Infrared and Visible Image Fusion Based on Innovation Feature Simultaneous Decomposition

Guiqing He, Dandan Dong*, Siyuan Xing*, Ximei Zhao†*
* Northwestern Polytechnical University, Xi'an, China
Email:guiqing_he@nwpu.edu.cn Tel:+86-15802991311
†Qilu Institute of Technology

Abstract— In the field of image fusion, for the problem of large difference between infrared and visible image, these two are decomposed by joint sparse representation and get the common feature and innovation feature; Furthermore, a new fusion method based on innovation feature simultaneous decomposition is proposed to solve the problem of feature rich in infrared and visible light images. Based on the joint sparse representation model, the multi- source image data is combined into a new signal at the corresponding position, and the innovation feature of the two is decomposed into the same atom by Simultaneous Orthogonal Matching Pursuit. Experimental results and analysis show that the proposed method is superior to the popular sparse representation method, the simultaneous orthogonal matching pursuit method and the traditional joint sparse representation method, and the subjective visual evaluation and the objective index evaluation is both superior. The relevant research can be extended to the specific feature rich remote sensing image or medical image fusion field.

Index Terms— infrared and visible image fusion, joint sparse representation, innovation feature and simultaneous orthogonal matching pursuit

I. INTRODUCTION

Because of the difference of imaging principle between infrared sensor and visible light sensor, the image information acquired by the two sensors is complementary and redundant. The integration of infrared and visible light images can provide more comprehensive information in all kinds of imaging conditions, and its fusion result is beneficial to target recognition, tracking and safety detection. Ref. [1] and [2] showed that it runs an important role in aviation, military, security monitoring and other fields.

Ref. [3] and [4] indicated that, in recent years, sparse representation theory has been widely used in the field of image fusion because it can sparse and adaptively represent the source image by learning over-complete dictionary .Subsequently, the researchers proposed a joint sparse representation model for the problem of redundancy and complementarity in multi-source images. The redundant information and complementary information were defined as common feature and innovation feature respectively. The joint sparse representation of infrared and visible images can have the same common sparse coefficients and different sparse coefficients, so that the common feature and innovation feature can be targeted fusion process and get a more reasonable fused image. In addition, in infrared and visible image fusion, we hope that infrared and visible images can be in the same dictionary on the atomic decomposition, because if different atoms on the decomposition, similar to the source image using different wavelet Base decomposition, that is, essentially two different decomposition methods, making it difficult to balance the coefficients and integration. Therefore, Ref. [3] showed that simultaneous orthogonal matching pursuit (SOMP) aroused the attention of researchers, the use of sparse representation model, multi-source image data in the corresponding position on the combined into a new signal, and then found in the dictionary the most suitable atoms to the multi-source image simultaneous decomposition.

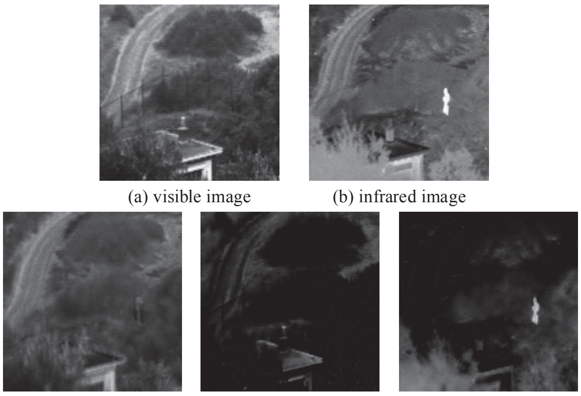
In this paper, a new fusion method based on the innovation feature simultaneous decomposition is proposed according to the characteristics of the target in the infrared image and the clear texture in the visible image. This method can be improved in the joint sparse representation model to fuse all the atomic representation coefficients involved in the image block. Namely, the innovation feature of the infrared and visible image is the same atom in the dictionary and the innovation sparse coefficients can be compared in the corresponding positions of infrared and visible light, so that the fusion rules can be guided more rationally and more accurately. Experimental results and analysis show that the proposed method has excellent fusion performance.

II. JOINT SPARSE REPRESENTATION OF IMAGES

Duarte et al. proposed three different joint sparse representation models in distributed compression perceptions. Ref. [4] introduced that the JSM-1 model (common feature + innovation feature) is applied to the image fusion problem (JSM-1 model is described below). The model considers that different data from the same scene have the same common sparse coefficient and the innovation sparse coefficient of the individual data under the dictionary representation.

A set of signals is now assumed $\boldsymbol{Y}=\{\boldsymbol{Y}_1,...,\boldsymbol{Y}_k,...\boldsymbol{Y}_K\}$ and redundant dictionaries $\boldsymbol{D}$, then all signals are composed of common sparse components and innovation sparse components, namely:

$$\boldsymbol{Y}_k=\boldsymbol{Y}^c+\boldsymbol{Y}_k^u=\boldsymbol{D}\boldsymbol{\alpha}^c+\boldsymbol{D}\boldsymbol{\alpha}_k^u,\quad k=1,2,...,K. \tag{1}$$



(a) visible image (b) infrared image

(c) common part (d) innovation part of visible image (e) innovation part of infrared image

Fig.1 Joint sparse representation of infrared and visible images

Where $\boldsymbol{Y}_k$ is the number k signal, $\boldsymbol{Y}^c$ is the common part, $\boldsymbol{Y}_k^u$ is the innovation part of the number k signal alone, $\boldsymbol{\alpha}^c$ and $\boldsymbol{\alpha}_k^u$ are the common sparse coefficient and the innovation sparse coefficient of the signal $\boldsymbol{Y}_k$ under the dictionary $\boldsymbol{D}$ respectively.

Fig.1 is a joint sparse representation of UNcamp data, the common sparse coefficient and the innovation sparse coefficient reconstructed the image of the common part and innovation part.

III. Infrared And Visible Image Fusion Based On Innovation Feature Simultaneous Decomposition

A. Innovation feature simultaneous decomposition

Fig.2 shows the internal sparse combination of the dictionary $\boldsymbol{D}$ in the joint sparse representation. Where the common atomic fraction corresponds to the common sparse coefficient, the innovation atom 1 corresponds to the innovation coefficient of the image 1, and the innovation atom 2 corresponds to the innovation coefficient of the image 2. Thus, the sparse representation indicates that after the iteration is completed, the innovation coefficient 1 and the innovation coefficient 2 do not have non-zero value on the same position.

In this paper, we consider decomposing the innovation feature into the same atom. Assuming that a piece of the signal to be decomposed is $[\boldsymbol{y}_1, \boldsymbol{y}_2]^T$. If the atom selected for this decomposition is atom 1, the method in this paper separates $\boldsymbol{y}_1$ and $\boldsymbol{y}_2$ on $\boldsymbol{d}_i$. So that the innovation feature can be decomposed on the same atom, and two innovation coefficients are obtained, as shown in Fig.3.

$$\underline{\boldsymbol{D}} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{D} & 0 \\ \boldsymbol{D} & 0 & \boldsymbol{D} \end{bmatrix} = \begin{array}{c} \begin{array}{ccc} \text{Common atom} & \text{Innovation atom 1} & \text{Innovation atom 2} \end{array} \\ \left[\begin{array}{c|c|c} \boldsymbol{d}_1\boldsymbol{d}_2 \ldots \boldsymbol{d}_P & \boldsymbol{d}_1\boldsymbol{d}_2 \ldots \boldsymbol{d}_P & 0\ 0\ \ldots 0 \\ \boldsymbol{d}_1\boldsymbol{d}_2 \ldots \boldsymbol{d}_P & 0\ 0\ \ldots\ 0 & \boldsymbol{d}_1\boldsymbol{d}_2 \ldots \boldsymbol{d}_P \end{array}\right] \end{array}$$

Fig.2 Joint sparse representation of the dictionary model

$$\underline{\boldsymbol{D}} = \left[\begin{array}{c|c|c|c|c} \boldsymbol{d}_1\boldsymbol{d}_i \ldots \boldsymbol{d}_P & \boldsymbol{d}_1 & \boxed{\boldsymbol{d}_i} & \ldots \boldsymbol{d}_P & 0\ 0\ \ldots 0 \\ \boldsymbol{d}_1\boldsymbol{d}_i \ldots \boldsymbol{d}_P & 0 & \boxed{0} & \ldots\ 0 & \boldsymbol{d}_1\boldsymbol{d}_i \ldots \boldsymbol{d}_P \end{array}\right] \xrightarrow{\text{Decomposition}} \begin{cases} \alpha_i \begin{bmatrix} \boldsymbol{d}_i \\ 0 \end{bmatrix} & \text{Traditional Method} \\ \alpha_i \begin{bmatrix} \boldsymbol{d}_i \\ 0 \end{bmatrix} + \beta_i \begin{bmatrix} 0 \\ \boldsymbol{d}_i \end{bmatrix} & \text{This article Method} \end{cases}$$

The selected atom

Fig.3 Innovation feature decomposition comparison chart

B. Fusion procedure

(1) Pretreatment.Read the source infrared and visible images and then divide the source images into P image blocks through sliding window, which size is $\sqrt{n} \times \sqrt{n}$ (n is dimension of the dictionary atoms) with sliding step is 1, and each image block is arranged in turn in column vector in order to get the infrared and visible images The data matrices are X_1 and X_2, and the matrix size is $n \times P$.

(2) Subtract the mean. The column vectors of X_1 and X_2 are averaged to get $m_1(t)$ and $m_2(t)$ respectively, where $t \in [1, P]$. The averaged image data $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ are achieved, and then new union data $\boldsymbol{Y} = \begin{bmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{bmatrix}$ is built.

(3) Dictionary learning. L samples are randomly and not repeatedly selected from $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ to form a training set. The redundant dictionary $\boldsymbol{D}$ is learned by the K-SVD algorithm,and then a joint dictionary $\underline{\boldsymbol{D}} = \begin{bmatrix} \boldsymbol{D} & \boldsymbol{D} & 0 \\ \boldsymbol{D} & 0 & \boldsymbol{D} \end{bmatrix}$ is created under the joint sparse model.

(4) Joint Sparse Decomposition. Via approximating the solution (1), the concatenated coefficient matrix $\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}^c \\ \boldsymbol{\alpha}_1^u \\ \boldsymbol{\alpha}_2^u \end{bmatrix}$ are achieved by the method described in Section A. And then the common coefficient $\boldsymbol{\alpha}^c$ and the innovation coefficients $\boldsymbol{\alpha}_1^u$ and $\boldsymbol{\alpha}_2^u$ of the infrared and visible image are obtained.

$$\min\{\|\boldsymbol{Y} - \underline{\boldsymbol{D}}\boldsymbol{\alpha}\|_F^2\} \text{ s.t. } \forall \text{t}, \|\boldsymbol{\alpha}(\text{t})\| \le T_0 \tag{2}$$

(5) The sparse coefficient is fused, and the first sparse coefficient is fused according to the following formula to obtain the fused sparse coefficient $\boldsymbol{\alpha}_f(t)$.

$$\begin{gathered} \boldsymbol{\alpha}_f(t) = \boldsymbol{\alpha}^c(t) + \boldsymbol{\alpha}_f^u(t), \\ \boldsymbol{\alpha}_f^u(t) = \left[\boldsymbol{\alpha}_{j_1}^u(t,1), \boldsymbol{\alpha}_{j_2}^u(t,2) \ldots, \boldsymbol{\alpha}_{j_i}^u(t,i)\right]^T, \\ j_i = \max\{|\boldsymbol{\alpha}_1^u(t,i)|, |\boldsymbol{\alpha}_2^u(t,i)|\}, \ j_i \in [1,2]. \end{gathered} \tag{3}$$

Where $\boldsymbol{\alpha}_1^u(t,i)$ is the i element value of the innovation sparse coefficient of the t block of the infrared image, j_i is the maximum value of the same location element modulus value of the infrared and visible image sparse coefficient, $\boldsymbol{\alpha}_f^u(t)$ is the maximum value of the t Block-innovation fusion coefficient.

(6) Mean fusion，The mean value of the fusion is obtained by fusing the mean value according to the following formula.

$$\begin{aligned} m(t) &= \tau m_1(t) + (1-\tau) m_2(t), \\ \tau &= 1/(1+\exp\{-\beta(\|\boldsymbol{Y}_1(t)\|_2 - \|\boldsymbol{Y}_2(t)\|_2)\}). \end{aligned} \quad (4)$$

(7) Image reconstruction. The fusion image is reconstructed to obtain the t fusion data. Finally, the final fusion image $\boldsymbol{I}_F$ is obtained by averaging $\boldsymbol{X}(t)$.

$$\begin{aligned} \boldsymbol{X}(t) &= \boldsymbol{D} \cdot \boldsymbol{\alpha}_f(t) + m(t) \times \boldsymbol{E}, \\ \boldsymbol{E} &= [1,1,...,1]^T, \boldsymbol{E} \in R^n. \end{aligned} \quad (5)$$

$\boldsymbol{E}$ is a column vector whose elements are all 1 and whose number of elements is equal to the size of the slider n in step (1).

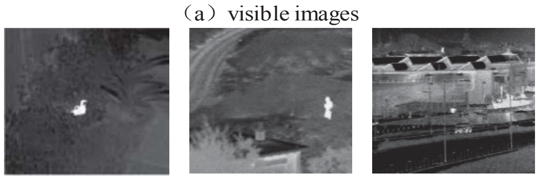
IV. Experimental Results And Analysis

In order to verify the effectiveness of this method, the classical multi-scale analysis method DWT, SR-OMP in [5], SR-SOMP in [6] based on sparse representation, JSR1 in [7], JSR2 in [8] and JSR-SOMP, which is based on the unique information synchronization decomposition method proposed in this paper. The experimental data is based on three closely matched infrared and infrared images in the image fusion database. Visible light images were recorded as: Duck group, UNcamp group and CCD group, as shown in Fig.4.

Experimental parameters and the experimental platform described below, the selected dictionary training sample size are P = 10000, slide window size $\sqrt{n}=8$， step size $step$ = 1， Dictionary size 64×256, K-SVD algorithm iteration 20, OMP algorithm sparse decomposition error = 0.1; JSR1, JSR2, JSR-SOMP dictionary training and sparse decomposition error and SR method exactly has the same settings. It should be noted that the image blocks in JSR2 and JSR-SOMP are de-averaged, and the mean value is $\beta = 0.01$ [5] in Fusing the mean on formula (4). This experiment takes Windows 7 operating system as the platform, CPU as G2030, frequency 3.00Ghz, RAM size 4.00GB, using MATLAB 2014a 7.10 emulation.

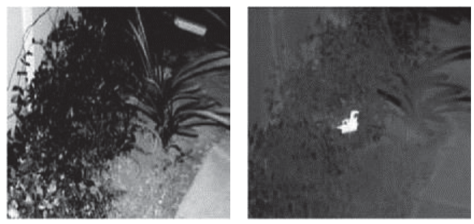
In the case of space, the fusion results show only the fusion images of the Duck group, and the other two sets of experimental results are similar. From the results of Duck's fusion in Fig.5, both sparse representations and three joint sparse representations can effectively integrate the ducks in the infrared image with the grasses in the visible image, and the image fusion quality is significantly higher than that of the traditional DWT method, the target area of DWT method is dim, which is mainly because the average fusion rule in the low frequency band makes the image smooth and reduces the contrast, which means that both the sparse representation and the joint sparse representation theory are suitable for infrared and Visible light fusion. Secondly, we compare the sparse representation and joint sparse representation: the SR-OMP method, the SR-SOMP method and the JSR1 method fuse the infrared image in the upper left part of the image, and there is a certain blur area, which shows that the above method will be part of the infrared image information Too much fusion into the results of the image, indicating that the sparse coefficient of the "activity level" to select the entire block (sliding window to get) the fusion method will be the existence of distortion, and JSR2 fusion rules will be unique sparse coefficient all remain , Which can extract the unique

（a）visible images

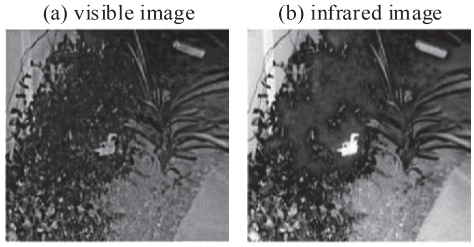


（b）infrared images

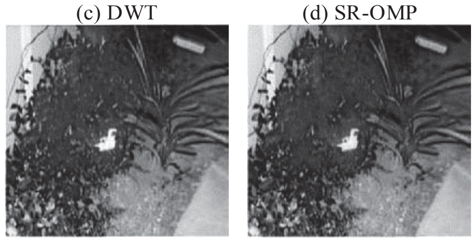
Fig.4 Three groups of infrared and visible images



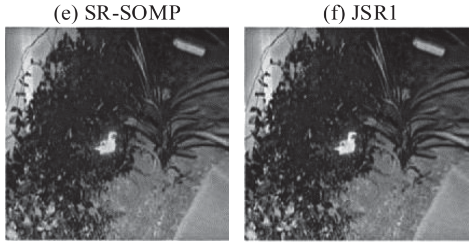
(a) visible image (b) infrared image



(c) DWT (d) SR-OMP



(e) SR-SOMP (f) JSR1



(g) JSR2 (h) JSR-SOMP

Fig.5 Six fusion methods subjective comparison of Duck group information of the visible light image, not only the effective fusion of the images, but also the fuzzy phenomena such as SR-SOMP and SR-JSR1.In the same way, JSR-SOMP preserves the unique information of JSR-SOMP because of the largest fusion rule of the atomic coefficients in the specific information. Compared with JSR2 method, the target area is brighter and the color of grass area is darker. Image contrast has improved.

TABLE I
THE AVERAGE PERFORMANCE OF SIX FUSION METHODS

method	Qo	Qw	Qe	Qabf
DWT	0.7476	0.7685	0.6045	0.6011
SR-OMP	0.7522	0.7927	0.5986	**0.6556**
SR-SOMP	0.7932	0.8247	0.6779	0.6549
JSR1	0.7837	0.8507	0.6997	0.6486
JSR2	0.8284	0.8661	0.7067	0.6369
JSR-SOMP	**0.8328**	**0.8761**	**0.7321**	0.6533

In order to quantitatively evaluate the performance of different fusion methods for infrared and visible image fusion, this paper uses the commonly used human visual system based on the evaluation index $Q_{AB/F}$ and Piella index(Q_0 、 Q_W 、 Q_E), these four indicators are in the [0,1], the closer to 1 the higher the quality of image fusion, among them, $Q_{AB/F}$ evaluates the goodness of the edge information retention, Q_0 reflects the overall similarity between the fused image and the source image, Q_W is the image fusion quality weighted by the window, and Q_E is another index for evaluating the edge of the fused image.

Table I shows the average values of the three sets of experimental data. Table I indicated that two sparse representations, SR-OMP, SR-SOMP, and two joint sparse representations are compared with the traditional transform domain DWT JSR1 and JSR2, the proposed method preserves the background information in visible light images and the target information in infrared images. The evaluation index has excellent performance on several parameters, which reflects the effectiveness and superiority of this algorithm.

V. CONCLUSIONS

In this paper, a new method of infrared and visible image fusion based on the unique decomposition of information is proposed. In the joint sparse representation model, the current fusion method based on image blocks is improved to fuse all the atomic representation coefficients in the image block. And because the atoms in the dictionary are based on the adaptive learning of the sample, so the atomic is actually the image structure of the essential response, the corresponding coefficient values reflect the image on the atomic structure strength.

In this paper, we can maximize the atomic characteristics of the infrared and visible images by maximizing the fusion rule. It not only integrates the specific information (target and background) of infrared and visible image, but also makes the fusion image more clear and coherent. The experimental results show that compared with the classical multi-scale analysis method, the sparse representation method and the joint sparse representation method, the proposed method has better fusion effect on subjective vision and objective evaluation index, and the related research can be generalized to unique information-rich areas of remote sensing images or medical image fusion.

ACKNOWLEDGMENT

This work was supported in part by the National Nature Science Foundation of China (NO.61402368), Aerospace Support Fund (2015-HT-XGD), Special Research Project of Shaanxi Education Department (16JK1774).

REFERENCES

[1]. Li Shu-tao, Kang Xu-dong, Fang Le-yuan, etc. Pixel-level image fusion: A survey of the state of the art. Information Fusion, 2017, 33: 100-112.

[2]. Z. Khan, F. Shafait, A. Mian. Joint group sparse PCA for compressed hyperspectral imaging. IEEE Transactions on Image Processing, 2015, 24(12): 4934-4942

[3]. M. J. Kim, D. K. Han, H. S. Ko. Joint patch clustering-based dictionary learning for multimodal image fusion. Information Fusion, 2016, 27: 198-214.

[4]. F. Frossi, R. Lanzarotti, J. Y. Lin. Orthogonal procrustes analysis for dictionary learning in sparse linear representation. Plos One, 12(1), 2017.

[5]. Yang B, Li S. Multifocus image fusion and restoration with sparse representation.[J]. IEEE Transactions on Instrumentation & Measurement, 2010, 59(4): 884-892.

[6]. Yang B, Li S. Pixel-level image fusion with simultaneous orthogonal matching pursuit[J]. Information Fusion, 2012, 13(1): 10-19.

[7]. Yu N, Qiu T, Bi F, et al. Image features extraction and fusion based on joint sparse representation[J]. IEEE Journal of Selected Topics in Signal Processing, 2011, 5(5): 1074-1082.

[8]. Li A S, Yin H. Multimodal image fusion with joint sparsity model[J]. Optical Engineering, 2011, 50(6): 409-421.