{c_i}\nabla_{MH} h_\mu(Y_i - MD_i MH_i)\right)$$

$$MD_{i+1} \in prox_{d_i}^{g}\left(MD_i - \frac{1}{d_i}\nabla_{MD} h_\mu(Y_i - MD_i MH_i)\right)$$

References

G. Moreau Envelopes

These properties support a method of approximating general functions in terms of certain 'envelope functions'.

1.22 Definition (Moreau envelopes and proximal mappings)**.** *For a proper, lsc function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ and parameter value $\lambda > 0$, the Moreau envelope function $e_\lambda f$ and proximal mapping $P_\lambda f$ are defined by*

$$e_\lambda f(x) \;:=\; \inf_w \Big\{ f(w) + \frac{1}{2\lambda}|w - x|^2 \Big\} \;\le\; f(x),$$
$$P_\lambda f(x) \;:=\; \operatorname{argmin}_w \Big\{ f(w) + \frac{1}{2\lambda}|w - x|^2 \Big\}.$$

Here we speak of $P_\lambda f$ as a mapping in the 'set-valued' sense that will later be developed in Chapter 5. Note that if f is an indicator function δ_C, then $P_\lambda f$ coincides with the projection mapping P_C in 1.20, while $e_\lambda f$ is $(1/2\lambda)d_C^2$ for the distance function d_C in 1.20.

In general, $e_\lambda f$ approximates f from below in the manner depicted in Figure 1–9. For smaller and smaller λ, it's easy to believe that, $e_\lambda f$ approximates f better and better, and indeed, $1/\lambda$ can be interpreted as a penalty parameter for a putative constraint $w - x = 0$ in the minimization defining $e_\lambda f(x)$. We'll apply Theorem 1.17 in order to draw exact conclusions about this behavior, which has very useful implications in variational analysis. First, though, we need an associated definition.

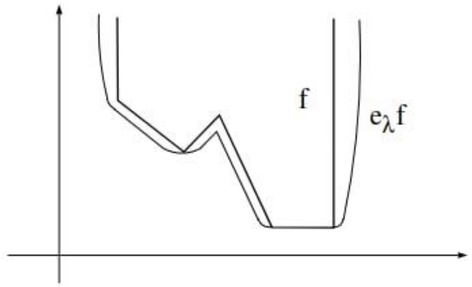


Fig. 1–9. Approximation by a Moreau envelope function.

$$A_{MH}^{i+1} := c_i(MH_i - MH_{i+1}) + \nabla_{MH} h_\mu(Y_i - MD_{i+1} MH_{i+1}) - \nabla_{MH} h_\mu(Y_i - MD_i MH_i)$$

$$A_{MD}^{i+1} := d_i(MD_i - MD_{i+1}) + \nabla_{MD} h_\mu(Y_i - MD_{i+1} MH_{i+1}) - \nabla_{MD} h_\mu(Y_i - MD_i MH_{i+1})$$

Results

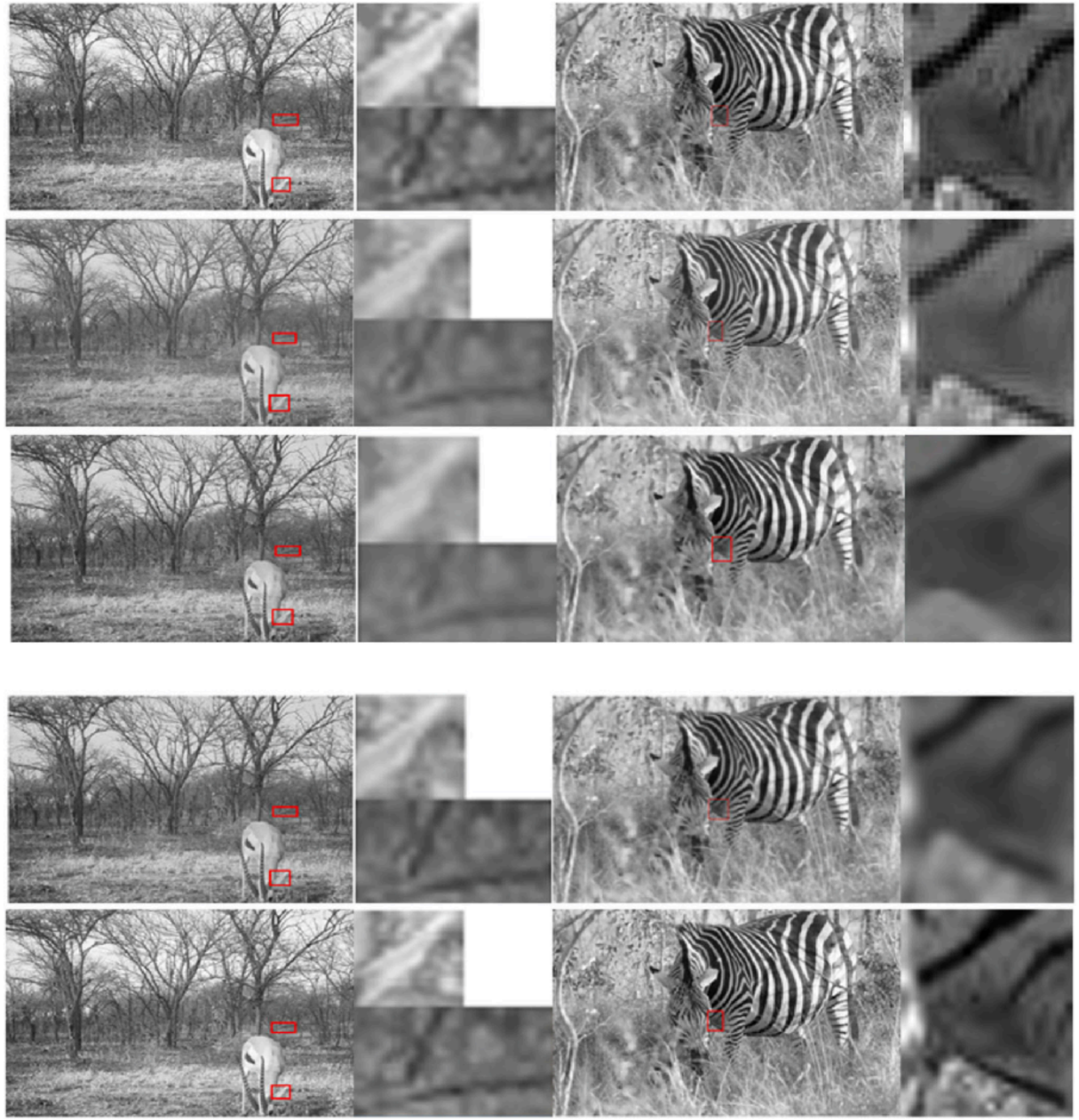


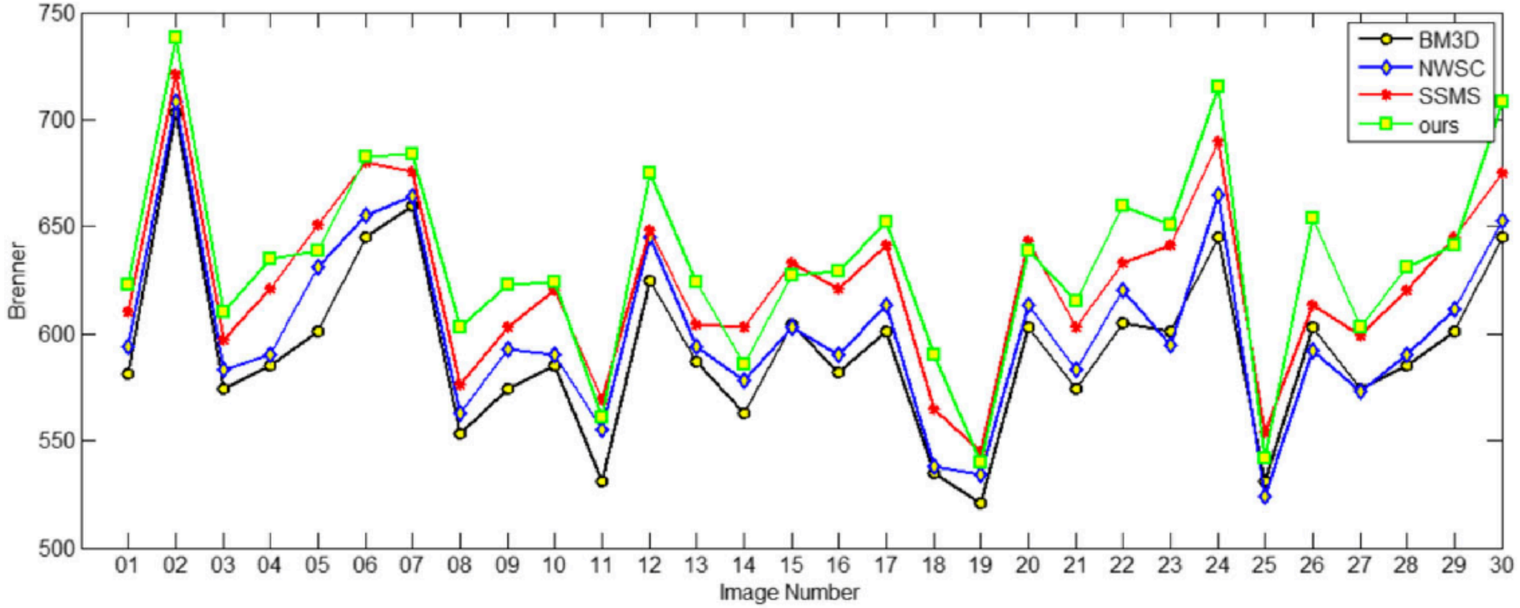
Fig. 4 Some restored results of the four methods for unknown random noise. Top row: original image; second row: BM3D [4]; third row: NWSC [19]; fourth Row: SSMS [13]; bottom row: our proposed method

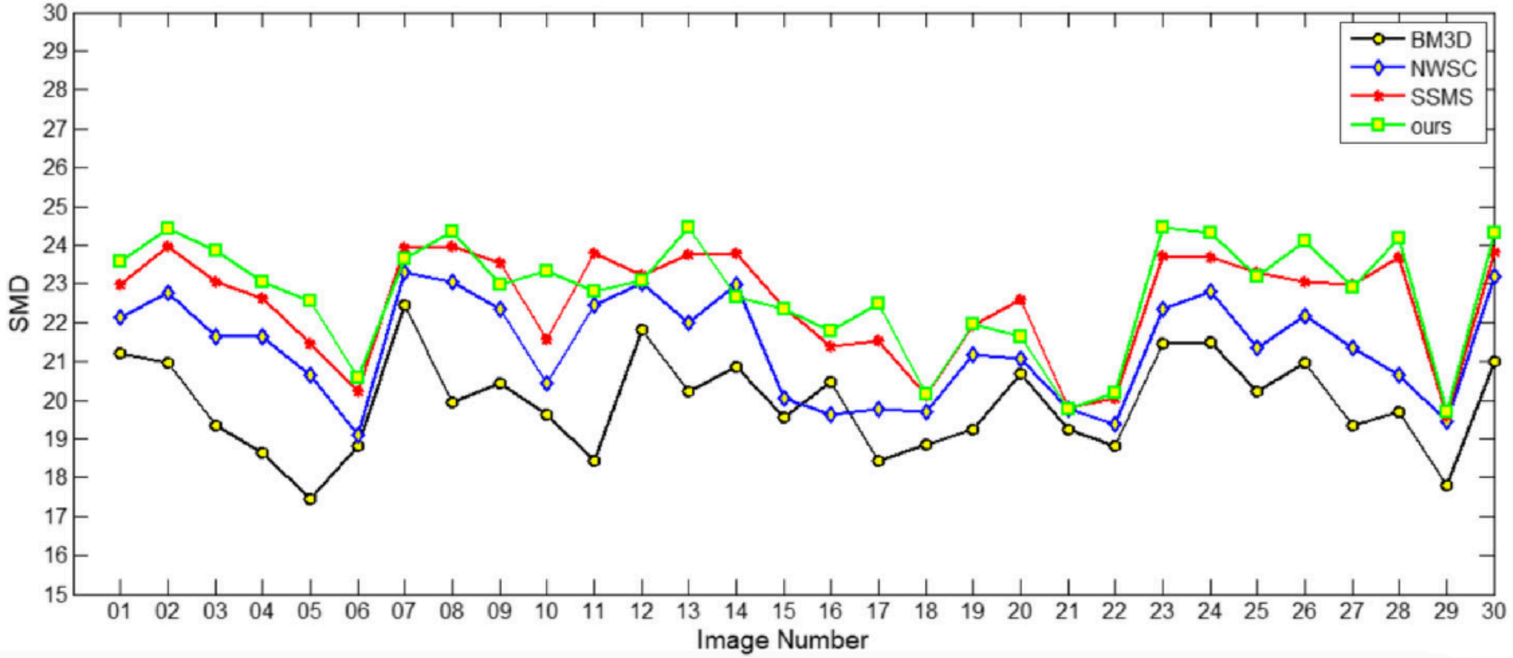
Sharpness of Edges

Save the structure and texture of skin

Due to the lack of reference images (i.e., pure images), we utilize the no-reference image quality evaluation, and two evaluation indices are Brenner gradient and SMD (Standardized mean difference). Brenner value uses the square of the gray difference between two adjacent pixels to judge the image clarity, and SMD takes the average degree of gray change as the evaluation standard of image clarity. The larger the two evaluation indices are, the clearer the image is, that is, the better the denoising effect is.

Figure 5 shows the Brenner and SMD values of the four methods used for removing unknown random noise of 30 wildlife images in the dataset. As seen from Fig. 4, in general, the proposed method can obtain higher Brenner and SMD values than the other three methods. Especially, the proposed algorithm shows good performance on the noise images with rich texture. However, it is slightly inferior to the method SSMS when facing to the images with a few textures and captured at night.


Brenner
Image Number
BM3D
NWSC
SSMS
ours


SMD
Image Number
BM3D
NWSC
SSMS
ours

3.3 Evaluation on the removal of the mixed noise added manually

Three types of the most common noises are Gaussian noise, impulse noise and salt-and-pepper noise. It is not easy to recover a latent image from the one corrupted by the mixed noise. We evaluate the performance of the proposed method with the mixed manual noise. Two objective quality indices, including peak signal-to-noise ratio (PSNR) and feature structural similarity (FSIM) index, are used in this paper. We employ the adaptive center-weighted median filter on visually pure image to obtain ground-truth images for computing PSRN and FSIM. The

Fig. 3 Some of the experimental samples. Top row: natural noise images; bottom row: visually pure images

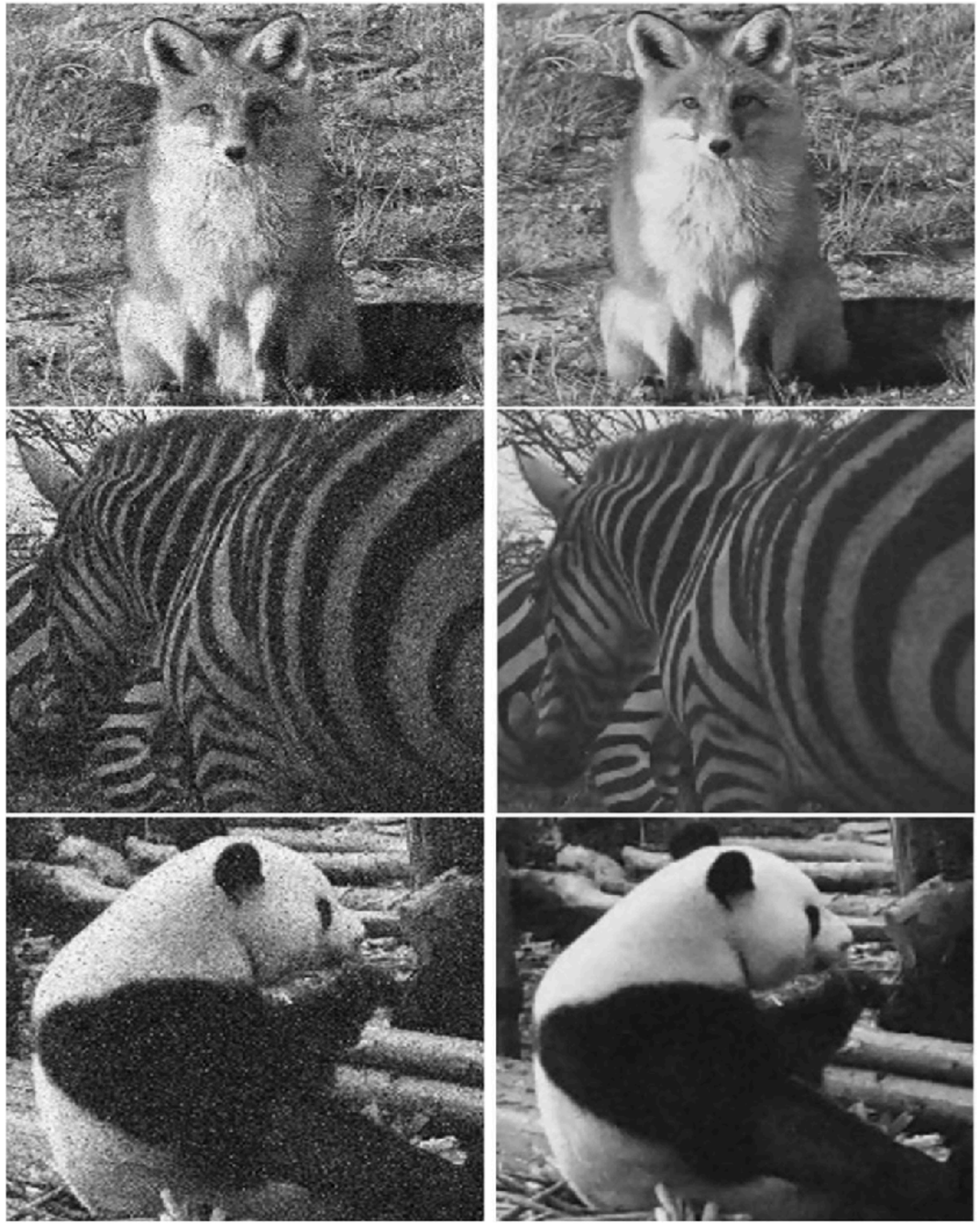


Fig. 6 Some denoising results of the proposed method for the manual noise with $(\sigma,\ sp,\ ri) = (10,\ 10\%,\ 5\%)$. Left: Corrupted images; Right: Restored images

parameter σ stands for the strength of Gaussian noise, and ri stands for the ratio of random value impulse noise, sp stands for the noise level of salt-and-pepper noise.

It is not easy to denoise the real-world images because of the difficulty of detecting random noise type. Figure 4 provides visual results of all competing methods. To have a visual comparison about the denoising behavior, the selected part is enlarged. The proposed method can not only reconstruct the geometry of the images but also preserve more details of images, especially textures and corners. Therefore, the proposed method is superior to the other three methods for the detail parts.

Table 1 PSNR、FSIM results of the mixed Gaussian noise and impulse noise removal

		PSNR (dB)				FSIM			
Image (σ,ri)		BM3D	NWSC	SSMS	Proposed	BM3D	NWSC	SSMS	Proposed
Image 1	(10, 10%)	26.13	26.98	27.32	**28.65**	0.901	0.898	0.906	**0.913**
	(20, 10%)	21.31	22.15	24.05	**24.46**	0.813	0.805	0.822	**0.848**
Image2	(10, 10%)	22.51	23.21	23.88	**25.46**	0.853	0.871	0.890	**0.892**
	(20, 30%)	18.77	20.07	20.14	**21.05**	0.769	0.798	0.809	**0.812**
Image3	(10, 10%)	28.78	29.24	**29.76**	29.20	0.872	0.897	0.903	**0.915**
	(20, 30%)	24.38	25.12	**24.32**	23.98	0.804	0.812	0.823	**0.844**
Image4	(10, 10%)	24.11	24.87	26.27	**28.56**	0.896	0.901	0.904	**0.908**
	(20, 30%)	21.54	21.97	23.20	**24.68**	0.832	0.841	0.857	**0.865**
Image5	(10, 10%)	27.35	27.32	28.32	**28.63**	0.905	0.903	**0.912**	0.901
	(20, 30%)	24.21	24.58	25.58	**26.11**	0.821	0.824	**0.851**	0.843
Image6	(10, 10%)	24.88	25.06	28.76	**29.01**	0.867	0.886	0.900	**0.904**
	(20, 30%)	20.79	21.27	25.73	**26.32**	0.785	0.790	0.826	**0.838**
Image7	(10, 10%)	22.13	23.29	23.76	**24.64**	0.860	0.845	0.889	**0.913**
	(20, 30%)	18.01	19.01	20.53	**21.03**	0.798	0.792	0.807	**0.853**
Image8	(10, 10%)	27.13	26.38	**28.25**	**27.65**	0.887	0.891	0.894	**0.897**
	(20, 30%)	24.15	22.17	**25.21**	**24.32**	0.812	0.822	0.813	**0.822**
Image9	(10, 10%)	26.25	25.25	28.25	**29.64**	0.885	0.895	0.901	**0.905**
	(20, 30%)	22.11	20.11	23.11	**25.33**	0.787	0.811	0.837	**0.853**
Image10	(10, 10%)	27.18	26.62	28.02	**29.49**	0.902	0.886	**0.911**	**0.906**
	(20, 30%)	24.13	23.05	25.90	**27.33**	0.813	0.793	0.842	**0.843**
Image11	(10, 10%)	23.25	24.57	26.76	**27.54**	0.854	0.852	0.879	**0.883**
	(20, 30%)	19.16	21.89	22.69	**23.23**	0.761	0.763	0.802	**0.803**
Image12	(10, 10%)	22.19	23.23	24.15	**24.47**	0.889	0.894	0.896	**0.913**
	(20, 30%)	18.93	20.18	**20.54**	20.32	0.796	0.787	0.785	**0.823**

Table 2 PSNR、FSIM results of the mixed impulse noise and salt-and-pepper noise removal

Image (ri,sp)		PSNR (dB)				FSIM			
		BM3D	NWSC	SSMS	Proposed	BM3D	NWSC	SSMS	Proposed
Image1	(10,10%)	25.40	25.72	26.24	**27.21**	0.885	0.876	0.893	**0.904**
	(30,15%)	20.22	21.09	22.01	**23.06**	0.803	0.791	0.815	**0.821**
Image2	(10,10%)	20.96	21.42	22.65	**23.26**	0.847	0.853	0.863	**0.875**
	(30,15%)	17.01	18.07	19.14	**20.05**	0.752	0.761	0.772	**0.798**
Image3	(10,10%)	26.32	28.83	28.76	**29.18**	0.881	0.884	0.899	**0.903**
	(30,15%)	22.38	23.09	23.32	**24.04**	0.814	0.812	0.827	**0.838**
Image4	(10,10%)	22.62	21.89	22.27	**22.62**	0.875	0.885	0.901	**0.902**
	(30,15%)	19.43	18.97	19.20	**20.68**	0.812	0.823	0.836	**0.841**
Image5	(10,10%)	26.35	27.12	27.32	**28.29**	0.896	0.895	0.891	0.901
	(30,15%)	22.37	23.06	24.58	**25.36**	0.820	0.813	0.804	0.825
Image6	(10,10%)	22.43	22.13	25.76	**26.21**	0.877	0.890	**0.908**	0.892
	(30,15%)	17.27	17.46	20.73	**21.32**	0.755	0.819	0.826	**0.828**
Image7	(10,10%)	24.03	25.13	23.76	**25.97**	0.881	0.858	0.889	**0.893**
	(30,15%)	20.36	21.02	20.53	**22.33**	0.756	0.786	0.815	**0.826**
Image8	(10,10%)	26.32	25.88	**28.25**	**27.35**	0.879	0.881	0.876	**0.887**
	(30,15%)	21.15	20.43	**25.21**	**24.39**	0.803	0.815	0.808	**0.812**
Image9	(10,10%)	23.65	23.72	24.25	**26.20**	0.892	0.901	0.895	**0.902**
	(30,15%)	20.52	20.15	21.11	**21.33**	0.801	0.814	0.812	**0.823**
Image10	(10,10%)	27.28	28.35	28.02	**28.40**	0.904	0.872	0.901	**0.906**
	(30,15%)	23.56	24.57	25.90	**26.30**	0.833	0.782	0.835	**0.837**
Image11	(10,10%)	25.34	25.78	26.76	**26.92**	0.898	0.861	0.904	**0.915**
	(30,15%)	21.26	22.09	**22.69**	22.43	0.802	0.775	**0.832**	0.826
Image12	(10,10%)	24.08	25.36	24.15	**25.98**	0.876	0.889	0.897	**0.913**
	(30,15%)	21.35	22.18	20.54	**23.39**	0.783	0.797	0.815	**0.823**

Table 3 PSNR、FSIM results of the mixed Gaussian noise and salt-and-pepper noise removal

Image (σ, sp)		PSNR (dB)				FSIM			
		BM3D	NWSC	SSMS	Proposed	BM3D	NWSC	SSMS	Proposed
Image1	(10, 10%)	27.26	26.84	28.08	**29.68**	0.907	0.901	**0.914**	0.912
	(20, 15%)	23.24	22.79	24.05	**25.31**	0.824	0.815	**0.822**	0.817
Image2	(10, 10%)	24.37	25.15	23.88	**26.78**	0.883	0.885	0.892	**0.905**
	(20, 15%)	20.76	21.73	20.14	**23.08**	0.819	0.808	0.803	**0.819**
Image3	(10, 10%)	28.21	28.24	**29.76**	28.80	0.895	0.902	0.897	**0.916**
	(20, 15%)	24.45	24.63	**25.32**	25.02	0.814	0.823	0.820	**0.847**
Image4	(10, 10%)	25.68	24.86	**26.27**	25.71	0.873	0.884	0.889	**0.914**
	(20, 15%)	20.49	20.01	21.20	**22.62**	0.822	0.830	0.835	**0.852**
Image5	(10, 10%)	26.72	27.22	**28.94**	28.33	0.901	0.886	**0.918**	0.906
	(20, 15%)	22.32	24.79	**25.58**	24.91	0.826	0.808	**0.841**	0.833
Image6	(10, 10%)	27.23	28.01	28.76	**29.15**	0.889	0.892	0.893	**0.902**
	(20, 15%)	23.66	24.93	25.73	**26.45**	0.825	0.810	0.826	**0.838**
Image7	(10, 10%)	24.13	22.56	24.76	**25.78**	0.879	0.861	**0.889**	0.881
	(20, 15%)	21.04	20.80	20.53	**21.79**	0.780	0.804	0.807	**0.813**
Image8	(10, 10%)	25.68	22.38	**28.25**	27.58	0.894	0.891	**0.894**	0.887
	(20, 15%)	21.51	21.19	**25.21**	24.92	0.808	0.815	**0.833**	0.820
Image9	(10, 10%)	28.65	27.33	28.25	**29.44**	0.902	0.893	0.901	**0.915**
	(20, 15%)	24.57	23.15	23.11	**25.63**	0.837	0.817	0.837	**0.842**
Image10	(10, 10%)	24.38	25.72	**26.02**	25.49	0.892	0.882	0.893	**0.910**
	(20, 15%)	20.13	21.57	**22.90**	22.33	0.825	0.803	**0.846**	0.833
Image11	(10, 10%)	25.19	27.23	26.76	**27.84**	0.871	0.897	**0.899**	0.892
	(20, 15%)	21.66	23.81	22.69	**24.09**	0.798	0.832	**0.842**	0.836
Image12	(10, 10%)	23.29	24.73	24.95	24.82	0.876	0.882	**0.896**	0.887
	(20, 15%)	19.58	20.94	**20.54**	20.12	0.802	0.798	0.815	**0.823**

Gaussian noise and impulse noise (Mixed noise1 for short), the mixed impulse noise and salt-and-pepper noise (Mixed noise2 for short), and the mixed Gaussian noise and salt-and-pepper noise (Mixed noise3 for short), respectively, where $\sigma = (10, \quad 20)$, $ri = (10\%, \quad 30\%)$, and $sp = (10\%, \quad 15\%)$. The best numerical results are highlighted. As shown from Tables 1, 2 and 3, the proposed method outperforms the other three methods for removing Mixed noise1 and Mixed noise2. For Mixed noise3, PSRN and FSIM values of the proposed method are comparable with SSMS. This shows that the proposed method has greater advantages in dealing with impulse noise.

Discussion

Springer Link