Fusing synergistic information from multi-sensor images: An overview from implementation to performance assessment ☆

Zheng Liu[*,a], Erik Blasch[b], Gaurav Bhatnagar[c], Vijay John[d], Wei Wu[e], Rick S. Blum[f]

[a] Faculty of Applied Science, University of British Columbia, BC, Canada
[b] US Air Force Research Laboratory, Rome, NY, USA
[c] Indian Institute of Technology Jodhpur, Rajasthan, India
[d] Toyota Technological Institute, Nagoya, Japan
[e] Sichuan University, Chengdu, China
[f] Lehigh University, Bethlehem, PA, USA

ARTICLE INFO

ABSTRACT

Image fusion is capable of processing multiple heterogeneous images acquired by single or multi-sensor imaging systems for an improved interpretation of the targeted object or scene. A diversity of applications have benefited from the fusion of multi-sensor images through a more reliable and comprehensive fused result. Likewise, numerous approaches to fuse multi-sensor images have been proposed and published in literature. However, due to a lack of benchmark resources and commonly accepted assessment measures, it is hard to identify the significance of new image fusion algorithms and implementations. This paper reviews and categorizes recent algorithms for image fusion and performance assessment based on reported comparative results. We recommend using non-parametric statistical tests to verify the performance of the pixel-level fusion algorithms. Furthermore, a comprehensive evaluation of 40 fusion algorithms from recently published results is conducted to demonstrate the significance of these algorithms in terms of statistical analyses within their respective applications. Although the results of these performance tests are limited by available data sets, baseline algorithms, and selected assessment metrics; it is a critical step for comparative image fusion research. This paper aims to advance image fusion development by creating a complete inventory of state-of-the-art image fusion techniques and advocating statistical comparison tests to avoid unnecessary duplication of development efforts. Establishing a benchmark study for image fusion is critical for performance comparisons of contemporary methods.

1. Introduction

Image fusion has benefited a diversity of applications, including medical diagnosis, security and surveillance, remote sensing, weather forecasting, industrial inspection, and biometrics, etc [1]. A fused image is characterized to provide more reliable information for understanding and perception of the scene or targeted object. Given the varied objectives of specific applications; image fusion, according to [2], is defined as "the process of combining information from two or more images of a scene into a single composite image that is more informative and is more suitable for visual perception or computer processing". Image fusion algorithms can operate on a single frame or on a video sequence, as illustrated in Fig. 1. In terms of the types of image data, image fusion can be categorized into: 1) temporal image fusion, which fuses images in video time sequences into one meaningful image;

2) spatial image fusion, which stitches images together for a wider field of view; 3) volumetric fusion, which creates a 3D object from image slices; or 4) connotative image fusion, which fuses multi-sensor (e.g., multi-spectral or multi-modal) images to integrate complementary information from inputs into the fused result.

The inputs for *temporal image fusion* are a sequence of images from video or still photographs of different exposure lengths. Temporal image fusion manipulates the temporal dynamic range of the fused images through integrating the details and structures available in the inputs [3]. Pixel-level fusion of images over time was applied to moving object tracking [4]. A temporal series of images can also be used to generate a "clean" scene image of high quality [5,6]. *Spatial image fusion* stitches images into a sharpened panoramic image of the scene even with the blurry inputs [7]. The fused image can give a large field of view or combine the salient results of the input images. The purpose

Temporal image fusion

Spatial image fusion

Complementary image fusion

Fig. 1. Types of image fusion: (let) temporal, (middle top) spatial, (middle bottom) volumetric, and (right) complementary fusion.

time

Image stitching

Volumetric fusion

of temporospatial fusion is to achieve a high spatial and temporal resolution simultaneously. In the application of remote sensing, a high-quality temporospatial image analysis is useful for natural resource management and monitoring of land-use and land-cover changes as well as ecosystem dynamics [8]. Examples include sharp panoramas generated from motion-blurred videos through joint global motion estimation and multi-frame deblurring [9]. A closely related topic is image super-resolution, which is beyond the scope of this paper. An excellent summary of the image super-resolution techniques is available in a book [10]. Volumetric fusion targets the reconstruction of the 3D object from 2D image slices [11]. All methods require image registration for the success of reconstruction, where registration of a nonrigid body within an image remains a challenge [12].
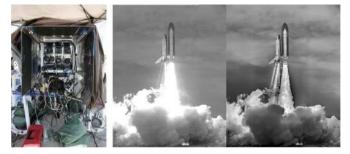
Complementary fusion aims at integrating harmonious features or information from the input image. The focus of this paper is to review the methods for fusing images collected by multiple types of imaging sensors. Recording the space shuttle STS-135 launch by NASA provides an example of a highly successful application of image fusion [13]. The camera setup for STS-135 included seven cameras: five visible spectrum black-and-white, a high speed, a high resolution, and two thermal
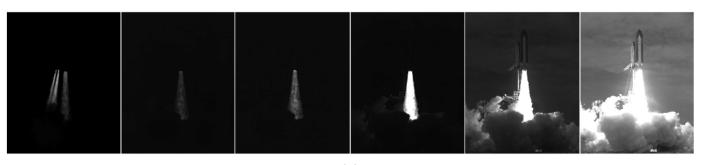
infrared cameras to capture temperature data (bottom left cameras) as shown in Fig. 2(a). One infrared camera did not function during the launch, so only six images in Fig. 2(b) are used to obtain the final fused image. According to [13], with the fused image, NASA researchers can better understand the structure of the plume when rockets fire, the motion of the flames flowing out of the rocket motor, and how to design optimal future motors.

A second example of fusing visual and thermal images is shown in Fig. 3 for context enhancement. The fused image can present both a clear foreground (human being) and background for easy scene interpretation. Other applications include medical imaging, remote sensing, and high dynamic range (HDR) image acquisition, which is achieved by multiple exposure fusion[14]. The most recent review of the state-of-the-art in medical image fusion can be found in [15]. Two reviews for remote sensing are published in literature [16] and [17] respectively. Readers are referred to these references for the details of the specific topics.

This paper focuses on general multi-sensor image fusion algorithms and performance assessment. Currently, the advantages of a fusion algorithm are judged by the improved values of selected fusion
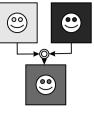

(a)

(b)

Fig. 2. Image fusion for STS-135 space shuttle launch (image credit: NASA/Louise Walker/J.T. Heineck). (a) Left: NASA camera array "Walle"; Middle: STS-135 view without fusion; Right: fused image. (b) Six images taken for STS-135 space shuttle (The first one is taken by a thermal infrared camera).

(a) Visible image   (b) Thermal image   (c) Fused image with the steerable pyramid   (d) Fused image with the approach in [18]
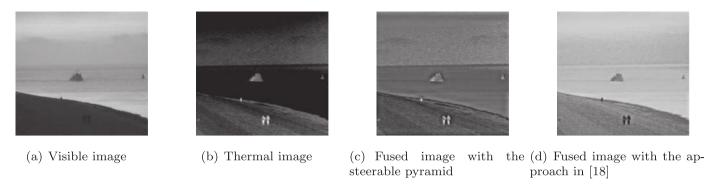
Fig. 3. Example of fusing TNO (in English: Netherlands Organisation for Applied Scientific Research) Kayak (frame 7118a) images [18].

performance metrics. There are numerous algorithms as well as fusion metrics that have been proposed and published. However, the comparative analysis of any reported performance improvements needs to be validated with statistical methods, such as significance tests. Otherwise, it is difficult to characterize better performance without understanding the statistical significance of the measured performance.

This paper creates an inventory of state-of-the-art fusion algorithms as well as fusion metrics, and conducts statistical comparisons between 40 algorithms. Both the data and codes for statistical comparison are available through GitHub. We demonstrate the need of significance tests for image fusion research in this paper. Section 2 presents the concept of multi-sensor image fusion while Section 3 categorizes the state-of-the-art in image fusion techniques at three levels, i.e. pixel-, feature-, and decision -level fusion. The currently available computational models or fusion metrics for pixel-level fusion performance assessment are summarized in Section 4. Statistical comparisons are performed and corresponding results are presented in Section 5. This paper provides a discussion in Section 6 and conclusions in Section 7, respectively.

2. Multi-sensor image fusion

Complementary image fusion can be categorized into *combinative* and *discriminative* fusion by its purposes. The combinative fusion integrates the complementary information from multiple input images into a composite image, which is more informative and better suited for further processing by a computer or exploitation by an end user. Typical applications of combinative fusion are listed in Tab. 1. In the discriminative fusion, the feature space constructed by the input images is expanded with multi-sensor inputs to enhance distinguishability from extra signatures. For example, a thematic map is derived from multiple inputs. Each pixel is characterized by individual entities or classes in a thematic map. Fig. 4 illustrates the concepts of the two categories. "A","B", and "C" in Fig. 4 represent the objects or features in a scene. "B" and "C", which are available in the separate input images, are integrated into the fused image through combinative fusion in Fig. 4(a). In Fig. 4(b), the discriminative fusion can identify "A","B", and "C" in the fused image, which is not possible in the input images.
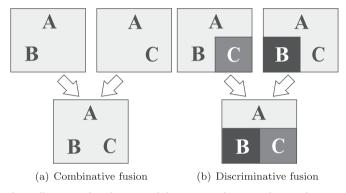

(a) Combinative fusion   (b) Discriminative fusion

Fig. 4. Illustration of combinative and discriminative fusion. Combinative fusion integrates complementary information while discriminative fusion identifies the different entities.

From an implementation perspective, image fusion can be accomplished at three levels, i.e. pixel level, feature level, or decision level as illustrated in Fig. 5. Pixel-level fusion generates a composite image, in which objects/regions of interest are enhanced. The fused image is further processed for object detection, target recognition, entity classification, or identity declaration (ID) for presentation to the end user. However, a single image may not be enough to present multiple attributes associated with each pixel object. Feature-level fusion associates attributes, intensities, or regions with an object extracted from one or more input images while at the decision level, categorical descriptors are generated for each object based on associated data [28]. If the categories are numerical, it is called score-level fusion.

The implementation of pixel-level fusion in the spatial domain is straightforward with logic, arithmetic, or pixel selection operations. Pixels' selection in the spatial domain is only applicable to homogenous images. Some measurements are considered in the decision of choosing to populate the fused image over an entire region with the pixels from one of the source images as opposed to some other measurements. One such measurement which is sometimes employed for multifocus images is the local frequency [29]. The pixels with higher local frequency are retained in the fused image. However, this method is not feasible for

Table 1
Sample applications of combinative fusion.

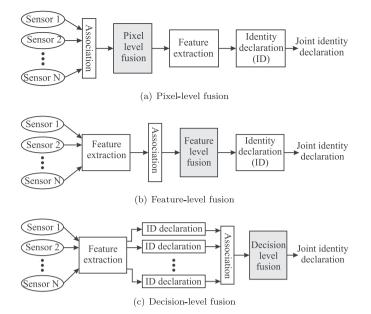| Multiple input images	Purpose of fusion	References
Multi-focus images	Obtain fully-focused image with complete information	[19]
Multi-exposure images	Obtain high dynamic range (HDR) image	[14,20,21]
Haze image processed by white balance and contrast enhancement	Enhance the visibility of a haze-degraded image	[22]
High-resolution panchromatic image low-resolution multispectral images	Obtain a high-resolution multispectral image suitable for both geometric analysis and thematic interpretation	[14,23,24]
Range image and infrared image	Face recognition	[25]
Transcranial color doppler and magnetic resonance images	Medical diagnosis	[26]
Computed tomography and magnetic resonance images	Medical diagnosis	[27]

Fig. 5. Image fusion at three different levels: a) pixel level, b) feature level, and c) decision level.

fusing infrared and visible images as they each keep a different intensity table. *Transform-based fusion approaches* present image features in a more accessible format of the coefficients in the transform domain. Thus, a fusion rule can be directly applied to the coefficients. Higher level measurements of features or information are formulated to guide the fusion process. The inverse transform will generate the fused image. Transform-based methods, such as wavelets, are the most popular method for pixel-level fusion and numerous studies have been reported. Besides the multiresolution analysis, *learning-based approaches* have recently been proposed to use data-derived features for image fusion [30–33].

Different from the pixel-level operations, feature-level fusion extracts and fuses features (e.g., shapes, edges, histogram of gradients) from the input images to meet the requirements for a deeper understanding of the captured information, which helps avoid the limitations of low-level fusion. As described, pixel-level fusion is not guided by the desired characteristics of the output, where the fused image is subject to the loss of information from the inputs [34]. The fused result at feature-level is presented as a silhouette, a contour map, or a extended feature vector [35,36]. As illustrated in Fig. 5(a), decision-level fusion combines the identity declarations or sensor confidence degrees from each image sources at a higher level of abstraction. The decision-fusion process deals with the symbolic representation of images. The outcome of the decision-level fusion could be binary, e.g. "Yes/No" or "accept/reject", or represented by a confidence degree value [37]. The identity could be the whole image, ROI (region of interest) and even a point in the image. More details are described in the following sections.

3. Image fusion algorithms

3.1. Pixel-level fusion

3.1.1. Multiresolution analysis based approach

The multiresolution analysis (MRA) approach represents an image I (x, y) with a set of basis functions $g_i(x, y)$, where i identifies the basis functions. The basic idea is to represented the image $I(x, y)$ as [38]:

$$I(x, y) = \sum_i t_i g_i(x, y) \tag{1}$$

$$t_i = \sum_{x,y} h_i(x, y) I(x, y) \tag{2}$$

where t_i are the transformed coefficients that can be obtained by projecting the image onto a set of projection functions $h_i(x, y)$. Different MRA approaches employ different basis functions. Here, t_i represents the corresponding coefficient.

A fusion rule is applied to each t_i based on the measurement of image features and characteristics of $g_i(x, y)$ to obtain a set of fused coefficients for each pixel (x, y). The process is illustrated with the diagram in Fig. 6. The key elements include the MRA algorithm and the coefficient combination rule or fusion rule. Thus, the two topics will be discussed in detail respectively.

3.1.1.1. MRA algorithms. The use of different MRA algorithms is motivated by the varied characteristics of those algorithms. The considerations may vary with the application requirements, such as the compactness of the representation, directional sensitivity, shift and rotation invariance. Table 2 summarizes the typical MRA algorithms used for image fusion. The coefficients in the transform domain are often more closely related to image features like edges and boundaries [39]. These coefficients will be combined and then an inverse transform is applied to generate the fused image.

3.1.1.2. Fusion rules. The combination of coefficients in the transform domain is implemented with a certain "fusion rule". Generally, the K input images I_k $(k = 1, \cdots, K)$ can be represented with their low- and high-pass components as $\{I_k^L, I_{j,k}^H\}$, where j refers to the scale or resolution level. The fused coefficient sets are $\{F^L, F_j^H\}$, with which the fused image F can be obtained. The well-adopted rule can be expressed as [53]:

$$\begin{cases} F_j^H & = \arg \max_{F_j^H, j, k} M(I_{j,k}^H) \\ F^L & = \frac{1}{K} \sum_k w_k I_k^L \end{cases} \tag{3}$$

where $M(\cdot)$ and w_k are a measurement function and weighting coefficients, respectively. For the low-pass component, the coefficients are often combined by averaging or weighted averaging. However, an averaging operation may introduce artifacts especially when input images are of different modalities. Thus, a compensation with some weights is desired. Maximum selection is usually applied to the high-pass and band-pass components, which is based on the assumption that larger coefficients correspond to salient features. The

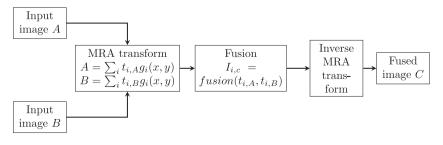Fig. 6. Image fusion in the MRA framework.

Table 2
Typical MRA algorithms used for image fusion.

MRA algorithm	Characteristics	References
Laplacian pyramid	Differences of the blurred versions between each levels	[40]
Steerable pyramid	Translation-invariant and rotation-invariant properties	[41,42]
Discrete wavelet	Compactness, orthogonality, and the availability of directional information	[43]
Discrete wavelet frame	Aliasing free and translation invariant (without down-sampling)	[44]
Multi-wavelet	Perfect reconstruction with orthogonality, linear-phase symmetry, and high order of approximation	[45]
Dual-tree complex wavelets	Over-complete wavelet transform with shift invariance, directional selectivity, and perfect reconstruction	[46,47]
Ridgelets	Orthonormality, invertibility and nonredundancy	[48,49]
Contourlet	Multiresolution directional tight frame	[50,51]
Curvelet	Multiresolution transform with frame elements indexed by scale and location parameters	[52]

Table 3
Typical measurement of image feature and activity level with coefficients in the transform domain.

Feature/activity-level measurement	Description	References
Coefficient-based methods	Luminance, chromaticity, and saliency weight maps	[22]
	Membership function of fused coefficients	[58,59]
Window-based methods	A small window (3 × 3 or 5 × 5 pixels) around the considered coefficient	[43]
	Sliding window with dense SIFT descriptor	[55]
	Statistical properties of a $w \times w$ window	[56]
Region-based methods	Region priority map	[46]
	Match and saliency measures for region	[47]

function $M(\cdot)$ may incorporate more relevant information. The typical computation methods for such information are summarized in Table 3. The key point is how to define the measurement function $M(\cdot)$ and weighting coefficients w_k. There are generally six categories of operations for the image fusion rule.

Activity-level measurement: The measurement of activity can be implemented in three different ways [54]:

- Coefficient-based activity;
- Window-based activity;
- Region-based activity.

A coefficient-based activity measure employs the absolute or square value of the individual coefficient while a small window (3 × 3 or 5 × 5 pixels) around the considered coefficient is used in the window-based method. The activity measure can be obtained by weighted summation of the coefficients within the window or by a rank filtering operation, which uses the maximum coefficients [43]. Region-based measurement differs from the window-based method with an automated segmentation approach to define the region [47]. Other methods include using the dense SIFT (scale invariant feature transform) to define the activity level or, statistical properties of the local neighborhood of a pixel [55,56].

Coefficient grouping: Another consideration in the fusion rule is the relation between the coefficients across frequency bands and decomposition levels [54]. There are basically two grouping methods, i.e. single-scale grouping and multi-scale grouping. In single-scale grouping, corresponding coefficients among different sub-bands at the same scale are fused in the same manner while multi-scale grouping method considers the coefficients at different scales. In [54], the weight average of the activity levels of the grouped coefficients was used in the fusion decision making in Eq. (3) above. A cross-band single-scale grouping scheme was proposed in [57], in which the fusion decision for coefficients in the band- or high-pass components was based on the sum of their activity levels and their corresponding coefficients at a higher

scale. In [58,59], the impact from the corresponding coefficients at adjacent scales was modeled with a membership function defined by a generalized random walk method.

Cross-scale coefficient selection: The purpose of cross-scale coefficient selection is to achieve intrascale and interscale consistencies so that most of the details can be preserved in the fused image while artifacts can be diminished. In [57], the selection decision at one scale also involves the corresponding coefficients at adjacent scales, i.e. one level up and down. In [58], a membership function is defined to choose of detailed information (coefficients in high-pass bands) and weight the approximate information (coefficients in low-pass band).

Coefficient modeling: The relationship of the coefficients can be represented with statistical models. In [60,61], a Markov model, more specifically hidden Markov tree, was adopted to model the interscale and/or intrascale relationships. The model was only applied to the high-frequency subbands. In [61], a binary segmentation map was generated to guide the selection of coefficients in high-frequency subbands.

Coefficient combination: The combination of coefficients follows the general fusion rule in Eq. (3). The choose-maximum strategy for coefficients in band- and/or high-pass bands considers the coefficient grouping method as well as activity-level measurements to determine the "maximum value" in the selection process. The approaches to generate weighting coefficients in the low-pass band include image similarity measurement, average local intensity, and pixel significance [62,63].

Consistency verification: Consistency verification ensures neighboring coefficients are fused in the same manner [54]. In [43], a window-based verification was implemented with a majority filtering operation at each scale. A cross-band verification was introduced in [57], in which the coefficient was recalculated if its "children" coefficients came from a different source. Herein the children coefficients refer to the corresponding coefficients at a lower resolution level.

Modifications of the MRA-based method include contrast pyramid and gradient-based multiresolution fusion [64,65], which makes more salient features available in the generated gradient maps. Thus, the fusion quality can be further improved through additional features.

3.1.2. Learning-based approach

3.1.2.1. Sparse representation and dictionary learning. Sparse representation (SR) obtained with a dictionary learning algorithm offers an innovative way to represent an image or signal. In conventional MRA methods, the image is represented by predefined basis functions, i.e. wavelet functions. In contrast, SR uses the basis vectors known as "dictionary" learned from the training data set to represent the image. More specifically, sparse representation represents a signal $\mathbf{Y} \in \mathbb{R}^n$ with a linear combination of basic elements, d, from a dictionary $\mathbf{D} = [\mathbf{d}_1, \cdots, \mathbf{d}_K] \in \mathbb{R}^{n \times K}$, whose is composed by atoms in K columns $\{\mathbf{d}_j\}_{j=1}^K$ [66]. Thus, there is $\mathbf{Y} = \mathbf{DX}$, where $\mathbf{X} \in \mathbb{R}^K$ is the sparse coefficient vector. The representation is to find the dictionary $\mathbf{D}$, which is overcomplete when $n < K$, and corresponding sparse coefficients $\mathbf{X}$. The solution is by solving the optimization [67]:
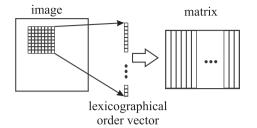
Fig. 7. Reorganize image patch into a lexicographic order vector and matrix.

$$\min_{\mathbf{X} \in \mathbb{R}^K} \left\| \mathbf{X} \right\|_p, \ s.\,t. \ \left\| \mathbf{Y} - \mathbf{DX} \right\|_p \leq \varepsilon \tag{4}$$

where ℓ^p-norms ($p = 0, 1, 2,$ and ∞) are used to measure the differences; ε is the threshold and $\|\mathbf{X}\|_p = \sqrt[p]{\sum_i |x_i|^p}$. ℓ^0 norm of $\mathbf{X}$ counts the nonzero elements in $\mathbf{x}$. Eq. (4) can be solved with approximation approaches like matching pursuit (MP) and orthogonal matching pursuit (OMP) [68,69].

To perform image fusion, the input images need to be reorganized as illustrated in Fig. 7. A sliding window of size $n \times n$ moves from the top left corner to the bottom right. Each patch is lexicographically ordered as a vector $\mathbf{y}^j$, which is written as [30]:

$$\mathbf{y}^j = \sum_{k=1}^{K} \mathbf{x}^j(k) \mathbf{d}_k \tag{5}$$

Eventually, the sparse matrix $\mathbf{X}$ can be obtained and $\mathbf{x}^j$ is the j column in the matrix. Thus, each input image i can be represented by $\mathbf{Y}_i = \mathbf{D}_i \mathbf{X}_i$.

A straightforward approach is to combine the sparse coefficients of the i input images, i.e. $\mathbf{X}_i$. In [30], a coefficient selection scheme is proposed, where an activity level measurement is defined for each column of $\mathbf{X}_i$, i.e. $A_i^j = \|\mathbf{X}_i^j\|_1$. The sparse coefficients associated with larger activity measure will be retained in the fused sparse matrix $\mathbf{X}_f$. Here, f refers to the fused result. This process can be expressed as [30]:

$$\mathbf{x}_f^j = \mathbf{x}_s^j, \ s = \arg \max_i (A_i^j) \tag{6}$$

$$\mathbf{Y}_f = \mathbf{DX}_f \tag{7}$$

A joint sparse representation scheme was introduced in [70], in which two joint sparsity models (JSM) were described. The first model (JSM-1) was adopted in [31] to implement image fusion. According to [70], a signal y_j can be represented by a shared common sparse component (z) of the signal ensemble together with an innovation component (z_j) for its own. Let $\underline{\mathbf{Y}} = \{\mathbf{Y}_1, \mathbf{Y}_2, \cdots, \mathbf{Y}_\Lambda\}$ represent an ensemble of input signals and $\mathbf{Y}_i \in \mathbb{R}^n$. Λ is the total number of the inputs. The signal $\mathbf{Y}_i$ can be represented as [32,70]:

$$\mathbf{Y}_i = \mathbf{Y}^c + \mathbf{Y}_i^u = \mathbf{DX}^c + \mathbf{DX}_i^u + \mathbf{n}_i \tag{8}$$

where $\mathbf{X}^c$ and $\mathbf{X}_i^u$ are the common and innovation component of the sparse coefficient matrix respectively. The signal noise is represented by $\mathbf{n}_i$. For all the signal in $\underline{\mathbf{Y}}$, Eq. (8) is represented as:

$$\underline{\mathbf{Y}} = \begin{bmatrix} \mathbf{Y}_1 \\ \vdots \\ \mathbf{Y}_\Lambda \end{bmatrix} = \begin{bmatrix} \mathbf{D} & \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{D} & \mathbf{0} & \mathbf{D} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{D} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{X}^c \\ \mathbf{X}_1^u \\ \vdots \\ \mathbf{X}_\Lambda^u \end{bmatrix} + \begin{bmatrix} \mathbf{n}_1 \\ \mathbf{n}_2 \\ \vdots \\ \mathbf{n}_\Lambda \end{bmatrix} \tag{9}$$

and can be further written as:

$$\underline{\mathbf{Y}} = \underline{\mathbf{D}}\mathbf{X} + \mathbf{n} \tag{10}$$

To obtain the dictionary and joint sparse coefficient matrix, requires solving the following optimization problem:

$$\min_{\mathbf{X} \in \mathbb{R}^{K \times \Lambda}} \left\| \mathbf{X} \right\|_0, \ s.\,t. \ \left\| \underline{\mathbf{Y}} - \underline{\mathbf{D}}\mathbf{X} \right\|_2 \leq \varepsilon \tag{11}$$

The solution to Eq. (11) includes the orthonormal matching pursuit

(OMP), K-singular value decomposition (K-SVD), and method of optimal directions (MOD) etc [67,68,71].

Based on the JSM-1 framework, several fusion approaches were proposed [31–33]. Similarly, the fusion of multiple images is processed with sliding windows and implemented by combining the sparse coefficients from the inputs. Another feature that distinguishes different fusion approaches is the way to build the dictionary. As far as the fusion rule is concerned, weighted averaging is often adopted. Fusion rules differ in the definition of weighting coefficients. In [33], the lexicographical order matrices were subtracted from their mean value (m_i, $i = 1, 2$) and then fed to the JSM-1 for representation. The fused image of two inputs was obtained by:

$$\mathbf{Y}_f = \mathbf{DX}^c + \mathbf{DX}_1^u + \mathbf{DX}_2^u + \tau m_1 + (1 - \tau)m_2 \tag{12}$$

where the weighting coefficient τ is defined as:

$$\tau = \frac{1}{1 + \exp(-\beta(\|\overline{\mathbf{Y}}_1\|_2 - \|\overline{\mathbf{Y}}_2\|_2))} \tag{13}$$

and the symbol $\overline{\mathbf{Y}}$ refers to the mean subtracted component from input and β is a positive constant. However, detailed information on how to select β is not available in [33]. A dictionary was built with K-SVD algorithm and images from USC-SIPI image database [72].

In [31], operation on mean substraction was not conducted and the fusion rule is written as:

$$\mathbf{X}_f = \mathbf{X}^c + \frac{1}{w_{\max}} \sum_{i=1}^{\Lambda} w_i \mathbf{X}_i^u \tag{14}$$

$$w_{\max} = \max_{1 < i < \Lambda} (w_i) \tag{15}$$

$$w_i = \|\mathbf{X}_i^u\|_1 \tag{16}$$

Thus, the fused image can be obtained by $\mathbf{Y}_f = \mathbf{DX}_f$. Algorithm K-SVD was employed to train an overcomplete dictionary. However, the authors did not clearly mention the image database used for training. The impact of the training data sets on the fusion result is unknown.

The fusion rule in Eq. (12) was modified for general case in [32] as:

$$\mathbf{Y}_f = \mathbf{DX}^c + \mathbf{DX}_f^u + m_{\hat{i}} \tag{17}$$

where $\mathbf{DX}_f^u$ is the second item in (14) and $\hat{i}$ corresponds to the maximum absolute value of m_i, i.e. $\hat{i} = \arg \max_i |m_i|$ with $i = \{1, \cdots, \Lambda\}$. This modification makes it possible to handle multiple inputs larger than two. This is a kind of "maximum coefficient selection" rule.

Sparse representation is relatively new to the image fusion community. The quality of training samples has a direct influence on the dictionary, which can be trained with either the input images or available image databases. The fused result may deteriorate with the degradation of the input images. When other image databases are used, how the multi-sensor input images can be accurately represented with the sparse representation becomes a major issue. Likewise, the training algorithm plays an important role as well. However, the available experimental results do not demonstrate significant improvements or advantages over the MRA approaches in terms of the assessment metrics selected in each study. The sparse representation introduces some complementary operations, such as denoising, which may improve the robustness of the fusion process [30,31,33].

When the pan-sharpening application is considered, images from a high-resolution (HR) panchromatic channel and several low-resolution (LR) multispectral channels are fused to generate an HR multispectral image. In [24], a sparse representation based approach named "SparseFI" was proposed. The fusion process is illustrated in Fig. 8. The HR panchromatic image X_0 is first down-sampled to image Y_0 of the same size as the LR image Y_k ($k = 1, \cdots, N$) from the N multispectral channels. Two dictionaries, namely LR dictionary $\mathbf{D}_l$ and HR dictionary $\mathbf{D}_h$, are built with the image patches $\mathbf{y}_0$ and $\mathbf{x}_0$ derived from Y_0 and X_0, respectively.
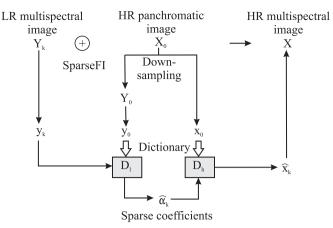
Fig. 8. The "sparseFI" algorithm for pan-sharpening application [24].

Fig. 9. The diagram for image fusion with image blocks and a neural network [75].

Each LR multispectral patch in the kth channel $\mathbf{y}_k$ is then re-presented as a linear combination of $\mathbf{y}_0$ with a coefficient vector $\hat{\boldsymbol{\alpha}}_k$, which is estimated by $L_1 - L_2$ minimization:

$$\hat{\boldsymbol{\alpha}}_k = \arg\min_{\alpha}\left\{\lambda\left\|\boldsymbol{\alpha}_k\right\|_1 + \frac{1}{2}\left\|\widetilde{\mathbf{D}}\boldsymbol{\alpha}_k - \widetilde{\mathbf{y}}_k\right\|_2^2\right\} \tag{18}$$

where there are:

$$\widetilde{\mathbf{D}} = \begin{bmatrix} \mathbf{D}_l \\ \beta\mathbf{P}\mathbf{D}_h \end{bmatrix} \qquad \widetilde{\mathbf{y}}_k = \begin{bmatrix} \mathbf{y}_k \\ \beta\mathbf{w}_k \end{bmatrix} \tag{19}$$

The standard Lagrangian multiplier, λ, is used to balance the sparsity of the solution and the fidelity of the approximation to $\mathbf{y}_k$. A SL1MMER algorithm can be applied to solve this optimization problem [24,73]. According to Zhu and Bamler [24], matrix $\mathbf{P}$ extracts the overlapped region between current target patch and previously reconstructed ones. And $\mathbf{w}_k$ contains the pixel values of the previously reconstructed HR multispectral image patch in the overlap region. β is a weighting coefficient, which balances the goodness of fit of the LR input and the consistency of reconstructed adjacent HR patches in the overlapping area.

Once the sparse vector $\hat{\boldsymbol{\alpha}}_k$ is obtained, then $\mathbf{y}_k = \mathbf{D}_l\hat{\boldsymbol{\alpha}}_k$. Thus, the HR multispectral patches can be obtained by $\hat{\mathbf{x}}_k = \mathbf{D}_h\hat{\boldsymbol{\alpha}}_k$. The key assumption here is that $\mathbf{y}_k$ and its corresponding $\mathbf{x}_k$ share the same sparse coefficients in $\mathbf{D}_l$ and $\mathbf{D}_h$ [24]. Thus, the HR multispectral image $\hat{X}$ can be obtained from the estimated patches $\hat{\mathbf{x}}_k$.

The sparseFI presents another way to fuse different-sized images. It overcomes some inherent problems with the sparse representation framework, such as inaccuracy in signal reconstruction and spatial inconsistency in the fused image [74]. Thus, combined approaches were proposed.

3.1.2.2. Combined fusion approach with MRA and SR. A general framework for image fusion, which combines MRA and SR methods, is introduced in [74]. In general, the MRA transform decomposes input images into low-pass and high-pass components (one-level decomposition). For the low-pass components of multi-input, sparse representation based fusion is applied by using the methods described in Section 3.1.2.1. The high-pass components are simply obtained by the maximum absolute value selection rule.

3.1.2.3. Artificial neural networks. Artificial neural networks (ANN) are another learning-based approach for image fusion. Fig. 9 illustrates the use of NN in image fusion [75]. A probabilistic neural network and a radial basis function network (RBFN) were used to select the input image blocks based on designated features. A cooperative neural fusion regularization algorithm was proposed in [76], which can achieve the optimal image estimate while mitigating loss of contrast information.
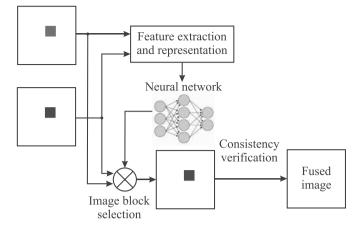
The convolutional neural network (CNN) is a recent advance in machine learning, which has shown great potential in various image analytic problems. In [77], a CNN was employed to first increase the resolution of multispectral images and then the Gram–Schmidt transform was used to fuse the enhanced multispectral image and panchromatic image. Although the CNN did not directly contribute to the fusion process in this study, it is worth more effort to investigate its benefits to image fusion research.

3.1.3. Fusion in color space

As the human visual system is sensitive to color, fusion in color space could achieve a more informative result. The idea of false color fusion is simple, where the different color channels are mapped to the processed input images [78], and a composite (fused) color image will be obtained. However, the false color image does not always have a natural appearance. Thus, a color transfer operation needs to be applied as illustrated in Fig. 10. In [79], two transfer methods, i.e. principal component transform and $l\alpha\beta$, were proposed to transfer the color characteristics from natural daylight imagery to false color night vision imagery.

A scheme to fuse visible and infrared (IR) images in the color space was presented in [80]. The fusion algorithm is illustrated in Fig. 11. Input images are manipulated in three color spaces, i.e., RGB (R: red, G: green, B : blue), HSV (H: hue, S : saturation, V: brightness value) and LAB (L: brightness, A: red–green chrominance, B: yellow–blue chrominance). The original infrared image and its reverse polarity are fused with the V component from the HSV space. The discrete wavelet frame (DWF) is used, which is a shift invariant transform. Following the flowchart in Fig. 11, seven fused images, {F_{1RGB}, F_{1LAB}, F_{2LAB}, F_{2RGB}, F_{2HSV}, F_{3HSV}, F_{3RGB}}, can be obtained. Among the fused results, F_{3RGB} is most similar to the original color image.

Another color fusion scheme is shown in Fig. 12 [81]. The reverse polarity of the infrared image was used in this method. A curvelet transform was employed to fuse the visible and infrared image as well
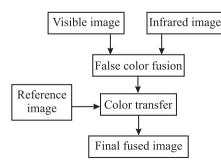

Fig. 10. Natural color mapping/transfering for false color fusion of visible and Infrared (IR) images.
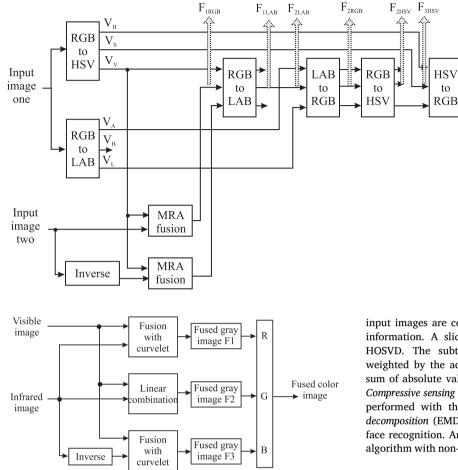
Fig. 11. Fusion of images in color space. The application is concealed weapon detection [80]. Image one is the visual image while image two is the infrared image.

Fig. 12. Fusion scheme for visible and infrared images in color space by using the reverse polarity of the infrared image [81].

as the inverse IR image. The fused results (F_1 and F_3) were fed into the R (red) and B (blue) channel respectively. The G (green) channel is the linear combination (see Fig. 10) result (F_2) of visible and infrared image. The RGB channels compose the fused result.

Fusion in color space can offer an intuitive result to the end user, which is more suitable for human perception. Zheng and coworkers [82,83] present the quantitative metrics for colorization evaluations. Another example is tone-mapping from which structural quality metrics are used to determine the color fusion of the dynamic range of input images [84]. However, it is not clear if the fused color image can benefit from computer-based enhancement and currently includes the investigation of matching objective machine metrics with that of human subjective metrics.

3.1.4. Other approaches
3.1.4.1. Other image transformations.
Besides the MRA and learned dictionaries; other transforms, which represent an image in the corresponding transform domain, have been employed as well. In [85], source images transformed by the fractional Fourier transform were then decomposed with self-fractional Fourier functions followed by selection operation based the maximum absolute value rule. Average filtering can also present an image in multiple scales, e.g. base layer and detail layer [86]. The sub-band images at the two layers are fused by a weighted averaging operation followed by a reconstruction procedure. The weighting coefficients are created by guided filtering. Higher order singular value decomposition (HOSVD), which is an efficient *tensor decomposition* technique to represent high-dimensional data and extract features, offers an alternative transform for image fusion [87]. The

input images are constructed into a tensor, which describes the local information. A sliding window is then applied to the tensor with HOSVD. The subtensors are combined with a sigmoid function weighted by the activity-level measurement, which is defined as the sum of absolute values of the coefficients from HOSVD of subtensors. *Compressive sensing* (CS) provides another domain, where fusion can be performed with the CS measurement values [88]. *Empirical mode decomposition* (EMD) based fusion was introduced in [89] to improve face recognition. An EMD pyramid was developed in [90]. The fusion algorithm with non-negative matrix factorization was described in [91].

3.1.4.2. Adaptive fusion.
The adaptive fusion approach is currently implemented by choosing the best one from a set of preselected fusion rules in terms of a predefined evaluation model. Zhang [92] comes up with a total variation minimizing process as the evaluation model for region-based fusion. However, this approach introduces with high computational cost and complexity as the final fusion result is selected for the evaluation. In other words, "adaptability" is achieved by picking up a better fusion rule based on the fused results, i.e. a post evaluation process.

3.1.4.3. Optimization-based fusion.
Image reconstruction and restoration are operations, which are often applied to derive a complete image from multiple input images. Variational models are successful in a wide variety of restoration problems [93,94]. The idea of image fusion with total variation is to preserve important image features while removing noises [93]. Such an example of multifocus image fusion can be found in [95], where weighted energy functions were proposed and minimized in the spatial and wavelet domain respectively. The energy functions are based on the local average modulus of gradients or intra- and inter-scale wavelet coefficients as well as the power transform [95].

The energy minimization problem can also be formulated as a maximum flow problem in a graph, which is defined as a minimal cut of the graph [96]. Thus, image fusion can be treated as a discrete multi-label optimization problem, which can be solved by a graph cuts approach via swap or alpha-expansion moves [27]. A total variation (TV) approach is implemented to fuse visible and infrared images as well as computed tomography (CT) and magnetic resonance imaging (MRI) at pixel level in [97].

In [98], a simultaneous registration and fusion based on the expectation-maximization (EM) algorithm is presented. The major advantage of the joint approach is the ability to automatically tune the

registration parameters so that an "optimal" (with respect to registration) fusion performance can be achieved.

3.1.4.4. Cascade fusion. Usually, a fusion operation is performed once in a process. A two-stage fusion scheme is presented in [99]. Multiple radar images were first fused by applying addition and multiplication or performing wavelet and principal component analysis (PCA) based fusion. A second stage applied fuzzy operators to achieve high target intensities and lower clutter levels [99]. The cascade scheme takes the advantages of multiple stages. Future research may study the structure of the cascade fusion and how to incorporate the feedback from fusion performance assessment for optimized image fusion results.

3.2. Feature-level fusion

Image features include edges, regions, shape, size, length and segments. The methodology for feature-level fusion depends on the nature of the image and varies with fusion application. A challenge remains for the fusion of features from different algorithms and data sources.

3.2.1. Image segmentations, contours and silhouettes

The fusion of multiple segmentations from reference images (atlases) has found great benefits for medical imaging [100]. Multiple atlases can provide more accurate and robust results than any individual. A shape-based averaging (SBA) was proposed in [101] to combine multiple segmentations of multidimensional images. The natural distance relationship between pixels of an *n*-dimensional image was utilized to average segmentations based on the signed Euclidean distance maps of the labels in each input segmentation [101]. The SBA method was further extended into a global weighted shape-based averaging and a local weighted shape-based averaging method by including the similarity information between the atlas and the target images with a Markov random field based neighborhood prior model [100].

In [102], edges extracted from the electro-optical (EO) image are blended with a registered infrared image with an alpha blending process, i.e. $c = (1 - \alpha) \times a + \alpha \times b$, for the purpose of highlighting the tiny objects in the scene. Herein, *a* and *b* refer to the two input images respectively. However, a complete quantitative assessment is not available in the report [102].

Logic operations, e.g. "OR" and "AND", were applied to objects detected from visible and infrared images in [35]. The targeted features were presented with centroid, size, aspect ratio and angular direction. The "OR" operation obtains the union of infrared tracking and visual tracking while the "AND" operation obtains the intersection of the two sensor tracks. In [103], the selection of relevant feature from the inputs was performed with mutual information.

In the application of land-cover classification with SAR (synthetic aperture radar) and optical data [104], the binary images from pre-classification for each data source were stacked together as a new multi-level feature to train a new classifier.

3.2.2. Image amplitude, phase, and eigen features

The amplitude and phase features of a facial image are independent of changes of illumination and contrast [105]. These features can be fused to improve the performance of face recognition. A dual *ν*-support vector machine (2*ν*-SVM) was trained to select the amplitude or phase feature. A match score was further obtained for the fused feature vector. In [106], genetic algorithm (GA) was employed to combine eigenfeatures from visible and infrared images. The eigenspaces for visible and infrared images are firstly built from which new inputs are projected onto the two eigenspaces respectively. GAs were applied to decide eigenfeature selection from each eigenspace. Face recognition was performed with the fused eigenfeatures.

3.2.3. Image statistical features

Images are characterized by its statistical features, such as properties of binary objects, histogram features, color features, and texture features etc. Usually, the extracted image feature is presented with a vector in a predefined feature space. When multiple input images are considered, those feature vectors can be simply stacked into a tensor or cascaded as a new vector, whose dimension or length is expanded. Principal component analysis (PCA) is often employed to reduce the size and keep the discriminatory information. In [107], a generalized *N*-dimensional independent component analysis (ICA) was proposed for multi-feature fusion for color image classification. The fusion of visible and infrared images was formulated as a statistical estimation problem in [108]. The mean square error between the fused image and the true scene is minimized by incorporating covariance intersection into the expectation maximization process. An iterative bootstrapping non-parametric expectation-maximization (BNEM) algorithm is proposed [109].

3.3. Decision-level fusion

Decision-level fusion as illustrated in Fig. 5(a) deals with the identity declaration (ID) directly. The initial IDs are derived from multiple input images with a classification process or statistical modeling. Fusion generates a joint ID from the input IDs. *Averaging* is the simplest form of decision-level fusion. In [110], pedestrian tracking was implemented in surveillance area with visible and infrared videos. Background and foreground (pedestrian) objects were segmented in each frame by applying a threshold to the averaged codewords from visible and infrared video. The codewords were defined for the visible and infrared images respectively. A Dezert–Smarandache (DSm) theory based *match score* fusion was proposed in [111] for face recognition with visible and infrared images. Match scores were computed by matching the global and local features for visible and infrared images respectively. The two match scores for each imaging modality were fused with DSm theory. The fused match scores for visible and infrared face images are combined again to get a composite match score for the multi-sensor input images. A final decision, accept or reject, is made with this composite score. The Bayesian rule was used to fuse the outputs from two classifiers, i.e. fast sparse representation and support vector machine, for a SAR image in [112]. However, it is not clear how the *priori* probability is defined. In [113], the classification of ground vehicles from SAR images were fused through multi-view, multi-feature, and multi-classifier decision fusion. Significant performance benefits for target recognition were reported. The posterior probabilities from multiple classifiers were fused with linear opinion pools and logarithmic opinion pools in [114] for hyperspectral image classification.

In [115], classifications of remote sensing image were fused with the iterative conditional modes algorithm to maximize the modified *a posteriori*. A so-called composite decision fusion, which combines SVM classification with the composite kernel method, was proposed in [116] to deal with the combined classification of a high spatial resolution color image and a lower spatial resolution hyperspectral image of the same scene.

Other approaches for decision-level fusion include random forest [117], fuzzy integral [118], Dempster–Shafer evidence theory [119], fuzzy logic [120], and support vector machine [121]. Typically, the approaches for combining multiple classifiers can be utilized in decision-level image fusion. A closely related concept in machine learning is ensemble method. More details about *ensembles of classifiers* are available in [122,123].

4. Fusion performance assessment

Generally, an image fusion assessment can be categorized into "**external**" and "**internal**" methods. The external or validation method is to check how the fusion meets the requirements of a specific
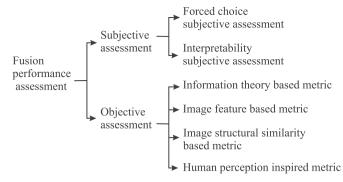
Fig. 13. Approaches for the assessment of combinative fusion.

application as an ultimate evaluation. The internal methods verify the quality of the fused image, which can be measured or quantified with certain metrics, such as entropy and information measurement. Performance assessment is critical to the fusion algorithm implementation. First, the effectiveness can be assessed against benchmark data. Secondly, the fusion algorithm can be further optimized by tuning the parameters in the algorithm based on the performance measurement. The current focus of research is on the quality of the fusion performance, which is represented by the measures of accuracy and reliability. This section considers the performance assessment of pixel-level fusion.

4.1. Combinative fusion assessment

As combinative fusion integrates complementary information, the assessment of the fusion's performance will count quantitatively how much information from inputs is transferred to the final fused image. There are different approaches to represent such information. As illustrated in Fig. 13, the assessment can be categorized into subjective and objective approaches. The *subjective methods* rely on human perception-based measures whereas *objective methods* are based on computational models, which count the amount of image features, contents, or information transferred from input data to the fused result [124]. Extensive subjective assessment is not always feasible because of the high cost and the difficulties in the control of varied human factors (e.g., individual differences, biases, and sensing abilities). The objective computational models are also known as fusion metrics, which can reveal certain inherent properties of the fusion process or the fused image. The summary of the available metrics is given in Table 4. A comparative study on these metrics was reported in [124]. The challenge for using these metrics is that it is difficult to know the

Table 4
Summary of fusion metric calculations [124].

Information theory based metric	Q_{MI}: Normalized mutual information [130]
	Q_{TE}: Mutual information based on Tsallias entropy [131]
	Q_{NCIE}: Nonlinear correlation information entropy [132]
Image feature based metric	Q_G: Gradient-based fusion metric [133]
	Q_M: Image fusion metric based a multiscale scheme [134]
	Q_{SF}: Image fusion metric based on spatial frequency [135]
	Q_P: Image fusion metric based on phase congruency [136,137]
Image structural similarity based metric	Q_S: Piella's metric [138]
	Q_C: Cvejie's Metric [139]
	Q_Y: Yang's Metric [140]
Human perception inspired fusion metric	Q_{CV}: Chen-Varshney metric [141]
	Q_{CB}: Chen-Blum metric [142]

significance between two metric values like 0.82 and 0.83.

Two examples on subjective fusion performance assessment were conducted [31,32]. One is the "forced choice subjective assessment", where passive, informal, and preference tests were designed to compare the fusion algorithms with "visual inspection." It can be applied when only a few choices are considered. It is not practical when multiple fusion algorithms need to be compared and ranked over large data sets. The work on using statistical methods to see if fusion measures match human ranks was reported in [125,126].

The assessment of multi-exposure image fusion is a special case of multi-modal image fusion as the images are of the same format or with the same intensity table. In [127], two perceptual quality measures, i.e. perceived local contrast and color saturation, were proposed for multi-exposure fusion assessment. Hassen et al. proposed to use the combination of image contrast preservation, sharpness, and structure preservation to evaluate the fused multi-exposure multi-focus image [128]. An objective image quality assessment (IQA), which utilizes multi-scale SSIM (structure similarity measure), was used to assess multi-exposure image fusion [129].

4.2. Discriminative fusion assessment

The assessment of discriminative fusion, whose result is often presented as a thematic map, can employ the similar criteria as to evaluate a classifier [143]. Given the ground truth reference, a confusion or error matrix X is employed to summarize the classification accuracy and highlight the thematic errors through comparison [143]. Let x_{ij} represent the number of pixels that actually belong to C_i but is classified into C_j for $i, j = 1, 2, \cdots, r$. The rows and columns correspond to the ground truth and calculated results respectively. Thus, the sample number for C_i, n_i, is the sum of row $x_{i+} \equiv \sum_{j=1}^{r} x_{ij}$ while the column total $x_{+j} \equiv \sum_{i=1}^{r} x_{ij}$ denotes the number of pixels classified into C_j [144].

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1r} \\ x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ x_{r1} & x_{r2} & \cdots & x_{rr} \end{bmatrix} \tag{20}$$

Thus, the class-averaged accuracy $C(X)$, overall accuracy $A(X)$, and kappa statistic coefficient $\kappa(X)$ are defined below for accuracy assessment [144]:

$$C(X) = \frac{1}{r} \sum_{i=1}^{r} \frac{x_{ii}}{n_i} \tag{21a}$$

$$A(X) = \frac{1}{N} \sum_{i=1}^{r} x_{ii} \tag{21b}$$

$$\kappa(X) = \frac{\sum_{i=1}^{r} x_{ii} - \frac{n_i x_{+i}}{N}}{N - \sum_{i=1}^{r} \frac{n_i x_{+i}}{N}} \tag{21c}$$

where $N \equiv \sum_{j=1}^{r} n_i$ is the total number of pixels. In [144], a coefficient based on Kullback–Leibler (KL) information in multinomial distributions was proposed to overcome the limitation of kappa coefficient, which is evaluated only by the diagonal elements, column and row totals of the error matrix X. The off-diagonal elements are neglected, and thus kappa is not an appropriate coefficient for accuracy assessment [144]. Moreover, two parameters quantity disagreement and allocation disagreement, were proposed in [145] to replace the kappa coefficient.

When the fused result gives a degree of a pixel belonging to a certain class or a confidence degree value, the calculations in Eq.(21a) to (21c) do not work. A fuzzy similarity based approach proposed in [146] is more feasible in such a situation. A comprehensive review on classification accuracy assessment is available in [147] for a more in-depth discussion.

5. Evaluation of the state-of-the-art

5.1. Significance test for statistical comparison

Most research on pixel-level image fusion usually seeks a comparison of the proposed algorithm with a baseline method in terms of selected fusion metrics. Even though variations in metric values are observed, the significance of such differences is not clear without a statistical analysis. The authors proposed conducting a non-parametric statistical test for comparison of fusion algorithms [148]. An *Image Fusion Toolbox Employing Significance Test* (ImTEST) was developed to facilitate such analysis [149]. The procedure for comparison is clearly defined in [149].

The significance is calculated based on the algorithm ranking rather than the metric values. The significance test is first conducted to check if there is a significant difference in the results. The null hypothesis (H_0) states that there is no effect or no difference, whereas the alternative hypothesis (H_1) claims the presence of an effect or a difference between algorithms. The Wilcoxon signed-ranks test and the Friedman test are applied to the cases of two or more than two algorithms respectively. The test reports "No" or "Yes" based on the *p*-value ($\alpha = 0.05$). For the tests of "Yes", a multiple ($1 \times N$) *post-hoc* procedure, i.e., Friedman test with FinnerâÇÖs correction, is followed to identify which pair of algorithms differ significantly. The Finner procedure is simple but powerful. Again, the ranks of algorithms are used. The newly proposed fusion algorithm is employed as a controlled reference. The results are presented with a "correlation plot" as shown in Figs. 14–16. The size and color of the circle show the fusion metric correlation values between the different algorithms. The approximated actual values are given in the lower part of the plot. If the difference between two algorithms is not significant, a red cross is placed in the corresponding square. For those (Algorithm 13 and 30) which only have two algorithms, the correlation plots are not presented.

5.2. Evaluation results and highlights

In this section, we survey the state-of-the-art in image fusion from 40 publications, which were screened based on specific image data sets as well as the selected fusion metrics in each study. The significance of 40 fusion algorithms is investigated with statistical analyses. Even though the comparison is not applicable across all the studies, the significance of the proposed algorithms can be understood through the comparison. The tests are categorized into three classes: 1) transform-based approach, 2) learning-based approach, and 3) other approaches. The abbreviations of the algorithms in Figs. 14–16 are provided in the Appendix of this paper for reference.

Fig. 14 shows the case of transform-based methods. Fig. 14(i) shows a performance comparison over ten methods, while Fig. 14(h) shows almost no differences. Since DWT is a common baseline, then comparisons shown an significance improvement with the sparse representation methods.

There are 19 algorithms in *transform-based approach* category as listed in Table 5. Among the 19 algorithms, six (32%) are not significant while the remaining 13 (68%) are significant in terms of the threshold value 0.05. Checking the twelve correlation plots with a significant indication in Fig. 14, we find five out of twelve have shown their significance in the *post-hoc* tests, which means the newly proposed algorithm is effective. However, we need to look at the context of the validation process for each new fusion algorithm. Three of them only conducted a limited comparison, e.g. compared with only three algorithms. The other two compared six other algorithms. Details are summarized in Table 8.

Similarly, for *leaning-based approaches*, there are six out of eight algorithms that demonstrate statistical significance (see Table 6). Among the six, five demonstrate their effectiveness in terms of the *post-hoc* test. The five proposed algorithms are compared with four, three,

and two algorithms respectively (see Fig. 15). In *other fusion approaches*, nine out of thirteen algorithms are significant and three among the nine do show the significance of the proposed algorithm. The details are available in Table 7 and Fig. 16.

Table 8 summarizes the statistic significant test results. In the all 40 algorithms, about 70% show the significant difference and about 50% indicate the difference coming from the proposed algorithms. In other words, about 35% of the total 40 algorithms demonstrated the effectiveness in the context of selected data sets, evaluation metrics, and algorithms to compare with. However, the insignificant test results do not necessarily mean the fusion method is not valid, just that: "**the proposed method is (or is not) significant in comparison with the selected fusion algorithms in terms of the specific data sets and selected fusion metrics**". Obviously, the impact on the statistical test originates from a number of factors. Currently, there are lack of benchmark data and commonly accepted protocols to conduct a comparative study, but this is critical for high quality research on image fusion.
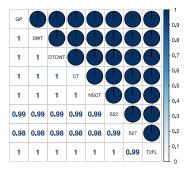
6. Discussions

One of the research efforts on image fusion is focused on representation of image features, which is typical an artful selection. The sparse representation is such an example to use the image's intrinsic features. It is worth more investigations in this direction. Recent advances in machine learning and deep neural network offer a mechanism to achieve the feature creation. In [28], an image is regarded as a distribution of attributes of objects (DAO) and image fusion becomes the matter of attribute composition. Thus, more attention should be paid to the physical implications of the image.
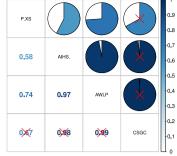
When a new fusion algorithm is proposed, the effectiveness and feasibility is usually validated or demonstrated by applying it to several different applications, where general fusion assessment metrics are adopted to evaluate the fusion performance. The disadvantages of this methodology is that the specific requirements of the application are not fully explored and considered. And the research may stay at the level of trying the different combinations for the image representation and no significant breakthroughs are achieved. Thus, the benefits offered by image fusion will not be fully understood. It is worth more efforts on the application itself when image fusion is considered as one of its solutions.

The fused image by combinative pixel-level fusion is intuitively meaningful. However, the significance offered by the fused image has not been fully revealed through quantitative analysis. Fusion performance assessment is indeed an application dependent issue. The general fusion metrics can reflect the certain aspect of the overall fusion process, e.g. information gain, image feature enrichment, etc; but is not always the complete performance.

Current performance assessment does not directly reflect the fusion objective. The assessments are concentrated on the fusion process, and thus a kind of "internal evaluation", which focuses on the quality of fused data. Most of the metrics assess the operational achievement, i.e. measuring how the entropy, energy, or features have been effectively transferred to the fused result. It is not clear how such achievements benefit the actual use of the fused image for an application involving humans, which is known as the bottom line "external evaluation". How the fused image can be further used is critical for evaluating the fusion performance. In [179], Hossny et al. introduced the concepts of Type-I and Type-II fusion errors, fusion capacity, and fusion worthiness, which are more general and target how the fusion outcome benefits the application. The use of human subjects is another practice. However, the challenges included in such campaign come from the neglected uncontrollable factors and the volume of the collected data. Limited scale of subjective assessment does not always assure a solid ground truth reference. Toet et al. [180] suggested the future research direction is "cognitive" fusion, which emphasizes the high-level image

(a) Type-2 fuzzy logic fusion in NSCT domain
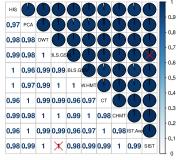
(b) Compressive sensing and graph cuts fusion

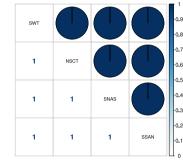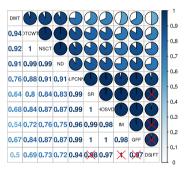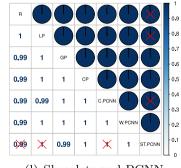(c) Shift-invariant shearlet transform (SIST) based fusion

(d) Directive contrast based fusion in NSCT domain

(e) Daubechies complex wavelet transform

(f) Hybrid multiresolution method

(g) Multicontourlet transform based fusion

(h) Support value transform

(i) Dense scale invariant feature transform

(j) Non-subsampled shearlet transform (NSST) and fast non-negative matrix factorization (NMF)

(k) Directive contrast in NSCT domain

(l) Shearlets and PCNN

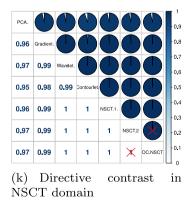**Fig. 14.** The correlation plots of the transform-based fusion algorithms. (The size and color of the circle show the fusion metric correlation values between the different algorithms. The approximated actual values are given in the lower part of the plot. If the difference between two algorithms is not significant, a red cross is put in the corresponding square.) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

(a) Group-sparse representation with dictionary learning

(b) Multi-task robust sparse representation

(c) Dictionary-based sparse representation

(d) Joint sparse representation

(e) Joint sparse model

(f) Joint patch clustering-based dictionary learning

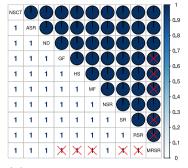**Fig. 15.** The correlation plots of the learning-based fusion algorithms (see Caption in Fig. 14 for explanation).

(a) Guided filtering fusion

(b) High order singular value decomposition (PKInf)

(c) Median filer and recursive filter

(d) Maximum likelihood

(e) Independent component analysis

(f) Energy evaluation model

(g) Image matting

(h) Edge-preserving decomposition and guided image filter

Fig. 16. The correlation plots of other fusion algorithms (see Caption in Fig. 14 for explanation). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 5
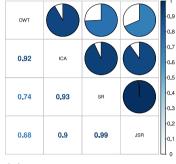Transform-based image fusion approaches.

No.	Algorithm	Data source	Significance
1	NSCT and intensity-hue-saturation transform	Table 1 in [51]	No (0.2160)
2	Improved additive-wavelet transform	Table 1 in [150]	No (0.8186)
3	Type-2 fuzzy logic (T2FL) in NSCT domain	Table 1 in [151]	Yes ($< 2.2e - 16$)
4	Compressive sensing and graph cuts	Table 3 in [152]	Yes (0.0099)
5	PCA/wavelet model-based fusion	Table 1 (top) in [153]	No (0.9891)
6	Complex wavelets transform (CWT)	Table 1 in [154]	No (0.05476)
7	$\grave{a}$ trous wavelet transform (AWT/UAW) based fusion	Table 2 in [155]	No (0.3468)
8	Shift-invariant shearlet transform (SIST) based fusion	Tables 1 and 2 in [156]	Yes ($6.7506e - 08$)
9	Directive contrast based fusion in NSCT domain	Table 1 in [157]	Yes ($1.5543e - 15$)
10	Daubechies complex wavelet transform (DCxWT)	Table 1–3 in [158]	Yes ($4.7696e - 08$)
11	Hybrid multiresolution method	Table 1 in [159]	Yes ($< 2.2e - 16$)
12	Multicontourlet transform based fusion (MCTF)	Table 3 in [160]	Yes ($4.4613e - 07$)
13	Discrete cosine transform (DCT)	Table 1 in [161]	Yes ($< 2.2e - 16$)
14	Support value transform (SVT)	Table 1 in [162]	Yes ($1.9250e - 11$)
15	Dense scale invariant feature transform (SIFT)	Table 1 in [74]	Yes ($< 2.2e - 16$)
16	Non-subsampled shearlet transform and fast non-negative matrix factorization (NF)	Table 1 and 2 in [163]	Yes ($1.4451e - 11$)
17	Directive contrast in NSCT domain (DC.NSCT)	Table 1 in [164]	Yes ($4.4542e - 13$)
18	Shearlets and PCNN (ST.PCNN)	Table 1 in [165]	Yes (0.0128)
19	Gradient entropy metric and p-Laplace diffusion constraint-based fusion algorithm	Table 2 in [166]	No (0.7975)

Table 6
Learning-based fusion approaches.

No.	Algorithm	Data source	Significance
20	Group-spare representation with dictionary learning (GSDL)	Table 3 in [167]	Yes ($7.1502e - 07$)
21	Multi-task sparse representation and spatial context	Table 2 in [168]	Yes ($1.8977e - 07$)
22	SparseFI	Table 1 in [24]	No (0.2013)
23	Online coupled dictionary learning	Table 2 in [169]	No (0.1361)
24	Dictionary-based sparse representation	Table 1 in [170]	Yes ($< 2.2e - 16$)
25	Joint sparse representation	Table 2 in [31]	Yes ($8.0329e - 10$)
26	Joint sparse model (JSM)	Table 2–4 in [33]	Yes ($< 2.2e - 16$)
27	Joint patch clustering-based dictionary learning (S.JCPD)	Table 1 in [171]	Yes ($3.8796e - 08$)

Table 8
Summary of the statistic significant tests for different fusion algorithms.

Statistic test results ($\alpha = 0.05$)		Transform-based approaches (19)	Learning-based approaches (8)	Other approaches (13)	Total (40)
Significance test	Yes	13 (68%)	6 (75%)	9 (69%)	28 (70%)
	No	6 (32%)	2 (25%)	4 (31%)	12 (30%)
post-hoc test	Yes	5 (42%)	5 (83%)	3 (37%)	14 (54%)
	No	7 (58%)	1 (17%)	5 (63%)	12 (46%)

7. Conclusion

Image fusion offers an effective way to integrate or combine images from multiple sources for a comprehensive perception of the target and/or scene. This paper screens and categorizes the state-of-the-art of fusion implementation and performance assessment. A complete process from fusion algorithm development to validation is presented. Moreover, an evaluation of 40 recently proposed fusion algorithms is conducted with significance test for statistical comparison. About 14 out of 40 algorithms (35%) show the significant differences from the others in the comprehensive comparison. No more than half of the algorithms are significant given the context of comparison settings, e.g. algorithms to compare with and image data sets in the experiments. Learning-based fusion is relatively new, but according to the *post-hoc* test, it appears to be more significant in its comparison.

understanding and interpretation. In the experiments described in [180], the human subjects were not asked to give the evaluation score directly. Instead, they were asked to find the contour of the target objective. The sketched contours were used as a subjective reference to evaluate the fusion performance. When the computer is used to further analyze or process the fused image, the classification/recognition/identification algorithms on the computer will tell which fused result is better.

Table 7
Other fusion approaches.

No.	Algorithm	Data source	Significance test (*p*-value)
28	Guided filtering fusion (GFF)	Table 1 in [86]	Yes ($< 2.2e - 16$)
29	High order singular value decomposition (PKInf)	Table 2 in [87]	Yes ($5.4057e - 13$)
30	Compressive sensing	Table 1 in [88]	Yes (0.003843)
31	Median filter and recursive filter (MFRF)	Table 1 in [172]	Yes ($2.1404e - 09$)
32	Intensity transformation function	Table 1 in [173]	No (0.06008)
33	Compressive measurement	Table 2 in [174]	No (0.09302)
34	Maximum likelihood (ML) fusion	Table 2 in [98]	Yes ($5.1993e - 06$)
35	Intuitionistic fuzzy sets	Table 5 and 7 in [175]	No (0.5270)
36	Independent component analysis (ICA)	Table 1 in [176]	Yes ($6.0862e - 13$)
37	Maximum local frequency	Table 4 in [29]	No (0.0828)
38	Energy evaluation model (EEM)	Table 1 in [92]	Yes (0.0031)
39	Image matting	Table 1 in [177]	Yes ($< 2.2e - 16$)
40	Edge-preserving decomposition and guided image filter (EPDGF)	Table 1 in [178]	Yes ($5.1503e - 12$)

The results suggest that fusion research should move from a feasibility study to problem-solving development. The research should demonstrate fusion as an effective solution in the context of a specific problem or application rather than simply showing a differently formulated fusion algorithm applicable to a number of general problems.

To facilitate future research on image fusion at its initial stage, benchmark data and assessment protocols need to be created and established to fill the gap for comparing different fusion methodologies. This remains a topic for our future work.

Appendix

Abbreviations of algorithms in the experiments

The abbreviations for algorithms that are used for the comparative study are given in Table 9.

Table 9
Abbreviations of algorithms in the comparative study.

AIHS	Adaptive intensity-hue-saturation (IHS) method	[181]
ASR	Adaptive sparse representation	[182]
AWLP	Proportional additive wavelet intensity-hue-saturation fusion	[183]
BLS.GSM	Bayesian-least squares-Gaussian-scale-mixture	[184]
CBF	Cross bilateral filter-based fusion	[185]
CHMT	HMT model in contourlet domain	[186]
Contrast	Contrast pyramid	[64]
CP	Contrast pyramid	[165]
C-PCNN	Contourlet-PCNN	[165]
CP.LS	Control-point based least square	[187]
CT, NSCT.1, NSCT.2	Contourlet transform	[50,188,189]
CVT	Curvelet transform	[190]
DCT	Discrete cosine transform	[161]
DCHWT	Multi-scale discrete cosine harmonic wavelet transform	[191]
DTCWT	Dual-tree complex wavelet transform	[192]
DWT/WTF	Discrete wavelet transform	[193]
GP	Gradient pyramid	[65]
Gradient	Gradient map image fusion	[65]
GRW	Generalized random walks	[194]
HOSVD	High order singular value decomposition	[87]
HMT	Hidden Markov tree	[186]
HS	Homogeneity similarity	[195]
iCAM06	Image color appearance model	[196]
I2CM	Improved intersecting cortical model	[197]
IG	Image gradient	[198]
LAP	Laplacian pyramid	[199]
Linear fusion	Linear mixing model	[200]
LWT	Lifting wavelet transform	[201]
MF/IFM	Image matting fusion	[170,177]
MI	Mutual information	[202]
MM	Mathematic morphology	[203]
MWT	Multiwavelet transform	[204]
ND	Neighbor distance	[205]
NSCT/CTF	Nonsubsampled contourlet transform	[188]
NSR	Sparse representation in NSCT	[74]
PCA	Principal component analysis	[206]
PCNN	Pulse coupled neural network	[207]
Photomatix	HDR images from uncompressed image data	[208]
P.XS	Variational model for panchromatic and multispectral images	[209]
Reinhard	Photographic tone reproduction	[210]
RSR	Robust sparse representation	[168]
RSSF	Region segmentation and spatial frequency	[211]
S22	Contrast based fusion in NSCT domain	[157]
S27	Neuro-fuzzy fusion	[212]
SAID	Saliency analysis and image decomposition	[213]
S.DCT, S.KSVD1, S.KSVD2	sparse representation-based method using the fixed or learned dictionaries	[171,214]
Sharp fusion	bilateral gradient-based sharpness criterion	[215]
SNAS	serial NSCT aiding SWT	[159]
SOMP	simultaneous orthogonal matching pursuit	[214]
SSAN	serial SWT aiding NSCT	[159]
ST	shearlet transform	[216]
SWT	stationary wavelet transform	[217]
UW.BLS.GSM	undecimated orthogonal wavelet transform	[218]
W.BLS.GSM	BLS.GSM for the orthogonal wavelet transform	[184]
WHMT	HMT model in wavelet domain	[219]
W-PCNN	wavelet-PCNN	[165]
WSSM	wavelet-based statistical sharpness measure	[220]

References

[1] R.S. Blum, Z. Liu, Multi-sensor Image Fusion and its Applications, Signal Processing and Communications, Taylor and Francis, 2005.

[2] A.A. Goshtasby, S. Nikolov, Image fusion: advances in the state of the art, Inf. Fus. 8 (2) (2007) 114–118. Special Issue on Image Fusion: Advances in the State of the Art

[3] F. Estrade, Temporal image fusion, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Columbus, Ohio, USA, (2014), pp. 1–9.

[4] S. Rokhsari, A.A. Elmdoust, M. Karimi, Multi times image fusion based on wavelet theory, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,XXXIX-B3, Melbourne, Australia, (2012), pp. 243–245.

[5] N. Joshi, M. Cohen, S.M. Rainier, Lucky imaging for multi-image denoising, sharpening, and haze removal, Computational Photography (ICCP), 2010 IEEE International Conference on, Cambridge, MA, USA, (2010), pp. 1–8.

[6] Z. Li, J. Zheng, Z. Zhu, S. Wu, Selectively detail-enhanced fusion of differently exposed images with moving objects, Image Process. IEEE Trans. 23 (10) (2014) 4372–4382.

[7] M.R.I.V.M. Group), Multi-image fusion, webpage, last accessed: January 2015, 2010. http://research.microsoft.com/en-us/um/redmond/groups/ivm/multiimagefusion/.

[8] D. Fu, B. Chen, J. Wang, X. Zhu, T. Hilker, An improved image fusion approach based on enhanced spatial and temporal the adaptive reflectance fusion model, Remote Sens. 5 (12) (2013) 6346–6360.

[9] Y. Li, S.B. Kang, N. Joshi, S. Seitz, D. Huttenlocher, Generating sharp panoramas from motion-blurred videos, Computer Vision and Pattern Recognition (CVPR), IEEE Conference on, San Francisco, CA, USA, 2010, (2010), pp. 2424–2431.

[10] P. Milanfar, Super-Resolution Imaging, CRC Press, 2010.

[11] F. Remondino, A. Roditakis, Human figure reconstruction and modeling from single image or monocular video sequence, 4th International Conference on 3D Digital Imaging and Modeling (3DIM), (2003), pp. 1–8.

[12] W.R. Crum, T. Hartkens, D.L.G. Hill, Non-rigid image registration: theory and practice, Br J Radiol 77 (2004) S140–S153. PMID: 15677356 . (suppl_2). doi:10.1259/bjr/25329214 .

[13] NASA, inside the flame: fusion imaging of the final space shuttle launch, last access: August 2016, 2011. http://www.nasa.gov/topics/shuttle_station/features/sts-135_launch_fused_imagery.html.

[14] M. Song, D. Tao, C. Chen, J. Bu, J. Luo, C. Zhang, Probabilistic exposure fusion, IEEE Trans. Image Process. 21 (1) (2012) 341–357.

[15] A.P. James, B.V. Dasarathy, Medical image fusion: a survey of the state of the art, information fusion, Inf. Fus. 19 (2014) 4–19. Special Issue on Information Fusion in Medical Image Computing and Systems

[16] C. Thomas, T. Ranchin, L. Wald, J. Chanussot, Synthesis of multispectral images to high spatial resolution: a critical review of fusion methods based on remote sensing physics, Geoscience and Remote Sensing, IEEE Transactions on 46 (5) (2008) 1301–1312.

[17] H. Ghassemian, A review of remote sensing image fusion methods, Inf. Fus. 32 (Part A) (2016) 75–89.

[18] Z. Liu, R. Laganiere, Context enhancement through infrared vision: a modified fusion scheme, Sig. Image Video Process. 1 (4) (2007) 293–301.

[19] A.A. Goshtasby, Fusion of multifocus images to maximize image information, in: K.L. Priddy, E. Ertin (Eds.), Proceedings of SPIE: Intelligent Computing: Theory and Applications IV, Orlando (Kissimmee), FL, USA, vol. 6229, 2006. 62290L–62290L–10

[20] T. Mertens, J. Kautz, F.V. Reeth, Exposure fusion, in: M. Alexa, S. Gortler, T. Ju (Eds.), Pacific Conference on Computer Graphics and Applications, Maui, Hawaii, USA, 2007.

[21] A.A. Goshtasby, Fusion of multi-exposure images, Image Vis. Comput. 23 (6) (2005) 611–618.

[22] C. Ancuti, C. Ancuti, Single image dehazing by multi-scale fusion, Image Process. IEEE Trans. 22 (8) (2013) 3271–3282.

[23] T. Ranchin, B. Aiazzi, L. Alparone, S. Baronti, L. Wald, Image fusion the {ARSIS} concept and some successful implementation schemes, {ISPRS}, J. Photogrammetry Remote Sensing 58 (12) (2003) 4–18.

[24] X. Zhu, R. Bamler, A sparse image fusion algorithm with application to pan-sharpening, Geosci. Remote Sensing IEEE Trans. 51 (5) (2013) 2827–2836.

[25] X. Chen, P. Flynn, K. Bowyer, Fusion of infrared and range data: Multi-modal face images, in: D. Zhang, A. Jain (Eds.), Lecture Notes in Computer Science: Advances in Biometrics, Lecture Notes in Computer Science, vol. 3832, Springer Berlin Heidelberg, 2005, pp. 55–63.

[26] M. Lagana, M. Preti, L. Forzoni, S. D'Onofrio, S. De Beni, A. Barberio, P. Cecconi, G. Baselli, Transcranial ultrasound and magnetic resonance image fusion with virtual navigator, Multimed. IEEE Trans. 15 (5) (2013) 1039–1048.

[27] B. Miles, I. Ayed, M. Law, G. Garvin, A. Fenster, S. Li, Spine image fusion via graph cuts, Biomed. Eng. IEEE Trans. 60 (7) (2013) 1841–1850.

[28] M. Zou, Y. Liu, Multi-sensor image fusion: Difficulties and key techniques, Image and Signal Processing, 2009. CISP '09. 2nd International Congress on, Tianjin, China, (2009), pp. 1–5.

[29] V.N. Gangapure, S. Banerjee, A.S. Chowdhury, Steerable local frequency based multispectral multifocus image fusion, Inf. Fus. 23 (2015) 99–115.

[30] B. Yang, S. Li, Multifocus image fusion and restoration with sparse representation, Instrum. Meas. IEEE Trans. 59 (4) (2010) 884–892.

[31] N. Yu, T. Qiu, F. Bi, A. Wang, Image features extraction and fusion based on joint sparse representation, IEEE J. Sel. Top. Sig. Process. 5 (5) (2011) 1074–1082.

[32] Q. Zhang, Y. Fu, H. Li, J. Zou, Dictionary learning method for joint sparse representation-based image fusion, Opt. Eng. 52 (5) (2013). 057006–057006

[33] H. Yin, S. Li, Multimodal image fusion with joint sparsity model, Opt. Eng. 50 (6) (2011). 067007-067007–10

[34] Z. Liu, W. Wu, Pattern recognition,Machine Intelligence and Biometrics, Fusion with Infrared Images for An Improved Performance and Perception, High Education Press and Springer, 2011, pp. 81–108. Ch. 4.

[35] R. Singh, M. Vatsa, A. Noore, Hierarchical fusion of multi-spectral face images for improved recognition performance, Inf. Fus. 9 (2008) 200–210.

[36] S.M. Desa, S. Hati, IR and visible face recognition using fusion of kernel based features, 19th International Conference on Pattern Recognition, Tampa, FL, USA, (2008), pp. 1–4.

[37] C.O. Conaire, N.E. O'Connor, A. Smeaton, Thermal-visual feature fusion for object tracking using multiple spatiogram trackers, Mach. Vis. Appl. 19 (2008) 483–494.

[38] D.J. Heeger, J.R. Bergen, Pyramid-based texture analysis/synthesis, Proceedings of the 22nd Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '95, ACM, New York, NY, USA, 1995, pp. 229–238.

[39] S. Li, B. Yang, J. Hu, Performance comparison of different multi-resolution transforms for image fusion, Inf. Fus. 12 (2) (2011) 74–84.

[40] P.J. Burt, R.J. Kolczynski, Enhanced image capture through fusion, Proceedings of 4th International Conference on Image Processing, (1993), pp. 248–251.

[41] E. Siomoncelli, W. Freeman, The steerable pyramid: A flexible architecture for multi-scale derivativecomputation, Proceedings of 2nd IEEE International Conference on Image Processing, Washington DC, (1995), pp. 444–447.

[42] Z. Liu, K. Tsukada, K. Hanasaki, Y.K. Ho, Y.P. Dai, Image fusion by using steerable pyramid, Pattern Recognit. Lett. 22 (2001) 929–939.

[43] H. Li, B.S. Manjunath, S.K. Mitra, Multisensor image fusion using the wavelet transform, Graph. Models Image Process. 57 (3) (1995) 235–245.

[44] S. Li, J.T. Kwok, Y. Wang, Using the discrete wavelet frame transform to merge landsat {TM} and {SPOT} panchromatic images, Inf. Fus. 3 (1) (2002) 17–23.

[45] H.-h. Wang, A new multiwavelet-based approach to image fusion, J. Math. Imaging Vis. 21 (2) (2004) 177–192.

[46] J.J. Lewis, R.J.O. Callaghan, S.G. Nikolov, D.R. Bull, N. Canagarajah, Pixel- and region-based image fusion with complex wavelets, Inf. Fus. 8 (2) (2007) 119–130. Special Issue on Image Fusion: Advances in the State of the Art.

[47] T. Wan, N. Canagarajah, A. Achim, Segmentation-driven image fusion based on alpha-stable modeling of wavelet coefficients, Multimed. IEEE Trans. 11 (4) (2009) 624–633.

[48] M.N. Do, M. Vetterli, The finite ridgelet transform for image representation, IEEE Trans. Image Process. 12 (1) (2003) 16–28.

[49] A. Krishn, V. Bhateja, A.S. Himanshi, PCA Based Medical Image Fusion in Ridgelet Domain, Springer International Publishing (2015) 475–482. Cham

[50] M.N. Do, M. Vetterli, The contourlet transform: an efficient directional multi-resolution image representation, IEEE Trans. Image Process. 14 (12) (2005) 2091–2106.

[51] W. Kong, Y. Lei, Y. Lei, X. Ni, Fusion technique for grey-scale visible light and infrared images based on non-subsampled contourlet transform and intensity-hue-saturation transform, IET Sig. Proc. 5 (1) (2011) 75–80, http://dx.doi.org/10.1049/iet-spr.2009.0263.

[52] M. Choi, R.Y. Kim, M.-R. Nam, H.O. Kim, Fusion of multispectral and panchromatic satellite images using the curvelet transform, IEEE Geosci. Remote Sens. Lett. 2 (2) (2005) 136–140.

[53] Z. Zhang, R.S. Blum, Image fusion for a digital camera application, Proceedings of 32nd Asilomar Conference on Signals Systems, and Computers, Monterey, CA, (1998), pp. 603–607.

[54] Z. Zhang, R. Blum, A categorization of multiscale-decomposition-based image fusion schemes with a performance study for a digital camera application, Proc. IEEE 87 (8) (1999) 1315–1326.

[55] Y. Liu, S. Liu, Z. Wang, Multi-focus image fusion with dense SIFT, Inf. Fus. 23 (0) (2015) 139–155.

[56] P. Shah, T. Srikanth, S. Merchant, U. Desai, Multimodal image/video fusion rule using generalized pixel significance based on statistical properties of the neighborhood, Sig. Image Video Process. 8 (4) (2014) 723–738.

[57] V.S. Petrovic, C.S. Xydeas, Cross-band pixel selection in multiresolution image fusion, in: B.V. Dasarathy (Ed.), Proceedings of SPIE, Sensor Fusion: Architectures, Algorithms, and Applications III, Orlando, FL, USA, vol. 3719, 1999, pp. 319–326.

[58] R. Shen, I. Cheng, A. Basu, Cross-scale coefficient selection for volumetric medical image fusion, Biomed> Eng. IEEE Trans. 60 (4) (2013) 1069–1079.

[59] R. Shen, I. Cheng, J. Shi, A. Basu, Generalized random walks for fusion of multi-exposure images, IEEE Trans. Image Process. 20 (12) (2011) 3634–3646.

[60] J. Yang, R.S. Blum, A statistical signal processing approach to image fusion using hidden Markov models, in: R.S. Blum, Z. Liu (Eds.), Multi-Sensor Image Fusion and Its Applications, Signal Processing and Communications, Taylor & Francis, Boca Raton, FL, USA, 2005, pp. 265–287.

[61] F. Flitti, C. Collet, E. Slezak, Image fusion based on pyramidal multiband multi-resolution markovian analysis, Sig. Image Video Process. 3 (3) (2009) 275–289.

[62] S.C. Nercessian, K.A. Panetta, S.S. Agaian, Multiscale image fusion using an adaptive similarity-based sensor weighting scheme and human visual system-inspired contrast measure, J. Electron. Imaging 21 (2) (2012). 021112-1–021112-13

[63] P. Shah, B. Reddy, S. Merchant, U. Desai, Context enhancement to reveal a camouflaged target and to assist target localization by fusion of multispectral surveillance videos, Sig. Image Video Process. 7 (3) (2013) 537–552.

[64] A. Toet, L.J. van Ruyven, J.M. Valeton, Merging thermal and visual images by a contrast pyramid, Opt. Eng. 28 (7) (1989). 287789-287789–

[65] V. Petrovic, C. Xydeas, Gradient-based multiresolution image fusion, IEEE Trans. Image Process. 13 (2) (2004) 228–237.

[66] H. Fu, E. Dellandrea, L. Chen, Reconstructive and discriminative sparse representation for visual object categorization, in: J. Hoey, S. McKenna, E. Trucco (Eds.), Proceedings of the British Machine Vision Conference, BMVA Press, 2011. 39.1–39.12

[67] M. Aharon, M. Elad, A. Bruckstein, K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation, IEEE Trans. Signal Process. 54 (11) (2006) 4311–4322.

[68] S. Mallat, Z. Zhang, Matching pursuits with time-frequency dictionaries, IEEE Trans. Signal Process. 41 (12) (1993) 3397–3415.

[69] Y. Pati, R. Rezaiifar, P. Krishnaprasad, Orthogonal matching pursuit: recursive function approximation with applications to wavelet decomposition, The Twenty-Seventh Asilomar Conference on Signals, Systems and Computers, vol. 1, (1993), pp. 40–44.

[70] M. Duarte, S. Sarvotham, D. Baron, M. Wakin, R. Baraniuk, Distributed compressed sensing of jointly sparse signals, Thirty-Ninth Asilomar Conference on Signals, Systems and Computers, (2005), pp. 1537–1541.

[71] K. Engan, S. Aase, J.H. Husoy, Method of optimal directions for frame design, IEEE International Conference on Acoustics, Speech, and Signal Processing, vol. 5, (1999), pp. 2443–2446.

[72] USC, The USC-SIPI image database, online, retrieved in november 2014, 2017. http://sipi.usc.edu/database/.

[73] X.X. Zhu, R. Bamler, Tomographic sar inversion by l_1 -norm regularization - the compressive sensing approach, IEEE Trans. Geosci. Remote Sens. 48 (10) (2010) 3839–3846.

[74] Y. Liu, S. Liu, Z. Wang, A general framework for image fusion based on multi-scale transform and sparse representation, Inf. Fus. 24 (2014) 147–164.

[75] S. Li, J.T. Kwok, Y. Wang, Multifocus image fusion using artificial neural networks, Pattern Recognit. Lett. 23 (8) (2002) 985–997.

[76] Y. Xia, M. Kamel, Novel cooperative neural fusion algorithms for image restoration and image fusion, Image Process. IEEE Trans. 16 (2) (2007) 367–381.

[77] J. Zhong, B. Yang, G. Huang, F. Zhong, Z. Chen, Remote sensing image fusion with convolutional neural network, Sens. Imaging 17 (1) (2016) 1–16.

[78] A. Toet, J. Walraven, New false color mapping for image fusion, Opt. Eng. 35 (3) (1996) 650–658.

[79] A. Toet, Natural colour mapping for multiband nightvision imagery, Inf. Fus. 4 (3) (2003) 155–166.

[80] Z. Xue, R. Blum, Y. Li, Fusion of visual and ir images for concealed weapon detection, Proceedings of ISIF 2002, (2002), pp. 1198–1205.

[81] S. Yin, L. Cao, Y. Ling, G. Jin, One color contrast enhanced infrared and visible image fusion method, Infrared Phys. Technol. 53 (2) (2010) 146–150.

[82] Y. Zheng, W. Dong, E.P. Blasch, Qualitative and quantitative comparisons of multispectral night vision colorization techniques, Opt. Eng. 51 (8) (2012) 087004–1–087004–16.

[83] Y. Zheng, K. Reese, E. Blasch, P. McManamon, Qualitative evaluations and comparisons of six night-vision colorization methods (2013).

[84] H. Yeganeh, Z. Wang, Objective quality assessment of tone mapped images, IEEE Trans. Image Process. 22 (2) (2013) 657–667.

[85] K. Sharma, M. Sharma, Image fusion based on image decomposition using self-fractional fourier functions, Sig. Image Video Process. 8 (7) (2014) 1335–1344.

[86] S. Li, X. Kang, J. Hu, Image fusion with guided filtering, IEEE Trans. Image Process. 22 (7) (2013) 2864–2875.

[87] J. Liang, Y. He, D. Liu, X. Zeng, Image fusion using higher order singular value decomposition, Image Process. IEEE Trans. 21 (5) (2012) 2898–2909.

[88] X. Li, S.Y. Qin, Efficient fusion for infrared and visible images based on compressive sensing principle, IET Image Proc. 5 (2) (2011) 141–147.

[89] H. Hariharan, A. Koschan, B. Abidi, A. Gribok, M. Abidi, Fusion of visible and infrared images using empirical mode decomposition to improve face recognition, Proceedings of International Conference on Image Processing, (2006), pp. 2049–2052.

[90] H. Li, Y. Zheng, Image fusion algorithm using pyramidal empirical mode decomposition, Proceedings of the 9th International Conference on Hybrid Intelligent Systems, Washington, DC, USA, (2009), pp. 152–157.

[91] V. Tsagaris, V. Anastassopoulos, Fusion of visible and infrared imagery for night color vision, Displays 26 (4–5) (2005) 191–196.

[92] Y. Zhang, Adaptive region-based image fusion using energy evaluation model for fusion decision, Sig. Image Video Process. 1 (3) (2007) 215–223.

[93] T. Chan, S. Esedoglu, F. Park, A. Yip, Recent developments in total variation image restoration, in: N. Paragios, Y. Chen, O.D. Faugeras (Eds.), Handbook of Mathematical Models in Computer Vision,Springer-Verlag US, 2006, pp. 17–31. Ch. 1.

[94] V. Caselles, Total variation based image denoising and restoration, Proceedings oh the International Congress of Mathematicians, Vol. 3, Madrid, Spain, (2006), pp. 1453–1472.

[95] W.-W. Wang, P.-L. Shui, X.-C. Feng, Variational models for fusion and denoising of multifocus images, Sig. Process. Lett. IEEE 15 (2008) 65–68.

[96] D.M. Greig, B.T. Porteous, A.H. Seheult, Exact maximum a posteriori estimation for binary images, J. R. Stat. Soc. Ser. B 51 (2) (1989) 271–279.

[97] M. Kumar, S. Dass, A total variation-based algorithm for pixel-level image fusion, IEEE Trans. Image Process. 18 (9) (2009) 2137–2143.

[98] S. Chen, Q. Guo, H. Leung, E. Bosse, A maximum likelihood approach to joint image registration and fusion, Image Process. IEEE Trans. 20 (5) (2011) 1363–1372.

[99] C.H. Seng, A. Bouzerdoum, M. Amin, S.L. Phung, Two-stage fuzzy fusion with applications to through-the-wall radar imaging, Geosci. Remote Sensing Lett. IEEE 10 (4) (2013) 687–691.

[100] S. Gorthi, M.B. Cuadra, P.-A. Tercier, A. Allal, J.-P. Thiran, Weighted shape-based averaging with neighborhood prior model for multiple atlas fusion-based medical image segmentation, Sig. Process. Lett. IEEE 20 (11) (2013) 1034–1037.

[101] T. Rohlfing, J. Maurer, C.r., shape-based averaging, Image Process. IEEE Trans. 16 (1) (2007) 153–161.

[102] X. Huang, R. Netravali, H. Man, V. Lawrence, Multi-sensor fusion of infrared and electro-optic signals for high resolution night images, Sensors 12 (8) (2012) 10326–10338.

[103] V. Sharma, J.W. Davis, Feature-level fusion for object segmentation using mutual information, Adv. Pattern Recognit. (2009) 295–320. Springer-Verlag London Limited

[104] B. Waske, S. van der Linden, Classifying multilevel imagery from sar and optical sensors by decision fusion, Geosci. Remote Sensing IEEE Trans. 46 (5) (2008) 1457–1466.

[105] H.G. Chew, C.C. Lim, R.E. Bogner, Optimization and Control with Applications, An Implementation of Training dual-nu Support Vector Machine, kluwer, dordrecht, 2004. Ch.

[106] G. Liu, Z. Jing, S. Sun, Image fusion based on an expectation maximization algorithm, Opt. Eng. 44 (7) (2005) 077001.

[107] D. Ai, G. Duan, X. Han, Y.-W. Chen, Generalized n-dimensional independent component analysis and its application to multiple feature selection and fusion for image classification, Neurocomputing 103 (2013) 186–197.

[108] M. Zribi, Non-parametric and region-based image fusion with bootstrap sampling, Information Fusion 11 (2) (2010) 85–94.

[109] J. Yang, R.S. Blum, A Region-based Image Fusion Method Using the Expectation-maximization Algorithm, 2006 40th Annual Conference on Information Sciences and Systems, (2006), pp. 468–473.

[110] R. Singh, M. Vatsa, A. Noore, Integrated multilevel image fusion and match score fusion of visible and infrared face images for robust face recognition, Pattern Recognit. 41 (2008) 880–893.

[111] R. Singh, M. Vatsa, A. Noore, Face recognition with disguise and single gallery images, Image Vis. Comput. 27 (3) (2009) 245–257.

[112] H. Liu, S. Li, Decision fusion of sparse representation and support vector machine for {SAR} image target recognition, Neurocomputing 113 (2013) 97–104.

[113] R. Huan, Y. Pan, Decision fusion strategies for sar image target recognition, Radar Sonar Navig. IET 5 (7) (2011) 747–755.

[114] W. Li, S. Prasad, J.E. Fowler, Decision fusion in kernel-induced spaces for hyperspectral image classification, IEEE Trans. Geosci. Remote Sens. 52 (6) (2014) 3399–3411, http://dx.doi.org/10.1109/TGRS.2013.2272760.

[115] R. Nishii, A Markov random field-based approach to decision-level fusion for remote sensing image classification, IEEE Trans. Geosci. Remote Sens. 41 (10) (2003) 2316–2319.

[116] G. Thoonen, Z. Mahmood, S. Peeters, P. Scheunders, Multisource classification of color and hyperspectral images using color attribute profiles and composite decision fusion, Sel. Top. Appl. Earth Obs. Remote Sensing IEEE J. 5 (2) (2012) 510–521.

[117] N. Kausar, A. Majid, Random forest-based scheme using feature and decision levels information for multi-focus image fusion, Pattern Anal. Appl. 19 (1) (2016) 221–236.

[118] N. Mitrakis, C. Topaloglou, T. Alexandridis, J. Theocharis, G. Zalidis, Decision fusion of ga self-organizing neuro-fuzzy multilayered classifiers for land cover classification using textural and spectral features, geoscience and remote sensing, IEEE Trans. 46 (7) (2008) 2137–2152.

[119] A. Wang, J. Jiang, H. Zhang, Multi-sensor image decision level fusion detection algorithm based on d-s evidence theory, Instrumentation and Measurement, Computer, Communication and Control (IMCCC), 2014 Fourth International Conference on, Harbin, China, (2014), pp. 620–623.

[120] Y. Wang, W. Chen, S. Mao, Multi-sensor decision level image fusion based on fuzzy theory and unsupervised FCM, 2006.

[121] P. Du, W. Zhang, S. Zhang, J. Xia, Hyperspectral remote sensing image classification based on decision level fusion, 2009 IEEE International Geoscience and Remote Sensing Symposium, vol. 4, (2009). IV–940–IV–943

[122] Z. Zhou, Ensemble methods: foundations and algorithms, Machine Learning and Pattern Recognition Series, Chapman & Hall /CRC, Boca Raton, FL, USA, 2012.

[123] L. Rokach, Pattern Classification Using Ensemble Methods, World Scientific Publishing Co., Inc., River Edge, NJ, USA, 2010.

[124] Z. Liu, E. Blasch, Z. Xue, J. Zhao, R. Laganière, W. Wu, Objective assessment of multiresolution fusion algorithms for context enhancement in night vision: a comparative study, IEEE Trans. Pattern Anal. Mach. Intell. 34 (1) (2012) 94–109.

[125] C. Wei, L. Kaplan, S. Burks, R. Blum, Diffuse prior monotonic likelihood ratio test for evaluation of fused image quality measures, IEEE Trans. Image Process. 20 (2) (2011) 327–344.

[126] L.M. Kaplan, S.D. Burks, R.S. Blum, R.K. Moore, Q. Nguyen, Analysis of image quality for image fusion via monotonic correlation, IEEE J. Sel. Top. Sig. Process. 3 (2) (2009) 222–235.

[127] R. Shen, I. Cheng, A. Basu, QoE-Based multi-exposure fusion in hierarchical multivariate gaussian CRF, IEEE Trans. Image Process. 22 (6) (2013) 2469–2478.

[128] R. Hassen, Z. Wang, M.M.A. Salama, Objective quality assessment for multi-exposure multifocus image fusion, IEEE Trans. Image Process. 24 (9) (2015) 2712–2724, http://dx.doi.org/10.1109/TIP.2015.2428051.

[129] K. Ma, K. Zeng, Z. Wang, Perceptual quality assessment for multi-exposure image fusion, IEEE Trans. Image Process. 24 (11) (2015) 3345–3356.

[130] M. Hossny, S. Nahavandi, D. Vreighton, Comments on 'information measure for performance of image fusion', Electron. Lett. 44 (18) (2008).

[131] N. Cvejic, C.N. Canagarajah, D.R. Bull, Image fusion metric based on mutual information and tsallis entropy, Electron. Lett. 42 (11) (2006).

[132] Q. Wang, Y. Shen, J. Jin, Performance evaluation of image fusion techniques, in:

T. Stathaki (Ed.), Image Fusion: Algorithms and Applications, Elsevier, 2008, pp. 469–492. Ch. 19

[133] C.S. Xydeas, V. Petrovic, Objective image fusion performance measure, Electron. Lett. 36 (4) (2000) 308–309.

[134] P. Wang, B. Liu, A novel image fusion metric based on multi-scale analysis, Proceedings of IEEE International Conference on Signal Processing, (2008), pp. 965–968.

[135] Y. Zheng, E.A. Essock, B.C. Hansen, A.M. Haun, A new metric based on extended spatial frequency and its application to DWT based fusion algorithms, Inf. Fus. 8 (2) (2007) 177–192.

[136] J. Zhao, R. Laganiere, Z. Liu, Performance assessment of combinative pixel-level image fusion based on an absolute feature measurement, Int. J. Innov. Comput. Inf. Control 3 (6(A)) (2007) 1433–1447.

[137] Z. Liu, D.S. Forsyth, R. Laganiere, A feature-based metric for the quantitative evaluation of pixel-level image fusion, Comput. Vis. Image Understand. 109 (1) (2008) 56–68.

[138] G. Piella, H. Heijmans, A new quality metric for image fusion, Proceedings of International Conference on Image Processing, Bacelona, (2003).

[139] N. Cvejic, A. Loza, D. Bul, N. Canagarajah, A similarity metric for assessment of image fusion algorithms, Int. J. Sig. Process. 2 (3) (2005) 178–182.

[140] C. Yang, J. Zhang, X. Wang, X. Liu, A novel similarity based quality metric for image fusion, Inf. Fus. 9 (2008) 156–160.

[141] H. Chen, P.K. Varshney, A human perception inspired quality metric for image fusion based on regional information, Inf. Fus. 8 (2007) 193–207.

[142] Y. Chen, R.S. Blum, A new automated quality assessment algorithm for image fusion, Image Vis. Comput. 27 (2009) 1421–1432.

[143] G.M. Foody, The evaluation and comparison of thematic maps derived from remote sensing, 7th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, (2006), pp. 18–31.

[144] R. Nishii, S. Tanaka, Accuracy and inaccuracy assessments in land-cover classification, IEEE Trans. Geosci. Remote Sens. 37 (1) (1999) 491–498.

[145] R.G.P. Jr, M. Millones, Death to kappa: birth of quantity disagreement and allocation disagreement for accuracy assessment, J. Remote Sens. 32 (15) (2011) 4407–4429. arXiv:https://doi.org/10.1080/01431161.2011.552923 .

[146] G. Jager, U. Benz, Measures of classification accuracy based on fuzzy similarity, IEEE Trans. Geosci. Remote Sens. 38 (3) (2000) 1462–1467.

[147] G.M. Foody, Status of land cover classification accuracy assessment, Remote Sens. Environ. 80 (1) (2002) 185–201.

[148] Z. Liu, E. Blasch, Statistical analysis of the performance assessment results for pixel-level image fusion, 17th International Conference on Information Fusion, Salamanca, Spain, (2014), pp. 1–8.

[149] Z. Liu, E. Blasch, V. John, Statistical comparison of pixel-level fusion algorithms: recommendations, Inf. Fus. 36 (2017) 251–260.

[150] Y. Kim, C. Lee, D. Han, Y. Kim, Y. Kim, Improved additive-wavelet image fusion, IEEE Geosci. Remote Sens. Lett. 8 (2) (2011) 263–267.

[151] Y. Yang, Y. Que, S. Huang, P. Lin, Multimodal sensor medical image fusion based on type-2 fuzzy logic in nsct domain, IEEE Sens. J. 16 (10) (2016) 3735–3745.

[152] V. Harikumar, P. Gajjar, M. Joshi, M. Raval, Multiresolution image fusion: use of compressive sensing and graph cuts, selected topics in applied earth observations and remote sensing, IEEE Journal of 7 (5) (2014) 1771–1780.

[153] F. Palsson, J. Sveinsson, M. Ulfarsson, J. Benediktsson, Model-based fusion of multi- and hyperspectral images using PCA and wavelets, IEEE Transactions on Geosci. Remote Sens. 53 (5) (2015) 2652–2663.

[154] N. Cvejic, T.S.A.E. nen, S. Godsill, A nonreference image fusion metric based on the regional importance measure, IEEE J. Sel. Top. Sig. Process. 3 (2) (2009) 212–221.

[155] Y. Byun, J. Choi, Y. Han, An area-based image fusion scheme for the integration of sar and optical satellite imagery, Sel. Top. Appl. Earth Obs. Remote Sens. IEEE J. 6 (5) (2013) 2212–2220.

[156] L. Wang, B. Li, L.F. Tian, EGGDD: an explicit dependency model for multi-modal medical image fusion in shift-invariant shearlet transform domain, Inf. Fus. 19 (0) (2014) 29–37. Special Issue on Information Fusion in Medical Image Computing and Systems.

[157] G. Bhatnagar, Q. Wu, Z. Liu, Directive contrast based multimodal medical image fusion in NSCT domain, IEEE Trans. Multimed. 15 (5) (2013) 1014–1024.

[158] R. Singh, A. Khare, Fusion of multimodal medical images using daubechies complex wavelet transform a multiresolution approach, Inf. Fus. 19 (0) (2014) 49–60. Special Issue on Information Fusion in Medical Image Computing and Systems

[159] S. Li, B. Yang, Hybrid multiresolution method for multisensor multimodal image fusion, Sens. J. IEEE 10 (9) (2010) 1519–1526.

[160] X. Chang, L. Jiao, F. Liu, F. Xin, Multicontourlet-based adaptive fusion of infrared and visible remote sensing images, Geosci. Remote Sens. Lett. IEEE 7 (3) (2010) 549–553.

[161] L. Cao, L. Jin, H. Tao, G. Li, Z. Zhuang, Y. Zhang, Multi-focus image fusion based on spatial frequency in discrete cosine transform domain, Sig. Process. Lett. IEEE 22 (2) (2015) 220–224.

[162] S. Zheng, W.z. Shi, J. Liu, G.x. Zhu, J.-W. Tian, Multisource image fusion method using support value transform, Image Processing, IEEE Transactions on 16 (7) (2007) 1831–1839.

[163] W. Kong, Y. Lei, H. Zhao, Adaptive fusion method of visible light and infrared images based on non-subsampled shearlet transform and fast non-negative matrix factorization, Infrared Phys. Technol. 67 (2014) 161–172.

[164] G. Bhatnagar, Q.M.J. Wu, Z. Liu, A new contrast based multimodal medical image fusion framework, Neurocomputing 157 (2014) 143–152.

[165] S. Cheng, M. Qiguang, X. Pengfei, A novel algorithm of remote sensing image fusion based on shearlets and {PCNN}, Neurocomputing 117 (2013) 47–53.

[166] W. Zhao, Z. Xu, J. Zhao, Gradient entropy metric and p-laplace diffusion constraint-based algorithm for noisy multispectral image fusion, Inf. Fus. 27 (2016) 138–149.

[167] S. Li, H. Yin, L. Fang, Group-sparse representation with dictionary learning for medical image denoising and fusion, Biomed. Eng. IEEE Trans. 59 (12) (2012) 3450–3459.

[168] Q. Zhang, M.D. Levine, Robust multi-focus image fusion using multi-task sparse representation and spatial context, IEEE Trans. Image Process. 25 (5) (2016) 2045–2058, http://dx.doi.org/10.1109/TIP.2016.2524212.

[169] M. Guo, H. Zhang, J. Li, L. Zhang, H. Shen, An online coupled dictionary learning approach for remote sensing image fusion, IEEE Journal of Sel. Top. Appl. Earth Obs. Remote Sens. 7 (4) (2014) 1284–1294.

[170] M. Nejati, S. Samavi, S. Shirani, Multi-focus image fusion using dictionary-based sparse representation, Inf. Fus. 25 (2015) 72–84.

[171] M. Kim, D.K. Han, H. Ko, Joint patch clustering-based dictionary learning for multimodal image fusion, Inf. Fus. 27 (2016) 198–214.

[172] S. Li, X. Kang, Fast multi-exposure image fusion with median filter and recursive filter, IEEE Trans. Consumer Electron. 58 (2) (2012) 626–632.

[173] L.C. Jiao, Q.X. Tang, Fusion for visual context enhancement using intensity transformation function of infrared images, Electron. Lett. 49 (12) (2013) 751–752.

[174] B. Kang, W.P. Zhu, J. Yan, Fusion framework for multi-focus images based on compressed sensing, IET Image Proc. 7 (4) (2013) 290–299.

[175] P. Balasubramaniam, V. Ananthi, Image fusion using intuitionistic fuzzy sets, Inf. Fus. 20 (0) (2014) 21–30.

[176] N. Cvejic, D. Bull, N. Canagarajah, Region-based multimodal image fusion using ica bases, Sens. J. IEEE 7 (5) (2007) 743–751.

[177] S. Li, X. Kang, J. Hu, B. Yang, Image matting for fusion of multi-focus images in dynamic scenes, Inf. Fus. 14 (2) (2013) 147–162.

[178] W. Gan, X. Wu, W. Wu, X. Yang, C. Ren, X. He, K. Liu, Infrared and visible image fusion with the use of multi-scale edge-preserving decomposition and guided image filter, Infrared Phys. Technol. 72 (2015) 37–51.

[179] M. Hossny, S. Nahavandi, D. Creighton, A. Bhatti, Towards autonomous image fusion, Control Automation Robotics Vision (ICARCV), 2010 11th International Conference on, Singapore, (2010), pp. 1748–1754.

[180] A. Toet, M. Hogervorst, S. Nikolov, J. Lewis, T. Dixon, D. Bull, C. Canagarajah, Towards cognitive image fusion, Inf. Fus. 11 (2) (2010) 95–113.

[181] S. Rahmani, M. Strait, D. Merkurjev, M. Moeller, T. Wittman, An adaptive IHS pan-sharpening method, IEEE Geosci. Remote Sens. Lett. 7 (4) (2010) 746–750.

[182] Y. Liu, Z. Wang, Simultaneous image fusion and denoising with adaptive sparse representation, IET Image Proc. 9 (5) (2015) 347–357.

[183] X. Otazu, M. Gonzalez-Audicana, O. Fors, J. Nunez, Introduction of sensor spectral response into image fusion methods. application to wavelet-based methods, IEEE Trans. Geosci. Remote Sens. 43 (10) (2005) 2376–2385.

[184] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of gaussians in the wavelet domain, IEEE Trans. Image Process. 12 (11) (2003) 1338–1351.

[185] B.K.S. Kumar, Image fusion based on pixel significance using cross bilateral filter, signal, Image Video Process. 9 (5) (2015) 1193–1204, http://dx.doi.org/10.1007/s11760-013-0556-9.

[186] D.D.Y. Po, M.N. Do, Directional multiscale modeling of images using the contourlet transform, IEEE Trans. Image Process. 15 (6) (2006) 1610–1620.

[187] B. Zitova, J. Flusser, Image registration methods: a survey, Image Vis. Comput. 21 (11) (2003) 977–1000.

[188] Q. Zhang, B.l. Guo, Multifocus image fusion using the nonsubsampled contourlet transform, Sig. Process. 89 (7) (2009) 1334–1346.

[189] Y. Chai, H. Li, X. Zhang, Multifocus image fusion based on features contrast of multiscale products in nonsubsampled contourlet transform domain, Optik 123 (7) (2012) 569–581.

[190] L. Tessens, A. Ledda, A. Pizurica, W. Philips, Extending the depth of field in microscopy through curvelet-based frequency-adaptive image fusion, IEEE International Conference on Acoustics, Speech and Signal Processing, vol. 1, (2007). I–861–I–864

[191] B.K.S. Kumar, Multifocus and multispectral image fusion based on pixel significance using discrete cosine harmonic wavelet transform, Sig. Image Video Process. 7 (6) (2013) 1125–1143, http://dx.doi.org/10.1007/s11760-012-0361-x.

[192] N.G. Kingsbury, The dual-tree complex wavelet transform: A new technique for shift invariance and directional filters, Proceedings of 8th IEEE Digital Signal Processing Workshop, (1998), pp. 319–322.

[193] S.M.M. Rahman, M.O. Ahmad, M.N.S. Swamy, Contrast-based fusion of noisy images using discrete wavelet transform, IET Image Proc. 4 (5) (2010) 374–384.

[194] R. Shen, I. Cheng, J. Shi, A. Basu, Generalized random walks for fusion of multiexposure images, IEEE Trans. Image Process. 20 (12) (2011) 3634–3646, http://dx.doi.org/10.1109/TIP.2011.2150235.

[195] H. Li, Y. Chai, H. Yin, G. Liu, Multifocus image fusion and denoising scheme based on homogeneity similarity, Opt. Commun. 285 (2) (2012) 91–100.

[196] J. Kuang, G.M. Johnson, M.D. Fairchild, Icam06: a refined image appearance model for {HDR} image rendering, J. Vis. Commun. Image Represent. 18 (5) (2007) 406–414. Special issue on High Dynamic Range Imaging

[197] W.W. Kong, Multi-sensor image fusion based on NSST domain I2CM, Electron. Lett. 49 (13) (2013) 802–803.

[198] S. Li, J.T. Kwok, Y. Wang, Combination of images with diverse focuses using the spatial frequency, Inf. Fus. 2 (3) (2001) 169–176.

[199] P. Burt, E. Adelson, The laplacian pyramid as a compact image code, IEEE Trans. Commun. 31 (4) (1983) 532–540.

[200] J.G.P.W. Clevers, R. Zurita Milla, Multisensor and multiresolution image fusion

using the linear mixing model, Image Fusion: Algorithms and Applications, Elsevier Academic Press, 2008, pp. 67–84.

[201] S. Kor, U.S. Tiwary, Feature level fusion of multimodal medical images in lifting wavelet transform domain, 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS'04), vol. 1, (2004), pp. 1479–1482.

[202] H.m. Chen, P.K. Varshney, Mutual information-based ct-mr brain image registration using generalized partial volume joint histogram estimation, IEEE Trans. Med. Imaging 22 (9) (2003) 1111–1119.

[203] I. De, B. Chanda, B. Chattopadhyay, Enhancing effective depth-of-field by image fusion using mathematical morphology, Image Vis. Comput. 24 (12) (2006) 1278–1287.

[204] Y.L.. Yang, J. Sun, PET/CT medical image fusion algorithm based on multiwavelet transform, Second International Conference on Advanced Computer Control (ICACC), vol. 2, (2010), pp. 264–268.

[205] H. Zhao, Z. Shang, Y.Y. Tang, B. Fang, Multi-focus image fusion based on the neighbor distance, Pattern Recognit. 46 (3) (2013) 1002–1011.

[206] H. Yesou, Y. Besnus, J. Rolet, Extraction of spectral information from landsat tm data and merger with spot panchromatic imagery a contribution to the study of geological structures, ISPRS J. Photogramm. Remote Sens. 48 (5) (1993) 23–36.

[207] Z. Wang, Y. Ma, Medical image fusion using m-pcnn, Inf. Fus. 9 (2) (2008) 176–185.

[208] G. Joffre, W. puech, F. Comby, J. Joffre, High dynamic range images from digital cameras raw data, Proceedings of ACM SIGGRAPH, (2005), p. 72.

[209] C. Ballester, V. Caselles, L. Igual, J. Verdera, B. Rougé, A variational model for p+XS image fusion, Int. J. Comput. Vis. 69 (1) (2006) 43–58.

[210] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photographic tone reproduction for digital images, Proceedings of the 29th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '02, ACM, New York, NY, USA, (2002), pp. 267–276.

[211] A. Agarwala, M. Dontcheva, M. Agrawala, S. Drucker, A. Colburn, B. Curless, D. Salesin, M. Cohen, Interactive digital photomontage, ACM Trans. Graph. 23 (3) (2004) 294–302.

[212] S. Das, M. Kundu, A neuro-fuzzy approach for medical image fusion, Biomed. Eng. IEEE Trans. 60 (12) (2013) 3347–3353.

[213] J. Zhao, Q. Zhou, Y. Chen, H. Feng, Z. Xu, Q. Li, Fusion of visible and infrared images using saliency analysis and detail preserving based image decomposition, Infrared Phys. Technol. 56 (2013) 93–99.

[214] B. Yang, S. Li, Pixel-level image fusion with simultaneous orthogonal matching pursuit, Inf. Fus. 13 (1) (2012) 10–19.

[215] J. Tian, L. Chen, L. Ma, W. Yu, Multi-focus image fusion using a bilateral gradient-based sharpness criterion, Opt. Commun. 284 (1) (2011) 80–87.

[216] Q.g. Miao, C. Shi, P.f. Xu, M. Yang, Y.b. Shi, A novel algorithm of image fusion using shearlets, Opt. Commun. 284 (6) (2011) 1540–1547.

[217] O. Rockinger, Image sequence fusion using a shift-invariant wavelet transform, Proceedings of International Conference on Image Processing, vol. 3, (1997), pp. 288–301.

[218] Y. Kim, M.S. Nadar, A. Bilgin, Wavelet-based compressed sensing using a gaussian scale mixture model, IEEE Trans. Image Process. 21 (6) (2012) 3102–3108.

[219] M.S. Crouse, R.D. Nowak, R.G. Baraniuk, Wavelet-based statistical signal processing using hidden markov models, IEEE Trans. Sig. Process. 46 (4) (1998) 886–902.

[220] J. Tian, L. Chen, Adaptive multi-focus image fusion using a wavelet-based statistical sharpness measure, Sig. Process, 92 (9) (2012) 2137–2146.