Multi-focus Source Images Reconstruction based on Adaptive Regional Data Hiding

Meng-yao Liu¹,a, Quan Zhou¹, Yi Zhang¹, Yan-lang Hu¹, Juan-ni Liu¹

¹National Key Laboratory of Science and Technology on Space Microwave, China Academy of Space Technology, Xi’an, Shaanxi, China

aemail: liumy504@dingtalk.com

Abstract: Preserving the information of multi-focus source images was neglected in previous image fusion schemes since source image data discarding will happen during the fusion process. Data hiding technology can utilize the redundant bits to embed essential secret data. Inspired from that we proposed a multi-focus image reconstruction algorithm using adaptive regional data hiding. Integrating the Human Visual System (HVS) into the reconstruction module, we further proposed a visual grayscale information entropy operator, which is implemented to segment fused images into texture and flat regions for adaptive data hiding after unfocused region data compression. Our method achieves excellent performances in reconstruction Peak signal-to-noise Ratio (PSNR) above 43dB and maintains the satisfying visual effect of the fused images.

1. Introduction

Multi-focus images are series of pictures taken in the same scene by different mechanisms, at different times, about different focus areas. Fusion methods of multi-focus images can improve the decision accuracy, which has attracted much attention in tracking, detection, recognition, information analysis, industrial control, environmental detection, intelligent system design, and so on [1][2][3][4]. It is necessary to preserve all source images data for some temporal applications like target tracking and for some regional sensitive applications like detection. Analyzing only source image data can improve the efficiency and accuracy of image processing algorithms. However, image fusion algorithms lose much source image information and drop the demand to maintain the original data.

Data hiding can increase information amount due to the cover image redundancy [5][6]. Data hiding is widely used for image compression, image transmission, and image enhancement. For instance, Cui Tao et al. [5] proposed an image compression method based on data hiding. After dividing the image into benchmark blocks and similar blocks, the position indexes of similar blocks are embedded into corresponding benchmark blocks. Information hiding can be divided into spatial-based and transform-based algorithms. Spatial domain methods have much more secret data capacity than transform domain’s[7][8][9]. However, the secret data capacity is inversely proportional to the visual quality, thus it is necessary to increase the hiding capacity while maintaining the imperceptibility. According to the visual human system (HVS), when the texture is complex, the impact on the visual quality is small and more secret information can be embedded in the cover image [10][11].

To reconstruct the source image information after fusion, this paper proposed a multi-focus source images reconstruction algorithm based on adaptive regional data hiding. Inspired by the characteristic of data hiding technology, the compressed data of source images which are regarded as secret...
information will be embedded adaptively into the fused image which is regarded as a cover image, and the compression data can be extracted to reconstruct the source images when it is necessary. We also proposed a visual grayscale information entropy according to the human visual system to segment fused images into texture and flat regions, which can adjust the embedding ratio in different image blocks. In this way, the source images can be reconstructed accurately and the fused image also keeps the excellent visual quality.

2. Proposed Method

Multi-focus source images reconstruction algorithm based on adaptive regional data hiding scheme is presented as follows:

Step 1: Generating the segmentation feature map $M_S$ with guided filter; (Section 2.1)

Step 2: Fusing the different focus regions into full-focus image $F$. Compressing the defocused regions data $A_{d1}$ and $A_{d2}$ using the Karhunen-Loève Transform (KLT); (Section 2.2.1)

Step 3: Generating the flat&texture region map $T_S$ according to the proposed visual grayscale information entropy $H_q$; (Section 2.2.2)

Step 4: The compressed defocused region data, segmentation labels, and flat&texture region labels are hidden in the corresponding full-focus image using adaptive regional data hiding; (Section 2.2.3)

Step 5: When reconstructing the source images, we can extract the compressed defocused region data from the full-focus fusion image according to the segmentation feature map and flat&texture region map, and integrating the corresponding focus-defocus regions to reconstructing source images $R_1$ and $R_2$. (Section 2.3)

Fig 1. Flowchart of adaptive regional data hiding and source images reconstruction

2.1. Segmentation feature map generation

The proposed method can be applied to different multi-focus image fusion algorithms. In this section, the multi-focus fusion algorithm based on guided filtering is used as an instance.

Guided filtering has always been utilized to extract texture features from the focus region[12]. The first guided filtering result is conducted with the original source image as the guidance image. Then the second guided filtering result is conducted with the first filtering result as the guidance image. The texture feature map $A_e$ is defined as:

$$A_e = f(A, A) - f(A, f(A, A))$$  \hspace{1cm} (1)

Where $f(p_1, p_2)$ is the guided filtering process. The first parameter is the input image and the second one is the corresponding guidance image. Then the segmentation feature map $M_S$ is generated by local energy comparison which is shown in Eq.2.


\[
E_1(x, y) = \frac{1}{9} \sum_{i=[x-1,x+1] \cap j=[y-1,y+1]} A_{d1}(i, j)^2 \\
E_2(x, y) = \frac{1}{9} \sum_{i=[x-1,x+1] \cap j=[y-1,y+1]} A_{d2}(i, j)^2
\]  

(2)

If \( E_1(x, y) > E_2(x, y) \), \( M_S(x, y) = 1 \); else \( M_S(x, y) = 0 \). Then we generate final segmentation feature map \( M_S \) after morphological filtering.

Then defocused region information \( A_{d1} \) and \( A_{d2} \) would be extracted according to \( M_S \) and the pixel value of the focus region is set to 0, which is illustrated as follows:

\[
A_{d1}(x, y) = \begin{cases} 
I_1(x, y) & \text{if } M_S(x, y) = 0 \\
0 & \text{if } M_S(x, y) = 1
\end{cases} \\
A_{d2}(x, y) = \begin{cases} 
I_2(x, y) & \text{if } M_S(x, y) = 1 \\
0 & \text{if } M_S(x, y) = 0
\end{cases}
\]  

(3)

Where \( I_1(x, y) \) and \( I_2(x, y) \) are pixel values at coordinate \( (x, y) \).

2.2. Data hiding for defocused image

2.2.1. Defocused region data compression and data size analysis

We need to guarantee that the defocused region information \( A_{d1} \) and \( A_{d2} \) can be embedded into the corresponding focus regions after compression, so we choose block-based compression algorithm KLT to reduce information redundancy. The visual quality of decoded images depends on the number of preserved eigenvalues.

Denote the uncompressed image sized \( M \times N \), and it is divided into non-overlapping \( a \times a \) blocks. For simplicity, we assume that \( M \) and \( N \) can be divided by \( a \) with no remainder. Assume that we preserve \( k \) eigenvalues, thus the compression data is \( U_k^T X P \) according to KLT. \( U_k \) is eigenvector matrix of covariance matrix consist of raster-scanning ordered image blocks, with row number is \( k \) and column number is \( MN/a^2 \). \( X \) is the first \( k \) rows of raster-scanning ordered image blocks matrix, \( X_P \) is the \( P \)th column of \( X \). The compressed data is a vector sized \( k \times 1 \), so the compressed results of \( A_{d1} \) and \( A_{d2} \) are 3D matrices sized \( (M/a) \times (N/a) \times k \).

2.2.2. Flat&Texture region map generation based on visual grayscale information entropy

Different image regions have different secret data capacities. Fig 2. illustrates the visual effect comparison of texture region and flat region with lowest 3 bits data embedded. It shows that the visual effect is not obvious after hiding 3bits secret data in the texture area, and the secret data capacity in the flat region is smaller in the texture one. Therefore, we need to generate flat&feature region map and choose to embed more secret data in the texture regions while less in the flat regions.

![Fig 2. Comparison of visual quality after data hiding in texture and flat regions](a)after data hiding  (b)original texture region
(c)after data hiding  (d)original flat region

Fig 2. Comparison of visual quality after data hiding in texture and flat regions

To segment the texture and flat area accurately, we propose a visual grayscale information entropy operator \( H_q \) for local image blocks according to the definition of image two-dimensional information entropy. This operator can describe the intensity of grayscale changing in one block. Firstly, the maximum gray value of the whole image \( Q_1 \) and minimum \( Q_2 \) is calculated. Then we need to quantize the gray values of the whole image into 32 intervals, the quantization steps of each layer are as follows:

The pixel values in each image block are quantified into corresponding intervals after quantization. The entropy $H_q$ of pixel $P_i$ in each image block is:

$$H_q = -\frac{1}{16} \sum P_i \log P_i$$

$P_i$ is the probability of pixel value $i$ in the whole image block.

The texture & flat region map $M_b$ is generated according to the grayscale information entropy. If the grayscale information entropy $H_q(x, y) < 0.5$, this image block is determined as flat region and $M_b(x, y) = 0$; when $H_q(x, y) > 0.5$, this image block is determined as texture region and $M_b(x, y) = 1$.

### 2.2.3. Compressed data optimization and adaptive information hiding

We will analyze the compressed data optimization in this section. Assume that we preserve 4 eigenvalues to represent the defocused image data $A_d1$ and $A_d2$. From $6 \times 8$ blocks which sized $4 \times 4$ from texture region and flat region, we can find that the 2nd - 4th rows values of the compressed data matrix in the texture region fluctuate greatly, while the values in the flat region fluctuate very little. Therefore, we can reduce carrier bits in the flat region in order to reduce the visual quality impact of the full-focus image while retaining high-quality decoded data of the defocused region information $A_d1$ and $A_d2$.

Fig 3. Compressed data in texture and flat regions

In flat regions, each block is assigned 24bits to embed the secret data: 1bit indicates source image label, 1bit indicates texture & flat region label, 10bits are binary secret data corresponding to the maximum eigenvalue, 12bits are binary secret data corresponding to the last three eigenvalues. In the texture area, each image block is assigned 36bits to embed the secret data, the number of bits assigned to labels and the maximum eigenvalue data is the same, and 24bits are the carrier of the secret data corresponding to the last three eigenvalues. The distribution of bit spatial location is shown in the following figure.

Fig 4. Cover image blocks bits assignment of texture & flat region

### 2.3. Source images reconstruction

The steps of source image reconstruction are illustrated as follow:

Step 1: According to the lowest bit of the upper left corner pixel of each block, which source image's defocus compression data is embedded in this image block is determined: When this bit is 0, the image block belongs to the focus area of the source image 2, and the defocus area data of the source image 1 is embedded; When this bit is 1, the image block belongs to the focus area of the
source image 1 and embed the defocus area data from the source image 2;

Step 2: According to the second lower bit of the upper left corner of each image block, whether the compressed data is from a flat area or texture area is determined, then we can extract the secret information bit by bit according to Section 2.2.3.

Step 3: We conduct inverse KLT on compressed data to get source image defocused region data. According to the segmentation feature map, we can generate reconstruction images by combining focus region data and defocused region data.

3. Experimental results and analysis

In this section, the experimental results of source image data hiding and reconstruction are analyzed. We evaluate: 1)The impact of the preserved eigenvalues numbers on reconstruction image visual quality; 2)The Peak signal-to-noise Ratio (PSNR) between full-focused secret image and the fused image without data hiding; 3)The PSNR between reconstructed source images and original source images. We choose 4 pairs of grayscale images from Lytro Multi-focus Dataset[13] which consists of 27 pairs of multi-focus images to evaluate the performance of the proposed algorithm.

3.1. Relationship between eigenvalues reserved number and defocused region decompression data quality

The PSNR of the reconstructed image and the source image is increasing with the enhancement of retained eigenvalue number. In order to verify the optimal eigenvalue number, we use several groups of multi-focus images to carry out experiments. The following figure shows the change of the image PSNR when different number of eigenvalues are retained. It shows that the signal-to-noise ratio of the defocused region of the two source images can reach more than 40dB when the 4 eigenvalues are retained.

![Fig 5. PSNR of decompression data according to number of reserved eigenvalues](image)

3.2. Fused image embedding result

The fusion algorithm based on guided filter is used to fuse 4 groups of test images, and then the compressed data of defocus region is embedded into the fused image. The pictures from the first line of Fig.6 are the fusion results, and from the second line are the fusion images result with hidden data. It shows that the visual effect is harmonious, and there is no obvious noise and block effect in them. The calculated PSNR is more than 40dB, which is imperceptible to human vision.

![Fig 6. Fused images and steganographic images](image)
### 3.3. Source images reconstruction

Then the reconstruction results of the source image are analyzed. The following figures illustrate the result of the reconstructed source image. It shows that this algorithm can reconstruct the multi-focus source image with high quality after obtaining the fused full focus image.

![Source image1](image1.png)

![Source image2](image2.png)

Fig 7. reconstruction source images

| Test Image | Leaves | Pepsi | Clocks | Calendar |
|------------|--------|--------|--------|----------|
| Ste Image  | 41.4875| 42.5549| 42.3662| 42.5460  |
| PSNR(dB)   |        |        |        |          |

### 4. Conclusion

In this paper, we have proposed a multi-focus source images reconstruction algorithm based on adaptive regional data hiding. Through this method, the source images can be reconstructed effectively only from the significant data embedded in the full-focus fusion image. And the visual quality of the full-focus fusion image maintains a promising accuracy for the follow-up decisions.

The proposed algorithm utilizes visual grayscale information entropy to segment the texture and flat area of the fusion image accurately. According to the human visual system, an adaptive regional data hiding algorithm is proposed. It can reconstruct the source images in high quality without affecting the visual effect of the fused full-focus image. Our simulation result shows that the PSNR of the reconstruction image can achieve more than 43dB.

Although the proposed method can keep excellent visual quality of the fused image, the low bit planes of the fused image has been changed. In the further research, we can focus on reversible data hiding to keep the fused image data undistorted.

### Acknowledgments

This work is supported by the National Key Laboratory Foundation (No.6142411204306) and the National Key Laboratory Foundation (No.6142411192205).

### References

[1] Carlos Castillo. J, Fernandez-Caballero. A, Serrano-Cuerda. J, et al. (2017) Smart environment architecture for robust people detection by infrared and visible video fusion. J. Journal of Ambient Intelligence & Humanized Computing, 8(2):223-237.
[2] Xiao H C, Zhou Q, Zheng X S. (2016) A Fusion Method of Satellite Remote Sensing Image Based on IHS Transform and Curvelet Transform. J. Journal of South China University of Technology (Natural Science Edition), 044(001):58-64.

[3] Ren Y M, Yang J, Guo Z Q, et al. (2020) Self-adaptive Entropy Weighted Decision Fusion Method for Ship Image Classification Based on Multi-scale Convolutional Neural Network. J. Journal of Electronics & Information Technology, 43(5):1424-1431.

[4] Akhloufi. M A, Porcher. C, Bendada. A. (2014) Fusion of thermal infrared and visible spectrum for robust pedestrian tracking. In. ‘Spie Defense + Security’ International Society for Optics and Photonics.

[5] Cui T, Zhou Q, Li J, et al. (2018) Block Compression of Remote Sensing Image based on Data Hiding. J. Chinese Space Science and Technology, 038(003):54-59.

[6] Zhu L H, Zhou Q. (2015) A Robust Lossless Data Hidden Transmission Method for Satellite Remote Sensing Image. J. Journal of Astronautics, 36(003):315-323.

[7] Chan C K, Cheng L M. (2007) Hiding data in images by simple LSB substitution. J. Pattern Recognition, 37(3):469-474.

[8] Fateh M, Rezvani M, Irani Y. (2021) A New Method of Coding for Steganography Based on LSB Matching Revisited. J. Security and Communication Networks, 2021(5):1-15.

[9] Cao F, Qian X, An B. (2020) Low-complexity reversible data hiding in encrypted image via MSB hierarchical coding and LSB compression. J. Multimedia Systems, 2020(3):1-14.

[10] Wang P, Cai B, Xu S, et al. (2020) Reversible Data Hiding Scheme Based on Adjusting Pixel Modulation and Block-Wise Compression for Encrypted Images. J. IEEE Access, PP (99):1-1.

[11] Shi X Y, Ma S L, Hu Y J, et al. (2015) A Pixel Block-based Adaptive Steganographic Algorithm with Embedding Priority Given to Image Texture. J. Acta Electronica Sinica, 43(006):1094-1100.

[12] Li S, Kang X, Hu J. (2013) Image Fusion with Guided Filtering. J. IEEE Transactions on Image Processing, 22(7):2864-2875.

[13] M. Nejati, S. Samavi, S. Shirani. (2015) Multi-focus Image Fusion Using Dictionary-Based Sparse Representation. J. Information Fusion, 25:72-84.