



OPEN ACCESS

EDITED BY
Xiaomin Yang,
Sichuan University, China

REVIEWED BY
Kaining Han,
University of Electronic Science
and Technology of China, China
Teng Li,
Anhui University, China

*CORRESPONDENCE
Weiwei Sun
✉ sunww@cqupt.edu.cn
Shixin Huang
✉ 15803659045@163.com

†These authors share first authorship

SPECIALTY SECTION
This article was submitted to
Neural Technology,
a section of the journal
Frontiers in Neuroscience

RECEIVED 17 November 2022
ACCEPTED 07 December 2022
PUBLISHED 06 January 2023

CITATION
Zhang G, Nie X, Liu B, Yuan H, Li J,
Sun W and Huang S (2023) A
multimodal fusion method
for Alzheimer's disease based on DCT
convolutional sparse representation.
Front. Neurosci. 16:1100812.
doi: 10.3389/fnins.2022.1100812

COPYRIGHT
© 2023 Zhang, Nie, Liu, Yuan, Li, Sun
and Huang. This is an open-access
article distributed under the terms of
the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution
or reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

A multimodal fusion method for Alzheimer's disease based on DCT convolutional sparse representation

Guo Zhang^{1,2†}, Xixi Nie^{3†}, Bangtao Liu², Hong Yuan², Jin Li²,
Weiwei Sun^{4*} and Shixin Huang^{1,5*}

¹School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China, ²School of Medical Information and Engineering, Southwest Medical University, Luzhou, China, ³Chongqing Key Laboratory of Image Cognition, College of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China, ⁴School of Optoelectronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing, China, ⁵Department of Scientific Research, The People's Hospital of Yubei District of Chongqing City, Yubei, China

Introduction: The medical information contained in magnetic resonance imaging (MRI) and positron emission tomography (PET) has driven the development of intelligent diagnosis of Alzheimer's disease (AD) and multimodal medical imaging. To solve the problems of severe energy loss, low contrast of fused images and spatial inconsistency in the traditional multimodal medical image fusion methods based on sparse representation. A multimodal fusion algorithm for Alzheimer's disease based on the discrete cosine transform (DCT) convolutional sparse representation is proposed.

Methods: The algorithm first performs a multi-scale DCT decomposition of the source medical images and uses the sub-images of different scales as training images, respectively. Different sparse coefficients are obtained by optimally solving the sub-dictionaries at different scales using alternating directional multiplication method (ADMM). Secondly, the coefficients of high-frequency and low-frequency subimages are inverse DCTed using an improved L1 parametric rule combined with improved spatial frequency novel sum-modified SF (NMSF) to obtain the final fused images.

Results and discussion: Through extensive experimental results, we show that our proposed method has good performance in contrast enhancement, texture and contour information retention.

KEYWORDS

Alzheimer, multimodal, sparse, fusion, convolutional

1. Introduction

Alzheimer's disease (AD) is a common neurodegenerative disease with concealed onset and incurable in the elderly. In clinical, AD is characterized by general dementia such as cognitive decline and memory loss (Dubois et al., 2021). Advanced multimodal neuroimaging techniques, such as magnetic resonance imaging (MRI) (Thung et al., 2014; Liu M. et al., 2016; Lian et al., 2018; Fan et al., 2019) and positron emission tomography (PET) (Chetelat et al., 2003; Liu M. et al., 2017) use different imaging mechanisms to reflect the location of organs or lesions from different angles, and can clearly show human tissue or metabolism and blood flow of organs. This technique has the complementarity and irreplaceability of medical information, which provides a good prospect for the early diagnosis of AD (Perrin et al., 2009; Mohammadi-Nejad et al., 2017; Wang et al., 2021).

Medical image fusion includes image decomposition, fusion rules, and image reconstruction. The main purpose of image decomposition is to extract the feature information from the image. The effectiveness of feature extraction determines the quality of fusion results. The current image fusion algorithms can be divided into three categories. The first kind of image fusion is based on wavelet and pyramid transform (Da Cunha et al., 2006; Yang et al., 2010; Miao et al., 2011; Liu S. et al., 2017; Liu X. et al., 2017). Among them, the Laplace pyramid transform has the best robustness in the sampling operator. Wang and Shang (2020) proposed a fast image fusion method based on discrete cosine transform (DCT), which decomposes each source image into a base layer and a detail layer for image fusion. And optimize the calculation method of the base layer to better preserve the structure of the image. In addition, non-subsampled shear wave transforms (NSST) (Kong et al., 2014) are also widely used in AD diagnosis because of their translation invariance and multidirectional. The second kind of image fusion is based on edge-preserving filtering (Farbman et al., 2008; Xu et al., 2011; He et al., 2012; Hu and Li, 2012; Zhang et al., 2014; Kou et al., 2015). This method can filter the image while erasing the details and retaining its strong edge structure. It can decompose the input image into smooth layers and detail layers. The smooth layer contains the main energy information of the image; the detail layer contains texture features. The third type is the feature selection method based on sparse learning, for example, the multiplier alternating directional multiplication method (ADMM) algorithm (Liu and Yan, 2011) organizes the whole learning and decomposition process into vectors, and iterates with a sliding window to achieve the convergence effect.

Sparse representation (SR) is a widely used image representation theory. It deals with the natural sparsity of signals according to the physiological characteristics of the human visual system. It is widely used in image classification (He et al., 2019), image recognition (Liu H. et al., 2016), image feature extraction (Liu et al., 2014), and multimodal image

fusion (Zhu et al., 2016). The fusion method based on SR and dictionary learning is widely used in image fusion proposed by compressed sensing theory (Donoho, 2006), and it is generally better than most traditional fusion methods (Zhang and Patel, 2017). It usually represents the source image in the form of a linear combination of overcomplete dictionaries and sparse coefficients. Because the weighted coefficients obtained are sparse, the significant information of the source image can be represented by a small number of non-zero elements in the sparse coefficients. In the methods based on SR, sparse coding is usually based on local image blocks. Yang and Li (2009) first introduces SR into image fusion. This method uses sliding window technology to make the fusion process robust to noise and registration errors. Because the adjacent image blocks overlap each other, the result of each pixel is multi-valued. Ideally, multiple values of each pixel should be equal to maintain the consistency of overlapping image blocks (Gu et al., 2015). However, sparse coding is performed independently on each image block. The correlation between image pixels is ignored, resulting in multiple unequal values for each pixel. At the same time, most fusion methods adopt the strategy of aggregation and averaging to obtain the final value of each pixel, which will cause the image details to be smoothed or even lost in fusion (Rong et al., 2017). Yin et al. (2016) obtained a joint dictionary by using the source image as training data and then fused the image using the maximum weighted multi-norm fusion rule. But the problem of missing details still exists. Zong and Qiu (2017) proposed a fusion method based on classified image blocks, which uses directional gradient histogram (HOG) features to classify image blocks to establish a sub-dictionary. Although the problem of loss of details has been reduced, it still inevitably leads to some details being smoothed. Zhang and Levine (2016) proposed an improved fusion method of multitasking robust SR combined with spatial context information. Like most methods based on SR, this method encodes for local image blocks rather than for the whole image. As a result, it can still lead to poor preservation of details. And usually, the fusion methods based on sparse coding use only one dictionary to represent the different morphological parts of the source image, which is easy to cause the loss of image information.

Therefore, we propose a multimodal fusion method for Alzheimer's disease based on DCT convolution SR to solve the above problems. It was evaluated on the neuroimaging database of Alzheimer's disease (ADNI) (Veitch et al., 2022), and its effectiveness was verified by experiments.

The contribution of this paper has the following three aspects:

1. An improved multiscale decomposition method of DCT is proposed. Firstly, the $M \times N$ size image is divided into blocks of 8×8 size, and then the DCT transform is applied to each small block separately. The DCT coefficients of each image block are normalized separately and their low-order DCT coefficients

are calculated. The ratio of the energy of the higher-order DCT coefficients to the energy of the lower-order DCT coefficients is used as the focus evaluation function. To solve the problem of fused image capability loss and contrast reduction.

2. A convolutional SR method is proposed to solve the problem of spatial inconsistency of multimodal image fusion by combining the high-frequency and low-frequency components obtained from the multiscale decomposition and adopting the improved rules of spatial frequency and L_1 parametric combination according to the characteristics of AD multimodal images.

3. To address the problem of limited detail preservation capability of medical image fusion methods based on SR and the lack of expression capability of single dictionary, the detail texture and contour of the fused image are enhanced by constructing multiple sub-dictionaries, and finally the fused detail layer image is fused and reconstructed with the fused base layer image to obtain the fused image. The fused AD medical information features are preserved.

2. Materials and methods

In the process of image fusion, the most important thing is how to extract low-high-frequency coefficients and the fusion criteria of low-high-frequency coefficients. First of all, the DCT transform is used to decompose the MRI image in multi-scale; the DCT coefficients of each image block are normalized respectively, and its low-order DCT coefficients are calculated. The ratio of the energy of the higher-order DCT coefficient to that of the lower-order DCT coefficient is used as the focusing evaluation function. Then, the sub-images on each scale are convoluted sparsely encoded, and the sparse coefficients of different sub-images are obtained. The high-frequency sub-image coefficients are combined with the improved L_1 norm and the novel sum-modified SF (NMSF), and the low-frequency sub-images are fused with the improved L_1 norm and regional energy. Finally, the fused low-frequency sub-band and high-frequency sub-band are transformed by inverse DCT transform to get the final fused image. The principle of the image fusion algorithm based on DCT transform is shown in **Figure 1**:

2.1. DCT decomposition

2.1.1. Low-frequency component

The most important part of information for vision is concentrated in the low frequencies of the image. Low frequencies represent the slow variation between image pixels. It is the large flat area of the image that describes the main part of the image and is a comprehensive measure of the intensity of the whole image. In order to maintain the visibility of the image, the low-frequency part of the image is preserved, and

changes in the low-frequency part may cause large changes in the image. The low-frequency coefficients of the fused image based on the DCT transform are averaged, assuming that there are p multi-exposure images, which can be defined as:

$$G(i, j) = \sum_{k=1}^p w_k G_k(i, j) \tag{1}$$

$$\sum_{k=1}^p w_k = 1 \tag{2}$$

$$w_1 = w_2 = \dots = w_n = \dots = w_p = 1/p \tag{3}$$

where $G_k(i, j)$ is the low-frequency coefficients extracted from the source image after DCT transformation; $G(i, j)$ is the fused low-frequency coefficients; and w_k is the weighting factor.

2.1.2. High-frequency component

The high-frequency coefficients correspond to detailed information in the image, such as edges, and are extracted from the 8×8 chunked image after the DCT transform. The standard deviation of the high-frequency coefficients $D(i, j)$ in the $(2k + 1) \times (2k + 1)$ neighborhood centered on pixel (i, j) is calculated separately.

$$C(i, j) = \sqrt{\frac{\sum_{m=i-k}^{i+k} \sum_{n=j-k}^{j+k} [D(m, n) - \bar{M}(m, n)]^2}{d - 1}} \tag{4}$$

where, D is the number of pixel points in the region $(2k + 1) \times (2k + 1)$; D is the value of the high frequency coefficient corresponding to the (m, n) point; $\bar{M}(m, n)$ is the average value of pixels in the region, which can be defined as:

$$\bar{M}(i, j) = \frac{1}{d} \sum_{m=i-k}^{i+k} \sum_{n=j-k}^{j+k} D(m, n) \tag{5}$$

The regional standard deviation of the high-frequency coefficients for each of the P multi-exposure images is $[C_1(i, j), C_2(i, j), \dots, C_p(i, j)]$, then the weight coefficients corresponding of the extracted high-frequency coefficient is:

$$w_k(i, j) = \frac{C_k(i, j)}{\sum_{k=1}^p C_k(i, j)} \tag{6}$$

where, the weights of the P multi-exposure images are compared to them, the fused high-frequency coefficient $D(i, j)$ is the high-frequency coefficient corresponding to the largest weighting factor.

$$w_k(i, j) = \max [w_1(i, j), w_2(i, j), \dots, w_k(i, j), \dots, w_p(i, j)] \tag{7}$$

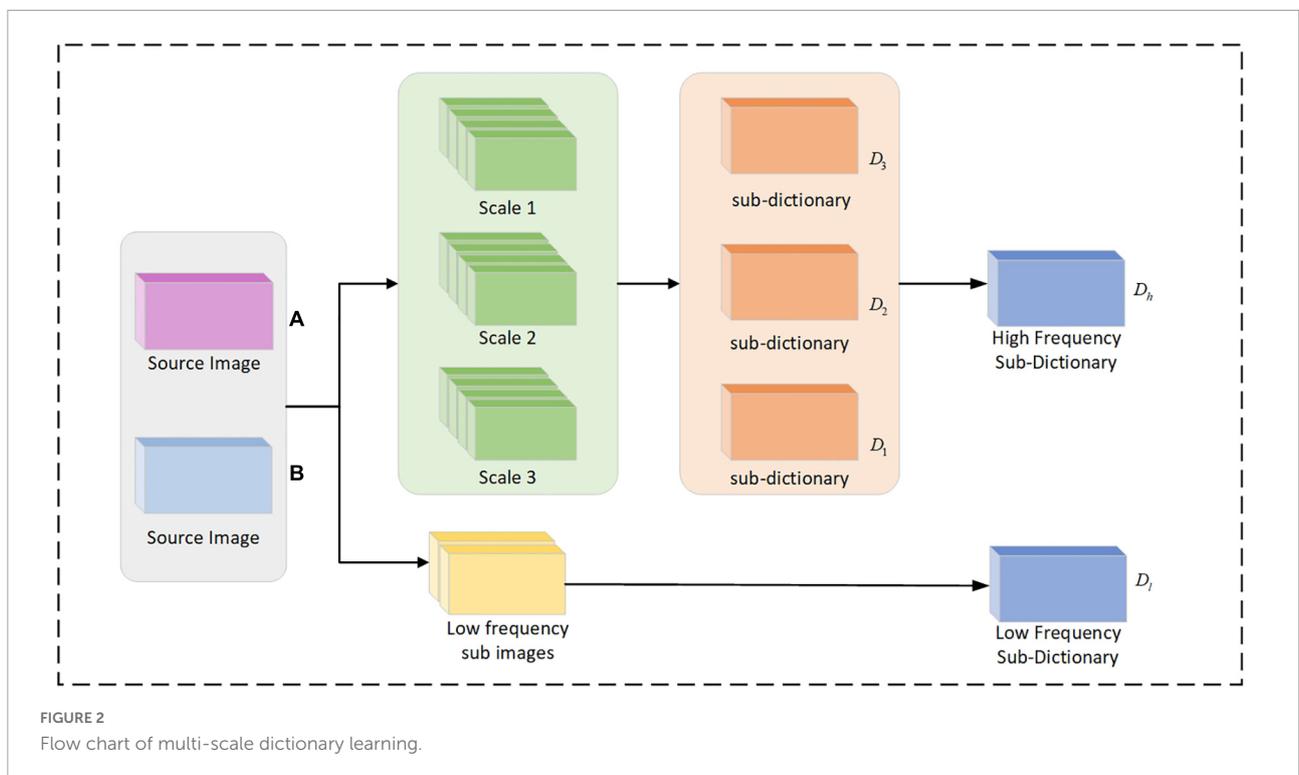
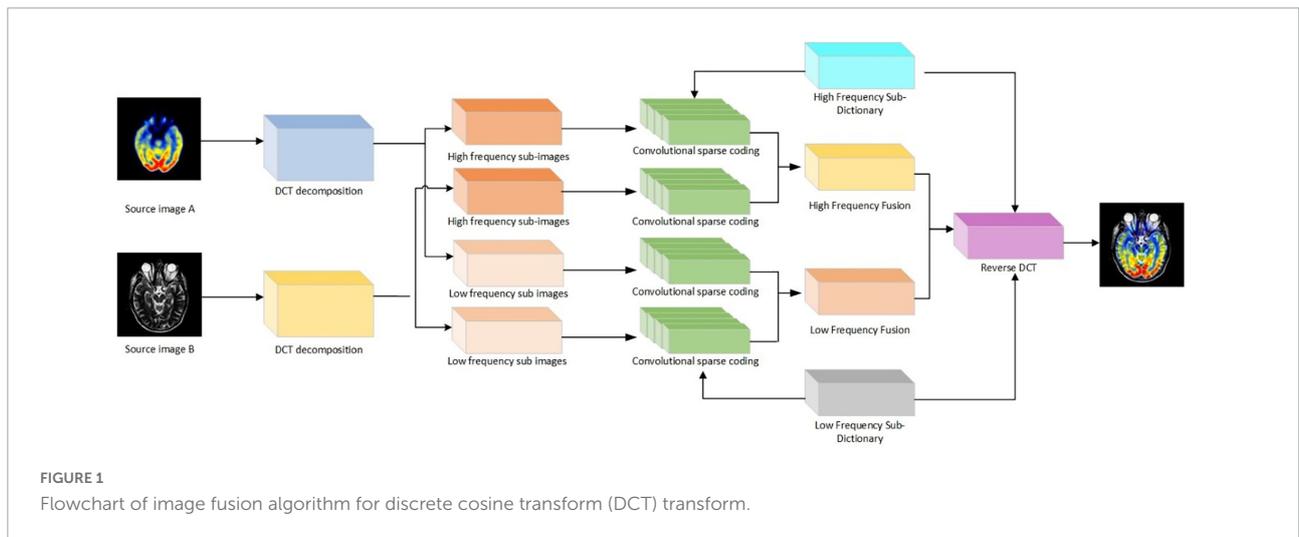
$$D(i, j) = D_k(i, j) \tag{8}$$

2.2. Sparse representation

The medical image fusion method consists of the following four parts: (1) Multi-scale dictionary learning to train sub-dictionaries on different scales of sub-images as training images. (2) Convolutional sparse coding of the dictionaries at different scales to find their convolutional sparse coefficients. (3) Low-frequency sub-band coefficient fusion rules for low-frequency sub-images are fused according to the set fusion rules. (4) High-frequency sub-band coefficient fusion rules fuse high-frequency sub-images at different scales.

2.2.1. Multi-scale dictionary learning

The source images A and B are firstly decomposed by l -level DCT to obtain their decomposition coefficients $\{H_A^{l,k}, L_A\}$ and $\{H_B^{l,k}, L_B\}$, respectively. Where, $H_A^{l,k}$ and $H_B^{l,k}$ denote the high-frequency sub-band coefficients of source images A and B at decomposition scale l and orientation k . L_A and L_B are the low-frequency sub-band coefficients of images A and B, respectively. The sub-band images of each scale are used as training images to train the corresponding convolutional sparse sub-dictionaries. The different convolutional sparse sub-dictionaries capture the



features of the sub-images at different scales. Finally, the low-frequency and high-frequency sub-dictionaries are formed by combining the sub-dictionaries at different scales. The high-frequency first-scale images $H_A^{l,k}$ and $H_B^{l,k}$ of source images A and B are used as training images $\{y_m\}_{m=1}^M$, and the corresponding convolutional sparse dictionary learning models are built. The formula is as follows:

$$\arg \min_{d,x} \frac{1}{2} \sum_{m=1}^M \left\| y_m - \sum_{k=1}^K d_k \times x_{m,k} \right\|_2^2 + \lambda \sum_{m=1}^M \sum_{k=1}^K \|x_{m,k}\|_1 \quad (9)$$

where, $x_{m,k}$ is the sparse coefficient corresponding to the m th training image; d_k is the corresponding filter; * denotes the two-dimensional convolution operation; λ is the regularization parameter; and $\|\cdot\|_1$ denotes the l_1 parametric number, which represents the sum of the absolute values of the elements.

(1) Dictionary update phase. By keeping the sparse coefficients constant, each filter is optimally updated with the following equation:

$$\arg \min_{d_k} \frac{1}{2} \sum_{m=1}^M \left\| y_m - \sum_{k=1}^K d_k \times x_{m,k} \right\|_2^2 \quad (10)$$

To optimize the filter in the discrete Fourier domain, the filter d_k needs to be zero-filled to the same size as $x_{m,k}$. Taking into account the normalization of d_k with zero padding, the formula is as follows:

$$\arg \min_{\{d_m\}, \{g_m\}} \frac{1}{2} \sum_k \left\| \sum_m x_{m,k} \times d_m - s_k \right\|_2^2 + \sum_m l_{CPN}(g_m) \quad (11)$$

The ADMM algorithm shows that $CPN = \{x \in R^N : (I - PP^T)x = 0, \|x\|_2 = 1\}$ represents the

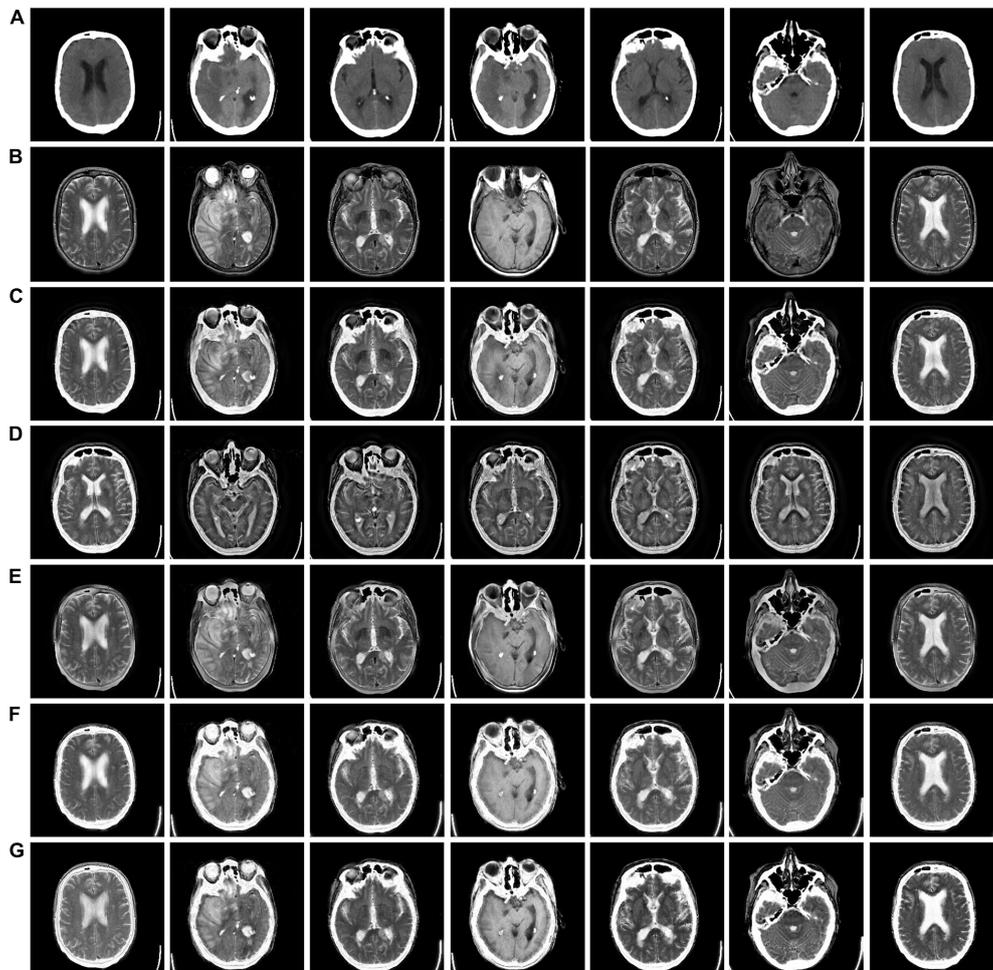


FIGURE 3 CT-MRI fusion images obtained by five fusion methods. (A) Computed tomography (CT) source image; (B) magnetic resonance imaging (MRI) source image; (C) nonsubsampling contourlet (NSCT) fusion result; (D) non-subsampling shear wave transform (NSSST) fusion result; (E) guided filter fusion (GFF) fusion result; (F) Laplacian redecomposition (ReLP) fusion result; (G) fusion result of the proposed method.

constraint range, the indicator function is defined as $l_s(X) = \begin{cases} 0 & \text{if } X \in S \\ \infty & \text{if } X \notin S \end{cases}$, and g_m is an auxiliary variable introduced to facilitate optimal derivation. The resulting updated convolutional filter d_k can be obtained.

(2) Convolutional sparse coefficient update phase. Update the coefficients by keeping the filter unchanged:

$$\arg \min_{x_{m,k}, z_{m,k}} \frac{1}{2} \sum_{m=1}^M \left\| y_m - \sum_{k=1}^K d_k \times x_{m,k} \right\|_2^2 + \lambda \sum_{m=1}^M \sum_{k=1}^K \|z_{m,k}\|_1 \quad (12)$$

where, $z_{m,k}$ is the introduced auxiliary variable. We obtain the updated convolutional sparse coefficients by alternating iterative solutions.

In **Figure 2**, the cyclic execution of the dictionary and the convolutional sparse coefficients are updated to a predetermined maximum number of cycles or a set parameter threshold to stop. The convolutional sub-dictionary D_1 for the first scale of the high-frequency sub-image is output. The second and third scales of the high-frequency sub-images are dictionary learned separately to obtain the convolutional sub-dictionaries D_2, D_3 . By combining each high-frequency sub-dictionary, and the high-frequency dictionary $D_h = [D_1, D_2, D_3]$ is obtained. The low-frequency sub-images are subjected to dictionary learning, and the low-frequency dictionary D_l is obtained.

2.2.2. Convolutional sparse coding

To better capture the detailed texture information of medical images and reduce the influence of artifacts, a high-frequency dictionary D_h and a low-frequency dictionary D_l are obtained by learning. Convolutional sparse coding is performed on the decomposition coefficients $\{H_A^{l,k}, L_A\}$ of the source image A. The TV regularization is then incorporated into the convolutional sparse coding model. The formula is as follows:

$$\arg \min_{\{x_k\}} \frac{1}{2} \left\| s - \sum_k d_k \times x_k \right\|_2^2 + \lambda \sum_k \|x_k\|_1 + \lambda_1 TV \left(\sum_{k=1}^K d_k * x_k \right) \quad (13)$$

where, $TV(X) = \|g_0 \times x\|_1 + \|g_1 \times x\|_1$, g_0 and g_1 are the filters used to calculate the gradients along the rows and columns of the image, respectively. The sparsity coefficients $\{x_{m,L}^A, x_{m,l,k,H}^A\}$ and $\{x_{m,L}^B, x_{m,l,k,H}^B\}$ of the coefficients in the sub-bands of the source images A and B, respectively, are obtained by optimal solution in the discrete Fourier domain. Where, m denotes the number of filters and convolutional sparse coefficient maps; L denotes the low frequency image; H denotes the high frequency image; and l and k denote the scale and orientation of the corresponding sub-bands, respectively.

2.2.3. Low-frequency coefficient fusion rules

After DCT decomposition, the energy information of the image is contained in the low-frequency sub-bands L_A and L_B of the source images A and B, which are displayed as basic information such as the contour and brightness of the image. The averaging strategy generally used for low-frequency coefficient fusion tends to lead to a reduction in the contrast of the image. In the case of fusion using the $Max - L_1$ rule with SR, the reduction of contrast can be effectively avoided. However, it can lead to the problem of spatial inconsistency of the image. At the same time, because the region energy can better reflect the energy and significant features of the image, and the convolution sparse coefficients of the L_1 parameter averaging strategy can effectively reduce the effect of misalignment. Therefore, a combination of region energy and averaged L_1 parameter is used to fuse the low-frequency sparse coefficients.

$$L_F = \sum_{m=1}^M d_m^l \times x_{m,L}^F \quad (14)$$

2.2.4. High-frequency coefficient fusion rules

The high frequency sub-bands $H_A^{l,k}$ and $H_B^{l,k}$ of the source images A and B contain a large amount of information such as texture details of the images. The convolutional SR of the fusion method has good performance in preserving detail information, and the improved spatial frequency and can well reflect the gradient changes of the image texture. Therefore, the improved spatial frequency combined with the average L_1 parameter strategy is used to fuse the high frequency sparse coefficients.

$$H_F^{l,k} = \sum_{m=1}^M d_m^h \times x_{m,l,k,H}^F \quad (15)$$

2.2.5. Multimodal medical image fusion

The problem of spatial inconsistency in multimodal images is caused when the L_1 parametric maximum fusion rule is used in traditional SR-based fusion methods. Therefore, we decompose the source image by performing DCT on it. Different sub-dictionaries are trained for features of different scales. A rule combining region energy and activity coefficients is used for fusion of the low frequency component coefficients, and a modified rule combining spatial frequency and activity

TABLE 1 Average values of index evaluation of different fusion methods for CT-MRI.

Methods	SF	SD	RMSE	GSM	VIF
NSCT	23.5217	1.4275	0.1617	0.9631	0.2376
NSST	23.3269	1.4729	0.1678	0.9620	0.2582
GFF	23.8825	1.4681	0.1604	0.9715	0.2265
ReLP	24.2733	1.4953	0.1539	0.9743	0.2674
Proposed	25.1246	1.5203	0.1482	0.9776	0.2743

coefficients is used for fusion of the high frequency component coefficients. The problems of reduced contrast, blurred details and inadequate information extraction are avoided. The specific steps are as follows:

(1) The DCT decomposition of source images A and B is performed to obtain the respective decomposition coefficients $\{H_A^{l,k}, L_A\}$ and $\{H_B^{l,k}, L_B\}$.

(2) In the dictionary learning stage, the images at different scales corresponding to the multimodal source images are used as training sets, and the sub-dictionaries D_0, D_1, D_2, D_3 corresponding of each scale is derived. The low-frequency dictionary is $D_l = D_0$. The high-frequency dictionary is $D_h = [D_1, D_2, D_3]$.

(3) Sparse coding stage. Convolutional sparse coding is performed on the sub-images of different orientations at each scale to obtain the corresponding convolutional sparse $\{x_{m,L}^A, x_{m,l,k,H}^A\}$ and $\{x_{m,L}^B, x_{m,l,k,H}^B\}$.

(4) Low-frequency component fusion stage. The regional energies E_A and E_B of L_A and L_B , and the active level maps $\bar{\alpha}_A$ and $\bar{\alpha}_B$ are calculated. The convolution sparsity coefficients

$x_{m,L}^F$ are obtained after fusion. Finally, the convolution sparse coefficients are reconstructed with the low-frequency dictionary convolution to obtain the low-frequency sub-band image L_F .

(5) High-frequency component fusion stage. The fused convolutional sparse coefficients C are obtained by calculating the improved spatial frequencies of $H_A^{l,k}$ and $H_B^{l,k}$. Then the high-frequency sub-band images $H_F^{l,k}$ are obtained by convolutional fusion with the high-frequency dictionary D_h .

(6) Finally, the fused image F is obtained by performing inverse DCT on the fused sub-band image $\{H_F^{l,k}, L_F\}$.

3. Results and discussion

3.1. Data set and training parameter settings

(1) Experimental settings

All our experiments are conducted on a computer with Intel Core i7-10750H CPU 2.60 GHz, 16 GB RAM, NVIDIA GeForce

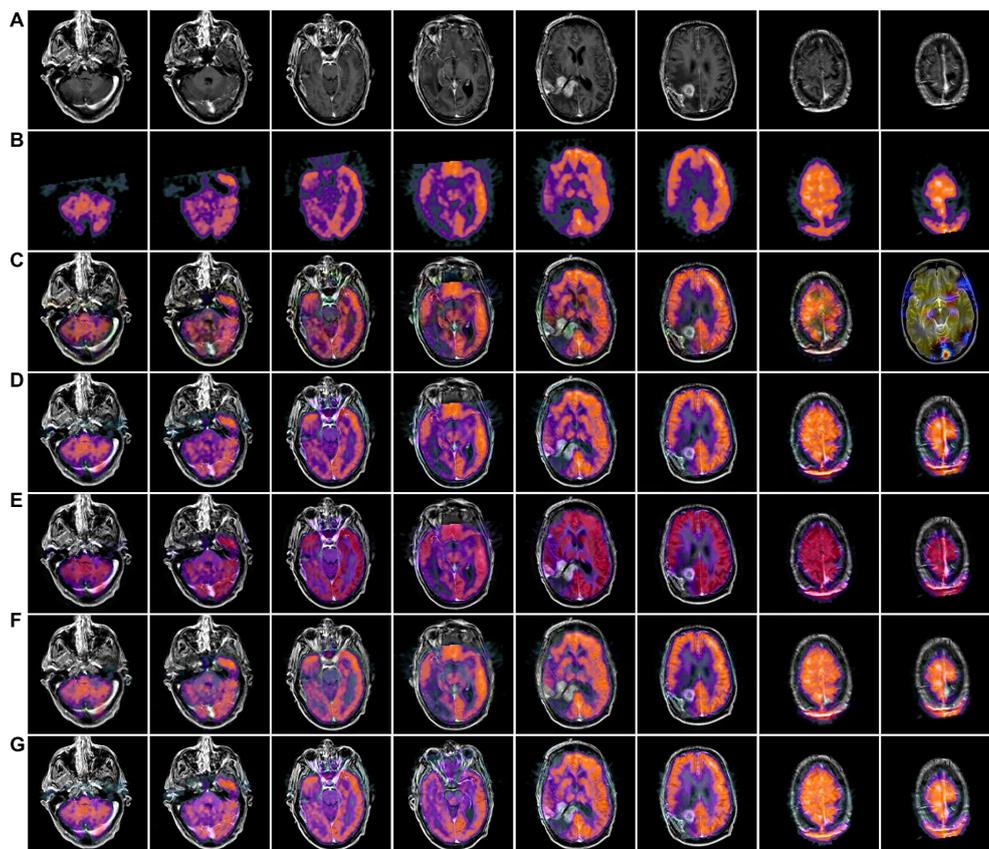


FIGURE 4

MRI-PET fusion images obtained by five methods in Paras1. (A) Magnetic resonance imaging (MRI) source image; (B) positron emission tomography (PET) source image; (C) nonsubsampling contourlet (NSCT) fusion result; (D) non-subsampling shear wave transform (NSST) fusion result; (E) guided filter ng fusion (GFF) fusion result; (F) Laplacian redecomposition (ReLP) fusion result; (G) fusion result of the proposed method.

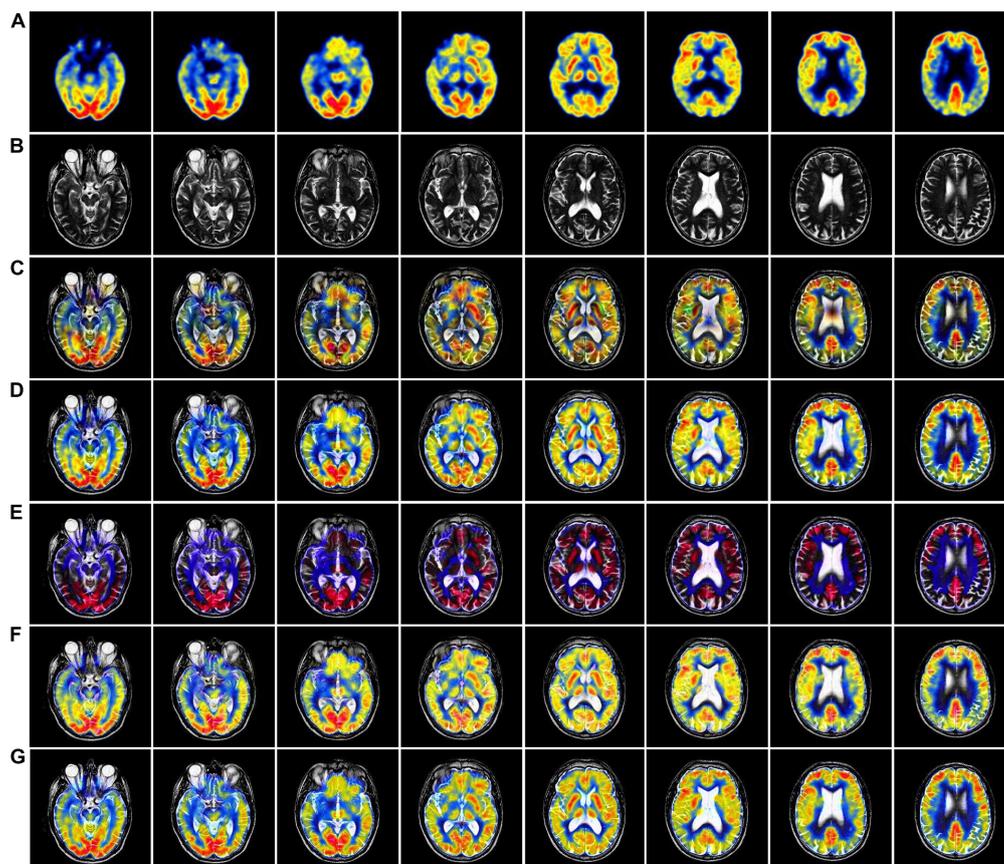


FIGURE 5 MRI-PET fusion images of five methods in Paras2. **(A)** Positron emission tomography (PET) source image; **(B)** magnetic resonance imaging (MRI) source image; **(C)** nonsubsampled contourlet (NSCT) fusion result; **(D)** non-subsampled shear wave transform (NSST) fusion result; **(E)** guided filtering fusion (GFF) fusion result; **(F)** Laplacian redecomposition (ReLP) fusion result; **(G)** is the fusion result of the proposed method.

TABLE 2 Average values of index evaluation of different fusion methods for MRI-PET.

CT-MRI	Methods	MI	SF	AG	EI	NIQE	TMQI
Paras1	NSCT	3.7269	15.2214	6.1476	45.3264	4.5158	0.8123
	NSST	3.8261	15.3842	6.1842	47.8139	4.5927	0.8546
	GFF	4.1027	16.7253	6.0878	46.9367	4.5231	0.7712
	ReLP	4.1503	17.2636	6.2913	48.4811	4.6080	0.8582
	Proposed	4.2018	17.3059	6.3682	50.6531	4.6675	0.8654
Paras2	NSCT	3.6535	16.1356	6.0631	46.6235	4.6825	0.7403
	NSST	3.7282	16.8575	6.0051	45.3315	4.7057	0.6790
	GFF	3.7451	17.2243	6.2527	46.8728	4.6817	0.7052
	ReLP	3.7871	17.7991	6.3517	46.3297	4.7354	0.7106
	Proposed	3.8010	18.1052	6.4063	47.2031	4.7521	0.7149

GTX 3090 Ti. We train the convolutional sparse and low-rank dictionary with sliding step size set to 1, sliding window size set to 8×8 , dictionary size set to 64×512 , error set to $\mathcal{E} = 0.03$, and decomposition level set to 3.

(2) Data sets and comparison methods

To validate the performance of the proposed method. We selected 136 sets of aligned AD brain medical images (image size of 256×256 pixels) as the source images to be fused. All image slices were obtained from the Harvard Whole Brain Atlas database (Johnson and Becker, 2001), and the three AD

medical image types included 42 sets of CT-MRI images; 42 sets of MRI-PET images; and 52 sets of MRI-SPECT images. Four contrast algorithms were adopted for comparison, including nonsubsampling contourlet (NSCT) (Li and Wang, 2011), NSST (Kong et al., 2014), guided filter ng fusion (GFF) (Li et al., 2013), and Laplacian redecomposition (ReLP) (Li et al., 2020).

(3) Objective evaluation metrics

We selected 10 metrics for objective index evaluation and analysis: mutual information (MI) (Xydeas and Petrovic, 2000), natural image quality evaluator (NIQE) (Mittal et al., 2012), average gradient (AG) (Du et al., 2017), edge intensity (EI) (Wang et al., 2012), tone-mapped image quality index (TMQI) (Yeganeh and Wang, 2012), spatial frequency (SF) (Eskicioglu and Fisher, 1995), SD (Li et al., 2008), root mean square error (RMSE) (Zhang et al., 2018), gradient similarity mechanism (GSM) (Liu et al., 2011), and VIF (Sheikh and Bovik, 2006). SF is the spatial frequency, which is a measure of the richness of image detail and reflects the sharpness of image detail. A larger value means that the image detail is richer. SD is the standard deviation, which measures the contrast of the image;

a larger value indicates a better contrast of the image. RMSE is the root mean square error, which measures the difference between the fused image and the source image; a smaller value indicates that the fused image information is closer to the source image. GSM measures the gradient similarity between images; a larger value indicates that the gradient information of the fused image is closer to the source image. NIQE index, the smaller the value, the smaller the distortion. VIF is an image information measure that quantifies the information present in the fused image; larger values indicate better fusion. NIQE measures the simple distance between the model statistic and the distorted image statistic. AG indicates the average gradient, which is used to extract the contrast and texture change features of the image. EI reflects the sharpness of the edges. TMQI index measures the significant features of brightness and contrast between the reference image and the fused image, and measures the structural fidelity of the fused image. Larger values of MI, SF, AG, EI, and TMQI indexes indicate better fusion.

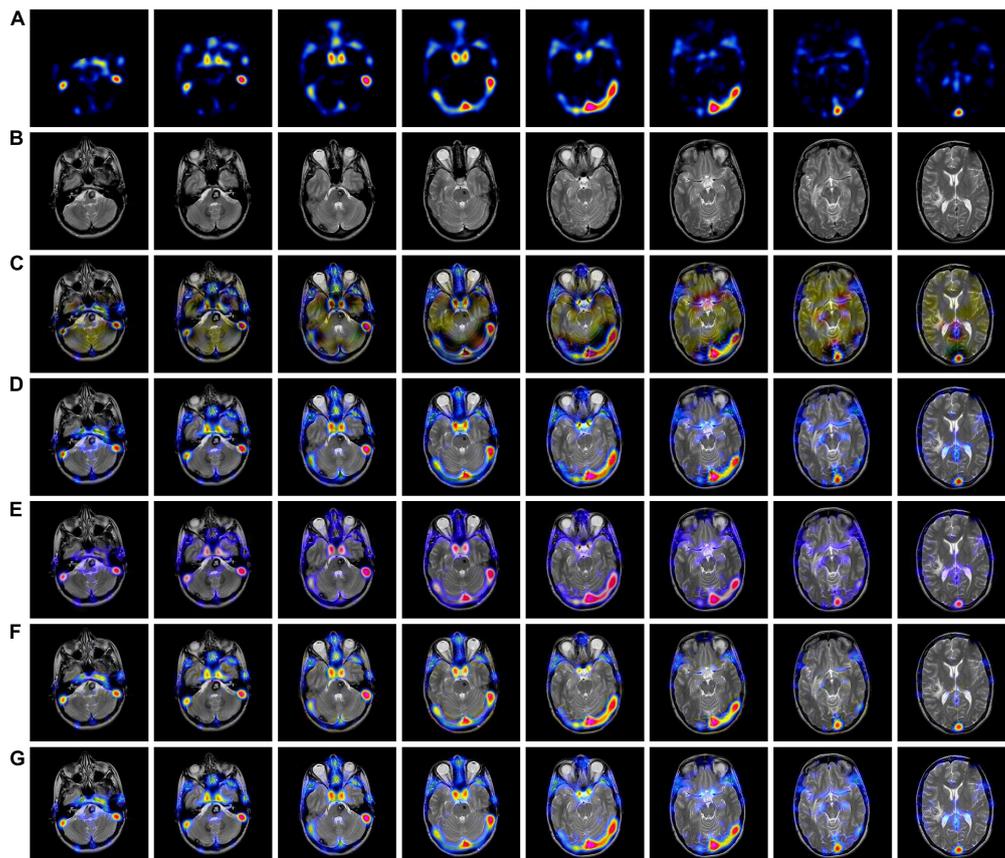


FIGURE 6

MRI-SPECT fusion images of the five methods in Paras1. (A) Single photon emission computed tomography (SPECT) source image; (B) magnetic resonance imaging (MRI) source image; (C) nonsubsampling contourlet (NSCT) fusion result; (D) non-subsampling shear wave transform (NSST) fusion result; (E) guided filter ng fusion (GFF) fusion result; (F) Laplacian redecomposition (ReLP) fusion result; (G) fusion result of the proposed method.

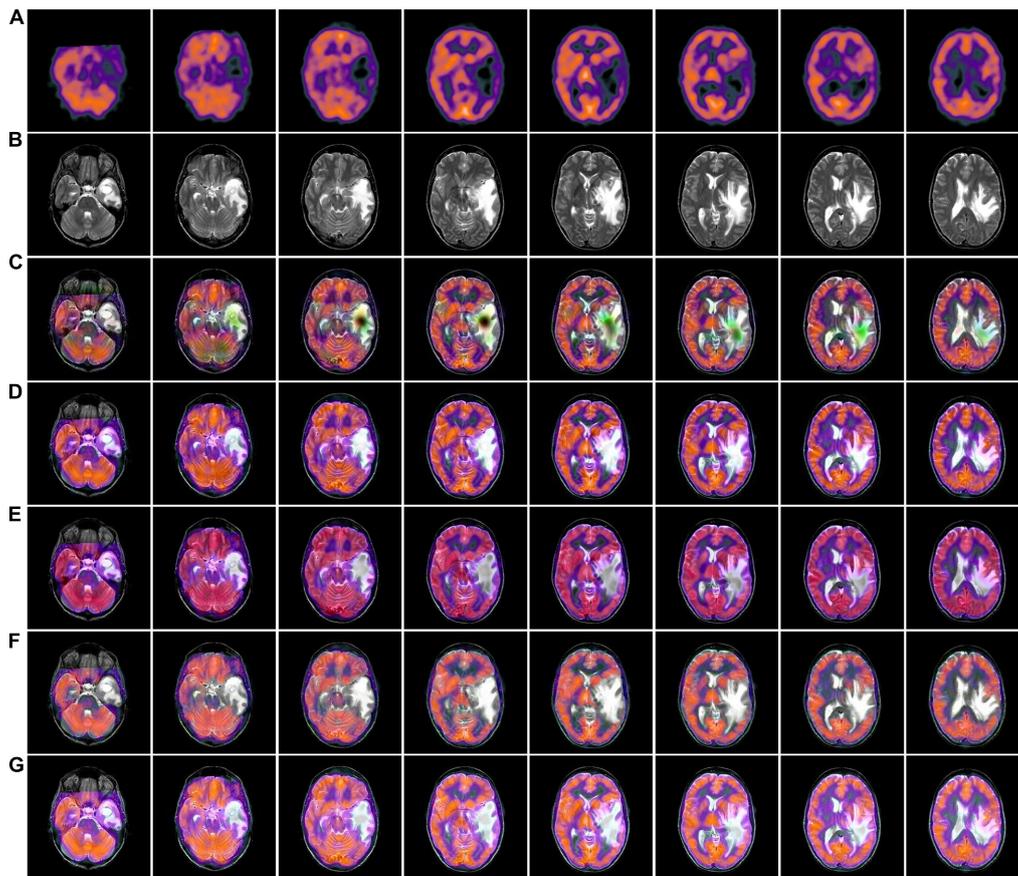


FIGURE 7
MRI-SPECT fusion images of the five methods in Paras2. (A) Single photon emission computed tomography (SPECT) source image; (B) magnetic resonance imaging (MRI) source image; (C) nonsampled contourlet (NSCT) fusion result; (D) non-sampled shear wave transform (NSST) fusion result; (E) guided filter ng fusion (GFF) fusion result; (F) Laplacian redecomposition (ReLP) fusion result; (G) fusion result of the proposed method.

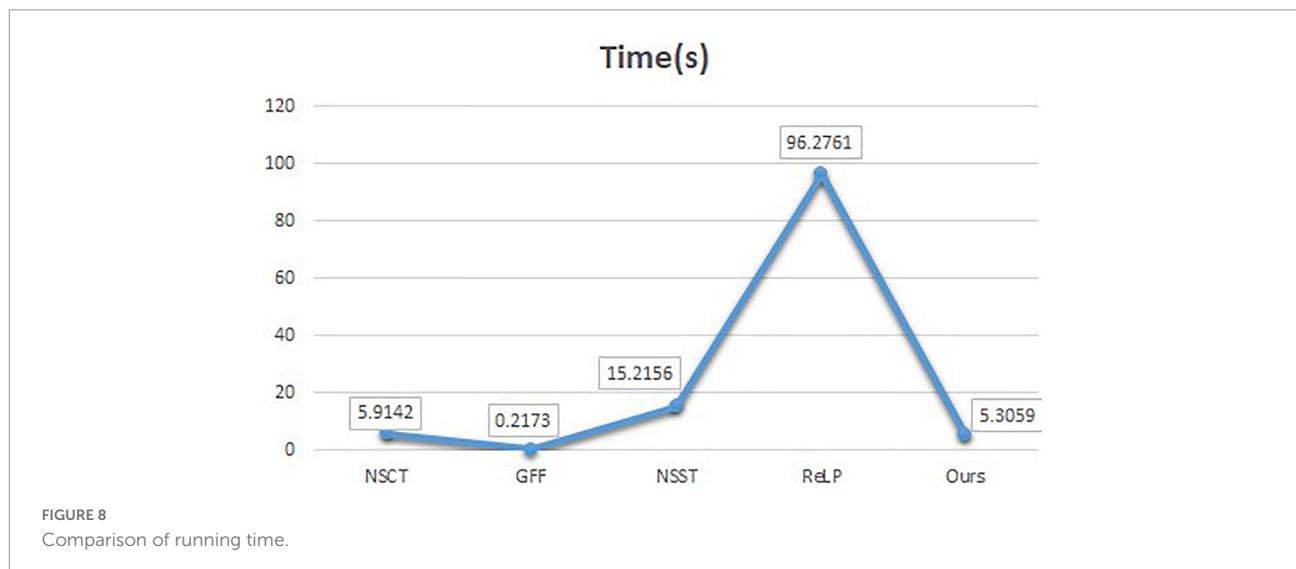
TABLE 3 Average values of index evaluation of different fusion methods for MRI-SPECT.

MRI-SPECT	Methods	MI	SF	AG	EI	NIQE	TMQI
Paras1	NSCT	3.6302	16.4106	6.2750	46.6732	4.2038	0.7931
	NSST	3.6231	15.8942	6.1773	47.0792	4.3107	0.7785
	GFF	3.7047	16.5378	6.2035	47.6470	4.2619	0.8063
	ReLP	3.8503	17.0346	6.3106	48.5718	4.3250	0.8307
	Proposed	3.8826	17.2069	6.4209	49.5014	4.4030	0.8526
Paras2	NSCT	3.6057	15.7431	6.1821	47.1003	4.3717	0.7352
	NSST	3.6183	15.7215	6.1040	46.6993	4.4602	0.7075
	GFF	3.6903	15.4903	6.2183	47.2854	4.3736	0.7202
	ReLP	3.7681	16.3107	6.3608	48.5810	4.4071	0.7464
	Proposed	3.7916	16.5012	6.4112	48.7106	4.5306	0.7503

3.2. CT-MRI fusion results comparison

In **Figure 3**, we used 42 sets of CT-MRI fused images and randomly selected seven fused images for comparison. From

the figure, we can see that the images fused by NSCT and GFF algorithms are too dark. The images fused by NSST are not only darker but also distorted. The images fused by ReLP algorithm have better brightness but not enough texture details. Our fusion



algorithm performs best in terms of brightness, detail texture and edge contour.

The mean values of objective evaluation metrics for fusion results corresponding to different rules are given in **Table 1**, where bold indicates that the method ranks best in the metrics. NSCT is low in SD in terms of metrics, indicating insufficient image contrast. NSST is lowest in SF in terms of metrics, with poor image details. ReLP is less distorted with our method in terms of RMSE and GSM. Our proposed method performs better performance in terms of color retention, contrast, and detail retention, and achieves the optimum.

3.3. MRI-PET fusion results comparison

Figures 4, 5 use 42 sets of MRI-PET images from MRI-PET datasets Paras1 and Paras2, respectively. We randomly selected eight fused images. From **Figure 4**, it can be seen that NSCT and GFF show severe distortion, and the fused images of NSST and ReLP algorithms are too dark and have loss of detail information. In **Figure 5**, NSCT and NSST have dark luminance and GFF has distortion, while ReLP and our fusion algorithm have better visual effect.

In **Table 2**, by comparing 42 sets of fused images on the MRI-PET dataset, our proposed algorithm has the best mean value in objective evaluation metrics. Higher contrast, sharper edges and finer details were obtained. The subjective results of the fused images of the two algorithms, NSCT and GFF, were not satisfactory. NSCT and GFF had more color distortion. NSST showed abnormal brightness. ReLP performed better and was close to our average value. So far, it is easy to see that the multi-objective evaluation index of the integrated information is consistent with the conclusions of the subjective analysis. Our proposed algorithm significantly outperforms the average

of all algorithms. In summary, we have a more comprehensive advantage over existing algorithms in the evaluation of objective metrics.

3.4. MRI-SPECT fusion results comparison

Figures 6, 7 we used MRI-SPECT datasets Paras1 and Paras2, respectively 52 sets of AD MRI-PET image fusion images for comparison. It can be seen from the figures that NSCT shows severe distortion, NSST fused images are too dark, GFF shows brightness anomalies, and ReLP does not perform well in terms of detail texture. Our fusion algorithm performs best in brightness, detail texture and edge contour.

In **Table 3**, we use 52 sets of fused images on the MRI-SPECT dataset for comparison, and our proposed DCT multiscale decomposition obtains sharper edges and finer details. The improved NMSF fusion rule obtains better brightness and contrast. The superiority of our method over other algorithms is demonstrated.

To compare the advantages of the proposed methods more comprehensively, we calculate the running times of the comparison methods on the same pair of images of 256×256 size. **Figure 8** shows the line graphs of the average running times of our method and the four comparison methods. The ReLP method has the longest running time and GFF has the shortest running time. From the line graph, it can be seen that our fusion method has the second best running speed than most of the other algorithms. However, medical image fusion is used to assist in diagnosis and treatment, and the effectiveness of the proposed method is demonstrated from objective and subjective evaluations. Therefore, the proposed method guarantees the quality of fusion results within an acceptable time consumption.

TABLE 4 Subjective quality evaluation of different fusion algorithms.

Method	CT-MRI				MRI-SPECT			
	Contrast ratio↑	Detail↑	Invariance↑	Acceptability↑	Contrast ratio↑	Detail↑	Invariance↑	Acceptability↑
NSCT	3.8 ± 0.52	3.6 ± 0.25	0.5 ± 0.39	3.6 ± 0.97	3.7 ± 0.85	3.8 ± 0.66	0.6 ± 0.14	3.7 ± 0.72
NSST	3.7 ± 0.68	3.7 ± 0.13	0.6 ± 0.37	3.5 ± 0.55	3.7 ± 0.63	3.9 ± 0.45	0.5 ± 0.97	3.6 ± 0.26
GFF	3.9 ± 0.43	3.6 ± 0.07	0.5 ± 0.93	3.7 ± 0.70	3.9 ± 0.51	4.0 ± 0.23	0.7 ± 0.41	3.7 ± 0.91
ReLP	4.0 ± 0.24	3.8 ± 0.64	0.7 ± 0.86	3.8 ± 0.61	4.1 ± 0.07	4.1 ± 0.58	0.7 ± 0.80	4.0 ± 0.96
Proposed	4.1 ± 0.19	3.9 ± 0.40	0.8 ± 0.75	4.0 ± 0.63	4.1 ± 0.79	4.2 ± 0.11	0.8 ± 0.05	4.1 ± 0.18

3.5. Comparison of subjective evaluation

Computed tomography (CT) or MRI unimodal imaging can no longer meet the demand for precise diagnosis in neurosurgery. A multimodality imaging technique that can clearly, visually, and holistically show AD brain atrophy and its association with surrounding cerebral vessels, nerves, and brain tissues can only accommodate the development of neurosurgical precision surgery. Quality assessment of multimodal fusion requires additional medical expertise. Therefore, we invited six chief neurosurgeons with more than 5 years of experience, and we randomly selected a test sample of 10 groups, each group including five fusion images. The subjective evaluation criteria were double stimulus continuous quality scale (DSCQS) including contrast, detail and invariance and acceptability scores of [1 (worst) to 5 (best)]. 1 indicates a non-diagnostic image and 5 indicates a good quality diagnostic image. Pathological invariance was scored as 0 (change) or 1 (no change). **Table 4** shows the ratings of six clinicians, and the optimal values are shown in bold.

In **Table 4**, the subjective physician evaluations of CT-MRI and MRI-SPECT fusion are presented. The NSCT and GFF contrast and brightness were insufficient and therefore rated low. The GFF showed the worst distortion acceptability evaluation. The ReLP was very close to our evaluation among the four evaluation metrics. Our algorithm has the best performance in edge detail, luminance, contrast and spatial coherence, and received the best physician evaluation.

4. Conclusion

Multimodal neuroimaging data have high dimensionality and complexity, and seeking efficient methods to extract valuable features in complex datasets is the focus of current research. To address the shortcomings of AD multimodal fusion images such as contrast reduction, detail blurring and color distortion, we propose a multimodal fusion algorithm for Alzheimer's disease based on DCT convolutional SR. The DCT multi-scale decomposition of the source medical image is performed to obtain the basic layer, local average energy layer and texture layer of the input image, and then the sub-images of different scales are used as training images respectively. The sub-dictionaries at different scales are optimally solved using the ADMM algorithm, and then convolutional sparse coding is performed, and the inverse DCT transform of the subimage coefficients is performed using a combination of improved L_1 parameters and improved NMSF rules to obtain the multimodal fusion images. We experimentally demonstrate that the algorithm has sharper edge details, better color and spatial consistency than other algorithms by fusing medical images in three modalities, CT-MRI, MRI-PET,

and MRI-SPECT. This proves that our algorithm outperforms existing state-of-the-art algorithms. In the future, we will use deep learning models for medical image multimodality classification and prediction, and apply them to early clinical diagnosis of AD.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

Author contributions

GZ and XN: investigation, methodology, software, validation, visualization, and writing—original draft. BL, JL, and HY: investigation, methodology, software, and supervision. WS and SH: conceptualization, data curation, formal analysis, funding acquisition, methodology, project administration, resources, supervision, validation, and writing—review and editing.

Funding

This work was supported by the Doctoral Innovative Talents Project of Chongqing University of Posts and Telecommunications (BYJS202107 and BYJS202112), the

open project program of Chongqing Key Laboratory of Photo-Electric Functional Materials [K(2022)215], Postdoctoral Science Foundation of Chongqing (cstc2021jcyj-bsh0218), Special financial aid to post-doctor research fellow of Chongqing (2011010006445227), the National Natural Science Foundation of China (U21A20447 and 61971079), the Basic Research and Frontier Exploration Project of Chongqing (cstc2019jcyjmsxmX0666), Chongqing technological innovation and application development project (cstc2021jscx-gksbx0051), the Science and Technology Research Program of Chongqing Municipal Education Commission (KJZD-k202000604), the Innovative Group Project of the National Natural Science Foundation of Chongqing (cstc2020jcyj-cxttX0002), and the Regional Creative Cooperation Program of Sichuan (2020YFQ0025).

Acknowledgments

We thank the School of Optoelectronic Engineering of Chongqing University of Posts and Telecommunications for their assistance in the research.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

- Chetelat, G., Desgranges, B., De La Sayette, V., Viader, F., Eustache, F., and Baron, J. C. (2003). Mild cognitive impairment: Can FDG-PET predict who is to rapidly convert to Alzheimer's disease? *Neurology* 60, 1374–1377. doi: 10.1212/01.wnl.0000055847.17752.e6
- Da Cunha, A. L., Zhou, J., and Do, M. N. (2006). The nonsubsampled contourlet transform: Theory, design, and applications. *IEEE Trans. Image Process.* 15, 3089–3101. doi: 10.1109/TIP.2006.877507
- Donoho, D. L. (2006). Compressed sensing. *IEEE Trans. Inform. Theory* 52, 1289–1306. doi: 10.1109/TIT.2006.871582
- Du, J., Li, W., and Xiao, B. (2017). Anatomical-functional image fusion by information of interest in local Laplacian filtering domain. *IEEE Trans. Image Process.* 26, 5855–5866. doi: 10.1109/TIP.2017.2745202
- Dubois, B., Villain, N., Frisoni, G. B., Rabinovici, G. D., Sabbagh, M., Cappa, S., et al. (2021). Clinical diagnosis of Alzheimer's disease: Recommendations of the international working group. *Lancet Neurol.* 20, 484–496. doi: 10.1016/S1474-4422(21)00066-1
- Eskicioglu, A. M., and Fisher, P. S. (1995). Image quality measures and their performance. *IEEE Trans. Commun.* 43, 2959–2965.

- Fan, J., Cao, X., Yap, P. T., and Shen, D. (2019). BIRNet: Brain image registration using dual-supervised fully convolutional networks. *Med. Image Anal.* 54, 193–206. doi: 10.1016/j.media.2019.03.006
- Farbman, Z., Fattal, R., Lischinski, D., and Szeliski, R. (2008). Edge-preserving decompositions for multi-scale tone and detail manipulation. *ACM Trans. Graph.* 27, 1–10. doi: 10.1145/1360612.1360666
- Gu, S., Zuo, W., Xie, Q., Meng, D., Feng, X., and Zhang, L. (2015). “Convolutional sparse coding for image super-resolution,” in *Proceedings of the international conference on computer vision, Santiago, Chile* (Manhattan, NY: IEEE), 1823–1831.
- He, K., Sun, J., and Tang, X. (2012). Guided image filtering. *IEEE Trans. Pattern Anal.* 35, 1397–1409. doi: 10.1109/TPAMI.2012.213
- He, Y., Li, G., Liao, Y., Sun, Y., Kong, J., Jiang, G., et al. (2019). Gesture recognition based on an improved local sparse representation classification algorithm. *Cluster Comput.* 22, 10935–10946. doi: 10.1007/s10586-017-1237-1
- Hu, J., and Li, S. (2012). The multiscale directional bilateral filter and its application to multisensor image fusion. *Inform. Fusion* 13, 196–206. doi: 10.1016/j.inffus.2011.01.002
- Johnson, K. A., and Becker, J. A. (2001). *The whole brain atlas*. Available online at: <https://pesquisa.bvsalud.org/portal/resource/pt/lis-LISBR1.1-5215>
- Kong, W., Zhang, L., and Lei, Y. (2014). Novel fusion method for visible light and infrared images based on an improved local sparse representation classification algorithm. *Infrared Phys. Techn.* 65, 103–112. doi: 10.1016/j.infrared.2014.04.003
- Kou, F., Chen, W., Wen, C., and Li, Z. (2015). Gradient domain guided image filtering. *IEEE Trans. Image Process.* 24, 4528–4539. doi: 10.1109/TIP.2015.2468183
- Li, S., Hong, R., and Wu, X. (2008). “A novel similarity based quality metric for image fusion,” in *Proceedings of international conference on audio, language and image processing, Shanghai, China* (Manhattan, NY: IEEE), 167–172.
- Li, S., Kang, X., and Hu, J. (2013). Image fusion with guided filtering. *IEEE Trans. Image Process.* 22, 2864–2875. doi: 10.1109/TIP.2013.2244222
- Li, T., and Wang, Y. (2011). Biological image fusion using a NSCT based variable-weight method. *Inform. Fusion* 12, 85–92. doi: 10.1016/j.inffus.2010.03.007
- Li, X., Guo, X., Han, P., Wang, X., Li, H., and Luo, T. (2020). Laplacian redecomposition for multimodal medical image fusion. *IEEE Trans. Instrum. Meas.* 69, 6880–6890. doi: 10.1109/TIM.2020.2975405
- Lian, C., Zhang, J., Liu, M., Zong, X., Hung, S. C., Lin, W., et al. (2018). Multi-channel multi-scale fully convolutional network for 3D perivascular spaces segmentation in 7T MR images. *Med. Image Anal.* 46, 106–117. doi: 10.1016/j.media.2018.02.009
- Liu, A., Lin, W., and Narwaria, M. (2011). Image quality assessment based on gradient similarity. *IEEE Trans. Image Process.* 21, 1500–1512. doi: 10.1109/TIP.2011.2175935
- Liu, G., and Yan, S. (2011). “Latent low-rank representation for subspace segmentation and feature extraction,” in *Proceedings of 2011 international conference on computer vision, Barcelona, Spain* (Manhattan, NY: IEEE), 1615–1622. doi: 10.1109/ICCV.2011.6126422
- Liu, H., Liu, Y., and Sun, F. (2014). Robust exemplar extraction using structured sparse coding. *IEEE Trans. Neur. Net. Lear.* 26, 1816–1821. doi: 10.1109/TNNLS.2014.2357036
- Liu, H., Yu, Y., Sun, F., and Gu, J. (2016). Visual-tactile fusion for object recognition. *IEEE Trans. Autom. Sci. Eng.* 14, 996–1008. doi: 10.1109/TASE.2016.2549552
- Liu, M., Zhang, D., and Shen, D. (2016). Relationship induced multi-template learning for diagnosis of Alzheimer’s disease and mild cognitive impairment. *IEEE Trans. Med. Imaging* 35, 1463–1474. doi: 10.1109/TMI.2016.2515021
- Liu, M., Zhang, J., Yap, P. T., and Shen, D. (2017). View-aligned hypergraph learning for Alzheimer’s disease diagnosis with incomplete multi-modality data. *Med. Image Anal.* 36, 123–134. doi: 10.1016/j.media.2016.11.002
- Liu, S., Shi, M., Zhu, Z., and Zhao, J. (2017). Image fusion based on complex-shearlet domain with guided filtering. *Multidim. Syst. Sign. Process.* 28, 207–224. doi: 10.1007/s11045-015-0343-6
- Liu, X., Mei, W., and Du, H. (2017). Structure tensor and nonsubsampling shearlet transform based algorithm for CT and MRI image fusion. *Neurocomputing* 235, 131–139. doi: 10.1016/j.neucom.2017.01.006
- Miao, Q. G., Shi, C., Xu, P. F., Yang, M., and Shi, Y. B. (2011). A novel algorithm of image fusion using shearlets. *Opt. Commun.* 284, 1540–1547. doi: 10.1016/j.optcom.2010.11.048
- Mittal, A., Soundararajan, R., and Bovik, A. C. (2012). Making a “completely blind” image quality analyzer. *IEEE Sig. Proc. Lett.* 20, 209–212. doi: 10.1109/LSP.2012.2227726
- Mohammadi-Nejad, A. R., Hossein-Zadeh, G. A., and Soltanian-Zadeh, H. (2017). Structured and sparse canonical correlation analysis as a brain-wide multimodal data fusion approach. *IEEE Trans. Med. Imaging* 36, 1438–1448. doi: 10.1109/TMI.2017.2681966
- Perrin, R. J., Fagan, A. M., and Holtzman, D. M. (2009). Multimodal techniques for diagnosis and prognosis of Alzheimer’s disease. *Nature* 461, 916–922. doi: 10.1038/nature08538
- Rong, Y., Xiong, S., and Gao, Y. (2017). Low-rank double dictionary learning from corrupted data for robust image classification. *Pattern Recogn.* 72, 419–432. doi: 10.1016/j.patcog.2017.06.038
- Sheikh, H. R., and Bovik, A. C. (2006). Image information and visual quality. *IEEE Trans. Image Process.* 15, 430–444. doi: 10.1109/TIP.2005.859378
- Thung, K. H., Wee, C. Y., Yap, P. T., Shen, D., and Alzheimer’s Disease Neuroimaging Initiative. (2014). Neurodegenerative disease diagnosis using incomplete multi-modality data via matrix shrinkage and completion. *NeuroImage* 91, 386–400. doi: 10.1016/j.neuroimage.2014.01.033
- Veitch, D. P., Weiner, M. W., Aisen, P. S., Beckett, L. A., DeCarli, C., Green, R. C., et al. (2022). Using the Alzheimer’s disease neuroimaging initiative to improve early detection, diagnosis, and treatment of Alzheimer’s disease. *Alzheimer’s Dement.* 18, 824–857. doi: 10.1002/alz.12422
- Wang, G., Li, W., and Huang, Y. (2021). Medical image fusion based on hybrid three-layer decomposition model and nuclear norm. *Comput. Biol. Med.* 129:104179. doi: 10.1016/j.compbiomed.2020.104179
- Wang, M., and Shang, X. (2020). A fast image fusion with discrete cosine transform. *IEEE Sig. Process. Lett.* 27, 990–994. doi: 10.1109/LSP.2020.2999788
- Wang, Y., Du, H., Xu, J., and Liu, Y. (2012). “A no-reference perceptual blur metric based on complex edge analysis,” in *Proceedings of international conference on network infrastructure and digital content, Beijing, China* (Manhattan, NY: IEEE), 487–491.
- Xu, L., Lu, C., Xu, Y., and Jia, J. (2011). “Image smoothing via L_0 gradient minimization,” in *Proceedings of the 2011 SIGGRAPH Asia conference, Beijing, China* (New York, NY: ACM), 1–12.
- Xydeas, C. S., and Petrovic, V. (2000). Objective image fusion performance measure. *Electron. Lett.* 36, 308–309. doi: 10.1049/el:20000267
- Yang, B., and Li, S. (2009). Multifocus image fusion and restoration with sparse representation. *IEEE T. Instrum. Meas.* 59, 884–892. doi: 10.1109/TIM.2009.2026612
- Yang, S., Wang, M., Jiao, L., Wu, R., and Wang, Z. (2010). Image fusion based on a new contourlet packet. *Inform. Fusion* 11, 78–84. doi: 10.1016/j.inffus.2009.05.001
- Yeganeh, H., and Wang, Z. (2012). Objective quality assessment of tone-mapped images. *IEEE Trans. Image Process.* 22, 657–667. doi: 10.1109/TIP.2012.2221725
- Yin, H., Li, Y., Chai, Y., Liu, Z., and Zhu, Z. (2016). A novel sparse-representation-based multi-focus image fusion approach. *Neurocomputing* 216, 216–229. doi: 10.1016/j.neucom.2016.07.039
- Zhang, H., and Patel, V. M. (2017). Convolutional sparse and low-rank coding-based image decomposition. *IEEE Trans. Image Process.* 27, 2121–2133. doi: 10.1109/TIP.2017.2786469
- Zhang, K., Huang, Y., and Zhao, C. (2018). Remote sensing image fusion via RPCA and adaptive PCNN in NSST domain. *Int. J. Wavelets Multiresolut. Inf. Process.* 16:1850037. doi: 10.1142/S0219691318500376
- Zhang, Q., and Levine, M. D. (2016). Robust multi-focus image fusion using multi-task sparse representation and spatial context. *IEEE Trans. Image Process.* 25, 2045–2058. doi: 10.1109/TIP.2016.2524212
- Zhang, Q., Shen, X., Xu, L., and Jia, J. (2014). “Rolling guidance filter,” in *European conference on computer vision, Zurich, Switzerland: ECCV*, eds D. Fleet, T. Pajdla, B. Schiele, and T. Tuytelaars (Cham: Springer), 815–830.
- Zhu, Z., Chai, Y., Yin, H., Li, Y., and Liu, Z. (2016). A novel dictionary learning approach for multi-modality medical image fusion. *Neurocomputing* 214, 471–482. doi: 10.1016/j.neucom.2016.06.036
- Zong, J. J., and Qiu, T. S. (2017). Medical image fusion based on sparse representation of classified image patches. *Biomed. Signal Process.* 34, 195–205. doi: 10.1016/j.bspc.2017.02.005