

Multi-Sensor Image Fusion: A Survey of the State of the Art

Bing Li¹, Yong Xian¹, Daqiao Zhang¹, Juan Su¹, Xiaoxiang Hu², Weilin Guo³

¹Xi'an High-Tech Research Institute, Xi'an, China

²School of Automation, Northwestern Polytechnical University, Xi'an, China

³Xi'an Satellite Control Center, Xi'an, China

Email: libingbenyi@163.com

How to cite this paper: Li, B., Xian, Y., Zhang, D.Q., Su, J., Hu, X.X. and Guo, W.L. (2021) Multi-Sensor Image Fusion: A Survey of the State of the Art. *Journal of Computer and Communications*, 9, 73-108. <https://doi.org/10.4236/jcc.2021.96005>

Received: April 27, 2021

Accepted: June 18, 2021

Published: June 21, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative

Commons Attribution International

License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Image fusion has been developing into an important area of research. In remote sensing, the use of the same image sensor in different working modes, or different image sensors, can provide reinforcing or complementary information. Therefore, it is highly valuable to fuse outputs from multiple sensors (or the same sensor in different working modes) to improve the overall performance of the remote images, which are very useful for human visual perception and image processing task. Accordingly, in this paper, we first provide a comprehensive survey of the state of the art of multi-sensor image fusion methods in terms of three aspects: pixel-level fusion, feature-level fusion and decision-level fusion. An overview of existing fusion strategies is then introduced, after which the existing fusion quality measures are summarized. Finally, this review analyzes the development trends in fusion algorithms that may attract researchers to further explore the research in this field.

Keywords

Multi-Sensor Image Fusion, Fusion Strategy, Feature Enhancement, Fusion Performance Assessment

1. Introduction

In the late 1970s, with the emergence and development of image sensors, multi-sensor information fusion facilitated the emergence of branch-image fusion, an emerging research field combining sensors, signal processing, image processing, and artificial intelligence while using images as the research object in the field of information fusion. This approach combines the image information about the same scene obtained either by multiple image sensors or by the same image sensor in different working modes to obtain a new and more accurate description of

the scene [1].

Different sensor images have different advantages and disadvantages. For example, visible images can provide texture detail, along with high spatial resolution and high definition in a manner consistent with the human visual system, but do not work well in all-day/night conditions. By contrast, infrared images can distinguish targets from their backgrounds using differences in radiation, and work well in all-day/night conditions, but the infrared images are more blurred [2]. Moreover, Synthetic Aperture Radar (SAR) can better reflect the texture characteristics and structural information of features, and also work well in all weather and all-day/night conditions, but its visual effect is relatively poor when compared to the visible images [3]. In addition, multispectral (MS) images have low resolution and high spectral density, while panchromatic (PAN) images have high resolution and low spectral density [4]. Due to the redundancy and complementarity between the image information obtained by different image sensors (or the same image sensor in different working modes), when compared to any of the individual input remote images, a more comprehensive and accurate image description of a certain scene can be obtained through the fuse of multiple source images (as illustrated in **Figure 1**) [5]; this approach overcomes the limitations of and differences between the geometric, spectral and spatial resolution of single sensor images, improves image clarity and comprehensibility, and provide more effective information for subsequent image processing task (e.g. image segmentation [6] [7], classification [8], saliency [9] [10], target detection and recognition [11] [12], localization [13], medical diagnosis [14], surveillance [15], energy monitoring [3] [16], agricultural applications [17] and military applications [18]).

Significant research efforts have accordingly been dedicated to the development of image fusion, with a great number of image fusion algorithms having been proposed in the literature. As illustrated in **Figure 2**, the number of scientific

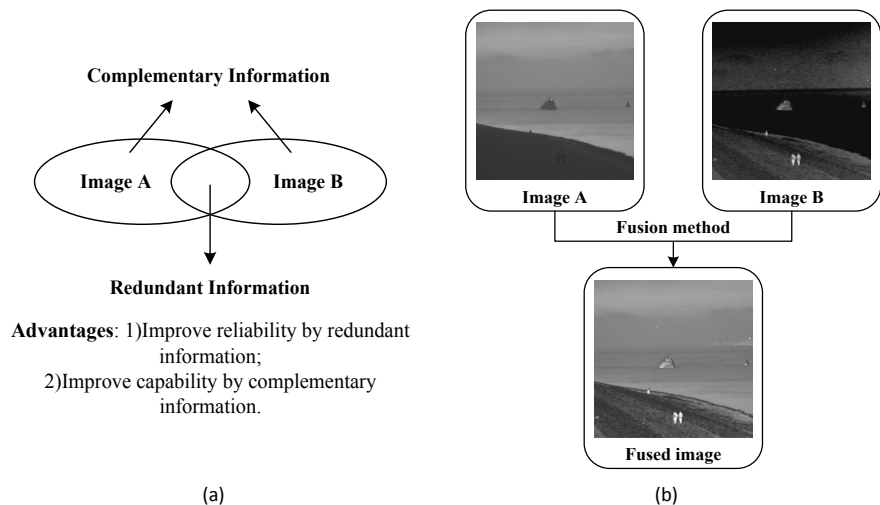


Figure 1. Image fusion example. (a) Information diagram between heterogeneous images; (b) Multi-sensor image fusion example.

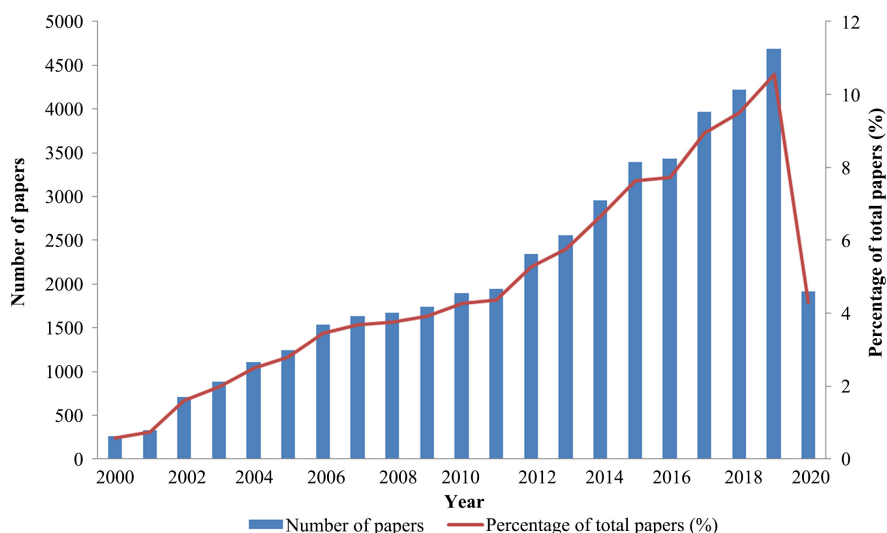


Figure 2. The number of scientific journal and conference papers published on the topic of image fusion (Statistics on Web of Science, June 20, 2020 [19]).

papers published on the topic of image fusion in international journals and at conferences has increased dramatically since 2010: these papers mainly relate to the engineering electrical electronic, computer science artificial intelligence, remote sensing, imaging science photographic technology, optics, computer science information systems, computer science interdisciplinary applications, and telecommunications research fields.

This rapid growth trend can be attributed to three major factors. 1) There has been increased social demand for developing low-cost and high-performance imaging technologies, although the design of high-quality sensors may be limited by technical factors. By combining the images obtained by different sensors, image fusion offers a powerful solution to this problem. 2) The development of signal processing and analysis theory. For example, powerful signal processing tools such as multi-scale decomposition, sparse representation, and neural networks offer opportunities for further improving the performance of image fusion. 3) The increase in both the number and diversity of complementary images obtained in different applications. For instance, in remote sensing applications, more and more satellites are acquiring remote sensing images of observation scenes with differing spatial, spectral and temporal resolutions. A similar situation has arisen in other applications such as medical imaging [20]. After investigating the early proposed image fusion methods, Burt and Adelson introduced a novel image fusion algorithm based on layered image decomposition [21]. In order to further improve stability and noise resistance, as well as to resolve the pathological condition arising from patterns with opposite contrast, [22] improved the pyramid fusion method. Pohl and Genderen subsequently described and explained the mainly pixel-based image fusion of Earth observation satellite data as a contribution to multi-sensor integration-oriented data processing [5]. A categorization of the multi-scale decomposition-based image

fusion schemes is provided in [23]. According to the adapted transformation strategy, Li *et al.* investigated various pixel-level image fusion algorithms, summarized the existing fusion performance evaluation methods and unresolved issues, and analyzed and summarized the main challenges encountered in different image fusion applications. Different from previous surveys, the purpose of this paper is to introduce the relevant fusion methods and applications recently introduced, which can provide new insights into the development of current image fusion theory and application [20]. In addition, the special issue [24] published in the Information Fusion journal by Goshtasby and Nikolov is an excellent source that tracks the development of image fusion methods.

Furthermore, some research applying image fusion methods to certain specific application fields has also been published in recent years [25]-[30]. Taking remote sensing as an example, [25] summarized the early proposed remote sensing image fusion algorithms, [26] conducted a critical comparison of recently proposed remote sensing image fusion methods, [27] reviewed the current multi-source remote sensing data fusion techniques and discussed their future trends and challenges. In other applications, such as the medical imaging field, a practical list of methods was provided in [28], which also summarized the broad scientific challenges facing the field of medical image fusion. Based on the reported comparative results, recent image fusion and performance assessment algorithms were reviewed and categorized, after which a comprehensive evaluation of 40 fusion algorithms from recently published results was conducted to demonstrate their significance in terms of statistical analyses within their respective applications [29].

The remainder of this paper is organized as follows. Section 2 briefly reviews the popular and state-of-the-art fusion methods at different levels (namely the pixel-level, feature-level, and decision-level). Section 3 presents an overview of the fusion strategy. Moreover, an overview of the fusion performance assessment metrics is introduced in Section 4. Finally, some future trends and conclusions are summarized in Section 5.

The novelty of the work in this paper can be summarized as follows: 1) this paper summarizes the existing multi-sensor image fusion algorithms, fusion strategies and evaluation indicators relatively completely, which has a reliable reference value for the subsequent image fusion researches; 2) unlike other image fusion reviews, this paper also summarizes image fusion algorithms based on the emerging theory-deep learning; 3) coupled with the analysis of the development trends in the field of image fusion, the paper provides a reference for researchers to further explore the research in this direction, which can promote the innovative development of this field.

2. Multi-Sensor Image Fusion

According to the fusion stage in the processing flow, image fusion can be broadly categorized into pixel-level fusion, feature-level fusion and decision-level fu-

sion depending on the degree of information abstraction. In this section, we present a comprehensive review of multi-sensor image fusion methods of the pixel-level fusion, feature level fusion, and decision-level fusion varieties.

2.1. Pixel-Level Fusion

Of the three levels of image fusion, pixel-level image fusion is the lowest level. Compared with feature-level and decision-level fusion, pixel-level image fusion involves the direct fusion of the source image pixels under strict registration conditions according to a fusion rule. This approach retains as much scene information as possible of the source images, and also has high precision; accordingly, it can be used to improve the sensitivity and signal-to-noise ratio of the signal, thereby facilitating subsequent image analysis, processing and understanding (e.g. feature extraction, image segmentation, scene analysis/monitoring, target recognition, image recovery, etc.). For example, in remote sensing applications, fused images with a high spatial resolution, along with the spectral content of the multi-spectral (MS) images, could be obtained by fusing low MS images and high resolution panchromatic (PAN) images, which would be useful for land-cover classification.

However, there are several key disadvantages of pixel-level fusion. Firstly, pixel-level image fusion requires high registration accuracy between source images, and should generally be achieved pixel-level registration. In addition, the pixel-level image fusion processing has a large amount of data, and is also encumbered by a slow processing speed, and poor real-time performance.

Due to the diversity of source images, along with the diversity of practical fusion applications, it is not possible to design a general method that is suitable for all image fusion tasks. However, the majority of pixel-level image fusion methods can be expressed in terms of three main stages: image transform, fusion of the transform coefficients, and inverse transform. Based on the transformation strategy, moreover, the existing pixel-level image fusion algorithms can be divided into four major categories: 1) multi-scale decomposition-based fusion methods; 2) sparse representation (SR)-based fusion methods; 3) methods for fusing the image pixels directly, or in other transform domains such as HSI or PCA; 4) methods combining multi-scale decomposition, sparse representation, principal component analysis, and other types of transformation [20]. **Table 1** presents the summary of the major pixel-level image fusion methods, along with the transforms and fusion strategies adopted; more details are presented below.

2.1.1. Multi-Scale Decomposition-Based Fusion Methods

Multi-scale decomposition-based fusion methods use a multi-scale decomposition method to decompose multisource images into different scales and different resolutions, and consequently, to obtain low-frequency sub-bands containing image energy information and high-frequency sub-bands containing detailed information of different scales of images; in the next step, the fusion is performed according to different fusion rules on the low-frequency sub-band and

Table 1. The major pixel-level image fusion methods, the adopted transforms and fusion strategies [20].

Methods	Transforms	Fusion strategies
Multi-scale decomposition	pyramid, wavelet, complex wavelet, curvelet, contourlet, shearlet, ripplelet, edge-preserving decompositions, anisotropic heat diffusion, log-Gabor transform, support value transform, optimal scale transform, etc.	coefficient, window, and region based activity level measurement (CAM, WAM, RAM), choose-max (CM), optimization based method (OP), weighted-average, and substitution based coefficients combining (WACC, SCC), window and region based consistency verification (WRCV), cross-scale fusion rule (CSF), guided filtering based weighted average (GFWA), etc.
Sparse representation (SR)	orthogonal matching pursuit, group SR, gradient constrained SR, simultaneous OMP (SOMP), joint sparsity model, SR with over-complete dictionary, structural sub-dictionary, spectral and spatial details dictionary, etc.	WAM, choose-max and weighted average based coefficients combining (CM-WACC), substitution of sparse coefficients (SSC), spatial context based weighted average (SCWA), etc.
Methods in other domains	spatial domain (non-transforms), hue saturation-intensity transform (HSI), principal component analysis (PCA), Gram-Schmidt (GS) transform, matting decomposition, independent component analysis (ICA), gradient domain, and fuzzy theory, etc.	machine learning based weighted average (MLWA), block and region based activity level measurement (BRAM), SCWA, model based method (MM), component substitution (CS), etc.
Combination of different transforms	hybrid wavelet-contourlet, multi-scale transform-SR, morphological component analysis-SR, contourlet-SR, IHS-retina inspired models, IHS-wavelet, etc.	coefficient and window based activity level measurement (CWAM), CM-WACC, CS, integration of component substitution and weighted average (ICSWA), etc.

high-frequency sub-bands respectively; finally, the multi-scale reconstruction is performed through the fused sub-band image to obtain the final fused image. A schematic diagram of the images fusion scheme based on general multi-scale decomposition is illustrated in **Figure 3**. Another key factor that affects the fusion results is the fusion strategy. This is a process that determines the formation of the fused image, based on coefficients or pixels of the source images.

Classical multi-scale decomposition methods include pyramid transform and wavelet transform. Common pyramid transformations include the Laplacian pyramid (LP) [22], contrast pyramid (CP) [31] [32] [33], steerable pyramid [34], and gradient pyramid (GP) [35]. Moreover, wavelet transforms include: discrete wavelet transform (DWT) [36] [37] [38], discrete wavelet packet [39], binary wavelet, lifting wavelet [40], dual tree complex wavelet (DTCWT) [41], multi-wavelet [42], à trous wavelet [43] [44], discrete stationary wavelet transform (DSWT) [45] etc. The contrast pyramid transform is derived from the Laplacian pyramid transform, which involves calculating the ratio of two adjacent low-pass filtering images of the Gaussian pyramid; therefore, the model becomes able to consider local contrast [31]. The work of [32] proposed an infrared and visible

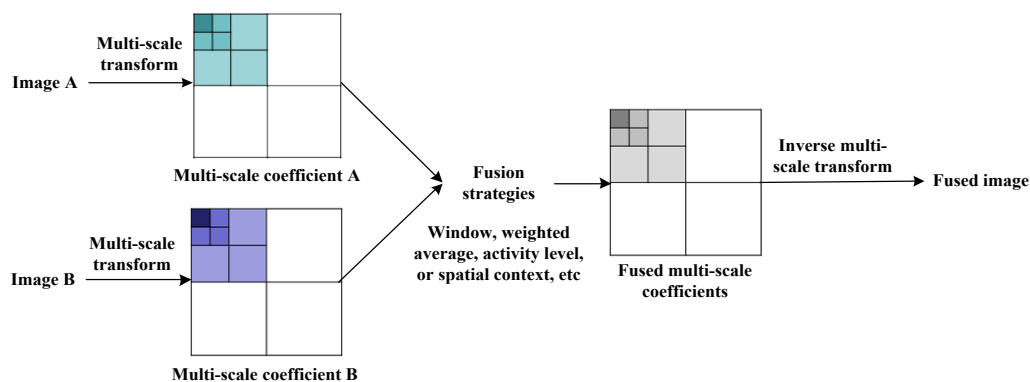


Figure 3. The schematic diagram of image fusion scheme based on general multi-scale decomposition.

image fusion method that was based on the contrast pyramid and multi-objective evolutionary algorithm, in which the multi-objective evolutionary algorithm was utilized to optimize fusion coefficients.

The image fusion method based on pyramid transform is simple to implement; however, when the gray value difference between images is large, square traces will appear in the fusion result due to the correlation of features between adjacent scales after the image is transformed by the pyramid. The wavelet transform has a good ability to analysis local time domain and frequency domain for the signal, while the optimal representation of the signal point singularity can be obtained for the one-dimensional segmented smooth signal. For the two-dimensional image signal, however, the commonly used two-dimensional separable wavelet is a “non-sparse” image representation: this is a tensor product of one-dimensional wavelets, with only a finite direction, that is unable to optimally represent a two-dimensional image containing line singularities. In short, a common limitation of the methods in the wavelet family is that they find it difficult to accurately represent the curves and edges of the images.

In order to solve these problems and better represent the images, many multi-scale decomposition methods have been proposed and applied to the image fusion problem. Representative examples include Bandlet transform, which is based on image edge adaptive construction, Ridgelet transform that has a good approximation effect on linear singularity, curvelet transforms (CVT) [46] [47], which can produce a good approximation of curve singularity, contourlet transforms [48] [49], nonsubsampling contourlet transform (NSCT) [50], shearlets which allow for the efficient encoding of anisotropic features in multivariate problem classes and place no restrictions on the 110 directions for the shearing or the size of the supports [51], and so on. Through the use of the parabolic scaling law, the curvelet and contourlet can resolve the dichotomous points along a smooth curve. Furthermore, the multi-scale decomposition methods described above have also been successfully applied in multi-sensor image fusion tasks [46] [49] [50]. However, these multi-scale decomposition methods cannot sparsely represent any anisotropic structure; therefore, the ripplelet transform is proposed [52]. In [53], moreover, a remote sensing image fusion method based

on ripplet transform and compressed sensing theory was proposed in order to minimize the spectral distortion of the pansharpened MS bands relative to the original ones. The primary advantage of these multi-scale decomposition methods is their good ability to preserve the details of different source images. However, since spatial consistency is not adequately considered during the fusion process, the fusion of these methods may result in distortion of the brightness and color values [54].

Recently, edge-preserving filters [55] have been actively researched within the field of image processing, and have also been successfully applied in multi-sensor image fusion. In [55], Farbman *et al.* introduce a novel method for constructing edge-preserving multi-scale image decomposition, in which the edge-preserving multi-scale image decomposition is accomplished using an alternative edge-preserving smoothing operator based on a weighted least squares optimization framework, which is well-suited to progressive coarsening and the multi-scale detail extraction of images. Li *et al.* introduced a novel two-scale fusion method that does not need to rely heavily on a specific image decomposition method; moreover a new guided filtering-based weighted average technique is utilized to make full use of spatial consistency when fusing the base layers and detail layers [54]. Hu and Li introduced a non-downsampling direction filter bank into a multi-scale bilateral filter, further proposing a new multi-scale geometric analysis method—the multi-scale bidirectional filter (MDBF)—which can better represent the intrinsic geometrical structure of images and achieves better performance than the traditional fusion methods [56]. In addition, the combination of Gaussian and bilateral filters is also successfully applied in the fusion of infrared and visible images [57]. Furthermore, some computer vision tools—such as support value transform [58], log-Gabor transform [59], and anisotropic heat diffusion [60]—have also been applied for multi-scale decomposition-based fusion.

As the authors of [20] point out, the spatial quality of the fused images may be less satisfactory if fewer decomposition levels are applied, while both the performance and computational efficiency of the method may also be reduced if too many decomposition levels are applied. Thus, some researchers have attempted to determine the optimal number of decomposition levels that will yield optimal fusion quality. For example, the authors in [61] compared various multi-resolution decomposition algorithms, focusing on the influence of the decomposition level and filters on fusion performance. They subsequently concluded that a short filter usually provides better fusion results than a long filter, while the most appropriate number of decomposition levels is four. The work of [62] estimated the optimal number of decomposition levels for multi-spectral and panchromatic image fusion with a specific resolution ratio.

2.1.2. Sparse Representation Methods

The research hotspots in the field of sparse representation include the approximate representation of models, uniqueness and stability of model solutions,

performance analysis of sparse representations, model-solving algorithms, dictionary learning algorithms, sparse decomposition algorithms, super-complete atomic dictionaries, specific applications of sparse representations, and close contact compression sensing, among other aspects. Among them, specific applications include: image processing (such as compression, enhancement and super-resolution), audio processing (such as blind source separation) and pattern recognition (such as face and gesture recognition). From a practical perspective, targeted flexible models, computational speed, adaptive and high-performance representation results are key ways in which sparse representation methods can achieve advantage in the application domain.

The purpose of the signal sparse representation is to represent the signal with as few atoms as possible in a given over-complete dictionary: this allows a more concise representation of the signal to be obtained, making it easier and more convenient for us to obtain the information contained in the signal.

Utilizing the characteristics of sparse coefficients, Yang and Li were the first to apply sparse representation theory to image fusion. Firstly, the multi-source input images are divided into many overlapping patches in order to capture local salient features and maintain shift invariance. In the next step, in order to obtain the corresponding sparse coefficients, the overlapping patches from multi-source images are decomposed in the overcomplete dictionary. Subsequently, the sparse coefficients from multiple source images are applied to the fusion process. Finally, the image is reconstructed using the fusion coefficients and dictionary [63]. In [64], a novel pan-sharpening method (HR panchromatic and corresponding LR spectral channel fusion), named Sparse Fusion of Image (SparseFI), is proposed. Based on the theory of compressed sensing, it utilizes the sparse representation of the HR/LR multispectral image block in the panchromatic image, along with its down-sampled LR version. Yu *et al.* used the first model (JSM-1) from the joint sparse representation method proposed in [65] to achieve image fusion [66]. In [67], Liu *et al.* proposed an improved, compressed sensing-based image fusion scheme. In this method, the low and high sparse coefficients of the source images obtained by discrete wavelet transform are fused by means of an improved entropy-weighted fusion rule and max-abs-based fusion rule respectively. After using the local linear projection of random Gaussian matrix, the fusion image is reconstructed using a compression sampling matching tracking algorithm. Compared with the traditional transform-based image fusion methods, the proposed approach can retain more details, such as edges, lines and contours. Ma *et al.* proposed an image fusion algorithm based on sparse representation and guided filtering. Firstly, a sparse representation (SR)-based method was utilized to construct the initial fused image from the source images. Then, the spatial domain (SF)-based fused image was obtained in order to make full use of the spatial information of the source image. Finally, the final fused image was obtained by using guided filtering to reconstruct the SR-based fusion image and the SF-based fusion image [68]. In [69], a noisy remote sensing image fusion method based on joint sparse representation

(JSR) was proposed to fuse SAR images with other source images. Firstly, the redundant complementary sub-images were obtained by the JSR method, and then the complementary sparse coefficients were fused together by using an improved fusion rule based on pulse coupled neural network (PCNN). At the same time, due to the different types of noises between the SAR image and other source images, these noises were treated as complementary information in the source images and suppressed in this step. Finally, the fused image is reconstructed by adding the fused complementary sub-images and the redundant information.

2.1.3. Methods in Other Domains

In addition to those based on multi-scale decomposition and sparse representation, there are also many fusion methods based on other theoretical knowledge.

For instance, [70] proposed a novel fusion framework based on the total spectrum change (TV) method and image enhancement. For the multi-scale decomposition of the spatial variation generated by the spectral variational framework, these authors verified that the decomposition components could be effectively modeled by the tail Rayleigh distribution (TRD) rather than the commonly used Gaussian distribution. Therefore, TRD-based saliency and matching measures were proposed to fuse each sub-band decomposition, while spatial intensity information is also employed to fuse the remaining image decomposition components. Zhang *et al.* reconstruct the infrared background using quad-tree decomposition and Bessel interpolation. Secondly, by subtracting the reconstructed background from the infrared image, the infrared brightness feature is extracted, and then refined by reducing the redundant background information. In order to resolve the overexposure problem, the fine infrared features are first adaptively suppressed and then added to the visual image, enabling the final fused image to be obtained [71]. Using fuzzy logic and population-based optimization, Kumar and Dass devised a new fusion method. In a departure from the weighted average of pixels method, the author proposed a method based on total variation to fuse the multiple sensor images. Under the proposed approach, the imaging process was modeled as a local affine model, while the fusion problem was framed as an inverse problem [72]. With the aim of further improving the fuse results, Shen *et al.* used maximum a posteriori (MAP) estimation in the hierarchical multivariate Gaussian conditional random field model to derive the optimal fusion weights [73]. In [74], the source images were first decomposed into a principal component matrix and a sparse matrix by means of robust principal component analysis; after that, the weights were estimated by taking the local sparse feature as the input of the pulse coupled neural network. [75] further proposed using a linear regression model to generate synthetic components, which were only partially replaced based on the correlation between the intensity component and the panchromatic image. In [76], firstly, the salient structure of the input image was fused in the gradient domain. In the next step, the fusion image was reconstructed by solving the Poisson equation, which ensures that the

gradient of the fusion image approaches the fusion gradient. In [77], the intuitionistic fuzzy set theory was applied to image fusion, which involved transferring input image to the fuzzy domain, after which the maximum and minimum fusion operations were used to achieve fusion processing. In [78], Liu *et al.* proposed a novel focus evaluation operator based on max-min filter. In the proposed focus metric, the max-min filter was combined with the average filter and the median filter (MMAM) to evaluate the focus degree of the source images. After that, based on the structure-driven fusion region and the depth information of the blurred images, MMAM is utilized to achieve high-quality multi-focus fused image.

2.1.4. Combination of Different Transforms

The most commonly used multi-scale decomposition methods are pyramid transform, wavelet transform, curved wave transform, non-downsampling contour wave transform, contourlet transform, Retinex transform, and morphological operators-based methods. Different methods have their own advantages and disadvantages in different application areas. Various fusion algorithms combining different transforms were subsequently proposed with the aim of fully utilizing these advantages, reducing defects, and ensuring that the methods employed complemented each other.

For example, in [79], a new image fusion methodology based on pulse couple neural networks and shearlets was proposed; this approach can not only extract more important visual information from the source images, but can also effectively prevent the introduction of manual information. Based on morphological component analysis and sparse representation, Jiang and Wang devised a novel fusion method in which the morphological component was utilized to decompose the source images, and then fused with the sparse representation-based method [80]. Moreover, with the goal of simultaneously preserving the appearance information of the visible image and the thermal radiation of the infrared image, Ma *et al.* put forward a novel fusion algorithm of visible and infrared images—referred to as Gradient Transfer Fusion (GTF)—based on gradient transfer and total variation (TV) minimization [81]. In [82], a Gaussian filter is utilized to decompose multi-source images, with a new combination of total variation rules employed to combine the base layers and a novel weight map construction method being proposed based on the saliency analysis. Furthermore, by using Nonsubsampled Contourlet Transform (NSCT) and sparse K-SVD dictionary learning to obtain the prominent features of the source images, Cai *et al.* proposed a novel fusion method which named as NSCT_SK_SVD [83]. In order to promote the performance of infrared and visual image fusion and provide better visual effects, Jin *et al.* proposed a novel image fusion method integrating DSWT, discrete cosine transform (DCT) and local spatial frequency (LSF). In this proposed methodology, DSWT was utilized to decompose the significant features of the source images into a series of sub-images, each with different levels and spatial frequencies; moreover, DCT was applied to separate the

important details of the sub-images on the basis of their different energy frequencies while LSF was employed to enhance the regional features of DCT coefficients that could be useful for image feature extraction [45]. The work of [84] further proposed a new method for fusing visible and infrared images, referred to as DTCWT-ACCD, based on DTCWT and an adaptive combined clustered dictionary. Yang *et al.* put forward a novel remote sensing image fusion method based on adaptively weighted joint detail injection. In the proposed method, firstly, the spatial details were extracted from MS and PAN images through à trous wavelet transform and multi-scale guided filter; after that, the extracted details were sparsely represented to generate the main joint details by dictionary learning from the sub-images themselves; subsequently, in order to obtain the refined joint details information, an adaptive weight factor was designed considering the correlation and difference between the previous joint details and PAN image details; finally, the fused image was obtained by injecting refine joint details into the MS image using modulation coefficients [85]. In [86], the PAN and MS image pansharpening algorithm was proposed based on adaptive neural network and sparse representation in the nonsubsampling shearing domain. Aiming at obtaining better fusion performance in multi-focus image fusion based on transform domain, Liu *et al.* proposed a novel multi-focus image algorithm, which combined with the adaptive dual-channel spiking cortical model (SCM) in NSST domain and the differential image. Due to the global coupling of dual SCMs, the synchronization characteristics of pulses, and the multi-resolution and direction of NSST, the proposed algorithm can well retain the information of the source image and fuse a clear image that is more in line with human visual effects [87].

2.1.5. Deep Learning-Based Fusion Methods

A new research direction and innovative idea for image fusion is Deep Learning-based methods which have achieved good results in image fusion.

To solve the problems of limited detail preservation and high sensitivity to false registration in most existing SR-based fusion methods, Liu *et al.* proposed a novel image fusion framework based on convolutional sparse representation (CSR), which was a new signal decomposition model. In this framework, each source image was decomposed into the base layer and the detail layer, so it could effectively overcome the above two problems, and then achieve high-quality multi-focus image fusion and multi peak image fusion [88].

In [89], Huang *et al.* proposed an improved sparse denoising auto-encoder (MSDA) algorithm, which was utilized to simulate the complex relationship between high resolution (HR) and low resolution (LR) PAN image blocks as a non-linear mapping. By connecting a series of MSDAs, a stacked MSDA (S-MSDA) which could effectively pre-train DNNs was obtained. In addition, the entire DNN was trained again through the back propagation algorithm after pre-training to better train the DNN. Finally, assuming that the relationship between HR/LR MS image blocks was the same as the relationship between HR/LR PAN

image blocks, the HR MS image would be reconstructed from the observed LR MS image using the trained DNN. In [90], a pan-sharpening method based on SRCNN model and Gran-Schmidt (GS) transform was proposed. In the proposed algorithm, the SRCNN model was used to enhance the spatial resolution of the MS image, and the average band of the enhanced MS image was calculated as a reference for histogram matching to modify the PAN image. Finally, the enhanced MS image and the modified PAN image were merged by GS transform to generate a high-resolution MS image. Another work that applies CNNs to pan-sharpening was introduced by Masi *et al.* [91], in which the SRCNN framework was also used to model the pan-sharpening process as an end-to-end mapping. In this algorithm, the input of the network was the superposition result of the up-sampled low-resolution MS image and PAN image, and the output of the network was the target high-resolution MS image.

In [92], Palsson *et al.* applied 3-D convolutional neural network to MS and hyperspectral (HS) image fusion, and obtained high-resolution HS image. The work of [93] put forward a superpixel-based multi-local convolutional neural network (SML-CNN) for panchromatic and MS image classification. To reduce the amount of the input data for CNN, a simple linear iterative clustering method was extended to segment MS images and generate superpixels, and replace pixels with superpixels as the basic analysis unit. In order to make full use of the spatial spectral information and environmental information of superpixels, a multi-local area joint representation method based on superpixels was proposed. Then, an SML-CNN model was established to extract effective joint feature representations, and a softmax layer was used to divide these features learned by multiple local CNNs into different categories. Finally, a multi-information modification strategy that combined the detailed information and semantic information was employed to eliminate the adverse effects on the classification results within and between superpixels, thereby improving the classification performance. Another work that applies CNNs to remote sensing image fusion was introduced by Liu *et al.* [94], in which an end-to-end learning framework based on deep multi-instance learning (DMIL) was utilized to classify MS and PAN images by using feature-based joint spectral and spatial information. The framework consists of two instances: one for capturing pan spatial information and one for describing MS spectral information, two examples of the characteristics obtained directly in line, could be regarded as simple fusion characteristics. In order to fully integrate spatial spectral information for further classification, the simple fusion feature is entered into a three-layer fully connected fusion network to learn the high-level fusion feature. And in [95], Ma *et al.* proposed a novel generation countermeasure network called FusionGAN to fuse infrared and visible images. In this network, the generator was utilized to generate the fused image with the main infrared intensity and visible gradient information. The discriminator aimed to prevent the fused image from containing more details in the visible image, so that the fusion image retains both the radiation information in the infrared image and the texture information in the visible image.

As the authors of [96] point out, a residual network (ResNet) was utilized to make full use of the high nonlinearity of deep learning model to realize image fusion task, and the proposed algorithm could achieve the highest spatial spectral unified accuracy. And in order to obtain clearer and richer texture feature panoramic images, Liu *et al.* proposed a novel multi focus image fusion algorithm, which combines NSST and ResNet. In the proposed algorithm, NSST was utilized to fully consider the high-frequency details in the image and low-frequency global features. For the high-frequency details, improved gradient sum of Laplace energy was employed to handle high frequency sub-band coefficients of different levels and directions by using different directional gradients. For the low-frequency details, ResNet with a deep network structure is used to obtain the spatial information characteristics of the low-frequency coefficient image [97]. For most fusion methods, the detection of the focus area is a key step. In view of this, Liu *et al.* proposed a multi-focus image fusion algorithm based on dual convolutional neural network (DualCNN). Firstly, the source images were input into the dual CNN to recover the details and structure from their super-resolution images and improve the contrast of the source images. Secondly, bilateral filtering was applied to reduce the noise of the fused image, and the guided filter was used to detect the focus area of the image and refine the decision map. Finally, the fusion image was obtained by weighting the source image according to the decision graph. Experimental results show that the algorithm can preserve image details well and maintain spatial consistency [98].

Another work that applies CNNs to image fusion was proposed by Zagoruyko and Komodakis, in which a CNN-based model was employed to directly learn a general similarity function for comparing image patches from the image data. Then, by exploring and studying a variety of neural network structures, it was proved that these network structures were particularly suitable for this task, which provides a new idea for studying the effectiveness of convolutional neural network design target fusion indicators [99].

2.2. Feature-Level Fusion

Feature-level image fusion involves the extraction of feature information (e.g. corners, edges, length, contours, shapes, textures, regions, etc.) from the source images, which is then comprehensively analyzed and processed. Feature-level image fusion is an intermediate-level form of information fusion; it not only retains the important information contained in the source image, but also compresses this information, which is beneficial to real-time processing. The choice of feature-level fusion method depends on the nature of the images and varies depending on fusion applications involved [29].

Feature-level fusion plays a significant role in information fusion processing. The key advantages of feature fusion can be summarized in terms of the following two aspects: on the one hand, it can obtain the most distinguishing information from the original multi-feature sets involved in the fusion; on the other hand, it can eliminate the redundant information arising from the correlation

between different feature sets, thereby facilitating real-time subsequent decision-making. In other words, feature fusion can enable the maximum efficiency and the minimum dimensional feature vector sets to be obtained, which facilitates the final decision [100].

2.2.1. Feature Selection-Based and Feature Extraction-Based Techniques

Generally speaking, existing feature fusion techniques can be subdivided into two key categories: namely, feature selection-based and feature extraction-based.

In feature selection-based fusion methods, all sets of feature vectors are first grouped together, after which an appropriate feature selection method is utilized. Battiti [101] proposed a method using supervised neural networks, [102] presented a fusion method based on dynamic programming, and Shi and Zhang provided a method based on support vector machines (SVM) [103].

In the feature extraction-based methods, moreover the multiple feature vector sets are combined into a single set of feature vectors, which are input into a feature extractor for fusion [100] [104]. Moreover, the classical feature combination algorithm is feature extraction-based, meaning that it groups multiple sets of feature vectors into one union-vector (or super-vector) [104].

In [100], the union-vector-based feature fusion method is defined as serial feature fusion, while the feature fusion method based on the complex vector is called parallel feature fusion. Qin and Yung devised a method that used localized maximum-margin learning to fuse different types of features during BOCVW modeling for eventual scene classification [105]. A new feature extraction method, based on feature fusion, is proposed in [106] based on the idea of canonical correlation analysis (CCA). In order to classify very high resolution (VHR) satellite imagery, Huang *et al.* [107] presented a multi-scale feature fusion methodology based on wavelet transform. Furthermore, in order to detect nighttime vehicles effectively, a novel bio-inspired image enhancement method with a weighted feature fusion technique was devised by Kuang and Zhang [108]. Yang *et al.* [109] presented a novel feature combination strategy, the idea behind which involves combining two sets of feature vectors using a complex vector instead of a real union-vector. Fernandezbeltran *et al.* proposed a novel pLSA-based image fusion method designed to reveal multimodal modal patterns in SAR and MSI data, thus effectively merging and classifying Sentinel-1 and Sentinel-2 remote sensing data [110]. The work of [111] proposed a novel label fusion method, referred to as Feature Sensitive Label Prior (FSLP), which takes both the variety and consistency of the different features into account as was utilized to gather atlas priors.

2.2.2. Feature-Level Fusion Methods in Other Domains

Classical feature fusion methods represent feature data as real numbers. These and include inference [109] [112] and estimation-based methods [113] [114], as well as methods that employ certain types of feature data [115] [116] [117] [118]. Peng *et al.* [119] presented a quantum-inspired feature fusion method that uses

the maximum mutual von Neumann entropy to define the relationship between quantized feature samples. Peng and Deng developed a quantum-inspired feature fusion method with collision and reaction mechanisms [120]. A new quantum-inspired feature fusion method, based around maximum fidelity in order to better improve the completeness and conciseness of existing feature data, was developed in [121]. In addition, a multi-modal feature fusion-based framework was proposed to improve geographic image annotation: this algorithm leverages a low-to-high learning flow for both the deep and shallow modality features, with the overall goal of achieving effective geographic images representation [122].

2.3. Decision-Level Fusion

The goal of decision-level image fusion is to obtain decisions from each source image, then combine these decisions into a global optimal decision reference to certain criteria and the credibility of each decision. Decision-level image fusion is the highest level of information fusion, and the results of this process provide a basis for command and control decisions. At this level of the fusion process, the initial judgments and conclusions for the same target are first established for each sensor image, after which the decision from each source image is processed: finally, the decision-level fusion process is executed in order to obtain the final joint image [4].

A variety of logical reasoning methods, statistical methods, information theory methods, and so forth can be used for decision-level fusion; these include Bayesian inference, Dempster-Shafer (D-S) evidence theory, consensus-based hybrid methods, joint measures methods, voting, fuzzy decision rules (such as fuzzy integral [123], and fuzzy logic [124]), rank-based, cluster analysis, Composite Decision Fusion (CDF), and neural networks. While decision-level convergence has good real-time and fault tolerance, its preprocessing cost is high and information loss is the most. **Table 2** illustrates several decision-level fusion methods which are presented in more details in the following.

Table 2. Different methods at decision-level fusion.

Algorithms	References
Bayesian inference	[125] [126]
Dempster-Shafer (D-S)	[127] [128]
Hybrid methods based on consensus	[129]
Joint Measures Method (JMM)	[130]
Voting	[131]
Fuzzy decision rule	[123] [124] [132]
Rank-based	[133]
Cluster analysis	[134]
Composite Decision Fusion (CDF)	[135]
Adaptive decision fusion based on the local scale of the structure	[136]
Neural network	[137] [138]

Bayesian inference is an abstract concept that provides only a probabilistic framework for recursive state estimation. Grid-based filters, particle filters (PFs), KF and EKF are all Bayesian-type methods [125]. Although Bayesian inference can effectively solve most fusion problems, the Bayesian method does not consider uncertainty; therefore, errors and complexity may be introduced into the posterior probability measurement [139]. In [126], Bayesian rules were utilized to fuse the results from the fast sparse representation classifier and support vector machine classifier for SAR image target recognition purposes.

The D-S evidence method, which is an extension of Bayesian inference, can be used without the need for prior probability distributions, and can thus being able to deal with uncertainty and overcome certain drawbacks. D-S reasoning clearly indicates that despite a lack of information about propositional probabilities, it can solve some problems that cannot be solved by probability theory.

A new joint measures-based approach to multi-system/sensor decision fusion was described in [130]. The authors extracted the mathematical properties of multi-sensor local classification results and used them to model the classifier performance in terms of plausibility and correctness. Subsequently, the plausibility and correctness distribution vectors and matrices are established to introduce two DS improvement methods, namely the DS (CM) and DS (PM) methods. After that, the authors introduced the joint measures decision fusion method, which is based on the combined use of these two measures. As stated in the paper, the proposed JMM can deal with any decision fusion problem that arises in the event of uncertain local classifiers results as well as clear local classifier results.

The work of [132] introduced a decision fusion method for the classification of urban remote sensing images, consisting of two key steps. In the first step, data was processed by each classifier separately, for each pixel, memberships degree for the considered classes were provided. In the second step, a fuzzy decision rule was used to aggregate the results provided by the algorithms depending on the function of the classifier. A general framework based on the definition of two accuracy measures was then proposed to combine the information of several individual classifiers via multi-class classification. The first measure was a point-by-point measure which estimates, the reliability of the information provided by each classifier for each pixel, while the second measure estimates the global accuracy of each classifier. Finally, the results were aggregated with an adaptive fuzzy operator governed by these two measures.

The ranking-based decision fusion algorithm is an example of a typical decision fusion algorithm. Huan and Pan proposed three different decision fusion strategies: namely, multi-view decision fusion strategy, multi-feature decision fusion strategy and multi-classifier decision fusion strategy. Their work proved that the performance of SAR image target recognition could be improved through the use of these strategies [133].

In [134], a strategy for the joint classification of multiple segmentation levels from multi-sensor imagery, using SAR and optical data, is introduced. Firstly,

the two data-sets were segmented separately to create independent aggregation levels at different scales; next, each individual level from the two data-sets was then preclassified using a support vector machine (SVM). Subsequently, the original outputs of each SVM (*i.e.*, images showing the distances of the pixels to the hyperplane fitted by the SVM) were used in a decision fusion to determine the final classes. The decision fusion strategy was based on the application of an additional classifier, which was applied to the pre-classification results.

The work of [135] proposed a Composite Decision Fusion (CDF) strategy. This approach combined a state-of-the-art kernel-based decision fusion technique with the popular composite kernel classification approach, enabling it to deal with the combined classification of a color image with high spatial resolution and a lower-spatial-resolution hyperspectral image of the same scene.

3. Fusion Strategy

Fusion strategies are important in the context of data fusion tasks from different sensors, and play a significant role in improving the quality of fused images. Therefore, the design of more advanced fusion strategies is anticipated to be another research direction in the image fusion field.

In [23], the authors reviewed some classical fusion strategies: namely, coefficient, window, and region based activity level measurement (CAM, WAM, RAM), window and region based consistency verification (WRCV), choose-max and weighted-average based coefficient combining methods (CM-WACC), etc. Moreover, these strategies are widely employed in multi-scale decomposition based image fusion algorithms. Moreover, in order to achieve better fusion performance, scholars have improved the traditional fusion rules and designed some novel one. In [140], image fusion was expressed as an optimization problem, while an information theory method was applied in a multi-scale framework to obtain fusion results. For their part, Zheng *et al.* use principal component analysis to fuse the basic components [141]. In this method, a choose-max (CM) scheme and a neighborhood morphological processing step were used to increase the consistency of coefficient selection, which reduced the distortion in the fused image. Through the use of a local optimization method, a novel guided filtering approach based on the weighted average method was presented to fuse the multi-scale decomposition of input images [54].

Generally speaking, most sparse representation-based image fusion methods are designed based on traditional fusion strategies, such as CM [88] [142], weighted average-based coefficient combination (WACC) [143] [144] [145], substitution of sparse coefficients (SSC) [146], and WAM [147]. To improve the fusion performance of the sparse representation-based methods, a spatial context-based weighted average (SCWA) was utilized on a sparse representation-based image fusion method; this approach considers not only the detailed information regarding each image patch, but also that regarding spatial neighbors [64].

Similar to the multi-scale decomposition-based fusion methods, weighted averages based on machine learning (MLWA) [148], block and region-based activity level measurement (BRAM) [149] [150] [151], model-based methods (MM) [152], SCWA [153], and component substitution (CS) [154] [155] [156] have been applied to fusion methods in other domains and in combination with different transforms, and have achieved the goal of improving fusion performance.

Many fusion strategies employ pixel-level image fusion. However, in most practical applications, individuals focus on the region-level of the image objects. Therefore, region-level information should also be considered during image fusion processing. Region-based rules are based on the spatial, inter-scale and intra-scale dependencies of images and can consider their low-level and middle-level structures of images. Thereby, region-based strategies have been widely used in image fusion applications.

4. Fusion Performance Evaluation

Generally speaking, a good fusion method should have the following characteristics: 1) the fused image should be able to preserve most of the complementary and useful information contained in the input images; 2) the fusion method should not produce visual artifacts that may distract the human observer or disrupt further processing tasks; 3) the fusion method should be robust to certain imperfect conditions, such as mis-registration and noise [20].

As illustrated in **Figure 4**, the quality and performance evaluation of image fusion can be divided into subjective and objective evaluation. The former can be divided into two main classes, namely interpretability subjective evaluation and forced-choice subjective evaluation. Moreover, objective evaluation can be categorized into information theory-based metrics, image feature-based metrics, image structural similarity-based metrics and human perception-inspired metrics.

The subjective evaluation method—that is, the subjective visual judgment method—evaluates the image quality according to the subjective feeling of the person. However, due to the high cost and the difficulty associated with controlling various human factors (e.g., individual differences, personal perception, biases, etc.), extensive subjective evaluation is not always feasible.

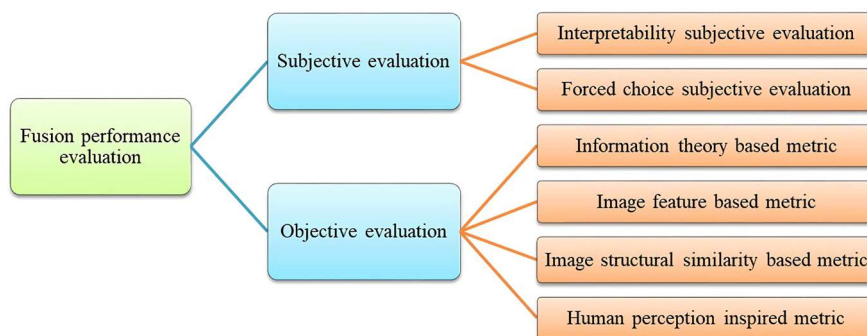


Figure 4. Image fusion example.

The objective evaluation method uses certain mathematical models to simulate the way the human eye perception the fused image and count the amount of image features, contents, or information transferred from the input images to the fused image, enabling it to quantitatively evaluate the quality of the fused image. Objective computational models, also known as fusion metrics, can reveal certain inherent properties of the fusion process or the fused image. The challenge associated with the use of these fusion metrics is as follows: while we can judge the quality of the fused image and source images by comparing the metric values, it is difficult to clarify the significance of the difference between the two index values, such as 0.78 and 0.79 [29]. A summary of the available fusion metrics is presented in **Table 3**, while these metrics are further summarized and compared in [155]. **Table 4** illustrates the expression of the specific metrics shown in **Table 3**. The specific meaning of each fusion evaluation metric and indicator expression are detailed in [155]-[170].

In **Table 4**, $MI(A, F)$ and $MI(B, F)$ are the mutual information between input image A and fused image F and the mutual information between input image B and fused image F , respectively. $H(A)$, $H(B)$ and $H(F)$ refer to the information entropy of image A , B and fused image F , respectively.

In $Q_{TE}^q = I^q(A, F) + I^q(B, F)$, $I^q(A, F)$ and $I^q(B, F)$ represent Tsallis entropy of input image A and fused image F , and Tsallis entropy of input image B and fused image F , respectively.

For $Q_{NCIE} = 1 + \sum_{i=1}^3 \frac{\kappa_i}{3} \log_l \frac{\kappa_i}{3}$, κ_i is the eigenvalue of the nonlinear correlation matrix R , where $R = \begin{pmatrix} NNC_{AA} & NNC_{AB} & NNC_{AF} \\ NNC_{BA} & NNC_{BB} & NNC_{BF} \\ NNC_{FA} & NNC_{FB} & NNC_{FF} \end{pmatrix}$.

Table 3. Summary of fusion evaluation metrics [4] [29] [155] [156].

Category	Specific Metrics
Information theory-based metric	Q_{MI} : Normalized mutual information [157]
	Q_{TE} : Mutual information based on Tsallias entropy [158] [159]
	Q_{NCIE} : Nonlinear correlation information [160]
	Q_{FMI} : Feature Mutual Information metric [156]
Image feature-based metric	Q_G : Gradient-Based Fusion Performance [161]
	Q_p : Image Fusion Metric-Based on Phase Congruency [162] [163]
	Q_M : Image Fusion Metric-Based on a Multi-scale Scheme [164]
Image structural similarity-based metric	Q_{SF} : Image Fusion Metric-Based on Spatial Frequency [165]
	Q_C : Cvejie’s Metric [166]
	Q_S : Piella’s Metric [167]
Human perception inspired metric	Q_Y : Yang’s Metric [168]
	Q_{CB} : Chen-Blum Metric [169]
	Q_{CV} : Chen-Varshney Metric [170]

Table 4. The expression of the evaluation indices.

Specific Metrics	Indicator expression
Q_{MI}	$Q_{MI} = 2 \left[\frac{MI(A, F)}{H(A) + H(F)} + \frac{MI(B, F)}{H(B) + H(F)} \right]$
Q_{TE}^g	$Q_{TE}^g = I^g(A, F) + I^g(B, F)$
Q_{TE}	$Q_{TE} = \frac{I^g(A, F) + I^g(B, F)}{H(A) + H(B) - I^g(A, B)}$
Q_{NCE}	$Q_{NCE} = 1 + \sum_{i=1}^3 \frac{\kappa_i}{3} \log_{\frac{\kappa_i}{3}}$
Q_{FMI}	$Q_{FMI} = \frac{I(F, A)}{H(A) + H(F)} + \frac{I(F, B)}{H(B) + H(F)}$
Q_G	$Q_G = \frac{\sum_n \sum_m [Q^{AF}(i, j)\omega^A(i, j) + Q^{BF}(i, j)\omega^B(i, j)]}{\sum_n \sum_m (\omega^A(i, j) + \omega^B(i, j))}$
Q_P	$Q_P = (\max(C_{AF}^p, C_{BF}^p, C_{SF}^p))^{\alpha} \cdot (\max(C_{AF}^m, C_{BF}^m, C_{SF}^m))^{\beta} \cdot (\max(C_{AF}^s, C_{BF}^s, C_{SF}^s))^{\gamma}$
Q_M	$Q_M = \prod_{s=1}^N \left(\frac{\sum_m \sum_n (EP_s^{AF}(m, n)\omega_s^A(m, n) + EP_s^{BF}(m, n)\omega_s^B(m, n))}{\sum_m \sum_n (\omega_s^A(m, n) + \omega_s^B(m, n))} \right)^{\alpha_s}$
Q_{SF}	$Q_{SF} = (SF_p - SF_r) / SF_R$
Q_C	$Q_C = \sum_{\omega \in W} sim(A, B, F \omega) Q(A, F \omega) + (1 - sim(A, B, F \omega)) Q(B, F \omega)$ $= \sum_{\omega \in W} sim(A, B, F \omega) (Q(A, F \omega) - Q(B, F \omega)) + Q(B, F \omega)$
Q_S	$Q_S = \frac{1}{ W } \sum_{\omega \in W} [\lambda(\omega) Q_0(A, F \omega) + (1 - \lambda(\omega)) Q_0(B, F \omega)]$
Q_Y	$Q_Y = \begin{cases} \lambda(\omega) SSIM(A, F \omega) + (1 - \lambda(\omega)) SSIM(B, F \omega), & SSIM(A, B \omega) \geq 0.75 \\ \max\{SSIM(A, F \omega), SSIM(B, F \omega)\}, & SSIM(A, B \omega) < 0.75 \end{cases}$
Q_{CB}	$Q_{CB} = \overline{\lambda_A(i, j) Q_{AF}(i, j) + \lambda_B(i, j) Q_{BF}(i, j)}$
Q_{CV}	$Q_{CV} = \frac{\sum_{i=1}^L (\lambda(I_A^{W_i}) D(I_A^{W_i}, I_F^{W_i}) + \lambda(I_B^{W_i}) D(I_B^{W_i}, I_F^{W_i}))}{\sum_{i=1}^L (\lambda(I_A^{W_i}) + \lambda(I_B^{W_i}))}$

For $Q_{FMI} = \frac{I(F, A)}{H(A) + H(F)} + \frac{I(F, B)}{H(B) + H(F)}$, $I(F, A)$ and $I(F, B)$ are the amount of feature information, which are individually measured by means of MI .

In $Q_G = \frac{\sum_n \sum_m [Q^{AF}(i, j)\omega^A(i, j) + Q^{BF}(i, j)\omega^B(i, j)]}{\sum_n \sum_m (\omega^A(i, j) + \omega^B(i, j))}$, the edge information preservation value $Q^{AF}(i, j) = \frac{\Gamma_g \cdot \Gamma_o}{\left(1 + e^{\kappa_g(G^{AF}(i, j) - \sigma_g)}\right) \cdot \left(1 + e^{\kappa_o(\Delta^{AF}(i, j) - \sigma_o)}\right)}$, here,

G^{AF} and Δ^{AF} represent the relative strength and orientation values between input image A and fused image F , respectively. And the constants Γ_g , κ_g , σ_g and Γ_o , κ_o , σ_o determine the shape of the sigmoid functions used to form

the edge strength and orientation preservation value. These equations are also valid for the joint distribution between image B and fused image F . In addition, the weighting coefficients are defined as $\omega^A(i, j) = [g_A(i, j)]^L$ and $\omega^B(i, j) = [g_B(i, j)]^L$, respectively. Herein, L is a constant value.

$$\text{In } Q_p = \left(\max(C_{AF}^p, C_{BF}^p, C_{SF}^p) \right)^\alpha \cdot \left(\max(C_{AF}^M, C_{BF}^M, C_{SF}^M) \right)^\beta \cdot \left(\max(C_{AF}^m, C_{BF}^m, C_{SF}^m) \right)^\gamma,$$

$$C_{xy}^l = \frac{\sigma_{xy}^l + C}{\sigma_x^l \sigma_y^l + C}$$

stands for the correlation coefficients between two sets x and y , and p, M, m represent phase congruency, maximum and minimum moments, respectively, and α, β, γ refer to index parameters which can be adjusted according to the importance of the three components.

For $Q_M = \prod_{s=1}^N \left(\frac{\sum_m \sum_n (EP_s^{AF}(m, n) \omega_s^A(m, n) + EP_s^{BF}(m, n) \omega_s^B(m, n))}{\sum_m \sum_n (\omega_s^A(m, n) + \omega_s^B(m, n))} \right)^{\alpha_s}$, the edge information

$$EP_s^{AF}(m, n) = \frac{e^{-|LH_s^A(m, n) - LH_s^F(m, n)|} + e^{-|HL_s^A(m, n) - HL_s^F(m, n)|} + e^{-|HH_s^A(m, n) - HH_s^F(m, n)|}}{3},$$

which is retrieved from the high-pass and band-pass components of the decomposition. $\omega_s^A(m, n) = LH_{A_s}^2(m, n) + HL_{A_s}^2(m, n) + HH_{A_s}^2(m, n)$ is obtained by the high-frequency energy of the input image A . These equations are also valid for the joint distribution between image B and fused image F .

In $Q_{SF} = (SF_F - SF_R) / SF_R$, $SF = \sqrt{(RF)^2 + (CF)^2 + (MDF)^2 + (SDF)^2}$, where RF, CF, MDF and SDF are the four first-order gradients along four directions. And distance weight $\omega_d = 1/\sqrt{2}$. These four reference gradients are obtained by calculating the maximum value of the absolute gradient value between the input images A and B in four directions:

$$Grad^D(R(i, j)) = \max\{\text{abs}[Grad^D(A(i, j))], \text{abs}[Grad^D(B(i, j))]\},$$

where $D = \{H, V, MD, SD\}$ represents horizontal, vertical, main diagonal, and secondary diagonal, respectively.

$$\text{For } Q_C = \sum_{\omega \in W} \text{sim}(A, B, F | \omega) Q(A, F | \omega) + (1 - \text{sim}(A, B, F | \omega)) Q(B, F | \omega)$$

$$= \sum_{\omega \in W} \text{sim}(A, B, F | \omega) (Q(A, F | \omega) - Q(B, F | \omega)) + Q(B, F | \omega),$$

the local $Q(A, F | \omega)$ and $Q(B, F | \omega)$ value were supposed to be calculated in

$$\text{the sliding window } W, \text{ and } \text{sim}(A, B, F | \omega) = \begin{cases} 0, & \frac{\sigma_{AF}}{\sigma_{AF} + \sigma_{BF}} < 0 \\ \frac{\sigma_{AF}}{\sigma_{AF} + \sigma_{BF}}, & 0 \leq \frac{\sigma_{AF}}{\sigma_{AF} + \sigma_{BF}} \leq 1. \\ 1, & \frac{\sigma_{AF}}{\sigma_{AF} + \sigma_{BF}} > 1 \end{cases}$$

$$\text{In } Q_S = \frac{1}{|W|} \sum_{\omega \in W} [\lambda(\omega) Q_0(A, F | \omega) + (1 - \lambda(\omega)) Q_0(B, F | \omega)],$$

$$\lambda(\omega) = \frac{s(A | \omega)}{s(A | \omega) + s(B | \omega)},$$

where $s(A | \omega)$ and $s(B | \omega)$ were local measure

of image salience.

$SSIM(A, F | \omega)$ and $SSIM(B, F | \omega)$ refer to the structural similarity index measure (SSIM) for input image A and the fused image F and the SSIM between input image B and image F , respectively.

In $Q_{CB} = \lambda_A(i, j)Q_{AF}(i, j) + \lambda_B(i, j)Q_{BF}(i, j)$, the saliency map for image A is defined as: $\lambda_A(i, j) = \frac{C'_A(i, j)}{C'_A(i, j) + C'_B(i, j)}$, while the information preservation

$$\text{value } Q_{AF}(i, j) = \begin{cases} \frac{C'_A(i, j)}{C'_F(i, j)}, & C'_A(i, j) < C'_F(i, j) \\ \frac{C'_F(i, j)}{C'_A(i, j)}, & \text{others} \end{cases} \quad \text{and } C'_A(i, j) = \frac{t(C_A)^p}{h(C_A)^q + Z}$$

where, t, p, h, q and Z are the real scalar parameters that determine the shape of the nonlinearity of the masking function. These equations are also valid for the joint distribution between image B and fused image F .

$$\text{For } Q_{CV} = \frac{\sum_{l=1}^L (\lambda(I_A^{W_l})D(I_A^{W_l}, I_F^{W_l}) + \lambda(I_B^{W_l})D(I_B^{W_l}, I_F^{W_l}))}{\sum_{l=1}^L (\lambda(I_A^{W_l}) + \lambda(I_B^{W_l}))}, \text{ where the saliency}$$

of the local region $\lambda(A^W) = \sum_{\omega \in W} G_A(\omega)^\alpha$, here, $G_A(\omega)$ is the edge strength in the local region and α is a constant. The local similarity measure

$$D(I_A^W, I_F^W) = \frac{1}{|W|} \sum_{\omega \in W} \hat{f}_r^W(i, j)^2 \text{ where, } |W| \text{ is the number of pixels in the local}$$

region W and r is the input images A and B .

5. Future and Conclusions

Although scholars have proposed a variety of image fusion and objective performance evaluation methods, there are still some problems with these approaches at present. Hence, it is still necessary to improve and innovate the image fusion algorithms in order to adapt to various applications. Potential future researches include the following:

1) Research on the application of image fusion. One of the key points is that, for different areas of image fusion application, the imaging mechanism of the corresponding imaging system and the physical characteristics of the imaging sensor should be analyzed so that a better fusion effect can be obtained. The multi-source image fusion algorithm must be effectively combined with the application in order to better analyze the fusion process and obtain improved results.

2) It is necessary to research multi-scale decomposition and reconstruction methods suitable for image fusion. For multi-scale decomposition based fusion methods, the fusion effect is largely dependent on which multi-scale decomposition methods are chosen. Hence, it is very important to improve the multi-scale decomposition methods available. In addition, it is necessary to study the influence of certain internal factors on the quality of the fused image obtained via the multi-scale decomposition method, as this will help to

find or improve the multi-scale decomposition method in order to improve the quality of the image fusion.

3) It is demanded to consider the improvements to the integration guidelines. The integration guidelines are key to this type of integration method. At present, the fusion criterion is not limited to simple methods such as coefficient selection and weighted averaging; people are studying ways to integrate human neural networks and other phenomena that can simulate human visual perception or reflect images.

4) It is essential to overcome the effects of mismatch and noise interference on the fusion results. In addition, most image fusion algorithms often assume that the source images have been accurately registered and that no noise pollution is received. In practical applications, however, the multi-sensor images are not only likely to contain mismatches, but may also be impacted by the effects of noise. Therefore, it is necessary to overcome the influence of mismatch and noise interference on the fusion results.

5) It is urgent to research new multi-scale fusion evaluation indices. At present, multi-scale decomposition-based image fusion algorithms are mostly based on the “energy” of pixels (or windows, or regions) as a fusion measure index, which is used to reflect the information contained in the coefficients at each resolution. However, the use of such a fusion measure is not always appropriate. To this end, it is necessary to combine the imaging characteristics of the source image to find a fusion measure index that can more accurately reflect the relative importance of the coefficients at each resolution.

6) It is needed to explore the application of deep learning (DL) in image fusion. In recent years, the use of DL has resulted in many breakthroughs in various computer vision and image processing problems, such as classification, segmentation and object detection. Deep learning-based research has also become an active topic in the field of image fusion over the past three years. The key issues and challenges associated with DL-based fusion algorithms are as follows: firstly, the design of network architecture, including the input, inner and output architecture; secondly, the generation of training datasets; thirdly, the application of conventional image fusion technology to specific fusion problems. At present, although DL-based image fusion research has achieved good results, but it still in the initial stage, there is remains a huge potential for its future development in this field. The application of DL in image fusion should be further explored from the three key problems and challenges of DL-based fusion algorithms.

Multi-sensor image fusion is an effective technology for use in fusing complementary information from multi-sensor images into fusion images. This complementary information can enhance the visibility of human eyes as while as complement each other's limitations, making multi-sensor image fusion a hot research topic in the field of remote sensing image processing. Due to increasing social demand and the rapid developments in science and technology, many experts and scholars have proposed a large number of fusion algorithms

that are well-suited to the corresponding application fields and obtain good fusion results. However, there are still many challenges associated with image fusion and objective fusion performance evaluation for a number of reasons; these include differences resolution between the source images, noise, imperfect environmental conditions, diversity of applications, computational complexity or the limitations of existing technologies. It is therefore, it is expected that new research and practical applications based on image fusion will continue to grow and develop over the next few years.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Liu, Y., Chen, X., Wang, Z., Wang, Z.J., Ward, R.K. and Wang, X. (2018) Deep Learning for Pixel-Level Image Fusion: Recent Advances and Future Prospects. *Information Fusion*, **24**, 158-173. <https://doi.org/10.1016/j.inffus.2017.10.007>
- [2] Ma, J., Ma, Y. and Li, C. (2018) Infrared and Visible Image Fusion Methods and Applications: A Survey. *Information Fusion*, **45**, 153-178. <https://doi.org/10.1016/j.inffus.2018.02.004>
- [3] Irwin, K., Beaulne, D., Braun, A. and Fotopoulos, G. (2017) Fusion of Sar, Optical Imagery and Airborne Lidar for Surface Water Detection. *Remote Sensing*, **9**, 890. <https://doi.org/10.3390/rs9090890>
- [4] Meher, B., Agrawal, S., Panda, R. and Abraham, A. (2019) A Survey on Region-Based Image Fusion Methods. *Information Fusion*, **48**, 119-132. <https://doi.org/10.1016/j.inffus.2018.07.010>
- [5] Pohl, C. and Van Genderen, J.L. (1998) Multisensor Image Fusion in Remote Sensing: Concepts, Methods and Applications. *International Journal of Remote Sensing*, **19**, 823-854. <https://doi.org/10.1080/014311698215748>
- [6] Jerripothula, K.R., Cai, J. and Yuan, J. (2016) Image Co-Segmentation Via Saliency Co-Fusion. *IEEE Transactions on Multimedia*, **18**, 1896-1909. <https://doi.org/10.1109/TMM.2016.2576283>
- [7] Gao, M., Chen, H., Zheng, S. and Fang, B. (2019) Feature Fusion and Non-Negative Matrix Factorization Based Active Contours for Texture Segmentation. *Signal Processing*, **159**, 104-118. <https://doi.org/10.1016/j.sigpro.2019.01.021>
- [8] Lu, Q., Huang, X., Li, J. and Zhang, L. (2016) A Novel Mrf-Based Multifeature Fusion for Classification of Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, **13**, 515-519. <https://doi.org/10.1109/LGRS.2016.2521418>
- [9] Li, J., Luo, L., Zhang, F., Yang, J. and Rajan, D. (2016) Double Low-Rank Matrix Recovery for Saliency Fusion. *IEEE Transactions on Image Processing*, **25**, 4421-4432. <https://doi.org/10.1109/TIP.2016.2588331>
- [10] Zhang, L. and Zhang, J. (2018) A Novel Remote-Sensing Image Fusion Method Based on Hybrid Visual Saliency Analysis. *International Journal of Remote Sensing*, **39**, 7942-7964. <https://doi.org/10.1080/01431161.2018.1479791>
- [11] Liang, G., Shivakumara, P., Lu, T. and Tan, C.L. (2015) Multi-Spectral Fusion-Based Approach for Arbitrarily Oriented Scene Text Detection in Video Image Source-map Rrx. *IEEE Transactions on Image Processing*, **24**, 4488-4501.

- <https://doi.org/10.1109/TIP.2015.2465169>
- [12] Tupin, F. and Roux, M. (2003) Detection of Building Outlines Based on the Fusion of SAR and Optical Features. *ISPRS Journal of Photogrammetry and Remote Sensing*, **58**, 71-82. [https://doi.org/10.1016/S0924-2716\(03\)00018-2](https://doi.org/10.1016/S0924-2716(03)00018-2)
- [13] Korus, P. and Huang, J. (2016) Multi-Scale Fusion for Improved Localization of Malicious Tampering in Digital Images. *IEEE Transactions on Image Processing*, **25**, 1312-1326. <https://doi.org/10.1109/TIP.2016.2518870>
- [14] Bhatnagar, G., Wu, Q.M.J. and Liu, Z. (2013) Directive Contrast Based Multimodal Medical Image Fusion in NSCT Domain. *IEEE Transactions on Multimedia*, **15**, 1014-1024. <https://doi.org/10.1109/TMM.2013.2244870>
- [15] Yamasaki, A., Takauji, H., Kaneko, S., Takeo, K. and Hidehiro, O. (2008) Denighting: Enhancement of Nighttime Images for A Surveillance Camera. *Proceedings of the 19th International Conference on Pattern Recognition*, Tampa, 8-11 December 2008, 1-4.
- [16] Errico, A., Angelino, C.V., Cicala, L., Persechino, G., Ferrara, C., Lega, M., et al. (2015) Detection of Environmental Hazards through the Feature-Based Fusion of Optical and SAR Data: A Case Study in Southern Italy. *Journal of Remote Sensing*, **36**, 3345-3367. <https://doi.org/10.1080/01431161.2015.1054960>
- [17] Du, X. and Zare, A. (2020) Multiresolution Multimodal Sensor Fusion for Remote Sensing Data with Label Uncertainty. *IEEE Transactions on Geoscience and Remote Sensing*, **58**, 2755-2769. <https://doi.org/10.1109/TGRS.2019.2955320>
- [18] Muller, A.C. and Narayanan, S. (2009) Cognitively-Engineered Multisensor Image Fusion for Military Applications. *Information Fusion*, **10**, 137-149. <https://doi.org/10.1016/j.inffus.2008.08.008>
- [19] Web of Science (2020). <http://www.webofknowledge.com>
- [20] Li, S., Kang, X., Fang, L., Hu, J. and Yin, H. (2017) Pixel-Level Image Fusion: A Survey of the State of the Art. *Information Fusion*, **33**, 100-112. <https://doi.org/10.1016/j.inffus.2016.05.004>
- [21] Tescher, A.G. (1984) Merging Images through Pattern Decomposition. *Proceedings of Spie the International Society for Optical Engineering*, **575**, 173.
- [22] Burt, P.J. and Kolczynski, R.J. (1993) Enhanced Image Capture through Fusion. *Proceedings of the 4th International Conference on Computer Vision*, Berlin, 11-14 May 1993, 173-182.
- [23] Zhang, Z. and Blum, R.S. (1999) A Categorization of Multiscale-Decomposition-Based Image Fusion Schemes with a Performance Study for a Digital Camera Application. *Proceedings of the IEEE*, **87**, 1315-1326. <https://doi.org/10.1109/5.775414>
- [24] Goshtasby, A. and Nikolov, S.G. (2007) Guest Editorial: Image Fusion: Advances in the State of the Art. *Information Fusion*, **8**, 114-118. <https://doi.org/10.1016/j.inffus.2006.04.001>
- [25] Thomas, C., Ranchin, T., Wald, L. and Chanussot, J. (2008) Synthesis of Multispectral Images to High Spatial Resolution: A Critical Review of Fusion Methods Based on Remote Sensing Physics. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, 1301-1312. <https://doi.org/10.1109/TGRS.2007.912448>
- [26] Vivone, G., Alparone, L., Chanussot, J., Mura, M.D., Garzelli, A., Licciardi, G., et al. (2015) A Critical Comparison among Pansharpening Algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, **53**, 2565-2586. <https://doi.org/10.1109/TGRS.2014.2361734>

- [27] Zhang, J. (2010) Multi-Source Remote Sensing Data Fusion: Status and Trends. *International Journal of Image and Data Fusion*, **1**, 5-24. <https://doi.org/10.1080/19479830903561035>
- [28] James, A.P. and Dasarathy, B.V. (2014) Medical Image Fusion: A Survey of the State of the Art. *Information Fusion*, **19**, 4-19. <https://doi.org/10.1016/j.inffus.2013.12.002>
- [29] Liu, Z., Blasch, E., Bhatnagar, G., John, V., Wu, W. and Blum, R.S. (2018) Fusing Synergistic Information from Multi-Sensor Images: An Overview from Implementation to Performance Assessment. *Information Fusion*, **42**, 127-145. <https://doi.org/10.1016/j.inffus.2017.10.010>
- [30] Azarang, A., Manoochehri, H.E. and Kehtarnavaz, N. (2019) Convolutional Autoencoder-Based Multispectral Image Fusion. *IEEE Access*, **7**, 35673-35683. <https://doi.org/10.1109/ACCESS.2019.2905511>
- [31] Jin, H., Jiao, L., Liu, F. and Qi, Y. (2008) Fusion of Infrared and Visual Images Based on Contrast Pyramid Directional Filter Banks Using Clonal Selection Optimizing. *Optical Engineering*, **47**, Article ID: 027002. <https://doi.org/10.1117/1.2857417>
- [32] Jin, H., Xi, Q., Wang, Y. and Hei, X. (2015) Fusion of Visible and Infrared Images Using Multi-Objective Evolutionary Algorithm Based on Decomposition. *Infrared Physics & Technology*, **71**, 151-158. <https://doi.org/10.1016/j.infrared.2015.01.022>
- [33] Xu, H., Wang, Y., Wu, Y. and Qian, Y. (2016) Infrared and Multi-Type Images Fusion Algorithm Based on Contrast Pyramid Transform. *Infrared Physics & Technology*, **78**, 133-146. <https://doi.org/10.1016/j.infrared.2016.07.016>
- [34] Liu, Z., Tsukada, K., Hanasaki, K., Ho, Y. and Dai, Y. (2001) Image Fusion by Using Steerable Pyramid. *Pattern Recognition Letters*, **22**, 929-939. [https://doi.org/10.1016/S0167-8655\(01\)00047-2](https://doi.org/10.1016/S0167-8655(01)00047-2)
- [35] Petrovic, V. and Xydeas, C. (2004) Gradient-based Multiresolution Image Fusion. *IEEE Transactions on Image Processing*, **13**, 228-237. <https://doi.org/10.1109/TIP.2004.823821>
- [36] Li, H., Manjunath, B.S. and Mitra, S.K. (1995) Multisensor Image Fusion Using the Wavelet Transform. *Graphical Models and Image Processing*, **57**, 235-245. <https://doi.org/10.1006/gmip.1995.1022>
- [37] Chipman, L.J., Orr, T.M. and Graham, L.N. (1995) Wavelets and Image Fusion. *Proceedings of the International Conference on Image Processing*, Vol. 3, 248-251.
- [38] Zhang, Y., De Backer, S. and Scheunders, P. (2009) Noise-Resistant Wavelet-Based Bayesian Fusion of Multispectral and Hyperspectral Images. *IEEE Transactions on Geoscience and Remote Sensing*, **47**, 3834-3843. <https://doi.org/10.1109/TGRS.2009.2017737>
- [39] Wang, H., Peng, J. and Wu, W. (2003) A Fusion Algorithm of Remote Sensing Image Based on Discrete Wavelet Packet. *Proceedings of the 2nd International Conference on Machine Learning and Cybernetics*, Xi'an, 2-5 November 2003, 2557-2562.
- [40] Zou, Y., Liang, X. and Wang, T. (2013) Visible and Infrared Image Fusion Using the Lifting Wavelet. *Indonesian Journal of Electrical Engineering and Computer Science*, **11**, 6290-6295. <https://doi.org/10.11591/telkomnika.v11i11.2898>
- [41] Lewis, J.J., Ocallaghan, R.J., Nikolov, S.G., Bull, D. and Canagarajah, N. (2007) Pixel- and Region-Based Image Fusion with Complex Wavelets. *Information Fusion*, **8**, 119-130. <https://doi.org/10.1016/j.inffus.2005.09.006>
- [42] Wang, Z. and Gong, C. (2017) A Multi-Faceted Adaptive Image Fusion Algorithm Using a Multi-Wavelet-Based Matching Measure in the PCNN Domain. *Applied*

- Soft Computing*, **61**, 1113-1124. <https://doi.org/10.1016/j.asoc.2017.02.035>
- [43] Shensa, M.J. (1992) The Discrete Wavelet Transform: Wedding the a Trouse and Mallat Algorithms. *IEEE Transactions Signal Process*, **40**, 2464-2482. <https://doi.org/10.1109/78.157290>
- [44] Joshi, M.V., Bruzzone, L. and Chaudhuri, S. (2006) A Model-Based Approach to Multiresolution Fusion in Remotely Sensed Images. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 2549-2562. <https://doi.org/10.1109/TGRS.2006.873340>
- [45] Jin, X., Jiang, Q., Yao, S., Zhou, D., Nie, R., Lee, S. and He, K. (2018) Infrared and Visual Image Fusion Method Based on Discrete Cosine Transform and Local Spatial Frequency in Discrete Stationary Wavelet Transform Domain. *Infrared Physics & Technology*, **88**, 1-12. <https://doi.org/10.1016/j.infrared.2017.10.004>
- [46] Nencini, F., Garzelli, A., Baronti, S. and Alparone, L. (2007) Remote Sensing Image Fusion Using the Curvelet Transform. *Information Fusion*, **8**, 143-156. <https://doi.org/10.1016/j.inffus.2006.02.001>
- [47] Ghahremani, M. and Ghassemian, H. (2015) Remote-Sensing Image Fusion Based on Curvelets and ICA. *Journal of Remote Sensing*, **36**, 4131-4143. <https://doi.org/10.1080/01431161.2015.1071897>
- [48] Do, M.N. and Vetterli, M. (2005) The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. *IEEE Transactions on Image Processing*, **14**, 2091-2106. <https://doi.org/10.1109/TIP.2005.859376>
- [49] Chang, X., Jiao, L., Liu, F. and Xin, F. (2010) Multicontourlet-Based Adaptive Fusion of Infrared and Visible Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, **7**, 549-553. <https://doi.org/10.1109/LGRS.2010.2041323>
- [50] Li, T. and Wang, Y. (2011) Biological Image Fusion Using a NSCT Based Variable-Weight Method. *Information Fusion*, **12**, 85-92. <https://doi.org/10.1016/j.inffus.2010.03.007>
- [51] Easley, G.R., Labate, D. and Lim, W. (2008) Sparse Directional Image Representations Using the Discrete Shearlet Transform. *Applied and Computational Harmonic Analysis*, **25**, 25-46. <https://doi.org/10.1016/j.acha.2007.09.003>
- [52] Xu, J., Yang, L. and Wu, D. (2010) Ripplet: A New Transform for Image Processing. *Journal of Visual Communication and Image Representation*, **21**, 627-639. <https://doi.org/10.1016/j.jvcir.2010.04.002>
- [53] Ghahremani, M. and Ghassemian, H. (2015) Remote Sensing Image Fusion Using Ripplet Transform and Compressed Sensing. *IEEE Geoscience and Remote Sensing Letters*, **12**, 502-506. <https://doi.org/10.1109/LGRS.2014.2347955>
- [54] Li, S., Kang, X. and Hu, J. (2013) Image Fusion with Guided Filtering. *IEEE Transactions on Image Processing*, **22**, 2864-2875. <https://doi.org/10.1109/TIP.2013.2244222>
- [55] Farbman, Z., Fattal, R., Lischinski, D. and Szeliski, R. (2008) Edge-Preserving Decompositions for Multi-Scale Tone and Detail Manipulation. *ACM Transactions on Graphics*, **27**, 67. <https://doi.org/10.1145/1360612.1360666>
- [56] Hu, J. and Li, S. (2012) The Multiscale Directional Bilateral Filter and Its Application to Multisensor Image Fusion. *Information Fusion*, **13**, 196-206. <https://doi.org/10.1016/j.inffus.2011.01.002>
- [57] Zhou, Z., Wang, B., Li, S. and Dong, M. (2016) Perceptual Fusion of Infrared and Visible Images through a Hybrid Multi-Scale Decomposition with Gaussian and Bilateral Filters. *Information Fusion*, **30**, 15-26.

- <https://doi.org/10.1016/j.inffus.2015.11.003>
- [58] Zheng, S., et al. (2007) Multisource Image Fusion Method Using Support Value Transform. *IEEE Transactions on Image Processing*, **16**, 1831-1839. <https://doi.org/10.1109/TIP.2007.896687>
- [59] Redondo, R., Sroubek, F., Fischer, S. and Cristobal, G. (2009) Multifocus Image Fusion Using the Log-Gabor Transform and a Multisize Windows Technique. *Information Fusion*, **10**, 163-171. <https://doi.org/10.1016/j.inffus.2008.08.006>
- [60] Wang, Q., Li, S., Qin, H. and Hao, A. (2015) Robust Multi-Modal Medical Image Fusion via Anisotropic Heat Diffusion Guided Low-Rank Structural Analysis. *Information Fusion*, **26**, 103-121. <https://doi.org/10.1016/j.inffus.2015.01.001>
- [61] Li, S., Yang, B. and Hu, J. (2011) Performance Comparison of Different Multi-Resolution Transforms for Image Fusion. *Information Fusion*, **12**, 74-84. <https://doi.org/10.1016/j.inffus.2010.03.002>
- [62] Pradhan, P.S., King, R.L., Younan, N.H. and Holcomb, D.W. (2006) Estimation of the Number of Decomposition Levels for a Wavelet-Based Multiresolution Multi-sensor Image Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 3674-3686. <https://doi.org/10.1109/TGRS.2006.881758>
- [63] Yang, B. and Li, S. (2010) Multifocus Image Fusion and Restoration with Sparse Representation. *IEEE Transactions on Instrumentation and Measurement*, **59**, 884-892. <https://doi.org/10.1109/TIM.2009.2026612>
- [64] Zhu, X.X. and Bamler, R. (2013) A Sparse Image Fusion Algorithm with Application to Pan-Sharpening. *IEEE Transactions on Geoscience and Remote Sensing*, **51**, 2827-2836. <https://doi.org/10.1109/TGRS.2012.2213604>
- [65] Duarte, M.F., Sarvotham, S., Baron, D., Wakin, M.B. and Baraniuk, R.G. (2005) Distributed Compressed Sensing of Jointly Sparse Signals. *Proceedings of the 39th Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, 28 October-1 November 2005, 1537-1541.
- [66] Yu, N., Qiu, T., Bi, F. and Wang, A. (2011) Image Features Extraction and Fusion Based on Joint Sparse Representation. *IEEE Journal of Selected Topics in Signal Processing*, **5**, 1074-1082. <https://doi.org/10.1109/JSTSP.2011.2112332>
- [67] Liu, Z., Yin, H., Fang, B. and Chai, Y. (2015) A Novel Fusion Scheme for Visible and Infrared Images Based on Compressive Sensing. *Optics Communications*, **335**, 168-177. <https://doi.org/10.1016/j.optcom.2014.07.093>
- [68] Ma, X., Hu, S., Liu, S., Fang, J. and Xu, S. (2019) Remote Sensing Image Fusion Based on Sparse Representation and Guided Filtering. *Electronics*, **8**, 303-319. <https://doi.org/10.3390/electronics8030303>
- [69] Ma, X., Hu, S., Liu, S., Wang, J. and Xu, S. (2020) Noisy Remote Sensing Image Fusion Based on JSR. *IEEE Access*, **8**, 31069-31082. <https://doi.org/10.1109/ACCESS.2020.2973435>
- [70] Zhao, W., Lu, H. and Wang, D. (2018) Multisensor Image Fusion and Enhancement in Spectral Total Variation Domain. *IEEE Transactions on Multimedia*, **20**, 866-879. <https://doi.org/10.1109/TMM.2017.2760100>
- [71] Zhang, Y., Zhang, L., Bai, X. and Zhang, L. (2017) Infrared and Visual Image Fusion through Infrared Feature Extraction and Visual Information Preservation. *Infrared Physics & Technology*, **83** 227-237. <https://doi.org/10.1016/j.infrared.2017.05.007>
- [72] Kumar, M. and Dass, S.C. (2009) A Total Variation-Based Algorithm for Pixel-Level Image Fusion. *IEEE Transactions on Image Processing*, **18**, 2137-2143. <https://doi.org/10.1109/TIP.2009.2025006>
- [73] Shen, R., Cheng, I. and Basu, A. (2013) Qoe-Based Multi-Exposure Fusion in Hie-

- rarchical Multivariate Gaussian CRF. *IEEE Transactions on Image Processing*, **22**, 2469-2478. <https://doi.org/10.1109/TIP.2012.2236346>
- [74] Zhang, Y., Chen, L., Zhao, Z., Jia, J. and Liu, J. (2014) Multi-Focus Image Fusion Based on Robust Principal Component Analysis and Pulse-Coupled Neural Network. *Optik*, **125**, 5002-5006. <https://doi.org/10.1016/j.ijleo.2014.04.002>
- [75] Choi, J., Yu, K. and Kim, Y.I. (2011) A New Adaptive Component-Substitution-Based Satellite Image Fusion by Using Partial Replacement. *IEEE Transactions on Geoscience and Remote Sensing*, **49**, 295-309. <https://doi.org/10.1109/TGRS.2010.2051674>
- [76] Sun, J., Zhu, H., Xu, Z. and Han, C. (2013) Poisson Image Fusion Based on Markov Random Field Fusion Model. *Information Fusion*, **14**, 241-254. <https://doi.org/10.1016/j.inffus.2012.07.003>
- [77] Balasubramaniam, P. and Ananthi, V.P. (2014) Image Fusion Using Intuitionistic Fuzzy Sets. *Information Fusion*, **20**, 21-30. <https://doi.org/10.1016/j.inffus.2013.10.011>
- [78] Liu, S., Lu, Y., Wang, J., Hu, S., Zhao, J. and Zhu, Z. (2020) A New Focus Evaluation Operator Based on Max-Min Filter and Its Application in High Quality Multi-Focus Image Fusion. *Multidimensional Systems and Signal Processing*, **31**, 569-590. <https://doi.org/10.1007/s11045-019-00675-2>
- [79] Geng, P., Wang, Z., Zhang, Z. and Xiao, Z. (2012) Image Fusion by Pulse Couple Neural Network with Shearlet. *Optical Engineering*, **51**, Article ID: 067005. <https://doi.org/10.1117/1.OE.51.6.067005>
- [80] Jiang, Y. and Wang, M. (2014) Image Fusion with Morphological Component Analysis. *Information Fusion*, **18**, 107-118. <https://doi.org/10.1016/j.inffus.2013.06.001>
- [81] Ma, J., Chen, C., Li, C. and Huang, J. (2016) Infrared and Visible Image Fusion via Gradient Transfer and Total Variation Minimization. *Information Fusion*, **31**, 100-109. <https://doi.org/10.1016/j.inffus.2016.02.001>
- [82] Ma, T., Ma, J., Fang, B., Hu, F., Quan, S. and Du, H. (2018) Multi-Scale Decomposition Based Fusion of Infrared and Visible Image via Total Variation and Saliency Analysis. *Infrared Physics & Technology*, **92**, 154-162. <https://doi.org/10.1016/j.infrared.2018.06.002>
- [83] Cai, J., Cheng, Q., Peng, M. and Song, Y. (2017) Fusion of Infrared and Visible Images Based on Nonsampled Contourlet Transform and Sparse k-svd Dictionary Learning. *Infrared Physics & Technology*, **82**, 85-95. <https://doi.org/10.1016/j.infrared.2017.01.026>
- [84] Aishwarya, N. and Thangammal, C.B. (2018) Visible and Infrared Image Fusion using DTCWT and Adaptive Combined Clustered Dictionary. *Infrared Physics & Technology*, **93**, 300-309. <https://doi.org/10.1016/j.infrared.2018.08.013>
- [85] Yang, Y., Lei, W., Haung, S., Wan, S. and Que, Y. (2018) Remote Sensing Image Fusion Based on Adaptively Weighted Joint Detail Injection. *IEEE Access*, **6**, 6849-6864. <https://doi.org/10.1109/ACCESS.2018.2791574>
- [86] Wang, X., Bai, S., Li, Z., Song, R. and Tao, J. (2019) The PAN and MS Image Pan-sharpening Algorithm Based on Adaptive Neural Network and Sparse Representation in the NSST Domain. *IEEE Access*, **7**, 52508-52521. <https://doi.org/10.1109/ACCESS.2019.2910656>
- [87] Liu, S., Wang, J., Lu, Y., Li, H., Zhao, J. and Zhu, Z. (2019) Multi-Focus Image Fusion Based on Adaptive Dual-Channel Spiking Cortical Model in Nonsampled Shearlet Domain. *IEEE Access*, **7**, 56367-56388.

- <https://doi.org/10.1109/ACCESS.2019.2900376>
- [88] Liu, Y., Chen, X., Ward, R.K. and Wang, Z.J. (2016) Image Fusion with Convolutional Sparse Representation. *IEEE Signal Processing Letters*, **23**, 1882-1886. <https://doi.org/10.1109/LSP.2016.2618776>
- [89] Huang, W., Xiao, L., Wei, Z., Liu, H. and Tang, S. (2015) A New Pan-Sharpener Method with Deep Neural Networks. *IEEE Geoscience and Remote Sensing Letters*, **12**, 1037-1041. <https://doi.org/10.1109/LGRS.2014.2376034>
- [90] Zhong, J., Yang, B., Huang, G., Zhong, F. and Chen, Z. (2016) A New Pansharpening Method with Deep Neural Networks. *Sensing Imaging*, **17**, 1-16. <https://doi.org/10.1007/s11220-016-0135-6>
- [91] Masi, G., Cozzolino, D., Verdoliva, L. and Scarpa, G. (2016) Pansharpening by Convolutional Neural Networks. *Remote Sensing*, **8**, 1-22. <https://doi.org/10.3390/rs8070594>
- [92] Palsson, F., Sveinsson, J.R. and Ulfarsson, M.O. (2017) Multispectral and Hyperspectral Image Fusion Using a 3-D-Convolutional Neural Network. *IEEE Geoscience and Remote Sensing Letters*, **14**, 639-643. <https://doi.org/10.1109/LGRS.2017.2668299>
- [93] Zhao, W., Jiao, L., Ma, W., Zhao, J., Zhao, J., Liu, H., Cao, X. and Yang, S. (2017) Superpixel-Based Multiple Local CNN for Panchromatic and Multispectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **55**, 4141-4156. <https://doi.org/10.1109/TGRS.2017.2689018>
- [94] Liu, X., Jiao, L., Zhao, J., Zhao, J., Zhang, D., Liu, F., et al. (2018) Deep Multiple Instance Learning-Based Spatial-Spectral Classification for PAN and MS Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 461-473. <https://doi.org/10.1109/TGRS.2017.2750220>
- [95] Ma, J., Yu, W., Liang, P., Li, C. and Jiang, J. (2019) FusionGAN: A Generative Adversarial Network for Infrared and Visible Image Fusion. *Information Fusion*, **48**, 11-26. <https://doi.org/10.1016/j.inffus.2018.09.004>
- [96] Wei, Y., Yuan, Q., Shen, H. and Zhang, L. (2017) Boosting the Accuracy of Multispectral Image Pansharpening by Learning a Deep Residual Network. *IEEE Geoscience and Remote Sensing Letters*, **14**, 1795-1799. <https://doi.org/10.1109/LGRS.2017.2736020>
- [97] Liu, S., Wang, J., Lu, Y., Hu, S., Ma, X. and Wu, Y. (2019) Multi-Focus Image Fusion Based on Residual Network in Non-Subsampled Shearlet Domain. *IEEE Access*, **7**, 152043-152063. <https://doi.org/10.1109/ACCESS.2019.2947378>
- [98] Liu, S., Ma, J., Yin, L., Li, H., Cong, S., Ma, X. and Hu, S. (2020) Multi-Focus Color Image Fusion Algorithm Based on Super-Resolution Reconstruction and Focused Area Detection. *IEEE Access*, **8**, 90760-90778. <https://doi.org/10.1109/ACCESS.2020.2993404>
- [99] Zagoruyko, S. and Komodakis, N. (2015) Learning to Compare Image Patches via Convolutional Neural Networks. *Proceedings of the 28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Portland, 23-28 June 2013, 4353-4361. <https://doi.org/10.1109/CVPR.2015.7299064>
- [100] Yang, J., Yang, J., Zhang, D. and Lu, J. (2003) Feature Fusion: Parallel Strategy vs. Serial Strategy. *Pattern Recognition*, **36**, 1369-1381. [https://doi.org/10.1016/S0031-3203\(02\)00262-5](https://doi.org/10.1016/S0031-3203(02)00262-5)
- [101] Battiti, R. (1994) Using Mutual Information for Selecting Features in Supervised Neural Net Learning. *IEEE Transactions on Neural Networks*, **5**, 537-550. <https://doi.org/10.1109/72.298224>

- [102] Zhang, X.H. (1998) An Information Model and Method of Feature Fusion. *International Conference on Signal Processing*, Vol. 2, 1389-1392.
- [103] Shi, Y. and Zhang, T. (2001) Feature Analysis: Support Vector Machine Approaches. In: *Image Extraction, Segmentation, and Recognition*, International Society for Optics and Photonics, Bellingham, 245-251. <https://doi.org/10.1117/12.441458>
- [104] Liu, C. and Wechsler, H. (2001) A Shape- and Texture-Based Enhanced Fisher Classifier for Face Recognition. *IEEE Transactions on Image Processing*, **10**, 598-608. <https://doi.org/10.1109/83.913594>
- [105] Qin, J. and Yung, N.H.C. (2012) Feature Fusion within Local Region Using Localized Maximum-Margin Learning for Scene Categorization. *Pattern Recognition*, **45**, 1671-1683. <https://doi.org/10.1016/j.patcog.2011.09.027>
- [106] Sun, Q., Zeng, S., Liu, Y., Heng, P. and Xia, D. (2005) A New Method of Feature Fusion and Its Application in Image Recognition. *Pattern Recognition*, **38**, 2437-2448. <https://doi.org/10.1016/j.patcog.2004.12.013>
- [107] Huang, X., Zhang, L. and Li, P. (2008) A Multiscale Feature Fusion Approach for Classification of Very High Resolution Satellite Imagery Based on Wavelet Transform. *Journal of Remote Sensing*, **29**, 5923-5941. <https://doi.org/10.1080/01431160802139922>
- [108] Kuang, H., Zhang, X., Li, Y., Chan, L.L.H. and Yan, H. (2017) Nighttime Vehicle Detection Based on Bio-Inspired Image Enhancement and Weighted Score-Level Feature Fusion. *IEEE Transactions on Intelligent Transportation Systems*, **18**, 927-936. <https://doi.org/10.1109/TITS.2016.2598192>
- [109] Yang, J. and Yang, J. (2002) Generalized k-l Transform Based Combined Feature Extraction. *Pattern Recognition*, **35**, 295-297. [https://doi.org/10.1016/S0031-3203\(01\)00152-2](https://doi.org/10.1016/S0031-3203(01)00152-2)
- [110] Fernandezbeltran, R., Haut, J.M., Paoletti, M.E., Plaza, J., Plaza, A. and Pla, F. (2018) Remote Sensing Image Fusion Using Hierarchical Multimodal Probabilistic Latent Semantic Analysis. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **11**, 4982-4993. <https://doi.org/10.1109/JSTARS.2018.2881342>
- [111] Bao, S. and Chung, A.C.S. (2017) Feature Sensitive Label Fusion with Random Walker for Atlas-Based Image Segmentation. *IEEE Transactions on Image Processing*, **26**, 2797-2810. <https://doi.org/10.1109/TIP.2017.2691799>
- [112] Guerriero, M., Svensson, L. and Willett, P. (2010) Bayesian Data Fusion for Distributed Target Detection in Sensor Networks. *IEEE Transactions on Signal Processing*, **58**, 3417-3421. <https://doi.org/10.1109/TSP.2010.2046042>
- [113] Zhao, H. and Wang, Z. (2012) Motion Measurement Using Inertial Sensors, Ultrasonic Sensors, and Magnetometers with Extended Kalman Filter for Data Fusion. *IEEE Sensors Journal*, **12**, 943-953. <https://doi.org/10.1109/JSEN.2011.2166066>
- [114] Prieto, J., Mazuelas, S., Bahillo, A., Fernandez, P., Lorenzo, R.M. and Abril, E.J. (2012) Adaptive Data Fusion for Wireless Localization in Harsh Environments. *IEEE Transactions on Signal Processing*, **60**, 1585-1596. <https://doi.org/10.1109/TSP.2012.2183126>
- [115] Byun, Y. (2014) A Texture-Based Fusion Scheme to Integrate High-Resolution Satellite SAR and Optical Images. *Remote Sensing Letters*, **5**, 103-111. <https://doi.org/10.1080/2150704X.2014.880817>
- [116] Wang, L., Li, B. and Tian, L. (2014) Full Length Article: Multi-Modal Medical Image Fusion using the Inter-Scale and Intra-Scale Dependencies between Image Shift-Invariant Shearlet Coefficients. *Information Fusion*, **19**, 20-28.

- <https://doi.org/10.1016/j.inffus.2012.03.002>
- [117] Liu, Y., Liu, S. and Wang, Z. (2015) Multi-Focus Image Fusion with Dense Sift. *Information Fusion*, **23**, 139-155. <https://doi.org/10.1016/j.inffus.2014.05.004>
- [118] Zhang, H., Lin, H. and Li, Y. (2015) Impacts of Feature Normalization on Optical and SAR Data Fusion for Land use Land Cover Classification. *IEEE Geoscience and Remote Sensing Letters*, **12**, 1061-1065. <https://doi.org/10.1109/LGRS.2014.2377722>
- [119] Peng, W. and Deng, H. (2014) Quantum Inspired Method of Feature Fusion Based on Von Neumann Entropy. *Information Fusion*, **18**, 9-19. <https://doi.org/10.1016/j.inffus.2013.10.003>
- [120] Peng, W. and Deng, H. (2015) A Collision and Reaction Model of Feature Fusion: Mechanism and Realization. *IEEE Intelligent Systems*, **30**, 56-65. <https://doi.org/10.1109/MIS.2015.38>
- [121] Peng, W., Chen, A. and Sun, Y. (2017) A Quantum-Inspired Feature Fusion Method Based on Maximum Fidelity. *IEEE Intelligent Systems*, **32**, 80-87. <https://doi.org/10.1109/MIS.2017.8>
- [122] Li, K., Zou, C., Bu, S., Liang, Y., Zhang, J. and Gong, M. (2018) Multi-Modal Feature Fusion for Geographic Image Annotation. *Pattern Recognition*, **73**, 1-14. <https://doi.org/10.1016/j.patcog.2017.06.036>
- [123] Mitrakis, N.E., Topaloglou, C., Alexandridis, T., Theocharis, J.B. and Zalidis, G. (2008) Decision Fusion of Ga Self-Organizing Neuro-Fuzzy Multilayered Classifiers for Land Cover Classification Using Textural and Spectral Features. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, 2137-2152. <https://doi.org/10.1109/TGRS.2008.916481>
- [124] Wang, Y., Chen, W. and Mao, S. (2006) Multi-Sensor Decision Level Image Fusion Based on Fuzzy Theory and Unsupervised FCM. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 62000J. <https://doi.org/10.1117/12.681719>
- [125] Mahmoudi, F.T., Samadzadegan, F. and Reinartz, P. (2015) Object Recognition Based on the Context Aware Decision-Level Fusion in Multiviews Imagery. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **8**, 12-22. <https://doi.org/10.1109/JSTARS.2014.2362103>
- [126] Liu, H. and Li, S. (2013) Decision Fusion of Sparse Representation and Support Vector Machine for SAR Image Target Recognition. *Neurocomputing*, **113**, 97-104. <https://doi.org/10.1016/j.neucom.2013.01.033>
- [127] Zhu, H., Basir, O.A. and Karray, F. (2002) Data Fusion for Pattern Classification via the Dempster-Shafer Evidence Theory. *IEEE International Conference on Systems, Man and Cybernetics*, **7**, 109-110.
- [128] Wang, A., Jiang, J. and Zhang, H. (2014) Multi-Sensor Image Decision Level Fusion Detection Algorithm Based on D-S Evidence Theory. *Proceedings of the 4th International Conference on Instrumentation and Measurement, Computer, Communication and Control*, Harbin, 18-20 September 2014, 620-623. <https://doi.org/10.1109/IMCCC.2014.132>
- [129] Benediktsson, J.A., Sveinsson, J.R. and Swain, P.H. (1997) Hybrid Consensus Theoretic Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **35**, 833-843. <https://doi.org/10.1109/36.602526>
- [130] Rashidi, A.J. and Ghassemian, M.H. (2004) A New Approach for Multi-System/Sensor Decision Fusion Based on Joint Measures. *International Journal of Information Acquisition*, **1**, 109-120. <https://doi.org/10.1142/S0219878904000148>
- [131] Jimenez, L.O., Moralesmorell, A. and Creus, A. (1999) Classification of Hyperdi-

- mensional Data Based on Feature and Decision Fusion Approaches Using Projection Pursuit, Majority Voting, and Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 1360-1366. <https://doi.org/10.1109/36.763300>
- [132] Fauvel, M., Chanussot, J. and Benediktsson, J.A. (2006) Decision Fusion for the Classification of Urban Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, **44**, 2828-2838. <https://doi.org/10.1109/TGRS.2006.876708>
- [133] Huan, R. and Pan, Y. (2011) Decision Fusion for the Classification of Urban Remote Sensing Images. *IEEE Transactions on Geoscience and Remote Sensing*, **5**, 747-755.
- [134] Waske, B. and Der Linden, S.V. (2008) Classifying Multilevel Imagery from SAR and Optical Sensors by Decision Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, **46**, 1457-1466. <https://doi.org/10.1109/TGRS.2008.916089>
- [135] Thoonen, G., Mahmood, Z.H., Peeters, S. and Scheunders, P. (2012) Multisource Classification of Color and Hyperspectral Images Using Color Attribute Profiles and Composite Decision Fusion. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **5**, 510-521. <https://doi.org/10.1109/JSTARS.2011.2168317>
- [136] Luo, B., Khan, M.M., Bienvenu, T., Chanussot, J. and Zhang, L. (2013) Decision-Based Fusion for Pansharpening of Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, **10**, 19-23. <https://doi.org/10.1109/LGRS.2012.2189933>
- [137] Benediktsson, A. and Kanellopoulos, I. (1999) Classification of Multisource and Hyperspectral Data Based on Decision Fusion. *IEEE Transactions on Geoscience and Remote Sensing*, **37**, 1367-1377. <https://doi.org/10.1109/36.763301>
- [138] Waske, B. and Benediktsson, J.A. (2007) Fusion of Support Vector Machines for Classification of Multisensor Data. *IEEE Transactions on Geoscience and Remote Sensing*, **45**, 3858-3866. <https://doi.org/10.1109/TGRS.2007.898446>
- [139] Ghassemian, H. (2016) A Review of Remote Sensing Image Fusion Methods. *Information Fusion*, **32**, 75-89. <https://doi.org/10.1016/j.inffus.2016.03.003>
- [140] Hamza, A.B., He, Y., Krim, H. and Willisky, A.S. (2005) A Multiscale Approach to Pixel-Level Image Fusion. *Computer-Aided Engineering*, **12**, 135-146. <https://doi.org/10.3233/ICA-2005-12201>
- [141] Zheng, Y., Essock, E.A., Hansen, B.C. and Haun, A.M. (2007) A New Metric Based on Extended Spatial Frequency and Its Application to DWT Based Fusion Algorithms. *Information Fusion*, **8**, 177-192. <https://doi.org/10.1016/j.inffus.2005.04.003>
- [142] Yin, H., Li, S. and Fang, L. (2013) Simultaneous Image Fusion and Super-Resolution using Sparse Representation. *Information Fusion*, **14**, 229-240. <https://doi.org/10.1016/j.inffus.2012.01.008>
- [143] Gao, Z., Yang, M. and Xie, C. (2017) Space Target Image Fusion Method Based on Image Clarity Criterion. *Optical Engineering*, **56**, Article ID: 053102. <https://doi.org/10.1117/1.OE.56.5.053102>
- [144] Lu, X., Zhang, B., Zhao, Y., Liu, H. and Pei, H. (2014) The Infrared and Visible Image Fusion Algorithm Based on Target Separation and Sparse Representation. *Infrared Physics & Technology*, **67**, 397-407. <https://doi.org/10.1016/j.infrared.2014.09.007>
- [145] Kim, M., Han, D.K. and Ko, H. (2016) Joint Patch Clustering-Based Dictionary Learning for Multimodal Image Fusion. *Information Fusion*, **27**, 198-214. <https://doi.org/10.1016/j.inffus.2015.03.003>
- [146] Wang, W., Jiao, L. and Yang, S. (2014) Fusion of Multispectral and Panchromatic

- Images via Sparse Representation and Local Autoregressive Model. *Information Fusion*, **20**, 73-87. <https://doi.org/10.1016/j.inffus.2013.11.004>
- [147] Yang, B. and Li, S. (2012) Pixel-level Image Fusion with Simultaneous Orthogonal Matching Pursuit. *Information Fusion*, **13**, 10-19. <https://doi.org/10.1016/j.inffus.2010.04.001>
- [148] Li, S., Kwok, J.T., Tsang, I.W. and Wang, Y. (2004) Fusing Images with Different Focuses Using Support Vector Machines. *IEEE Transactions on Neural Networks*, **15**, 1555-1561. <https://doi.org/10.1109/TNN.2004.837780>
- [149] Bai, X., Zhang, Y., Zhou, F. and Xue, B. (2015) Quadtree-Based Multi-Focus Image Fusion Using a Weighted Focus-Measure. *Information Fusion*, **22**, 105-118. <https://doi.org/10.1016/j.inffus.2014.05.003>
- [150] Li, S. and Yang, B. (2008) Multifocus Image Fusion Using Region Segmentation and Spatial Frequency. *Image and Vision Computing*, **26**, 971-979. <https://doi.org/10.1016/j.imavis.2007.10.012>
- [151] Wang, J., Peng, J., Feng, X., He, G., Wu, J. and Yan, K. (2013) Image Fusion with Nonsampled Contourlet Transform and Sparse Representation. *Journal of Electronic Imaging*, **22**, Article ID: 043019. <https://doi.org/10.1117/1.JEI.22.4.043019>
- [152] Shen, R., Cheng, I., Shi, J. and Basu, A. (2011) Generalized Random Walks for Fusion of Multi-Exposure Images. *IEEE Transactions on Image Processing*, **20**, 3634-3646. <https://doi.org/10.1109/TIP.2011.2150235>
- [153] Mitianoudis, N. and Stathaki, T. (2007) Pixel-Based and Region-Based Image Fusion Schemes Using ICA Bases. *Information Fusion*, **8**, 131-142. <https://doi.org/10.1016/j.inffus.2005.09.001>
- [154] Zhang, Y. and Hong, G. (2005) An IHS and Wavelet Integrated Approach to Improve Pansharpening Visual Quality of Natural Colour IKONOS and Quickbird Images. *Information Fusion*, **6**, 225-234. <https://doi.org/10.1016/j.inffus.2004.06.009>
- [155] Liu, Z., Blasch, E., Xue, Z., Zhao, J., Laganier, R. and Wu, W. (2012) Objective Assessment of Multiresolution Image Fusion Algorithms for Context Enhancement in Night Vision: A Comparative Study. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **34**, 94-109. <https://doi.org/10.1109/TPAMI.2011.109>
- [156] Haghghat, M., Aghagolzadeh, A. and Seyedarabi, H. (2011) A Non-Reference Image Fusion Metric Based on Mutual Information of Image Features. *Computers & Electrical Engineering*, **37**, 744-756. <https://doi.org/10.1016/j.compeleceng.2011.07.012>
- [157] Hossny, M., Nahavandi, S. and Creighton, D. (2008) Comments on Information Measure for Performance of Image Fusion. *Electronics Letters*, **44**, 1066-1067. <https://doi.org/10.1049/el:20081754>
- [158] Cvejic, N., Canagarajah, C.N. and Bull, D. (2006) Image Fusion Metric Based on Mutual Information and Tsallis Entropy. *Electronics Letters*, **42**, 626-627. <https://doi.org/10.1049/el:20060693>
- [159] Nava, R., Cristobal, G. and Escalante-Ramrez, B. (2007) Mutual Information Improves Image Fusion Quality Assessments. SPIE News Room. <https://doi.org/10.1117/2.1200708.0824>
- [160] Wang, Q., Shen, Y. and Jin, J. (2008) Performance Evaluation of Image Fusion Techniques. *Image Fusion: Algorithms and Applications*, **19**, 469-492. <https://doi.org/10.1016/B978-0-12-372529-5.00017-2>
- [161] Xydeas, C. and Petrovic, V. (2000) Objective Image Fusion Performance Measure. *Electronics Letters*, **36**, 308-309. <https://doi.org/10.1049/el:20000267>

- [162] Zhao, J., Laganieri, R. and Liu, Z. (2007) Performance Assessment of Combinative Pixel-Level Image Fusion Based on an Absolute Feature Measurement. *International Journal of Innovative Computing, Information and Control*, **3**, 1433-1447.
- [163] Liu, Z., Forsyth, D.S. and Laganieri, R. (2008) A Feature-Based Metric for the Quantitative Evaluation of Pixel-Level Image Fusion. *Computer Vision and Image Understanding*, **109**, 56-68. <https://doi.org/10.1016/j.cviu.2007.04.003>
- [164] Wang, P. and Liu, B. (2008) A Novel Image Fusion Metric Based on Multi-Scale Analysis. *Proceedings of the 9th International Conference on Signal Processing*, Beijing, 26-29 October 2008, 965-968.
- [165] Zheng, Y., Essock, E.A., Hansen, B.C. and Haun, A.M. (2007) A New Metric Based on Extended Spatial Frequency and Its Application to DWT Based Fusion Algorithms. *Information Fusion*, **8**, 177-192. <https://doi.org/10.1016/j.inffus.2005.04.003>
- [166] Cvejic, N., Loza, A., Bull, D. and Canagarajah, N. (2005) A Similarity Metric for Assessment of Image Fusion Algorithms. *International Journal of Signal Processing*, **2**, 178-182.
- [167] Piella, G. and Heijmans, H. (2003) A New Quality Metric for Image Fusion. *Proceedings of International Conference on Image Processing*, Barcelona, 14-18 September 2003, 111-173.
- [168] Yang, C., Zhang, J., Wang, X. and Liu, X. (2008) A Novel Similarity Based Quality Metric for Image Fusion. *Information Fusion*, **9**, 156-160. <https://doi.org/10.1016/j.inffus.2006.09.001>
- [169] Chen, Y. and Blum, R.S. (2009) A New Automated Quality Assessment Algorithm for Image Fusion. *Image and Vision Computing*, **27**, 1421-1432. <https://doi.org/10.1016/j.imavis.2007.12.002>
- [170] Chen, H. and Varshney, P.K. (2007) A Human Perception Inspired Quality Metric for Image Fusion Based on Regional Information. *Information Fusion*, **8**, 193-207. <https://doi.org/10.1016/j.inffus.2005.10.001>