Multimedia Tools and Applications (2021) 80:6369–6396
https://doi.org/10.1007/s11042-020-08834-5

Survey study of multimodality medical image fusion methods

Nahed Tawfik[1] · Heba A. Elnemr[1] · Mahmoud Fakhr[1] · Moawad I. Dessouky[2] · Fathi E. Abd El-Samie[2,3]

Received: 1 April 2019 / Revised: 13 January 2020 / Accepted: 6 March 2020 /
Published online: 15 October 2020

Abstract
Multimodality medical image fusion is the process of combining multiple images from single or multiple modalities of imaging. Medical image fusion methods are adopted to increase the quality of medical images by attaining the salient features in the fusion results. Hence, they raise the clinical applicability of medical images for appraisal and diagnosis problems. This purpose is achieved by capturing the complementary information presented in two or more images of different modalities in the fusion result. Medical image fusion is generally concerned with Magnetic Resonance Imaging (MRI), Magnetic Resonance Angiogram (MRA), Positron Emission Tomography (PET), Structural Positron Emission Tomography (SPET), Computerized Tomography (CT), and Single-Photon Emission Computed Tomography (SPECT) modalities. Each modality has its merits and drawbacks. This induces new fusion methods for merging information from multiple imaging modalities. Researchers have presented several methods for medical image fusion, and these methods achieved good results. However, medical image fusion is a resurgent field that needs to be enhanced to conquer the increasing challenges. This paper presents a comprehensive survey of some existing medical image fusion methods. It is expected that this study will be useful for the researchers scrutinizing medical image fusion. Furthermore, it is expected to establish a concrete foundation for developing more powerful fusion methods for medical applications.

Keywords Imaging modalities · Medical image fusion · MRI · CT · PET · SPECT · Clinical diagnosis

1 Introduction

Image fusion is the process of integrating a diversity of images obtained with different imaging modalities. The objective of the fusion process in medical applications is to improve the image

✉ Nahed Tawfik
nahedtawfik@eri.sci.eg

Extended author information available on the last page of the article

suitability for diagnosis and evaluation of medical problems by preserving and illuminating the salient features [25].

Commonly, medical image fusion is concerned with MRI, PET, CT and SPECT images. The MR images present a superior definition of soft tissues and a higher resolution in the spatial domain. Unfortunately, they provide little information about movements. On the other hand, CT has a significance as a three-dimensional imaging technique with the advantages of better image resolutions and short times for scans. Yet, tissue description is limited. Furthermore, PET provides highly-sensitive images owing to the molecular imaging procedure. However, the obtained images have low resolution. Finally, SPECT is a nuclear imaging technique that is used to explore the blood flow in tissues and organs. Each imaging technique has its practical characteristics and its limitations [25]. Hence, there is a need to find new imaging technologies or to develop new fusion methods that are able to merge information from several images with different modalities and take the privileges of the complementary characteristics exhibited in these different images. The multimodality medical image fusion generally focuses on three groups: MRI-CT, MRI-PET, and MRI-SPECT images. Figure 1 displays examples of the fusion groups and fusion results. The source images that are widely used in medical image fusion are listed in the Harvard Medical School www.med.harvard.edu/aanlib/home.html. Besides, the fusion groups can be downloaded from http://www.metapix.de/fusion.htm and www.imagefusion.org.

Generally, image fusion methods comprise input image decomposition, a fusion rule, image reconstruction, and image quality assessment. Initially, the source images are decomposed into sub-images by performing decomposition techniques such as Intensity-Hue-Saturation (IHS), pyramid, distinctive wavelet, Non-Subsampled Contourlet Transform (NSCT), shearlet transform, Sparse Representation (SR), and others. Then, fusion algorithms are applied to merge multiple features from these sub-images. The commonly-used fusion algorithms are fuzzy logic, Principal Component Analysis (PCA), Human Visual System (HVS) fusion, Artificial Neural Networks (ANNs), Pulse-Coupled Neural Networks (PCNNs), Mutual Information (MI) fusion, etc. Afterwards, image reconstruction algorithms are implemented to obtain the final fusion result. Finally, the image quality assessment is performed to evaluate the quality of the fusion result by using both subjective and objective assessment metrics. For subjective evaluation, the radiologists are asked to evaluate the fusion result, while in the objective evaluation, different parameters such as Peak Signal-to-Noise Ratio (*PSNR*), Root-Mean-Squared Error (*RMSE*), standard deviation (*std*), entropy, Spatial Frequency (*SF*), Universal image quality index (Q_o), Structural Similarity (*SSIM*) index, Edge-based similarity measure ($Q^{AB/F}$), Feature Similarity (*FSIM*) index, weighted fusion quality index Q_W, Average Gradient (*AG*), and natural image quality are used [14]. Figure 2 illustrates the main steps of the image fusion process.

Image fusion methods can be categorized into pixel-level, feature-level, and decision-level methods. Pixel-level image fusion is the process of directly merging the original information from the source images or from their multi-resolution transforms to create a more informative final image for visual perception. Feature-level fusion aims to extract salient features from the source images such as shape, length, edges, segments, and directions. The extracted features from input images are merged to form more meaningful features, which provide a better descriptive and comprehensive image. Decision-level fusion represents a high level of fusion that indicates the actual target. It joins the results from several algorithms to produce an ultimate fusion decision [43].

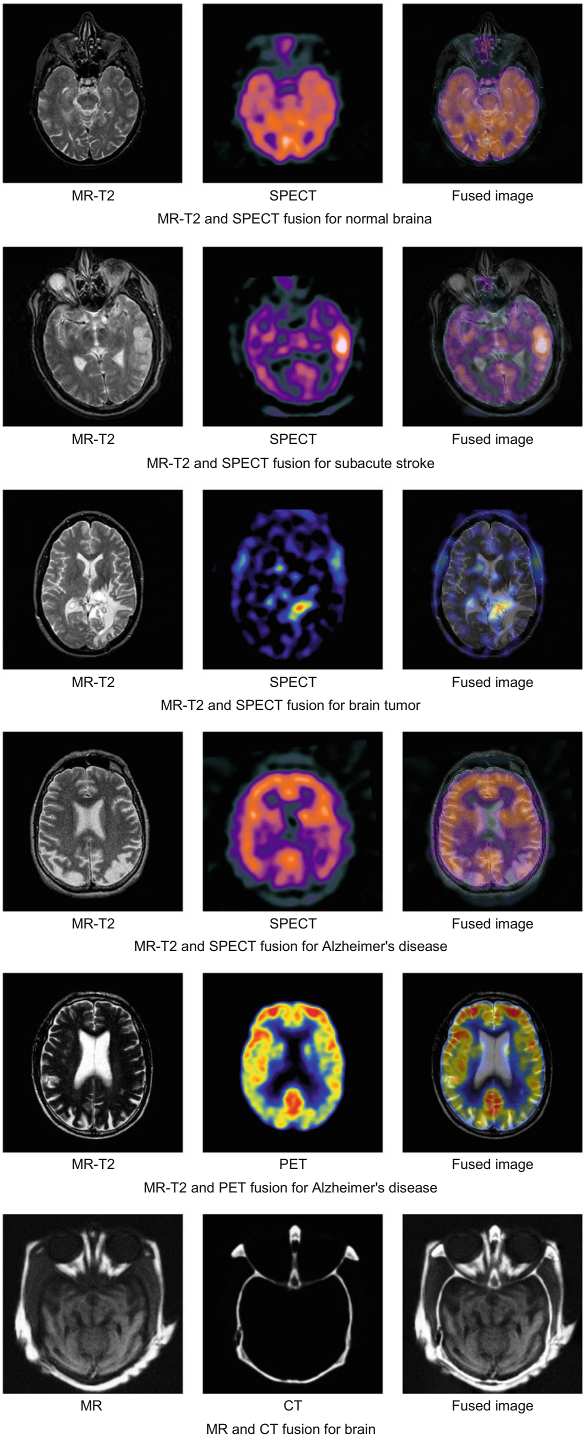

Fig. 1 Examples of multimodality medical image fusion (fusion of MR-CT, MR-PET and MR-SPECT images)

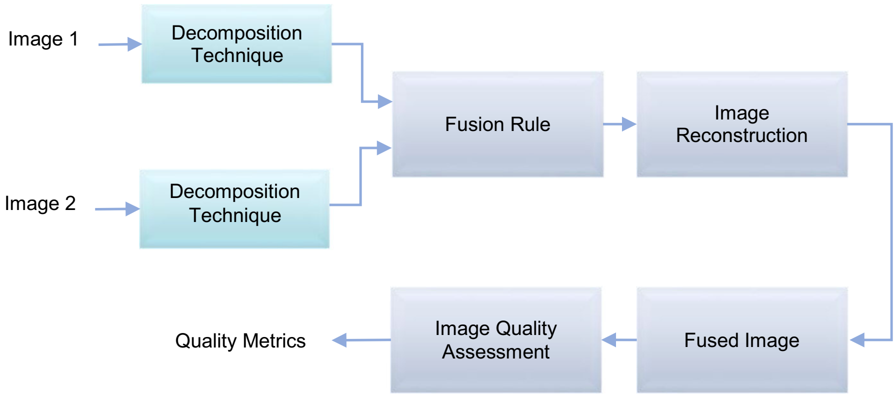

Fig. 2 Steps of image fusion

Researchers have addressed several methods for medical image fusion, and these methods give good results. Still, there is a scope for enhancement in the performance of medical image fusion. In this work, we present a detailed survey of some salutary image fusion methods.

2 Types of multimodality medical image fusion

Generally, image fusion methods are categorized into pixel-level fusion, feature-level fusion, and decision-level fusion. In the literature, the majority of reported medical image fusion methods are directed towards pixel-level fusion. On the other hand, some researches concentrated on feature-level fusion methods. In addition, several published works are based on combinations of pixel-level and feature-level fusion. Furthermore, some researchers considered decision-level fusion to accomplish a higher level of fusion. Figure 3 illustrates the commonly-used multimodality medical image fusion methods. In the following sections, we introduce some of these directions.

2.1 Pixel-level fusion

Pixel-level image fusion methods merge the original information from the source images directly or from their multi-resolution transforms. They aim to get more informative fusion results that are appropriate for visual comprehension and computer processing.

Pixel-level fusion methods can be implemented in the spatial domain as well as in different transform domains. In the spatial-domain methods, the pixel values of the images to be fused are merged linearly or non-linearly to obtain the fusion result. However, in the transform-domain methods, the images are firstly transformed to an appropriate transform domain. Thus, the fusion process is performed in this domain. Hence, the inverse transform is performed to obtain the fusion result. The pixel-level image fusion may be performed using different multi-scale decomposition techniques. Several researches adopted the Discrete Wavelet Transform (DWT) with its different types and different fusion rules. Other works dealt with the medical image fusion based on different transform techniques like NSCT, fuzzy transform, Non-Subsampled Shearlet Transform (NSST), curvelet transform, etc. Furthermore, guided-

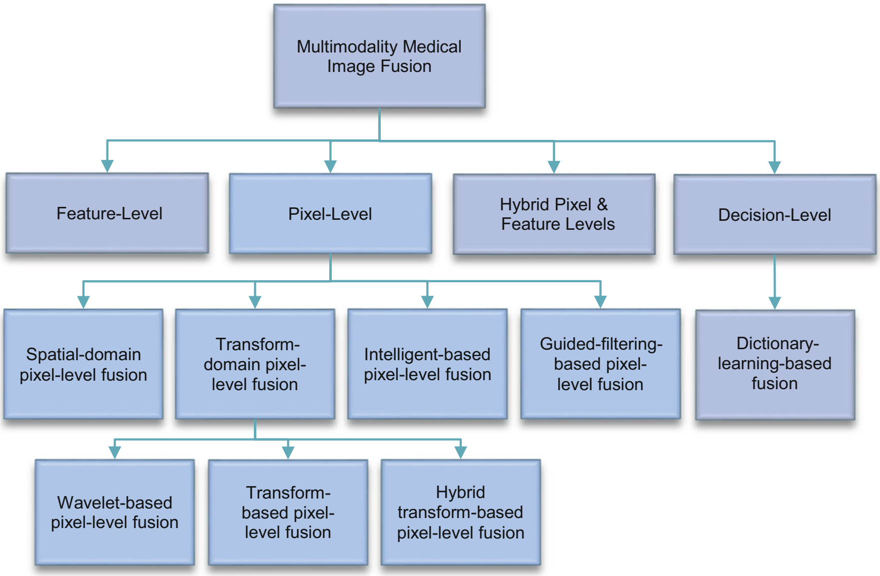

Fig. 3 Multimodality medical image fusion methods

filtering-based methods as well as intelligent-based methods have been addressed in some studies. In this section, we discuss some of these works.

2.1.1 Spatial-domain pixel-level fusion

The pixel-level fusion methods in the spatial domain directly manipulate the pixels of the input images to obtain the final fusion result. The average, maximum, minimum, PCA, IHS and Brovey are popular spatial-domain pixel-level fusion methods. In the average fusion method, the pixels of the fusion result represent the averages of the corresponding pixels of the input images [29]. In the maximum fusion method, the maximum pixel intensity values of the corresponding pixels for all input images are procured and used to represent the pixels of the output fusion result [53]. The minimum fusion method resembles the maximum fusion method, but the resultant pixels are the lowest corresponding pixel values of the input images to be fused [53]. The PCA is a statistical method that converts correlated variables into a set of uncorrelated variables (the principal components). This characteristic of the PCA is employed to implement image fusion [62]. The IHS is the most prevalent fusion method applied in remote sensing. In the IHS fusion method, the images with Red, Green and Blue (RGB) channels are converted into intensity, hue and saturation components. Thus, this method efficiently separates the spatial information (I) from the spectral information (H, S) of the standard RGB image. These components of the input images are fused and then transformed back to RGB to obtain the fusion result [18]. The Brovey method aims to reproduce a fusion result that merges a multi-spectral image and a panchromatic image. Each spectral band of the color image is normalized using all spectral bands, and then multiplied by the panchromatic image. This procedure annexes the spatial information to the output image, and keeps the spectral characteristics of each pixel [8].

He et al. [21] proposed a medical image fusion technique based on integrating IHS and PCA methods. The authors tried to consolidate the advantages of both the IHS and the PCA methods to enhance the output fusion result. The mutual information has been utilized to evaluate this system performance.

A comparative study of various methods for fusing medical images was presented in [44]. In this study, various spatial- and transform-domain pixel-level fusion methods have been investigated. The performance of these methods was assessed using quality metrics such as entropy, *PSNR*, *SNR*, and *SSIM*.

Generally, the direct image fusion methods depend on generating each pixel in the fusion result as the weighted average of the corresponding pixels in the source images. Spatial-domain pixel-level fusion methods have their advantages and limitations. The spatial-domain methods are fast and easy to implement. Nevertheless, they may create block artifacts on image boundaries. Moreover, they produce lower *SNR*s and several spatial distortions in the fusion results [41]. Besides, the fusion performance is quite limited, which may be unsuitable in some medical applications. For these reasons, and to overcome these drawbacks, researchers have carried out pixel-level medical image fusion in transform domains.

2.1.2 Transform-domain pixel-level fusion

In the transform-domain pixel-level fusion, the images are first turned into the transform domain, thereafter the fusion process is performed in this domain. Finally, an inverse transformation process is performed to produce the fusion result.

Wavelet-based pixel-level fusion The wavelet-based pixel-level fusion methods depend on applying the wavelet transform to attain the high- and low-frequency components that represent different perspective information of an image. Afterwards, different fusion rules are utilized to fuse the sub-images at different frequencies. Finally, the image is reconstructed by fusing these sub-images as shown in Fig. 4.

In [31], MR, and CT images are decomposed using the DWT. Then, pixel-by-pixel fusion is applied to merge the DWT coefficients. Finally, the Inverse Discrete Wavelet Transform

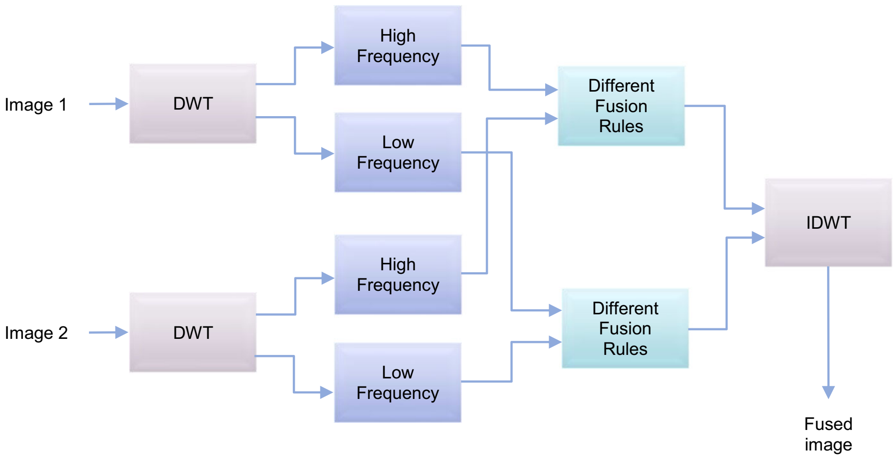

Fig. 4 Wavelet-based pixel-level fusion

(IDWT) is applied to obtain the fusion result. With the same previous steps, the authors presented another method by applying the Discrete Multi-Wavelet Transform (DMWT) for image fusion. The results illustrate that the DMWT method outperforms the DWT method in terms of *PSNR*, entropy, standard deviation and image quality.

Liu et al. [34] exhibited an efficient multi-modality medical image fusion method on the basis of compressive sensing for fusing CT and MR images. The multi-scale DWT is exploited to obtain significant sparse coefficients from the input images. The high-frequency coefficients are fused by applying an improved weighted fusion rule, while the PCNN fusion rule is used to combine the low-frequency coefficients. The Compressive Sampling Matched Pursuit (CoSaMP) algorithm and the IDWT are used for obtaining the fusion result. The authors employed the *AG*, *MI*, standard deviation, $Q^{AB/F}$, correlation and deviation to evaluate the fusion results.

Indira et al. [24] proposed a pixel-based medical image fusion method. In this method, multi-scale DWT and Stationary Wavelet Transform (SWT) are applied on the input images to obtain the detail and approximation coefficients. The approximation coefficients are fused using wavelet fusion. The average values of the pixels of the input images are computed; namely for particulars. The fused approximation coefficients and the particulars are concatenated forming a new coefficient matrix and the inverses of the DWT and the SWT are applied to obtain the final fusion result. The *RMSE*, *PSNR*, entropy, and correlation are the performance metrics used to assess the fusion performance.

Bhavana and Krishnappa [8] presented a method for pixel-level fusion of MR and PET images. In this method, both images are firstly pre-processed using spatial filtering with Gaussian filters. The enhanced images are, then, decomposed based on DWT to obtain the low- and high-frequency coefficients. The high-frequency coefficients are combined using the averaging method. The fused high-frequency output is obtained by performing the IDWT. In a similar way, the low-frequency coefficients are combined and the the IDWT is applied to get the low-frequency output. Finally, both outputs are combined to realize the fusion result. Performance indices of this method include *PSNR*, Mean Square Error (*MSE*), *AG*, and spectral discrepancy.

A multimodality medical image fusion method based on Discrete Fractional Wavelet Transform (DFRWT) was elucidated in [58]. The DFRWT is a fractional Fourier convolution procedure using various scales. The DFRWT technique decomposes the input images using different orders of the Fractional Fourier Transform (FRFT). The weighted regional variance fusion rule is used to fuse the whole coefficients of all sub-bands. Finally, the fusion result is acquired using the inverse DFRWT. The *MI*, $Q^{AB/F}$, and standard deviation are exploited as the fusion metrics.

The method suggested in [54] depends on principal component average fusion based on DWT for CT and MR images. Input images are decomposed to multi-scale images using DWT. Then, the PCA is applied on the approximation and detail coefficients. The average value of the principal components computed for all relevant decomposed elements create the weights used for the fusion rule. The performance of this method is evaluated by four metrics: Average Mutual Information (*AMI*), Average Quality Index (*AQI*), Average Peak Signal-to-Noise Ratio (*APSNR*) and *QHNC*, which is a metric based on *MI* and entropy.

The authors of [13] proposed an image fusion algorithm based on spatial- and transform-domain methods. The developed algorithm employs the IHS transformation in the wavelet domain. A pre-processing procedure based on discrete ripplet transform, rolling guidance filter, and Butterworth high-pass filter is performed on the input images before the fusion process, which improves the image fusion quality. The IHS transform is performed on the pre-processed images to obtain the intensity components. The high- and low-frequency coefficients for the intensity components are acquired using wavelet transform. The high- and low-frequency components of the images to be fused are merged using the wavelet fusion rule. The fusion result is obtained by applying the inverse IHS-wavelet transform. The amount of information transfer from the input images to the fusion result is measured using the objective evaluation metric $Q^{RS/F}$, which represents the image quality.

The wavelet transform has some drawbacks, including lack of phase information, shift sensitivity, and poor directionality. In order to overcome these weaknesses, the authors of [50] adopted the Daubechies Complex Wavelet Transform (D-CWT) to decompose the source images at multiple levels. Thereafter, the wavelet coefficients are fused using the maximum selection rule. Finally, Inverse D-CWT is applied to attain the fusion result. Edge strength, entropy, standard deviation, fusion symmetry, and fusion factor are used for evaluation. The obtained results proved that the utilization of the D-CWT for medical image fusion outperforms visually and quantitatively the wavelet-domain as well as spatial-domain fusion methods.

Transform-based pixel-level fusion A framework for multimodality medical image fusion based on the NSCT was reported in [5]. The NSCT decomposition technique is performed to acquire the low- and high-frequency sub-bands of the source images. Directive contrast and phase congruency fusion rules are applied to fuse the high- and low-frequency bands, respectively. The fusion result is finally reconstructed by performing the inverse NSCT. This method is evaluated using the normalized *MI*, *SSIM* and $Q^{AB/F}$.

Bhatnagar et al. [7] proposed a method for multimodality medical image fusion based on the NSCT. This method decomposes the source images into low- and high-frequency bands in the NSCT domain. The low-frequency bands are merged with an activity measure based on the normalized Shannon entropy, whereas the high-frequency bands are fused with the directive contrast. These fusion rules have the advantage of preserving more spatial features as well as functional information and transferring them to the fusion result. The fusion result is finally reconstructed by appling inverse NSCT on the obtained low- and high-frequency sub-images. This method was tested using two clinical bone images and four sets of the human brain. The objective performance evaluation metrics for this method include normalized *MI*, *SSIM* and $Q^{AB/F}$.

The authors of [46] offered a multimodality medical image fusion method based on local Laplacian energy and phase congruency. In this method, the NSCT is implemented on image pairs to decompose the input images into high-pass and low-pass sub-bands. The high-pass coefficients are fused by a phase-congruency-based fusion rule, whereas a local Laplacian energy-based fusion rule is applied for low-pass coefficient fusion. The inverse NSCT is implemented on the fused high-pass and low-pass sub-bands to attain the fusion result. Different parameters are estimated for evaluating the image quality, namely Q^{IE} and Q^{TE} which measure the nonlinear correlation information entropy and Tsallis entropy, $Q^{AB/F}$ and Q^{P} which evaluate the edge information, *MI* and Q^{Y} which estimate the similarity between the fusion result and the input images, and finally Q^{CB} and *VIF* which represent the human perception of the fusion result.

In [36], two multimodality medical image fusion methods that integrate the NSST and tensor structure into a united optimization model were suggested for fusing the CT and MR images. In the first method, the input images are combined utilizing weighted structure tensor fusion. In the second method, the NSST is used to decompose the source images to low- and high-frequency sub-bands. The average fusion rule is implemented for fusing the low-frequency sub-bands. On the other hand, the coefficients with the maximum sum of salient features are chosen to represent the fused high-frequency coefficients. A conjugate gradient optimization model is used to confirm that more detail and gradient information are injected into the final fusion result. Q_o, Q_W, edge-dependent fusion quality index (Q_E), and $Q^{AB/F}$ are utilized to evaluate the fusion results.

The authors of [52] introduced a fusion method based on curvelet transform. In the first step of this method, the input images are transformed into curvelet coefficients. Then, these coefficients are fused using local energy fusion rule. Finally, the inverse curvelet transform is applied to obtain the final fusion result. This method is assessed using entropy, standard deviation, $Q^{AB/F}$, sharpness and AG.

In [48], a medical image fusion method based on curvelet transform was presented. Each source image is first transformed into its curvelet coefficients. Afterwards, these coefficients are fused using PCA fusion rule. Eventually, the inverse curvelet transform is implemented to obtain the final fusion result. The fusion result is assessed using entropy, *RMSE*, and *PSNR*.

Manchanda et al. [41] presented a method of multimodality medical image fusion that depends on fuzzy transform. Firstly, the images are divided into blocks with the same size. Then, the fuzzy transform is applied to transform these blocks into different-size sub-blocks. Using the maximum-entropy fusion rule, these sub-blocks are fused. The inverse fuzzy transform is performed on the fused sub-blocks. The performance of this fusion method is assessed both subjectively and objectively. Objective measures include entropy, $Q^{AB/F}$, fusion artifacts, standard deviation, *MI*, fusion symmetry, *SSIM*, fusion loss, and *FSIM* index.

Hybrid transform-based pixel-level fusion The coefficients procured from the transform-based pixel-level fusion methods can be processed with further fusion methods to improve the validity of the pixel-level fusion methods based on discrete transforms.

In [15], several hybrid medical image fusion methods comprising Additive Wavelet Transform (AWT), curvelet, PCA, NSCT, and Dual-Tree Complex Wavelet Transform (DT-CWT) with high-pass sharpening were presented. The hybrid fusion methods are implemented and assessed on CT and MR images. Six combinations are deemed in this work, including DWT with PCA, DWT with curvelet transform, AWT with PCA, AWT with DT-CWT, AWT with NSCT, and NSCT with DT-CWT. By employing these different transform-domain methods, the input images are transformed into multi-scale or multi-resolution image representations. Afterwards, two different fusion rules are used: average fusion and maximum fusion to fuse low-frequency sub-bands and high-frequency sub-bands, respectively. Ultimately, an inverse transform is exploited to obtain the fusion result. The hybrid and traditional methods are assessed using different quality metrics comprising entropy, AG, local contrast, standard deviation, edge intensity, *SSIM*, Q_o, *FSIM* index, *PSNR*, *MI*, processing time, and $Q^{AB/F}$. Results demonstrate that the hybrid method of AWT and DT-CWT with a high-pass sharpening filter gives the best fusion results.

The work submitted in [2] introduced a two-stage multimodality fusion method that applies SWT and NSCT on the source images (i.e., MR and CT images). In the first stage, the SWT is used to decompose the input images into detail and approximation coefficients. Then, a PCA-based fusion rule is used to fuse these coefficients and the inverse SWT is exploited to reconstruct the fusion results. In the second stage, the NSCT decomposition technique is applied on the reconstructed images and the obtained coefficients are merged using the maximum fusion rule. The fusion result is obtained by applying the inverse NSCT. This method is assessed with Image Quality Assessment (*IQA*), *PSNR*, *SSIM*, entropy, standard deviation and Fusion Factor (*FF*).

The authors of [3] proposed a hybrid method of curvelet transform and PCA for medical image fusion. The curvelet transform is applied on the source images for decomposition. Then, the PCA technique is used for dimensionality reduction to keep the features that have significant contribution. With the maximum selection fusion rule, the coefficients are merged and the final image is created by performing the inverse curvelet transform. This work was implemented on CT and MR datasets and evaluated using different objective parameters, namely *FF* and *SSIM*.

In [60], a multimodality medical image fusion method that depends on both the NSCT and fuzzy logic techniques was presented. Firstly, the NSCT is utilized to get the high- and low-frequency coefficients. Then, a fuzzy logic fusion rule is used for the fusion of the high-frequency coefficients, while a local energy algorithm is used for the fusion of the low-frequency coefficients. Finally, the fusion result is reconstructed by performing the inverse NSCT. Both subjective and objective assessment indices are used to evaluate the fusion results. The objective indices include *MI*, $Q^{AB/F}$, and standard deviation.

A fusion method for the CT and MR medical images was presented in [51]. Initially, the source images are divided with the NSST into a single low-frequency sub-band and multiple high-frequency sub-bands. The regional energy is measured for the low-frequency component, and then a maximum selection rule is implemented to select the coefficients with the highest activity measure. The sum-modified Laplacian is computed for the high-frequency components. The sum-modified Laplacian is used to motivate the PCNN for high-frequency component fusion. Finally, the fusion result is obtained with the inverse NSST. Entropy, standard deviation, *SF*, *MI*, Image Quality Index (*IQI*) as well as $Q^{AB/F}$ are used as evaluation indices.

A method for multimodality medical image fusion based on a hybrid neuro-fuzzy strategy in the NSCT domain was presented in [11]. The source images are initially divided by applying the NSCT. Then, the linking strengths are computed for both the low- and high-frequency sub-bands based on fuzzy logic. Afterwards, the coefficients and the linking strengths are delivered as inputs to motivate the Reduced PCNNs (RPCNNs). The fused coefficients are defined using the RPCNNs. Eventually, the fusion result is acquired by implementing the inverse NSCT. Entropy, standard deviation and *SF* are used as performance metrics.

In [56], a medical image fusion method based on PCNN and sparse representation was presented. The NSCT is applied to obtain the low- and high-frequency coefficients of the source images. The low-frequency sub-bands are fused based on a sparse representation, while the high-frequency sub-bands are fused based on the PCNN model. The K-means generalized Singular Value Decomposition (K-SVD) method is performed for training with the low-frequency sub-bands to attain the over-complete dictionary, and then the Orthogonal Matching Pursuit (OMP) procedure is utilized to obtain the sparse coefficients. On the other hand, the spatial components of the high-frequency sub-bands are used to stimulate the PCNN, and the fusion coefficients are chosen depending on the count of ignition times. Finally, the inverse NSCT is applied to get the fusion result. Information Entropy (*IE*), *SF*, *AG*, clarity (*MC*), *MI*, standard deviation and $Q^{AB/F}$ are the parameters utilized for performance evaluation.

The study in [61] presented a multimodality medical image fusion method in the NSST domain. In this method, the multi-direction and multi-scale representations of the source images are obtained using the NSST decomposition. A parameter adaptive PCNN fuses the high-frequency sub-bands, whereas, the low-frequency sub-bands are combined by a method that addresses two important processes, namely detail extraction and energy preservation. To perform these processes, activity level measurement of the low-frequency band is used. Eventually, the final fusion result is obtained by applying the inverse NSST. Entropy, standard deviation, localized *MI*, Q_W, and the Visual Information Fidelity Fusion (*VIFF*) metric are used for image quality assessment.

Ramlal et al. [47] presented a medical image fusion method depending on a PCNN model and the NSST. The source images are firstly decomposed to detail and approximation coefficients by the NSST. The approximation sub-bands are merged using the regional energy activity measure. For detail sub-bands, a morphological gradient is applied, and then its output is fed as an external exciter to the PCNN to combine the detail sub-bands. Eventually, the inverse NSST is implemented on the fused low- and high-frequency sub-bands to acquire the fusion result. This method is evaluated with *MI*, $Q^{AB/F}$, *SF*, standard deviation and mean.

In [26], a method for PET, MR, and CT image fusion into a single image was proposed. In this method, the NSST and simplified PCNNs are used in the Hue, Saturation and Value (HSV) space. Initially, the NSST decomposition is applied on the MR and CT images to obtain the low- and high-frequency components. After that, the outlines and edges in a larger area of the high-frequency coefficients are extracted using Intersecting Cortical Models (ICMs), while in a smaller area, the detail information is described by S-PCNNs. Finally, diversity fusion rules are performed to fuse the low- and high-frequency coefficients of CT and MR images. For low-frequency components, the maximum absolute rule is utilized for fusion. The high-frequency coefficients are fed into the ICM and the S-PCNNs to obtain their Oscillation Frequency Graph (OFG). Thereafter, the high-frequency coefficients are merged using the maximum OFG. For the PET image, the HSV space transformation is applied, and then the V component is obtained. The NSST is performed on the V component of the PET image for decomposition into low- and high-frequency components. The same fusion process is applied for the low- and high-frequency components of the PET image. Finally, the fusion result is extracted by applying the inverse HSV and the NSST. This method is assessed using *AG*, *SF*, standard deviation and mean.

In summary, the transform-domain pixel-level fusion methods are applied to overcome the drawbacks of spatial-domain fusion methods. However, each transform has its merits and shortcomings. The DWT suffers from aliasing, shift variance, and lack of directionality [63]. The DT-CWT is implemented to avoid these drawbacks [32] due to its directional selectivity and shift-invariance. However, wavelet families cannot well represent the edges and curves of images. Therefore, some multi-scale methods are presented into image fusion to capture accurately the spatial structures of the images such as contourlet fusion [12]. For contourlet fusion implementation, the Laplacian pyramid is utilized first to represent the point discontinuities. After that, point discontinuities are linked into linear structures using a directional filter bank. Owing to the successful representation of spatial structures, the contourlet transform is successfully applied in medical imaging. However, it has no shift-invariant property, due to the down-sampling processes in the transform domain. The NSCT solves this problem, but it is time consuming. Recently, the shearlet transform has been applied more efficiently compared to the contourlet transform. In addition, there are no restrictions on the size of the supports, and the number of directions for shearing [55].

In the transform-domain fusion methods, the choice of the multi-scale decomposition technique is an important issue, in addition to the determination of the of decomposition levels. The number of levels affects the performance of the fusion process. If few levels are adopted, the spatial quality of the resultant image may be limited. On the other hand, if excessive decomposition levels are adopted, the efficiency of the fusion method may be reduced. Consequently, some researchers are directed to specify the optimum number of decomposition levels that enhances the quality of the fusion result. Furthermore, to improve the overall performance of the fusion methods, some researchers tried to combine different transforms in the fusion process. The decomposition techniques, as mentioned previously, have their advantages as well as limitations. To integrate the advantages of different transforms, different fusion methods have been proposed based on hybrid transforms.

2.1.3 Guided-filtering-based pixel-level fusion

The guided filter is considered as an edge-preserving smoothing filter with a superior behaviour near edges. Furthermore, it is capable of transferring the texture structures of the guided image to the output of the filter. Therefore, it has shown an efficient and effective performance in numerous computer vision applications including denoising, detail enhancement, dehazing, image matting, etc. [22].

The work in [23] presented an image fusion method that depends on guided filtering and pixel screening. At first, the source images are merged by weighted fusion and delivered as input images to the guided filtering stage. Then, the output of the filtering process is compressed using Dynamic Range Compression (DRC) to clarify edge details. Eventually, a pixel screening process is performed to complete the texture structure of the fusion result. The results in this study are evaluated using *MI*.

The authors of [49] concentrated on the pixel-level medical image fusion. They presented a medical image fusion method that depends on integrating a Spiking Cortical Model (SCM) and a Rolling Guidance Filter (RGF). Initially, the saliency (edge information) of the source images is extracted by the RGF. Then, a self-adaptive threshold of the SCM is captured by using the mean and variance of the input images. Finally, the fusion result is obtained by the SCM stimulated by the RGF coefficients. The performance of this method is evaluated by some objective parameters such as *MI*, $Q^{AB/F}$, $L^{AB/F}$; which is a metric that represents the information lost through the fusion procedure, and $N^{AB/F}$; which is a metric that denotes the artifacts fed into the fusion result.

The authors of [42] presented a fusion method that combines DWT, NSCT, and guided filter techniques. This method includes two stages. In the first stage, the DWT is applied on the source images, and then the detail and approximation coefficients are obtained. Afterwards, the approximation coefficients are combined using the maximum selection fusion rule, while the detail coefficients are merged using the average fusion rule. In the second stage, the NSCT decomposition technique is applied on the fused approximation and detail coefficients. After that, the obtained low-frequency coefficients are fused using the phase congruency fusion technique, owing to its insensitivity to illumination contrast. Then, the guided filter, due to its edge-preserving property, is applied as a fusion technique to fuse the high-frequency coefficients. Finally, the inverse NSCT is implemented to obtain the final fusion result. The quality of the resultant image is assessed by different objective metrics including entropy, *SSIM*, *MI*, and Cvejic's metric (Q_C), which determines the amount of distortion in the fusion result.

In recent years, edge-preserving filtering has become an active research area in image processing, especially in image fusion. Edge-preserving filters have been accurately implemented as multi-scale representation tools for image fusion. Generally, the main property of these methods is their ability to split successfully large-scale spatial structures, middle-scale edges, and fine-scale texture details of the images. In addition, they definitely retain the details of the original images. This advantage leads to decreasing the aliasing and halo artifacts in the fusion results, and consequently, the fusion performance is improved for human visual perception. However, the computation times of these methods are based on the filter size.

2.1.4 Intelligent pixel-level fusion

The application of computational intelligence techniques in the area of medical image fusion has been manifested as an active topic in the last years. Soft computing is the soul of image fusion applications based on intelligent computation techniques. Different soft computing techniques like Particle Swarm Optimization (PSO) [59], Neural Network (NN) [35], HVS [4, 6], fuzzy logic [27] and Grey Wolf Optimization (GWO) [9, 10] were implemented in pixel-level image fusion applications.

The authors of [59] introduced a method to fuse multimodality medical images based on an adaptive PCNN, which is optimized by the Quantum Particle Swarm Optimization algorithm (QPSO). In this method, the source images are processed using the QPSO-PCNN model. The PCNN model can find the optimal parameters of the source images through the QPSO algorithm. Then, the fusion result is obtained by a factor based on the firing maps of the two source images. The *MI*, *SSIM*, and image entropy are exploited as performance indices.

The work in [35] presented a multimodality medical image fusion method that incorporates a Gradient Minimization Smoothing Filter (GMSF) and a PCNN. The GMSF is utilized to decompose the source images into a series of detail images and single base images. The regional weighted sum of gradient energy and pixel energy are utilized to merge the base images. The detail images, on the other hand, are fused using the PCNN. The final fusion result is obtained by combining the fused approximation and the detail images. This method has been implemented on different datasets of CT and MR images. Five evaluation metrics including *MI*, Q_o, Q_W, Q_E, and $Q^{AB/F}$ are utilized to assess the quality of the fusion results of this method.

Medical image fusion methods that depend on the HVS were presented in [4, 6]. Framelet transform is utilized to decompose the original images into low- and high-frequency sub-bands. The low-frequency sub-bands are fused by implementing the adaptive weighted average local energy, whereas the high-frequency sub-bands are fused based on the texture attributes of the HVS model. In another method, two different fusion rules inspired by the HVS are implemented to fuse the low- and high-frequency coefficients, respectively. The premier coefficients are fused by conceiving the visibility as a fusion metric, and the latter coefficients are fused by perceiving the texture information. Mean, standard deviation, entropy, *SF*, *MI*, *SSIM*, saliency-based similarity measure and $Q^{AB/F}$ are used to assess these methods.

The application of fuzzy logic for multimodality medical image fusion was addressed in [27]. This method is based on a fuzzy inference system in which fuzzy rules stem from the pixel intensities of the image. Nine inference rules are applied in this method. The inference engine is based on the min-max method and the defuzzification relies on the centroid method. The *PSNR*, *MSE*, *SNR* are the parameters used for fusion result quality assessment in this method.

In [9], the authors depicted an Optimum Spectrum Mask (OSM) for fusion of medical images. In this method, the conventional GWO procedure is applied. At first, the input images are transformed into the frequency domain. Then, the GWO algorithm is used to obtain the optimum scale values. These values are combined using pixel-based average fusion rule, and the final fusion result is obtained by performing the inverse Fourier transform. This method has been evaluated for MR-CT, MRI-SPECT, MRI: T1-T2, and MR-PET brain images. *MI*, Entropy, standard deviation and $Q^{AB/F}$ are used for fusion evaluation.

An optimal complementary Laplacian and wavelet masking medical image fusion method was submitted in [10]. The wavelet and Laplacian filters are applied to filter the original images. Afterwards, a Hybrid Cuckoo Search Grey Wolf Optimization (HCS-GWO) model is implemented to select the optimal scale values for the Laplacian filtered source images and the approximation coefficients of the wavelet transformed source images. The pixel-level average fusion rule is applied to fuse the Laplacian and wavelet masks of the input images. Four fusion performance metrics, namely, entropy, *MI*, $Q^{AB/F}$, and standard deviation are utilized for quality assessment.

Recently, researchers have designed fusion methods built on different optimization techniques to improve the fusion results. However, optimization-based pixel-level fusion has a common shortcoming. The fusion performance is inefficient, since these methods need multiple iterations to obtain the global optimal solution. Furthermore, the optimization-based fusion methods have another drawback, which is over-smoothing of the output fusion results.

2.2 Feature-level fusion

The feature-level fusion methods aim to deal with the image features since the features and objects within a scene are more valuable than the individual pixels. In the feature-level fusion, the features are extracted separately from each source image, and then the fusion method is applied based on the features from the source images [25] .

The study in [16] offered a feature-based fusion method for multimodality medical images using sparse representation and feature extraction techniques. Firstly, the source images are divided into patches, and then the patches are arranged into vectors according to the positions of the patches in the original images, and a decision map is designed. Thirdly, the decision map is utilized to decide which vector of each group is selected. Fourthly, the other vector pairs are fused by the sparse representation method. Finally, the method produces the fusion result based on the decision map. The average is calculated for the overlapping patches. Five objective evaluation parameters are used to assess the fusion performance, including Q_o, Q_W, Q_E, $Q^{AB/F}$, and *MI*.

The authors in [38] introduced an image fusion method using the QPSO algorithm for clustering. Initially, the wavelet transform decomposition technique is applied on the input images. The source images in the feature space are segmented with a clustering algorithm, which consolidates the QPSO algorithm with Fuzzy C-Means (FCM) in the feature space that is composed by a multi-channel Gabor filter, thereafter the weighting factors are created. Eventually, the fusion result is obtained by applying the inverse wavelet transform. The performance of this image fusion method is assessed using five metrics, namely entropy, peak-to-peak signal-to-noise ratio, cross entropy, *MI*, and *RMSE*.

The authors of [17] presented an image fusion method that depends on a feature-stimulated adaptive PCNN in the NSST domain. The NSST is performed on the source images to attain low- and high-frequency sub-bands. The low-frequency sub-bands of the input images are combined to obtain the fused low-frequency sub-bands utilizing a PCNN-based fusion algorithm. These low-frequency sub-bands represent a smooth description of the image that comprises most of the energy. The total variation of the coefficient squares is provided as an external stimulus for PCNN processing of the low-frequency sub-bands. As well, the high-frequency sub-bands of the input images are merged using a PCNN-based fusion algorithm. Since the high-frequency sub-bands contain the detail information of the images, the absolute values of the coefficients in these sub-bands are fed to the PCNN. Finally, the inverse NSST of the consolidated sub-bands offers the fusion result. The *MI*, standard deviation, *SF*, $Q^{AB/F}$, Q_W, and Q_E are the evaluation parameters used to assess this method.

In [20], a feature-based medical image fusion method for MR and SPECT images was presented. First, the input images are transformed into coefficients using Hadamard transform. The obtained coefficients are divided into equal-size blocks. For each block, statistical feature values including standard deviation, entropy, visibility, SF and correlation are calculated to perform the fusion process. The fusion rule is performed depending on the selection of the maximum of the obtained features. Finally, the obtained features are combined together, and the inverse Hadamard transform is implemented to get the fusion result. The *MI* and $Q^{AB/F}$ are used for assessing this method.

A feature-level fusion method based on Uniform Discrete Curvelet Transform (UDCT) was addressed in [57]. The UDCT decomposition is applied to get the sub-band coefficients of the source images. The low-pass sub-band coefficients are fused using a fusion rule based on the *FSIM* index. A Complex Coefficient Feature Similarity (*CCFSIM*) index is used to fuse the high-pass sub-bands. The inverse UDCT is finally applied on the fused coefficients to attain the final fusion result. The *MI*, entropy and $Q^{AB/F}$ metrics are harnessed for performance evaluation.

The DT-CWT is exploited to achieve feature-level image fusion in [40] . The DT-CWT and the watershed transform are applied to extract the features of the images separately or jointly to create a region map. The approximation components are merged by the average fusion rule. The detail component fusion is performed using the maximum selection fusion rule. Eventually, the fusion result is obtained using the inverse DT-CWT. This method is assessed using entropy, standard deviation, *SF*, cross entropy, and *MI* metrics.

Feature-level image fusion methods that are based on lifting wavelet transform have been presented in [30, 39]. The authors in [30] used Daubechies wavelets, while in [39] they used Haar wavelet. The method in [30] has five steps. Firstly, the source images are decomposed using the lifting wavelet transform. Secondly, the coefficient gradients of the lifting wavelet transform are computed. Thirdly, the fusion process is performed based on the thresholded values of the obtained gradients of coefficients. Then, the inverse lifting wavelet transform is employed to acquire the fusion result. Finally, the image fusion performance is evaluated by gradient parameters, standard deviation, entropy, and cross entropy. On the other hand, in [39] the Haar lifting wavelet transform was exploited to decompose the input images. Then, the edge and boundary features are extracted using the maximum criteria for the wavelet transform modulus function. Eventually, the fusion result is reconstructed by performing the inverse Haar lifting wavelet transform. The performance of this fusion method is assessed using standard deviation, entropy and gradient metrics.

Feature-level fusion methods depend on a lower-level content of image features. These methods improve the quality of the fusion results in terms of average information content, detail information, and image contrast. Furthermore, the feature-based fusion methods increase the reliability of the fusion process, decrease the loss of information and avoid the artifacts. However, the fusion process at this level is hard to implement since features acquired from different imaging modalities may be heterogeneous.

2.3 Hybrid Pixel- and Feature-Level Fusion

The hybrid methods based on pixel- and feature-level image fusion merge the advantages of both of them. The hybrid methods avoid the pixel-level drawbacks such as high sensitivity to noise and blurring effects. Furthermore, they consider the information content correlated with each pixel.

The authors of [28] presented a method that depends on integrating both pixel- and feature-level image fusion methods. Images are firstly divided into detail and approximation sub-bands using Discrete Wavelet Frame Transform (DWFT). Then, the fusion process is implemented with two phases: pixel- and feature-level fusion. Initially, for the feature-level fusion, the approximation sub-bands are segmented into three regions using an FCM clustering method. Afterwards, they are fed into a fusion scheme, which relies on a fuzzy rule to obtain the fused approximation. For pixel-level fusion, the absolute maximum rule is implemented to fuse the detail sub-bands. Finally, the inverse DWFT is applied to merge the details and the approximation to extract the fusion result. The *IQI* is used as an evaluation metric for this method.

The authors of [33] presented a medical image fusion method that combines the pixel- and feature-level fusion. The source images are divided into several regions. Two image segmentation algorithms: edge detection and region growing are applied. Afterwards, the segmentation outputs and the source images are combined to define the regions each pixel belongs to. The features, namely salience and visibility, are extracted from each image region to obtain the feature vector. The fusion weight of each region of the image is determined based on the feature vector. Finally, according to the fusion weights, the fusion result is obtained. The performance of this method is evaluated using the *SF* metric.

2.4 Decision-level fusion

Decision-level fusion aims to merge higher-level aggregation of the results from several algorithms to get the final decision for the fusion process. Each image to be fused is first handled individually and then supplied to the fusion algorithm. The dictionary learning technique is commonly used for decision-level medical image fusion.

2.4.1 Dictionary-learning-based fusion

The dictionary-learning-based fusion methods have a large importance in the multimodality medical image fusion, for their good performance. The dictionary learning methods are applied to improve the detail information of the fusion result.

An image fusion method based on dictionary learning was given in [64]. This method includes three steps. Firstly, image patch sampling is performed to get informative patches. Then, a local density peak clustering procedure is applied to divide the image patches into different groups of patches with similar structural information. The K-SVD is applied to classify each patch group into a compact sub-dictionary. Finally, the sub-dictionaries are merged to get a compact, informative, and complete dictionary. This

Table 1 Summary of medical image fusion methods, procedures, modalities, and applications

Methods	Procedures	Modalities	Applications
Spatial-domain pixel-level fusion	Average, maximum, minimum, PCA, IHS and Brovey [21]	MR, PET	Medical diagnostics
Wavelet-domain pixel-level fusion	Wavelet-based pixel-level fusion	MR, CT	Enhancement of the quality of images
	DWT [8] DWT and SWT [24] DFRWT [58]	MR, PET CT, PET MR, CT MR, PET MR-T1, MR-T2	Diagnosis for a physician in the field of biomedical applications
	IHS-wavelet transform [13]	Osseous and vascular data	
	D-CWT [50]	MR-T1 and MRA MR, CT	
Transform-based pixel-level fusion	NSCT [5, 7, 46]	MR, CT MR-T1, MR-T2 MR, PET MR, SPECT	Medical diagnosis and treatment
	Structure tensor, NSST [36] and Curvelet transform [48]	MR, CT	
	Curvelet transform [52]	MR, MRA MR-T1, MR-T2 MR, CT	
	Fuzzy transform [41]	MR, PET MR, CT MR-T1, MR-T2	
Hybrid transform-based pixel-level fusion	AWT, curvelet, NSCT, and DT-CWT [15], Curvelet and PCA [3], and NSST and PCNN [51]	MR, CT	Clinical diagnosis
	SWT and NSCT [2]	MR-T1, CT MR-T2, CT MR-T1, MR-T2	
	NSCT and fuzzy logic [60]	MR-T1, MRA MR-T1, MR-T2 MR-Gad, MR-T2 MR, PET MR, SPECT	
	NSCT, neuro-fuzzy, and RPCNN [11]	CT, MRI MR-T1, MRA PET, MR PET, CT	
	NSCT with SR and PCNN [56]	MR, CT MR-T1, MR-T2 MR-PD/MR-T1 MR-PD/MR-T2	
	NSST and PCNN [47, 61]	MR, CT MR-T1, MR-T2 MR, PET MR, SPECT	
	NSST and SPCNN [26]	MR, CT, and PET	
Guided-filtering-based pixel-level fusion	Guided filtering and DRC [23] RGF and SCM [49]	MR, CT	Clinical diagnosis

Table 1 (continued)

Methods	Procedures	Modalities	Applications
		MR-T1, MR-T2 Ultrasound, SPECT MR, CT	
Intelligent-based pixel-level fusion	NSCT and Guided filter [42]	MR, CT	Clinical applications
	QPSO-PCNN [59]	MR, CT CT, MRI-T2 MR-T1, MRI, T2 MR, SPET MR-T1, FDG	
	GMSF and PCNN [35]	MR, CT CT, MR-T2	
	Framelet transform and HVS model [4], and Min-max and defuzzification with centroid [27]	CT, MR	
	Framelet transform and HVS model [6]	CT, MR MR, PET	
	Fourier transform with GWO [9] and wavelet with HCS-GWO [10]	MR, SPECT MR, CT MR, PET MR-T1, MR-T2	
Feature-level fusion	SR and decision map [16]	MR-T1, MR-T2 CT, MR CT, PET	Clinical applications.
	Wavelet transform with QPSO-FCM [38], UDCT [57], DT-CWT and watershed transform [40],	MR, CT	
	NSST and PCNN [17]	SPECT, MRI MR, CT	
	Hadamard transform [20]	SPET, MR	
	Lifting wavelet transform [30, 39]	MR-T1, MR-T2	
Hybrid Pixel- and feature-level fusion	DWFT and fuzzy c-means clustering [28], Region growing and edge detection segmentation [33]	CT, MR	Improving the fusion result quality
Dictionary-learning-based decision-level fusion	Image patch sampling and local density peak clustering [64]	MR, MRA MR, PET MR, CT	Improving the fusion result quality
	SR and ODL [65]	MR, SPECT CT, MR CT, PET	
	Gaussian filter with integrated dictionary learning [45]	CT, MR MR, PET MR, SPECT	
	MSF, OMP, and SR [1]	MR, CT MR-T1, MR-T2 MR, PET MR, SPECT	
	CS-MCA [37]	MR-T2, CT MR-T1, MR-T2	

Table 2 Advantages and disadvantages of the state-of-the-art methods

Methods	Advantages	Disadvantages
Spatial-domain pixel-level fusion [29, 53, 62, 18, 19, 21]	• Preserves more spatial features and functional information • Has low complexity and fast processing	• Generates block artifacts on boundaries • Fusion performance is quite limited. • Causes color distortion and spectral degradation
Wavelet-domain pixel-level fusion [31, 34, 24, 8, 58, 54, 13, 50]	• Reduces the color distortion problem and retains the statistical parameters. • Minimizes distortion of the spectral information and denoising effects	• The shift variance problem, lack of directionality, and aliasing • Cannot represent the curves and edges of images. • Less spatial resolution for the final fusion results
Transform-based pixel-level fusion [5, 7, 46, 36, 52, 48, 41]	• Represents the spatial structures in the images more accurately • Captures the intrinsic geometrical structure of the images • Represents well the curves and edges of images.	• Higher complexity in the computation and setting of parameters • More time consuming • The number of decomposition levels influences the performance of the fusion process. If few decomposition levels are applied, the spatial quality of the resultant image may be less satisfactory. If excessive levels are utilized, the performance and computational efficiency of the fusion method deteriorate.
Hybrid transform-based pixel-level fusion [15, 2, 3, 60, 51, 11, 56, 61, 47, 26]	• Experiments prove that the hybrid methods perform better than the individual ones • Avoids the weak points of a single fusion method • Combines the advantages of both transform-domain methods	• More time consuming • Complexity increases.
Guided-filtering-based pixel-level fusion [23, 49, 42]	• Their ability to accurately separate large-scale spatial structures, fine-scale texture details, and middle-scale edges of an image increases. • Assists to decrease halo and aliasing artifacts in the fusion result • The quality of the output fusion result is good for human visual perception	• Time consumption is high
Intelligent-based pixel-level fusion [59, 35, 4, 6, 27, 9, 10]	• More efficient than traditional methods	• Time consumption is high
Feature-level fusion [16, 38, 17, 20, 57, 40, 30, 39]	• Deals with the image features since the features and objects within a scene are more valuable than individual pixels • Decreases the loss of information • Avoids the artifacts • Improves the reliability of the fusion process	• The fusion process at this level is hard to implement since features acquired from different imaging modalities may be heterogeneous.
Hybrid pixel- and feature-level fusion [28, 33]	• Merges the advantages of the pixel-level and the feature-level fusion • The hybrid methods avoid the pixel-level drawbacks such as high sensitivity to noise and blurring effects. • Enhances the feature-level fusion by comprising the information content correlated with each pixel.	• More time consumption • Complexity of the method increases.
Dictionary-learning-based decision-level fusion [64, 65, 45, 1, 37]	• Reduces the redundancy and uncertain information.	• The time consumption and complexity of the method increase

method is evaluated using entropy, Edge Intensity (*EI*), $Q^{AB/F}$, *MI*, and Visual Information Fidelity (*VIF*).

A medical image fusion method based on obtaining sparse representation versions of classified image patches was presented in [65]. The source images are first decomposed into classified patches based on the geometrical direction of the patch. Then, the Online Dictionary Learning (ODL) algorithm and the Least Angle Regression algorithm (LARS) are applied to code every patch, sparsely. Afterwards, the maximum fusion rule is employed to fuse these sparse coefficients. Eventually, the fusion result is extracted from the corresponding sub-dictionary and the fused sparse coefficients. Six metrics including *SF*, standard deviation, *MI*, Xydeas's fusion metric and $Q^{AB/F}$ are used for evaluation.

The authors of [45] presented a medical image fusion method based on integrated dictionary learning. This method comprises three steps. Initially, a Gaussian filter is used. The source images are decomposed into low- and high-frequency coefficients. The weighted averaging algorithm is applied for merging low-frequency coefficients. The high-frequency components are combined by dictionary learning. The final image is obtained by merging the coefficients of low-frequency and high-frequency components. The quality of the fusion result is assessed by different parameters, namely the *AG*, $Q^{AB/F}$, *MI*, and *VIF*.

The authors of [1] introduced an adaptive dictionary learning method for medical image fusion. It has three steps. Firstly, the variance computation is used to discard zero informative patches of the input images. Secondly, by using the Modified Spatial Frequency (MSF), the structural information of the remaining patches is assessed. Eventually, a fusion rule is applied to select the more informative patches of the input images for performing the dictionary learning. In the fusion process, the OMP algorithm is used to evaluate the sparse coefficients. The fusion result is reconstructed with a fusion rule, which measures the activity level of both the transform and spatial domains, with the trained dictionary and the sparse vectors. Seven metrics including $Q^{AB/F}$, normalized *MI*, local quality index Q, Q_w, Q_e, Chen-Blum metric Q_{CB}, and Mean Gradient (*MG*) are used to measure the performance of this method.

A Convolutional Sparsity-based Morphological Component Analysis (CS-MCA) method for medical image fusion was introduced in [37]. Merging MCA with Convolutional Sparse Representation (CSR) into an integrated optimization model allows the CS-MCA model to jointly realize global as well as multi-component sparse representations of input images. For each image, the CS-MCA scheme with pre-learned dictionaries is applied to attain the sparse representations of the texture and cartoon components. Then, for all source images, the sparse coefficients are fused and the fused coefficients are recreated employing the corresponding dictionary. Eventually, the fusion result is obtained as the superposition of the combined texture and cartoon components. The objective assessment utilized to appraise the fusion result quality includes localized *MI*, $Q^{AB/F}$, Q_E, and Q_{HVS}.

Table 1 demonstrates the summary of the state-of-the-art medical image fusion methods, the addressed procedures as well as the utilized modalities, and their applications in medical diagnosis. Additionally, Table 2 addresses the advantages and disadvantages of the main medical image fusion methods. Moreover, Fig. 5 illustrates the outlines of the main stages involved in the multimodality medical image fusion methods.

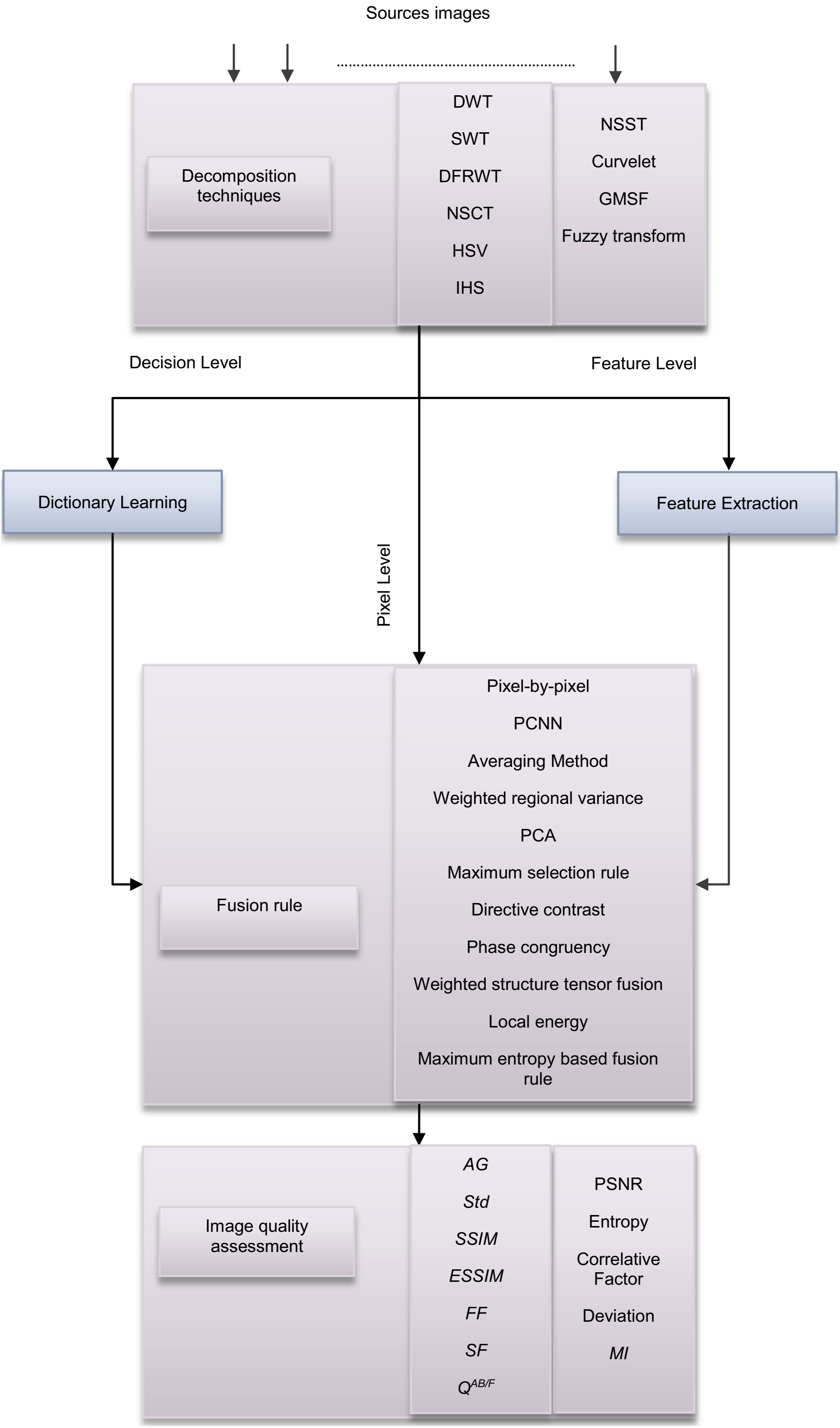

Fig. 5 Summary of different stages in the multimodality medical image fusion

3 Conclusion

Multimodality medical image fusion methods are adopted to merge information from a diversity of images to acquire a more informative image. This article provided an inclusive survey of the state-of-art research directions in the multimodality medical image fusion domain. A review of the literature shows that most of the researchers are directed towards transform-domain pixel-level fusion methods. Furthermore, it is also perceived that combining more than one pixel-level fusion method is effective in medical image fusion. Each step of the image fusion process is a challenge for the researchers to improve the overall performance of the fusion process. One of these challenges is that the appropriate image decomposition is required to extract features at different scales and orientations from the images. Besides, the most common image fusion rules that are simple to implement in a short time cause pseudo-Gibbs and obscure edges. To deal with this problem, effective fusion rules such as PCNN, self-generating neural network, and visual cortex model can be utilized. Other challenges in the image fusion field are the lack of sufficient features in each medical image modality, the image noise, the increased computational complexity, and the difference in dimensions between images in different modalities. Moreover, accurate registration is also a challenging issue since the medical images are usually acquired with different imaging modalities. Hence, it is believed that pioneering ideas and innovative research contributions will endure development of medical image fusion in the coming years.

A recommended future work may be directed towards enhancing the hybrid pixel- and feature-level fusion methods for multimodality medical images. Furthermore, the decision-level medical image fusion trend needs further extensive research and exploration. Another point of interest is the utilization of evolving techniques that enhance the regions of interest in images prior to the fusion process. The limitations of the imaging modalities and the noise increase the need for developing algorithms to improve the image quality of the fusion results. Developing novel multi-scale decomposition techniques for multimodality image fusion is a trend that deserves consideration.

References

1. Aishwarya N, Bennila Thangammal C (2018) A novel multimodal medical image fusion using sparse representation and modified spatial frequency. Int J Imaging Syst Technol 28:175–185. https://doi.org/10.1002/ima.22268
2. Bhateja V, Patel H, Krishn A et al (2015) Multimodal medical image sensor fusion framework using Cascade of wavelet and Contourlet transform domains. IEEE Sensors J 15:6783–6790. https://doi.org/10.1109/JSEN.2015.2465935
3. Bhateja V, Krishn A, Sahu A (2016) Medical image fusion in Curvelet domain employing PCA and maximum selection rule. In: Proceedings of the Second International Conference on Computer and Communication Technologies. pp 1–9
4. Bhatnagar G, Wu QMJ (2012) An image fusion framework based on human visual system in Framelet domain. Int J wavelets, multiresolution Inf process 10:–1250002. https://doi.org/10.1142/S0219691311004444
5. Bhatnagar G, Wu QMJ, Liu Z (2013) Directive contrast based multimodal medical image fusion in NSCT domain. IEEE Trans Multimed 15:1014–1024. https://doi.org/10.1109/TMM.2013.2244870
6. Bhatnagar G, Jonathan Wu QM, Liu Z (2013) Human visual system inspired multi-modal medical image fusion framework. Expert Syst Appl 40:1708–1720. https://doi.org/10.1016/j.eswa.2012.09.011
7. Bhatnagar G, Wu QMJ, Liu Z (2015) A new contrast based multimodal medical image fusion framework. Neurocomputing 157:143–152. https://doi.org/10.1016/j.neucom.2015.01.025

8. Bhavana V, Krishnappa HK (2015) Multi-modality medical image fusion using discrete wavelet transform. Procedia Comput Sci 70:625–631. https://doi.org/10.1016/j.procs.2015.10.057
9. Daniel E, Anitha J, Kamaleshwaran KK, Rani I (2017) Optimum spectrum mask based medical image fusion using gray wolf optimization. Biomed Signal Process Control 34:36–43. https://doi.org/10.1016/j.bspc.2017.01.003
10. Daniel E, Anitha J, Gnanaraj J (2017) Optimum laplacian wavelet mask based medical image using hybrid cuckoo search – grey wolf optimization algorithm. Knowledge-Based Syst 131:58–69. https://doi.org/10.1016/j.knosys.2017.05.017
11. Das S, Kundu MK (2013) A neuro-fuzzy approach for medical image fusion. IEEE Trans Biomed Eng 60: 3347–3353. https://doi.org/10.1109/TBME.2013.2282461
12. Do MN, Vetterli M (2002) Contourlets: A new directional multiresolution image representation. In: Proceedings. International Conference on Image Processing. pp 497–501
13. Dogra A, Goyal B, Agrawal S, Ahuja CK (2017) Efficient fusion of osseous and vascular details in wavelet domain. Pattern Recogn Lett 94:189–193. https://doi.org/10.1016/j.patrec.2017.03.002
14. Du J, Li W, Lu K, Xiao B (2016) An overview of multi-modal medical image fusion. Neurocomputing 215: 3–20. https://doi.org/10.1016/j.neucom.2015.07.160
15. El-hoseny HM, Rabaie EM El, Elrahman WA, El-samie FEA (2017) Medical image fusion techniques based on combined discrete transform domains. In: 34th National Radio Science Conference (NRSC 2017) IEEE. pp 471–480
16. Fei Y, Wei G, Zongxi S (2017) Medical image fusion based on feature extraction and sparse representation. Int J Biomed Imaging 2017:1–11. https://doi.org/10.1155/2017/3020461
17. Ganasala P, Kumar V (2016) Feature-motivated simplified adaptive PCNN-based medical image fusion algorithm in NSST domain. J Digit Imaging 29:73–85. https://doi.org/10.1007/s10278-015-9806-4
18. Genderen CP, JLvan To (1998) Review article multisensor image fusion in remote sensing: concepts, methods and applications. Int J Remote Sens 19:823–854. https://doi.org/10.1080/014311698215748
19. Gharbia R, El Baz AH, Aboul Ella Hassanien MFT (2014) Remote sensing image fusion approach based on Brovey and wavelets transforms. In: Proceedings of the Fifth International Conference on Innovations in Bio-Inspired Computing and Applications IBICA. pp 311–321
20. Haribabu M, Hima Bindu C (2017) Feature level based multimodal medical image fusion with Hadamard transform. Int J Control Theory Appl
21. He C, Liu Q, Li H, Wang H (2010) Multimodal medical image fusion based on IHS and PCA. Procedia Eng 7:280–285. https://doi.org/10.1016/j.proeng.2010.11.045
22. He K, Sun J, Tang X (2013) Guided image filtering. IEEE Trans Pattern Anal Mach Intell 35:1397–1409. https://doi.org/10.1109/TPAMI.2012.213
23. He C, Qin Y, Cao G and Lang F (2013) medical image fusion using guided filtering and pixel screening based weight averaging scheme. 77–85
24. Indira KP, Rani Hemamalini R, Indhumathi R (2015) Pixel based medical image fusion techniques using discrete wavelet transform and stationary wavelet transform. Indian J Sci Technol 8:1–7. https://doi.org/10.17485/ijst/2015/v8i26/56192
25. James AP, Dasarathy BV (2014) Medical image fusion: a survey of the state of the art. Inf Fusion 19:4–19. https://doi.org/10.1016/j.inffus.2013.12.002
26. Jin X, Chen G, Hou J et al (2018) Multimodal sensor medical image fusion based on nonsubsampled Shearlet transform and S-PCNNs in HSV space. Signal Process 153:379–395. https://doi.org/10.1016/j.sigpro.2018.08.002
27. Kaur H, and Kumar S (2018) Fusion of multi-modality medical images : a fuzzy approach. In: IEEE 3rd international conference on computing, Communication and Security (ICCCS). IEEE, pp 112–115
28. Kayani BN, Mirza AM, Bangash A, Iftikhar H (2007) Pixel & feature level multiresolution image fusion based on fuzzy logic. Innov Adv Tech Comput Inf Sci Eng Springer 129–132
29. Keyur N, Brahmbhatt RMM (2013) Comparative study on image fusion methods in spatial domain. Int J Adv Res Eng Technol 4:161–166
30. Kor S, Tiwary U (2004) Feature level fusion of multimodal medical images in lifting wavelet transform domain. In: Conf Proc IEEE Eng Med Biol Soc. IEEE, pp 1479–1482
31. Kumar A, Sachdeva S, Gupta L, Sharma A (2013) Design and implementation of pixel level medical image fusion. Int J Latest Res Eng Comput 1:26–32
32. Lewis JJ, O'Callaghan RJ, Nikolov SG et al (2007) Pixel- and region-based image fusion with complex wavelets. Inf Fusion 8:119–130. https://doi.org/10.1016/j.inffus.2005.09.006
33. Li M, Li G, Cai W, Li XY (2008) A novel pixel-level and feature-level combined multisensor image fusion scheme. Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics) 5264 LNCS:658–665. 10.1007/978-3-540-87734-9-75

34. Liu Z, Yin H, Chai Y, Yang SX (2014) A novel approach for multimodal medical image fusion. Expert Syst Appl 41:7425–7435. https://doi.org/10.1016/j.eswa.2014.05.043
35. Liu X, Mei W, Du H (2016) Multimodality medical image fusion algorithm based on gradient minimization smoothing filter and pulse coupled neural network. Biomed Signal Process Control 30:140–148. https://doi.org/10.1016/j.bspc.2016.06.013
36. Liu X, Mei W, Du H (2017) Structure tensor and nonsubsampled shearlet transform based algorithm for CT and MRI image fusion. Neurocomputing 235:131–139. https://doi.org/10.1016/j.neucom.2017.01.006
37. Liu Y, Chen X, Ward RK, Wang ZJ (2019) Medical image fusion via convolutional Sparsity based morphological component analysis. IEEE Signal Process Lett 26:1–1. https://doi.org/10.1109/lsp.2019.2895749
38. Luo X, Wu X, Zhang Z (2013) Image fusion of the feature level based on quantum-behaved particle swarm optimization algorithm. J Algorithm Comput Technol 7:101–112. https://doi.org/10.1260/1748-3018.7.1.101
39. Majumdar S, Bharadwaj J (2014) Feature level fusion of multimodal images using Haar lifting wavelet transform. Int J Comput Inf Eng 8:1023–1027
40. Malik SS, Shivprasad BJ, Maruthi GB (2013) Feature level image fusion. In: proceeding of CCSO. Pp 42–46
41. Manchanda M, Sharma R (2016) A novel method of multimodal medical image fusion using fuzzy transform. J Vis Commun Image Represent 40:197–217. https://doi.org/10.1016/j.jvcir.2016.06.021
42. Mehta N, Budhiraja S (2018) Multimodal medical image fusion using guided filter in NSCT domain. Biomed Pharmacol J 11:1937–1946. https://doi.org/10.13005/bpj/1566
43. Nirmala E, Vaidehi V (2015) Comparison of pixel-level and feature level image fusion methods. In: International Conference on Computing for Sustainable Global Development. pp 743–748
44. Parekh P, Patel N, Visavalia S (2014) Comparative study and analysis of medical image fusion techniques. Int J Comput Appl 90:12–16
45. Qi G, Wang J, Zhang Q et al (2017) An integrated dictionary-learning entropy-based medical image fusion framework. Futur Internet 9:1–25. https://doi.org/10.3390/fi9040061
46. Qi G, Wang D, Zhu Z et al (2019) A phase congruency and local Laplacian energy based multi-modality medical image fusion method in NSCT domain. IEEE Access 7:1–1. https://doi.org/10.1109/access.2019.2898111
47. Ramlal SD, Sachdeva J, Kamal C, Niranjan A (2018) Multimodal medical image fusion using non-subsampled shearlet transform and pulse coupled neural network incorporated with morphological gradient. Signal, Image Video Process 12:1479–1487. https://doi.org/10.1007/s11760-018-1303-z
48. P Rn, Desai U, Shetty VB (2014) Medical image fusion analysis using curvelet transform. In: Int. Conf. on Adv. in Comp., Comm., and Inf. Sci. (ACCIS-14). pp 1–8
49. Shuaiqi L, Jie Z, Mingzhu S (2015) medical image fusion based on rolling guidance filter and spiking cortical model. Comput Math Methods Med 2015:. https://doi.org/10.1155/2015/156043
50. Singh R, Khare A (2014) Fusion of multimodal medical images using Daubechies complex wavelet transform - a multiresolution approach. Inf Fusion 19:49–60. https://doi.org/10.1016/j.inffus.2012.09.005
51. Singh S, Gupta D, Anand RS, Kumar V (2015) Nonsubsampled shearlet based CT and MR medical image fusion using biologically inspired spiking neural network. Biomed Signal Process Control 18:91–101. https://doi.org/10.1016/j.bspc.2014.11.009
52. Srivastava R, Khare A, Prakash O (2016) Local energy-based multimodal medical image fusion in curvelet domain. IET Comput Vis 10:513–527. https://doi.org/10.1049/iet-cvi.2015.0251
53. Sruthy S, Parameswaran L, Sasi AP (2013) Image fusion technique using DT-CWT. In: IEEE international multi conference on automation, computing, control, communication and compressed sensing, iMac4s 2013. Pp 160–164
54. Vijayarajan R, Muttan S (2015) Discrete wavelet transform based principal component averaging fusion for medical images. AEU - Int J Electron Commun 69:896–902. https://doi.org/10.1016/j.aeue.2015.02.007
55. Wang L, Li B, Tian L (2014) Multi-modal medical image fusion using the inter-scale and intra-scale dependencies between image shift-invariant shearlet coefficients. Inf Fusion 19:20–28. https://doi.org/10.1016/j.inffus.2012.03.002
56. Xia J, Chen Y, Chen A (2018) Chen Y (2018) medical image fusion based on sparse representation and PCNN in NSCT domain. Comput Math Methods Med. https://doi.org/10.1155/2018/2806047
57. Xu L, Du J, Hu Q, Li Q (2013) Feature-based image fusion with a uniform discrete curvelet transform. Int J Adv Robot Syst 10: 10.5772/56345
58. Xu X, Wang Y, Chen S (2016) Medical image fusion using discrete fractional wavelet transform. Biomed Signal Process Control 27:103–111. https://doi.org/10.1016/j.bspc.2016.02.008
59. Xu X, Shan D, Wang G, Jiang X (2016) Multimodal medical image fusion using PCNN optimized by the QPSO algorithm. Appl Soft Comput J 46:588–595. https://doi.org/10.1016/j.asoc.2016.03.028

60. Yang Y, Que Y, Huang S, Lin P (2016) Multimodal sensor medical image fusion based on Type-2 fuzzy logic in NSCT domain. IEEE Sensors J 16:3735–3745. https://doi.org/10.1109/JSEN.2016.2533864
61. Yin M, Liu X, Liu Y, Chen X (2018) Medical image fusion with parameter-adaptive pulse coupled-neural network in nonsubsampled Shearlet transform domain. IEEE trans Instrum Meas 1–16. https://doi.org/10.1109/TIM.2018.2838778
62. Zhang WCBLY (2003) a Remote Sensing Image Fusion Method Based on Pca. In: IEEE Int. Conf. Neural Networks 8 Si nal Processing. pp 976–981
63. Zhang Z, Blum RS (1999) A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application. Proc IEEE 87:1315–1326. https://doi.org/10.1109/5.775414
64. Zhu Z, Chai Y, Yin H et al (2016) A novel dictionary learning approach for multi-modality medical image fusion. Neurocomputing 214:471–482. https://doi.org/10.1016/j.neucom.2016.06.036
65. Zong J, Qiu T (2017) Medical image fusion based on sparse representation of classified image patches. Biomed Signal Process Control 34:195–205. https://doi.org/10.1016/j.bspc.2017.02.005

Nahed Tawfik received B.Sc. (Honors) and MSc. From the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt in 2008 and 2015, respectively. She is a Research Assistant at Electronics Research Institute, Egypt. Her interests are image processing and biomedical engineering.

Heba A. Elnemr is an Associate Professor at the Electronics Research Institute, Cairo, Egypt. She received her B. Sc. degree, M. Sc. degree and Ph.D. degree in Electronics and Communications Engineering from the Faculty of Engineering, Cairo University, Egypt. She has supervised several master and PhD. students in the field of image processing. Her research interests include pattern recognition, signal processing, biometrics, computer vision and image processing.

Mahmoud Fakhr received the B.S. degree in automatic control from the Menoufia University in 1991 and the M.Sc. and PhD degree in Computer Engineering, from Electronics and Communications Department, Cairo University, Faculty of Engineering, in 1996 and 2004 respectively. He received the professor degree in 2018. He works as a professor of computer science at the Electronics Research Institute. He works as the head of the grants and international cooperation GICO office at ERI. He has also worked at the Faculty of Engineering, Abha, KSA and different universities in Egypt. He has supervised several master and PhD. students in the field of biomedical engineering. He works as a consultant at the ministry of communications and information technology, Egypt, from 2004 till now. He collected different experiences in information technology, while working in different governmental organizations in Egypt. His areas of interest are evolutionary computation, advanced automatic control, image processing and biomedical engineering.

Moawad I. Dessouky received the B.Sc. (Honors) and M.Sc. degrees from the Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1976 and 1981, respectively, and the Ph.D. from McMaster University, Canada, in 1986. He joined the teaching staff of the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University, Menouf, Egypt, in 1986. He was the vice dean of Faculty of Electronic Engineering, Menoufia University. He has received the most cited paper award from Digital Signal Processing journal for 2008. His current research areas of interest include spectral estimation techniques, image enhancement, image restoration, super resolution reconstruction of images, satellite communications, and spread spectrum techniques.

Fathi E. Abd El-Samie received the B.Sc., (Hons.), M.Sc., and Ph.D. degrees from Menuofia University, Menouf, Egypt, in 1998, 2001, and 2005, respectively. Since 2005, he has been a Teaching Staff Member with the Department of Electronics and Electrical Communications, Faculty of Electronic Engineering, Menoufia University. He was a researcher at KACST-TIC in Radio Frequency and Photonics for the e-Society (RFTONICs). His current research interests include image enhancement, image restoration, image interpolation, super-resolution reconstruction of images, data hiding, multimedia communications, medical image processing, optical signal processing, and digital communications. He was a recipient of the Most Cited Paper Award from the Digital Signal Processing journal in 2008.

Affiliations

Nahed Tawfik[1] • **Heba A. Elnemr**[1] • **Mahmoud Fakhr**[1] • **Moawad I. Dessouky**[2] • **Fathi E. Abd El-Samie**[2,3]

Heba A. Elnemr
Heba@eri.sci.eg

Mahmoud Fakhr
Mahmoud@eri.sci.eg

Moawad I. Dessouky
dr_moawad@yahoo.com

Fathi E. Abd El-Samie
fathi_sayed@yahoo.com

1 Computers and Systems Department, Electronics Research Institute (ERI), Cairo, Egypt

2 Department of Electronics and Electrical Communications Engineering, Faculty of Electronic Engineering, Menoufia University, Menouf 32952, Egypt

3 Department of Information Technology, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 84428, Saudi Arabia